NONCOMMUTATIVE CHRISTOFFEL-DARBoux KERNELS

SERBAN T. BELINSCHI, VICTOR MAGRON, AND VICTOR VINNIKOV

Abstract. We introduce from an analytic perspective Christoffel-Darboux kernels associated to bounded, tracial noncommutative distributions. We show that properly normalized traces, respectively norms, of evaluations of such kernels on finite dimensional matrices yield classical plurisubharmonic functions as the degree tends to infinity, and show that they are comparable to certain noncommutative versions of the Siciak extremal function. We prove estimates for Siciak functions associated to free products of distributions, and use the classical theory of plurisubharmonic functions in order to propose a notion of support for noncommutative distributions. We conclude with some conjectures and numerical experiments.

1. Introduction

The goal of this paper is to study the noncommutative analog of Christoffel-Darboux kernel associated to a certain class of distributions occurring in free probability, and to investigate the related asymptotic properties. In the classical commutative setting, one considers a finite signed Borel measure $\mu$ on $\mathbb{R}^n$. Integrating with respect to $\mu$ defines an inner product on the space of polynomials. The Christoffel-Darboux kernel [Sim08] is the reproducing kernel associated to the Hilbert space containing all polynomials up to a given degree. One way to compute this kernel is to rely on the elements of the orthonormal basis of this Hilbert space [DX14].

One appealing feature of the Christoffel-Darboux kernel and the related Christoffel function is their ability to capture some properties of $\mu$, such as the support or the density of its absolutely continuous component, from the only a priori knowledge of its moments. We refer the interested reader to [MN80, MNT91] for the univariate case. Recent research efforts [LP19] have focused on the multivariate case. When the measure $\mu$ is uniform or empirical and when the support of $\mu$ satisfies specific compactness conditions, the sequence of level sets of the Christoffel function associated to $\mu$ converges to its support with respect to the Hausdorff distance. The rate of convergence for estimating the support of a measure from a finite, independent, sample based on such empirical Christoffel-Darboux kernels is analyzed in [VBP19]. Current applications of the Christoffel function include statistical leverage scores [PBV18], sorting out typicality [PL16, LP19], and detection of outliers [BPSS20]. An extension to the case of certain singular measures is addressed in [PPL20], for instance when $\mu$ is the Hausdorff measure supported on the unit sphere. The framework from [MFH19] allows one to approximate attractors of dynamical systems, while relying on Christoffel-Darboux kernels associated to the...
moment matrix of the measure which is invariant with respect to the dynamics. Further applications [MPW+19] consider approximations of (possibly discontinuous) functions arising from weak (or measure-valued) solutions of optimal control problems or entropy solutions to non-linear hyperbolic PDEs. Note that in the two latter cases, the measure $\mu$ can be singular continuous.

One important motivation for the use of Christoffel-Darboux kernels is to extract the support of measures arising, e.g., in the above-mentioned dynamical systems or in a polynomial optimization problem (POP). In the latter case, one minimizes a polynomial over the intersection of finitely many level sets of polynomials, i.e., over a basic closed semialgebraic set. The minimizers of the POP belong to the support of atomic measures. Solving this problem is NP-hard in general [Lau09]. Lasserre’s hierarchy [Las01] is a well established framework to approximate the value of POPs. This methodology consists of approximating the optimal value of the initial POP by considering a hierarchy of semidefinite programs [AL12], involving moment matrices of growing sizes. By Putinar’s Positivstellensatz [Put93], if the quadratic module generated by the polynomials describing the semialgebraic set is archimedean, the hierarchy of semidefinite bounds converges from below to the minimum of the polynomial over this semialgebraic set. A somehow more delicate problem is to compute or at least approximate the minimizers of the POP. For this, one can extract the support of the atomic measure thanks to [HL05], which provides a numerical algorithm based on linear algebra. However, this procedure can be applied only if finite convergence occurs and assuming that the resulting obtained moment matrix is flat [CF98].

In the free noncommutative context, one can use sum of hermitian squares decompositions of positive polynomials [Hel02, McC01] to perform eigenvalue optimization of noncommutative polynomials over noncommutative semialgebraic sets. The noncommutative analogue of Lasserre’s hierarchy [HM04, NPA08, PNA10, CKP12, BCKP13] allow one to approximate as closely as desired the optimal value of such eigenvalue minimization problems. Further efforts [PNA10, CKP12, BCKP13] have been pursued to derive a hierarchy of semidefinite relaxations to optimize the trace of a given polynomial under positivity constraints. Sparsity exploiting hierarchies [KMP21, WM21] allow one to reduce the associated computational burden. The case of more general trace polynomials has been investigated in [KMV21]. Algorithms similar to the one from [HL05] allow one to extract optimizers of eigenvalue or trace minimization problems; see, e.g., [PNA10], [AL12, Chapter 21], [BKP16, Theorem 1.69] and [BCKP13]. As for the commutative case, when the extraction procedure fails (i.e., without flatness of the associated moment matrix), one still hopes to approximate such optimizers by considering them as elements in the support of a tracial noncommutative distribution. This support would be approximated by computing the levelsets of the noncommutative Christoffel-Darboux kernel associated to this distribution.

One of the equivalent definitions of the Christoffel-Darboux kernel is via a sum of orthonormal polynomials. This has been done before by Constantinescu [Con02] in the noncommutative context. Further explorations of systems of multivariate orthogonal and orthonormal noncommutative polynomials have been undertaken by several authors, for instance [Ans10, Ans08b, Ans08a, BC04]. We bring as novelty to this study the structure of operator spaces [Pau02] and the application...
of the classical analysis of plurisubharmonic functions [GZ17] – which we believe has never been considered before in the noncommutative context.

2. Preliminaries

2.1. Hermitian matrices, words and noncommutative polynomials. Given $k, n \in \mathbb{N}$, let us denote by $M_k(\mathbb{C})$ (resp. $S_k$) the space of all complex (resp. hermitian) matrices of order $k$, and by $S^S_k$ the set of $n$-tuples $\mathbf{A} = (A_1, \ldots, A_n)$ of hermitian matrices $A_i$ of order $k$. Let $I_k \in M_k(\mathbb{C})$ stand for the identity matrix.

For a fixed $n \in \mathbb{N}$, we consider a finite alphabet $\{X_1, \ldots, X_n\}$ and denote by $\langle \mathbf{X} \rangle$ the set of all possible words of finite length formed with the letters $X_1, \ldots, X_n$. The empty word is denoted by $1$. The length of a word $w$ is defined to be the number of letters (counted with repetition) that form $w$, and is denoted by $|w|$ (for example, $|X_1^2X_2X_1| = 7$). The length of the empty word is zero. For $d \in \mathbb{N}$, $\langle \mathbf{X} \rangle_d$ is the subset of all words of length at most $d$. We endow $\langle \mathbf{X} \rangle$ with the graded lexicographic order $\leq_{gl}$ and its strict version $<_{gl}$ (that is, $w <_{gl} w'$ iff $w \leq_{gl} w'$ and $w \neq w'$). The graded lexicographic order is defined as follows: first, $X_1 <_{gl} X_2 <_{gl} \cdots <_{gl} X_n$. Second, if $w, w' \in \langle \mathbf{X} \rangle$ are such that $|w| < |w'|$, then $w <_{gl} w'$. Third, if $|w| = |w'|$, then one applies the usual lexicographic order: if the first letter of $w$ is less than the first letter of $w'$, then $w <_{gl} w'$; if the first letters of $w$ and $w'$ coincide, one goes on to compare the second letters of $w$ and $w'$, and so on. For example, $X_n <_{gl} X_1^2 <_{gl} X_1X_2 <_{gl} X_2X_1 <_{gl} X_2^2 <_{gl} X_1^2 <_{gl} X_1^3 <_{gl} X_1^5X_n <_{gl} X_1^5X_nX_1$.

The set $\langle \mathbf{X} \rangle$ is in fact a monoid, known as the free monoid with $n$ generators, where the multiplication is the juxtaposition of words and the neutral element is the empty word $1$. Thus, the complex vector space spanned by it, which we denote by $\mathbb{C}(\langle \mathbf{X} \rangle)$, has a complex algebra structure. This algebra is known also as the algebra of polynomials in noncommutative indeterminates $\mathbf{X} = (X_1, \ldots, X_n)$ (that is, if we view the letters as indeterminates). We denote by $W_d(\langle \mathbf{X} \rangle)$ the vector of all words of $\langle \mathbf{X} \rangle_d$ ordered w.r.t. $\leq_{gl}$. The dimension of $\mathbb{C}(\langle \mathbf{X} \rangle)_d$ equals the length of $W_d(\langle \mathbf{X} \rangle)$, which is $\sigma(n, d) := \sum_{i=0}^{d} n^i = \frac{n^{d+1} - 1}{n-1}$. We equip the set $\mathbb{C}(\langle \mathbf{X} \rangle)$ with the involution $\star$ that conjugates the elements of $\mathbb{C}$, fixes $\{X_1, \ldots, X_n\}$ pointwise and reverses words, in such a way that $\mathbb{C}(\langle \mathbf{X} \rangle)$ is the $\star$-algebra freely generated by $n$ self-adjoint letters $X_1, \ldots, X_n$. We will systematically identify the set of words in $n$ letters and the set of monomials in $n$ self-adjoint indeterminates, and use this identification in notations as well. It will sometimes be convenient to write a word as $\mathbf{X}^w$ instead of just $w$; that is, if $w$ has $p$ letters, we let $w = \ell_1 \ell_2 \ell_3 \cdots \ell_p$, and we identify it with the degree $p$ monomial $\mathbf{X}^w = X_{\ell_1}X_{\ell_2}X_{\ell_3} \cdots X_{\ell_p}$. The set of all self-adjoint elements of $\mathbb{C}(\langle \mathbf{X} \rangle)$ is defined as $\text{Sym}\mathbb{C}(\langle \mathbf{X} \rangle) := \{f \in \mathbb{C}(\langle \mathbf{X} \rangle): f = f^*\}$.

2.2. Tracial functionals and moment matrices. Let $\tau: \mathbb{C}(\langle \mathbf{X} \rangle) \to \mathbb{C}$ be a positive tracial functional, that is, a $\mathbb{C}$-linear map such that $\tau(f^*f) \geq 0$ and $\tau(fg) = \tau(gf)$ for all $f, g \in \mathbb{C}(\langle \mathbf{X} \rangle)$. The functional $\tau$ is called faithful if $\tau(f^*f) = 0$ implies $f = 0$. Let $M_d(\tau)$ be the moment matrix of $\tau$, i.e., the symmetric matrix indexed by words of $\langle \mathbf{X} \rangle_d$ ordered according to $\leq_{gl}$, which is defined by $M_d(\tau)_{w,v} := \tau(w^*v)$. Since $\tau$ is positive, the moment matrix $M_d(\tau)$ is positive semi-definite. If, in addition, $\tau$ is faithful, then $M_d(\tau)$ is positive definite, and thus invertible. As some of the applications we envision are to the theory of von Neumann algebras [Tak79, Chapter 5], we consider a special subcategory of the set of such traces, which we shall call bounded traces. These are traces $\tau$ that satisfy the additional condition
that there exists a real number $M > 0$ such that $|\tau (w)| < M^d$ for all $d \in \mathbb{N}$, $w \in (X)_d$. If $M$ is known/given, we say that $\tau$ is bounded by $M$. If in addition $\tau (1) = 1$, then $\tau$ is called a \textit{bounded tracial state}. Bounded tracial states correspond to noncommutative probability distributions; see [AGZ10, Proposition 5.2.14 (d)] for more details about how they appear in the free probability literature.

2.3. \textbf{Plurisubharmonic functions.} One of the main tools we use in our paper is the theory of plurisubharmonic functions. We introduce here the basic notions necessary for our purposes. The main references we use are [GZ17, Kli91, ST97].

There are many equivalent characterizations of plurisubharmonic functions. We choose here the classical path: given a domain $G \subseteq \mathbb{C}$, a function $f : G \rightarrow [-\infty, +\infty)$ is \textit{subharmonic} if it is not identically equal to $-\infty$, it is upper semicontinuous, and it satisfies the sub-mean value inequality: $f(z) \leq (2\pi)^{-1} \int_{-\pi}^{\pi} f(z + re^{\theta}) d\theta$ for any $z \in G$ and $r > 0$ such that the closed disk of center $z$ and radius $r$ is included in $G$.

Given a domain $D \subseteq \mathbb{C}^n$, a function $u : D \rightarrow [-\infty, +\infty)$ is \textit{plurisubharmonic} if it is upper semicontinuous, not identically $-\infty$, and for any $z \in D$ and $b \in \mathbb{C}^n$, the function $\zeta \mapsto u(z + \zeta b)$ is subharmonic or identically $-\infty$ on each component of $\{ \zeta \in \mathbb{C} : z + \zeta b \in D \}$.

An extremely useful characterization of plurisubharmonic functions is given in terms of their second derivatives: essentially, $u$ is plurisubharmonic if and only if for any $\xi \in \mathbb{C}^n$, $\langle \mathcal{L}_u (\xi), \xi \rangle := \sum_{1 \leq j,k \leq n} \xi_j \xi_k \frac{\partial^2 u}{\partial z_j \partial z_k} \geq 0$ in the sense of (Schwartz) distributions – hence $\langle \mathcal{L}_u (\xi), \xi \rangle$ is a positive measure on $\mathbb{C}^n$. For the precise statement and proof of this result, see, for instance, [GZ17, Proposition 1.43], or [Kli91, Section 2.9].

Plurisubharmonic functions behave well with respect to taking limits. Two such results will be implicitly used later in our paper (see [GZ17, Propositions 1.28, 1.39, and 1.40]):

1. Let $\{u_k\}_{k \in \mathbb{N}}$ be a decreasing sequence of plurisubharmonic functions in $D$. If $\lim_{k \rightarrow \infty} u_k$ is not identically $-\infty$, then it is plurisubharmonic.

2. Let $(u_i)_{i \in I}$ be a family of plurisubharmonic functions in a domain $D$, which is locally uniformly bounded from above, and let $u = \sup_{i \in I} u_i$. Then the upper semicontinuous regularization $u^*$ of $u$ is plurisubharmonic in $D$ ($u^*$ can be obtained as $u^*(z) = \limsup_{\zeta \rightarrow z} u(\zeta)$).

The set $\{ u^* > u \}$ is generally small – always of zero Lebesgue measure. In cases of interest to us, this result can be strengthened. For this, we introduce the notion of pluripolar sets. A set $F \subseteq \mathbb{C}^n$ is called pluripolar if if for all $a \in F$ there exists a neighborhood $W$ of $a$ in $\mathbb{C}^n$ and a function $v$ which is plurisubharmonic on $W$ such that $F \cap W \subseteq \{ v = -\infty \}$ (see [ST97, Definition 1.4, Appendix B.1] and comments following it). It is known that the set on which a plurisubharmonic function is equal to $-\infty$ cannot be large. For instance, it has volume zero both in $\mathbb{C}^n$ and in any (maximal dimension) Euclidean sphere or torus included in $\mathbb{C}^n$ – see [GZ17, Proposition 1.34], and, for a much deeper understanding, [GZ17, Section 4.4]. A property is said to hold \textit{quasi-everywhere} on a set $S$ if it holds on $S \setminus F$ for a pluripolar set $F$. Theorem 1.7 of [ST97, Appendix B.1], due to Bedford and Taylor, states that if the functions $u_i$, $i \in I$, from (2) above are defined on all of $\mathbb{C}^n$ and have at most logarithmic growth at infinity, then the set $\{ u^* > u \}$ is actually pluripolar. Moreover, according to Theorem 1.6 from the same reference, $u^*$ itself has at most logarithmic growth at infinity.
3. Noncommutative Christoffel-Darboux Kernels

3.1. Orthonormal polynomials. Assume the linear functional $\tau: \mathbb{C}(\langle X \rangle) \to \mathbb{C}$ is positive. Then $(u, v) \mapsto \tau(v^* u)$ defines a positive sesquilinear form on $\mathbb{C}(\langle X \rangle)$ which transforms it into a pre-Hilbert space. Completing it with respect to the seminorm $\|u\| = \tau(u^* u)^{\frac{1}{2}}$ and factoring out the kernel $\{\|u\| = 0\}$ yields a Hilbert space which we denote by $L^2(\tau)$. Although most of the following statements still hold under weaker hypotheses, we assume from now on that $\tau$ is a faithful tracial state. In that case $\langle X \rangle$ is linearly independent and spans $L^2(\tau)$ as Hilbert space. For all $d \in \mathbb{N}$, let us define the family of orthonormal polynomials, $\{P_w\}_{w \in \langle X \rangle_d}$, ordered according to the lexicographic order, and satisfying for all $v, w \in \langle X \rangle_d$:

$$\tau(P_v^* P_w) = \delta_{v=w}, \quad \tau(P_v^* w) = 0, \text{ if } w <_{gl} v, \quad \tau(P_w^* w) > 0, \quad P_1 = 1.$$ (3.1)

Of course, span$\{w: w <_{gl} v\} = \text{span}\{P_w: w <_{gl} v\}$ for all $v \in \langle X \rangle$. Such a family can be constructed easily by using the classical Gram-Schmidt orthonormalization process applied to the basis of monomials $\langle X \rangle$. Written in the language of the $\tau$-induced inner product, it is given recursively by the initial condition $P_1(\langle X \rangle) = 1 \in \mathbb{C}$ and the general expression

$$P_w(\langle X \rangle) = \frac{w - \sum_{v <_{gl} w} \tau(P_v^* (\langle X \rangle) w) P_v(\langle X \rangle)}{\tau(w^* w) - \sum_{v <_{gl} w} |\tau(P_v^* w)|^2}^{\frac{1}{2}}, \quad w \in \langle X \rangle_d.$$ It is immediately clear from the construction that $\tau(P_w^* P_w) = 1$, $\tau(P_v^* P_w) = 0$ if $v <_{gl} w$, and

$$\tau(w^* P_w) = \frac{\tau(w^* w) - \sum_{v <_{gl} w} \tau(P_v^* (\langle X \rangle) w) \tau(w^* P_v(\langle X \rangle))}{\tau(w^* w) - \sum_{v <_{gl} w} |\tau(P_v^* w)|^2} > 0.$$ Of course, since $\tau(P_w^* w) = \tau(w^* P_w)$, we also have $\tau(P_w^* w) > 0$. Finally, if $v_0 <_{gl} w$, then $v_0$ is a linear combination of elements $P_v, v \leq_{gl} v_0$, so that $\tau(P_w v_0) = \sum_{v \leq_{gl} v_0} \tau(P_w P_v) = 0$. In the context, we should emphasize that faithfulness of $\tau$ automatically implies the denominator in the expression of $P_w(\langle X \rangle)$ above is nonzero. If $\tau$ is not faithful, it may well happen that $\tau(w^* w) = \sum_{v <_{gl} w} |\tau(P_v^* w)|^2$; in that case, we may perform the orthonormalization procedure above and the vast majority of the analysis that follows below on a quotient space, as in [BPSS20]: one “drops” in the Gram-Schmidt process any monomial $w$ which is a linear combination of elements $v <_{gl} w$ in the pre-Hilbert space induced by $\tau$.

**Remark 3.1.** (1) Since $\tau(P) = \overline{\tau(\langle P \rangle)}$, it follows immediately from the above that the family $\{P_w\}_{w \in \langle X \rangle_d}$ is itself an orthonormal basis for $\mathbb{C}(\langle X \rangle)_d$. Moreover, $\tau(P_w^* w) > 0$ if $\tau$ is faithful, and $\tau(P_w^* w) = 0$ whenever $w <_{gl} v$. However, generally $P_w \neq P_v$, and it is not even clear whether $P_v \in \{P_w: w \in \langle X \rangle_d\}$. It is nevertheless clear that span$\{P_w: w \in \langle X \rangle_d \setminus \langle X \rangle_{d-1}\} = \text{span}\{P_w^*: w \in \langle X \rangle_d \setminus \langle X \rangle_{d-1}\}$, $d \in \mathbb{N}$.

(2) The fact that both $\{P_w\}_{w \in \langle X \rangle_d}$ and $\{P_w^*\}_{w \in \langle X \rangle_d}$ are orthonormal bases implies that any correspondence sending one to the other is a unitary transformation of $L^2(\mathbb{C}(\langle X \rangle)_d, \tau)$. Thus, there exists a unitary matrix $U = (U_{v,w})_{v,w \in \langle X \rangle_d}$
such that $P_w(\mathcal{X}) = \sum_{v \in \langle \mathcal{X} \rangle} U_{v \omega, w} P_v(\mathcal{X})$. (As seen above, $U$ leaves $\text{span}\{P_w : w \in \langle \mathcal{X} \rangle_d \setminus \langle \mathcal{X} \rangle_{d-1}\}$ invariant for all $d$, so it is block-diagonal.) The matrix $U$ being unitary is equivalent to stating that $UU^* = U^*U = I_{\sigma(n,d)}$. In particular, $P_{v_0}(\mathcal{X}) = \sum_{v \in \langle \mathcal{X} \rangle} (U^*)_v w P_v(\mathcal{X}) = \sum_{v \in \langle \mathcal{X} \rangle} U_{v \omega, w} P_v(\mathcal{X})$, and $\delta_{v = v_0} = \sum_{v \in \langle \mathcal{X} \rangle_d} U_{v \omega, w}$. Moreover, the orthonormality relations yield $\tau(P_{v_0}(\mathcal{X}) P_{v_0}^*(\mathcal{X})) = \sum_{v \in \langle \mathcal{X} \rangle_d} U_{v \omega, v} \tau(P_v(\mathcal{X}) P_v^*(\mathcal{X})) = U_{v_0, v_0}$. Under the assumption of traciality for $\tau$, we obtain that $U_{v_0, v_1} = U_{v_1, v_0}$ for all $v_0, v_1 \in \langle \mathcal{X} \rangle_d$, so that the unitary matrix $U$ (and necessarily its adjoint $U^*$ too) is symmetric. This simple observation will be useful in the definition of the Christoffel-Darboux kernel.

Polynomials $P_w$ are usually not selfadjoint, and the construction offered above cannot be expected to provide selfadjoint orthonormal bases from a non-selfadjoint set of monomials. While the orthonormal basis $\{P_w : w \in \langle \mathcal{X} \rangle\}$ turns out to be sufficient for our purposes, we would like to specify as an aside that it is quite easy to create a family of selfadjoint orthonormal polynomials out of the hermitization of our monomials. This follows directly from the nature of the Gram-Schmidt process, if we place a convenient modification of the lexicographic order on the set of monomials. We proceed the following way: when we see a selfadjoint element, we leave it alone; when we see a non-selfadjoint monomial $X^w$, we replace it with its real part $\Re X^w = \frac{X^w + X^{w^*}}{2}$. At the same time, we replace $X^{w^*}$ (which, according to this procedure, must necessarily be greater than $X^w$) with the the imaginary part of $X^w$, namely $\Im X^w = \frac{X^w - X^{w^*}}{2i}$, to be placed in our ordered list in the slot previously occupied by $X^{w^*}$ (of course, the first element in the list is selfadjoint, namely $X^1$). This process yields a basis of $\text{Span}_C(\langle \mathcal{X} \rangle_d \setminus \langle \mathcal{X} \rangle_{d-1})$ which is entirely composed of selfadjoint elements, for each $d \in \mathbb{N}$ (the case $d = 0$ corresponds to the selfadjoint 1 and the case $d = 1$ to the selfadjoint $X_1, \ldots, X_n$). Thus, our leave it ordered according to the same lexicographic order, where the position of selfadjoint monomials remains unchanged, and if $w^{\text{gl}} \succ w$, then the position of $\Re X^w$ is at $w$, and the position of $\Im X^w$ is at $w^*$. The Gram-Schmidt procedure applied to this basis yields a family $\{S_w : w \in \langle \mathcal{X} \rangle\} \subset \text{SymC}(\mathcal{X}) \subset \mathbb{C}(\mathcal{X})$ of selfadjoint orthogonal polynomials. This is a known fact (it can be viewed as a reformulation of the general fact that the Gram-Schmidt procedure applied to a basis in a real Hilbert space with respect to the real inner product or with respect to its complexification yields the same result). However, it can also easily be argued by induction after the ordered set of words. With the convention, valid only in this paragraph, that $w \in \langle \mathcal{X} \rangle$ denotes not the monomial $X^w$, but the selfadjoint basis element it indexes according to the above procedure, the general formula of the $w_0^{th}$ orthonormal polynomial is, as before,

$$S_{w_0}(\mathcal{X}) = \frac{w_0 - \sum_{w^{\text{gl}} \succ w_0} \tau(S_v(\mathcal{X})w_0) S_v(\mathcal{X})}{\left(\tau(w_0^\ast w_0) - \sum_{w^{\text{gl}} \succ w_0} |\tau(S_v(\mathcal{X})w_0)|^2\right)^{1/2}}, \quad w_0 \in \langle \mathcal{X} \rangle_d.$$
By hypothesis, \( u_0 = w_0^* \), \( S_\tau(X) = S_{\tau}^*(X) \). Moreover, \( \tau(S_\tau(X)u_0) = \tau((S_\tau(X)w_0)^*) = \tau(w_0 S_{\tau}^*(X)) = \tau(w_0 S_{\tau}(X)) = \tau(S_\tau(X)w_0) \), so that \( \tau(S_\tau(X)w_0) \in \mathbb{R} \). Thus, \( S_{w_0}(X) = S_{w_0}^*(X) \). As \( S_1(X) = 1 \) is self-adjoint, this completes our argument.

Before going forward to the study of the Christoffel-Darboux kernel, let us list a few simple or (by now) well-known examples of orthogonal polynomials:

**Example 3.2.**

1. It is well-known, and easy to check, that any compactly supported Borel probability measure on \( \mathbb{R} \) whose support contains an infinity of points admits a family of orthonormal polynomials from \( C(X) = C[\mathbb{R}] \). The state \( \tau \) corresponding to it is simply the integration of the polynomial with respect to the given probability measure. \( \tau \) is in fact faithful if and only if the support of the corresponding probability measure is an infinite set. Otherwise, the number of linearly independent monomials equals the number of points in the support. The reader can find a vast and fascinating literature on the subject by searching in [DX14] and references therein.

2. Multivariate orthogonal polynomials are another obvious classical example, to which reference [DX14] is mainly dedicated. Given a Euclidean space \( \mathbb{R}^n \), a compactly supported Borel probability measure on it admits a family of orthonormal polynomials from \( C[X] = C(\mathbb{R})/\langle X_iX_j = X_jX_i : 1 \leq i, j, \leq n \rangle \) with respect to the trace \( \tau \) defined by the integration with respect to the given probability measure. If the support of our probability is not concentrated on any finite union of algebraic curves (for instance if its interior in \( \mathbb{R}^n \) is nonempty), then the trace is faithful on \( C[X] \). However, it is quite clear that such a trace does not satisfy our condition of faithfulness: since \( -(X_1X_2 - X_2X_1)^2 = (X_1X_2 - X_2X_1)(X_2X_1 - X_1X_2) = (X_2X_1 - X_1X_2)^*(X_2X_1 - X_1X_2) \) is a positive polynomial in \( C(X) \setminus \{0\} \) but it is the zero polynomial in \( C[X] \), we have \( \tau(-(X_1X_2 - X_2X_1)^2) = \tau(0) = 0 \), so \( \tau \) is far from being faithful on \( C(X) \) whenever \( n \geq 2 \). Thus, the noncommutative version of orthogonal polynomials differs drastically from the classical one in more than one variable.

3. An example from [Ans10, Section 3] will be useful later in the paper. Assume that \( \tau \) is bounded and the components \( X_1, \ldots, X_n \) of \( X \) are free [Voi85] with respect to \( \tau \). Assume moreover that \( \tau \) is faithful, which, in this case, is equivalent to requiring that the - classical - distribution of \( X_j \) with respect to \( \tau \) has infinite support in \( \mathbb{R} \). We denote by \( P_k^{(j)}(X_j), k \in \mathbb{N} \), the orthogonal polynomials associated to the distribution of \( X_j \) with respect to \( \tau \). Given \( w \in (X)_d \), we write it as \( w = (w_1, w_2, \ldots, w_d) \). We partition \( w \) in intervals given by consecutively repeating indices: \( w = (\pi_1, \pi_2, \ldots, \pi_\ell) \) for some \( \ell \in \{1, \ldots, d\} \). For example, if \( n = 3 \), \( d = 10 \), and \( w = (2, 2, 1, 3, 1, 1, 1, 1) \), then \( w = (\pi_1, \pi_2, \pi_3, \pi_4) \) where \( \pi_1 = (2, 2, 2), \pi_2 = (1), \pi_3 = (3, 3), \pi_4 = (1, 1, 1, 1) \). We associate to each \( \pi_j \) the orthonormal polynomial \( P_{\pi_j}^{(\ell_j)}(X_j) \), where \( |\pi_j| \) is the cardinality of the block \( \pi_j \), and \( \ell_j \in \{1, \ldots, n\} \) is the index that is contained by \( \pi_j \). Then \( P_w(X) = P_{\pi_1}^{(\ell_1)}(X_1)P_{\pi_2}^{(\ell_2)}(X_2) \cdots P_{\pi_\ell}^{(\ell_\ell)}(X_\ell), w \in (X), \) is a system of orthonormal polynomials for \( \tau \). (For example, with \( w = (2, 2, 1, 3, 3, 1, 1, 1) \), we would obtain \( P_w(X_1, X_2, X_3) = P_3^{(2)}(X_2)P_1^{(1)}(X_1)P_2^{(3)}(X_3)P_4^{(1)}(X_1). \) This
follows from the definition of free independence and the condition of orthonormality imposed on each family $P_k^{(j)}(X_j), 1 \leq j \leq n, k \in \mathbb{N}$. The case when, say, $\mathbb{C}\langle X_1, \ldots, X_p \rangle$ and $\mathbb{C}\langle X_{p+1}, \ldots, X_n \rangle$ are free with respect to $\tau$ is treated conceptually precisely the same way; and the orthonormal polynomials associated to $\tau$ on $\mathbb{C}\langle X \rangle$ are again alternating products of orthonormal polynomials corresponding to the restrictions of $\tau$ to $\mathbb{C}\langle X_1, \ldots, X_p \rangle$ and $\mathbb{C}\langle X_{p+1}, \ldots, X_n \rangle$, respectively, but the notation becomes more involved. Faithfulness of $\tau$ on either of $\mathbb{C}\langle X_1, \ldots, X_p \rangle$ or $\mathbb{C}\langle X_{p+1}, \ldots, X_n \rangle$ is not necessary for the above result to hold.

(4) Several other interesting examples can be found in, or adapted from, [Ans10, Ans08b], including free Meixner laws or orthogonal polynomials corresponding to other noncommutative independences.

3.2. Operator space structure and the Christoffel-Darboux kernel. For any two vector spaces $V, W$ over $\mathbb{C}$, the algebraic tensor product vector space over $\mathbb{C}$, $V \otimes W$, is well-defined. If $V, W$ are complex algebras, this vector space has itself two algebra structures: $V \otimes W$ with the multiplication $(\sum_i v_i \otimes w_i)(\sum_j v_j \otimes w_j) = \sum_i \sum_j v_i v_j \otimes w_i w_j$, and $V \otimes W^{op}$ with the multiplication $(\sum_i v_i \otimes w_i)(\sum_j v_j \otimes w_j) = \sum_i \sum_j v_i v_j \otimes w_j w_i$. This second multiplication is important because it allows in some circumstances the identification $V \otimes V^{op} \simeq \mathcal{L}(V, V)$, the space of linear operators from $V$ to itself, via $\sum_i v_i \otimes \tilde{v}_i \mapsto [x \mapsto \sum_i v_i x \tilde{v}_i]$. This identification works well for the space $M_k(\mathbb{C})$ and, when $V$ is a von Neumann algebra, there are completions of the range of $V \otimes V^{op}$ in $\mathcal{L}(V, V)$ that are quite important in the theory of operator spaces (we refer to [ER00, Pau02, Pis03] as fundamental sources for operator spaces theory, and as an example of the application of the usefulness of this view of $V \otimes V^{op}$ in free probability, to the now-classical paper [GS14]).

Let us apply the above to $V = M_k(\mathbb{C})$ and $W = \mathbb{C}\langle X \rangle$. We have the natural identification $M_k(\mathbb{C}) \otimes \mathbb{C}\langle X \rangle \simeq M_k(\mathbb{C}\langle X \rangle)$. This is a star algebra with the adjoint operation $[(P_{ij}(\mathbb{C}\langle X \rangle))_{i,j=1}^k]^* = (P_{ij}^*(\mathbb{C}\langle X \rangle))_{i,j=1}^k$. Writing $(P_{ij}(\mathbb{C}\langle X \rangle))_{i,j=1}^k$ as $(P_{ij}(\mathbb{C}\langle X \rangle))_{i,j=1}^k = \sum_w \alpha_{ij}^{(w)} P_w(\mathbb{C}\langle X \rangle)$, yields

$$
(P_{ij}(\mathbb{C}\langle X \rangle))_{i,j=1}^k = \left( \sum_w \alpha_{ij}^{(w)} P_w(\mathbb{C}\langle X \rangle) \right)_{i,j=1}^k = \sum_{i,j=1}^k \sum_w \alpha_{ij}^{(w)} e_{ij} \otimes P_w(\mathbb{C}\langle X \rangle)
$$

$$
= \sum_{i,j=1}^k \left( \sum_{i,j=1}^k \alpha_{ij}^{(w)} e_{ij} \right) \otimes P_w(\mathbb{C}\langle X \rangle)
$$

$$
= \sum_w C_w((P_{ij}(\mathbb{C}\langle X \rangle))_{i,j=1}^k) \otimes P_w(\mathbb{C}\langle X \rangle),
$$

where $C_w((P_{ij}(\mathbb{C}\langle X \rangle))_{i,j=1}^k) \in M_k(\mathbb{C}), C_w((P_{ij}(\mathbb{C}\langle X \rangle))_{i,j=1}^k) = \sum_{i,j=1}^k \alpha_{ij}^{(w)} e_{ij}$. Written in this form, the star operation is simply $(\sum_w C_w((P_{ij}(\mathbb{C}\langle X \rangle))_{i,j=1}^k) \otimes P_w(\mathbb{C}\langle X \rangle))^* = \sum_w C_w((P_{ij}(\mathbb{C}\langle X \rangle))_{i,j=1}^k)^* \otimes P_w^*(\mathbb{C}\langle X \rangle)$. We define an $M_k(\mathbb{C})$-valued sesquilinear form $\langle (P_{ij}(\mathbb{C}\langle X \rangle))_{i,j=1}^k, Q_{ij}(\mathbb{C}\langle X \rangle))_{i,j=1}^k \rangle = (\text{tr}_{M_k(\mathbb{C})} \otimes \tau)((P_{ij}(\mathbb{C}\langle X \rangle))_{i,j=1}^k)^* = \sum_w C_w((P_{ij}(\mathbb{C}\langle X \rangle))_{i,j=1}^k)^* C_w((Q_{ij}(\mathbb{C}\langle X \rangle))_{i,j=1}^k)^*$. With these notations, if $A \in M_k(\mathbb{C})$, $P, Q \in M_k(\mathbb{C}\langle X \rangle)$, then $\langle PA, Q \rangle = \langle P, QA^* \rangle, \langle AP, Q \rangle = A \langle P, Q \rangle, \langle P, AQ \rangle = \langle P, QA^* \rangle$. This makes $M_k(\mathbb{C}\langle X \rangle)$ into a $M_k(\mathbb{C})$-Hilbert module (see [Pau02, Chapter 14]). From now on, we denote matrix-valued polynomials by boldface letters.
Remark 3.3. For fixed $0 \neq k \in \mathbb{N}$, a polynomial $P \in \mathcal{M}_k(\mathbb{C}[X])$, $P(X) = \sum_{w \in \langle X \rangle} c_w \otimes P_w(X)$ is a noncommutative function [KVV14, Section 2.1]. Indeed, consider an arbitrary unital $C^*$-algebra $A$ (we will mostly need the case $A = M_l(\mathbb{C})$ for some $l$, possibly different from $k$) and an $\mathcal{M}_k(\mathbb{C})$-$A$-bimodule $B$ (usually $\mathcal{M}_{k \times l}(\mathbb{C})$). One defines the function

$$A^n \times B \ni (a, b) \mapsto P(a, b) = \sum_{w \in \langle X \rangle} c_w b P_w(a).$$

Amplification to $\iota \times \iota$ matrices is done the obvious way: if $(a, b) \in (\mathcal{M}_k(A))^n \times \mathcal{M}_k(B)$, then

$$P(a, b) = \sum_{w \in \langle X \rangle} (c_w \otimes I_c) b P_w(a) = \sum_{w \in \langle X \rangle} [\operatorname{diag}(c_{w_0}, \ldots, c_{w_l})] b P_w(a).$$

For reasons that will become clear shortly, we prefer to write $P(a)/(b)$ instead of $P(a, b)$. Indeed, evaluation of polynomials as above is essential in the rest of the paper. For a $P(X) \in \mathcal{M}_k(\mathbb{C}[X])$, the evaluation in an $n$-tuple of matrices $A \in \mathcal{M}_k(\mathbb{C})^n$ allows us, as already mentioned, to view $P$ as a linear map on $\mathcal{M}_k(\mathbb{C})$: if $P(X) = \sum_{w \in \langle X \rangle} c_w \otimes X^w$, then $P(A) = \sum_{w \in \langle X \rangle} c_w \otimes A^w$: $\mathcal{M}_k(\mathbb{C}) \to \mathcal{M}_k(\mathbb{C})$, $P(A)(C) = \sum_{w \in \langle X \rangle} c_w A^w C$. In this context, recall that the star operation on $\mathbb{C}[X]$ extends to $\mathcal{M}_k(\mathbb{C}[X])$ the obvious way: if, say, $P(X) = \sum_{w \in \langle X \rangle} c_w \otimes X^w$, then $P^*(X) = \sum_{w \in \langle X \rangle} c_w^* \otimes X^{w^*} = \sum_{w \in \langle X \rangle} c_w^* \otimes \overline{X}^w$. Thus, when performing evaluations on possibly non-selfadjoint tuples, $P(A)^* = \left(\sum_{w \in \langle X \rangle} c_w \otimes A^w\right)^* = \sum_{w \in \langle X \rangle} c_w^* \otimes (A^*)^w = \sum_{w \in \langle X \rangle} c_w^* \otimes (A^*)^w = P^*(A^*)$.

It is useful to record this equality, together with the one corresponding to scalar polynomials:

(3.2) $P(A)^* = P^*(A^*)$, $P(A)^* = P^*(A^*)$, $A \in \mathcal{M}_k(\mathbb{C})^n, k \in \mathbb{N}$.

It is remarkably convenient that this star operation coincides with the adjoint operation when $\mathcal{M}_k(\mathbb{C})$ is viewed as a Hilbert space with the Hilbert-Schmidt norm. Indeed, with $\Tr$ denoting the non-normalized trace,

$$\langle C, P(A)(D) \rangle_\text{HS} = \Tr_k \left( \sum_{w \in \langle X \rangle} c_w D A^w \right)^* C = \Tr_k \left( \sum_{w \in \langle X \rangle} (A^w)^* D^* c_w^* C \right)$$

$$= \Tr_k \left( D^* \sum_{w \in \langle X \rangle} c_w^* C (A^*)^w \right) = \Tr_k (D^* P^*(A^*)(C)) = \langle P^*(A^*)(C), D \rangle_\text{HS}, \quad C, D \in \mathcal{M}_k(\mathbb{C}), A \in \mathcal{M}_k(\mathbb{C})^n, k \in \mathbb{N}.$$

It should be emphasized that re-normalizing $\Tr$ changes nothing in the above. Also, the same algebraic computations apply equally well for a tracial state on a $C^*$-algebra. In particular, for a $\Pi_1$ factor $\mathcal{M}$ with normal faithful trace-state $\operatorname{tr}$, if $P(X) \in \mathcal{M} \otimes \mathbb{C}[X]$, then $[P(a)]^*(c) = P^*(a^*)(c)$ for any $a \in \mathcal{M}^n, c \in L^2(\mathcal{M}, \operatorname{tr})$.

When viewed as an operator on the Hilbert space $\mathcal{M}_k(\mathbb{C})$, the norm of $P(A)$ is defined the usual way: $\|P(A)\| = \sup\{\|P(A)(C)\|_\text{HS}: \|C\|_\text{HS} = 1\}$. This is essentially the same norm as the $C^*$-norm of $P(A)$ when viewed as acting on $\mathbb{C}^n \otimes \mathbb{C}^k$. 
Indeed, with \( P(A) = \sum_{w \in (X)} c_w \otimes A^w \), the following two quantities are equal:

\[
\sup_{\|C\|_{\text{HS}}=1} \|P(A)(C)\|_{\text{HS}}^2 = \sup_{\|C\|_{\text{HS}}=1} \text{Tr}_k \left( \sum_{v, w \in (X)} (A^v)^* C^* e_v e_w C A^w \right),
\]

(3.3)

\[
\sup_{i,j=1}^k \|P(A)\|_{i,j} = \sup_{v, w \in (X)} \sum_{i,j=1}^k \alpha_{i,j} \langle c_v e_i, c_w e_j \rangle \langle A^v, A^w \rangle,
\]

(3.4)

where the supremum in the second equality is taken after all \( \alpha_{i,j}, 1 \leq i, j \leq k \) such that \( \sum_{i,j=1}^k |\alpha_{i,j}|^2 = 1 \). (As before, the result in (3.3) does not change regardless of whether one uses \( \text{Tr}_k \) or \( \text{tr}_k \) in defining the HS norm.) The first expression identifies with the second by simply splitting \( C \) into matrix units. This form identifies \( P(A) \) acting on \((M_k(\mathbb{C}), \langle \cdot, \cdot \rangle_{\text{HS}})\) as an element in the tensor product of von Neumann algebras \( M_k(\mathbb{C}) \otimes M_k(\mathbb{C})^{op} \) and \( P(A) \) acting on \((C^k \otimes C^k, \langle \cdot, \cdot \rangle)\) as an element of \( M_k(\mathbb{C}) \otimes M_k(\mathbb{C}) \). In this case it follows that \( P(A) \) has the same norm regardless of whether it is viewed in one or in the other. It should also be noted that the spectral properties of \( P(A) \) do not change regardless of whether it is viewed as acting on \( C^k \otimes C^k \), on \( (M_k(\mathbb{C}), \| \cdot \|_{\text{HS}}) \) or on \( (M_k(\mathbb{C}), \| \cdot \|) \). However, the notion of positivity as an operator on the Hilbert space \( (M_k(\mathbb{C}), \| \cdot \|_{\text{HS}}) \) is different from the notion of positivity as an operator on the Banach space \( (M_k(\mathbb{C}), \| \cdot \|) \) endowed with the order determined by the cone of positive semidefinite matrices.

For every \( d \in \mathbb{N} \), we define the bi-variate polynomial \( \kappa_{\tau, d} \) as follows

\[
\kappa_{\tau, d}(X, Y) := \sum_{w \in (X)_d} P_w(X) \otimes P_w^*(Y).
\]

(3.5)

If \( \tau \) is not faithful, then the above notation \( w \in (X)_d \) should be understood as running only through a subset of linearly independent monomials in the completion \( L^2(\tau) := L^2((X), \tau) \) of \( (C(X), \langle \cdot, \cdot \rangle) \) with respect to \( \|P\|_2 = \tau(P^* P)^{1/2} \) which span \( C(X)_d \) – we do not “repeat” any summand. Placing the adjoints on the first or on the second tensor in the above definition is a matter of choice. Indeed, as noted in Section 3.1, \( \{P_v^*\}_{v \in (X)_d} \) is an orthonormal basis as well, so that the correspondence between the elements of \( \{P_w\}_{w \in (X)_d} \) and \( \{P_v^*\}_{v \in (X)_d} \) is unitary. With the notations from Section 3.1,

\[
\sum_{w \in (X)_d} P_w(X) \otimes P_w^*(Y) = \sum_{v \in (X)_d} \left( \sum_{v \in (X)_d} U_{w,v}(U^*) \right)^* P_v^*(X) \otimes P_v(Y) + \sum_{v \in (X)_d} U_{w,v}(U^*) P_v^*(X) \otimes P_v(Y)
\]

(3.6)

(As shown in Remark 3.1, \( U \) is symmetric, so that \( U_{w,v}(U^*) = U_{v,w}(U^*) \)).
One may view \( \kappa_{\tau,d} \) as a reproducing kernel for the finite-dimensional space \( L^2(\mathbb{C}(\langle X \rangle)_d, \tau) \): we tautologically have

\[
(\text{Id}_{\mathbb{C}(\langle X \rangle)} \otimes \tau)(\kappa_{\tau,d}(X, Y)(1 \otimes P(Y))) = \sum_{v \in \langle X \rangle_d} (\text{Id}_{\mathbb{C}(\langle X \rangle)} \otimes \tau)(P_v(X) \otimes P_w^*(Y) P(Y)) = \sum_{v \in \langle X \rangle_d} P_v(X) \otimes \tau(P_v^*(Y) P(Y)) 1 = \sum_{v \in \langle X \rangle_d} \tau(P_v^*(Y) P(Y)) P_v(X) \otimes 1 = P(X) \otimes 1,
\]

for any \( P \in \mathbb{C}(\langle X \rangle)_d \). More remarkably, with the previously introduced Hilbert space \( L^2(\tau) := L^2(\mathbb{C}(\langle X \rangle), \tau) \) as the completion of \( \mathbb{C}(\langle X \rangle) \) with respect to \( \| P \|_2 = \tau(P^* P)^{\frac{1}{2}} \), by letting \( d \to \infty \) in the above, we may write the \( L^2 \)-limit

\[
(3.7) \quad \lim_{d \to \infty} (\text{Id}_{\mathbb{C}(\langle X \rangle)} \otimes \tau)(\kappa_{\tau,d}(X, Y)(1 \otimes f(Y))) = f(X) \otimes 1, \quad f \in L^2(\tau).
\]

In this (very weak) sense, one may state that \( \kappa_{\tau,d}(X, Y) = \lim_{d \to \infty} \kappa_{\tau,d}(X, Y) \) exists. The reproducing kernel \( \kappa_{\tau,d}(X, Y) \) is called the \textit{noncommutative Christoffel polynomial} associated to \( \tau \) and \( d \). Note that

\[
(\tau \otimes \tau)(\kappa_{\tau,d}(X, Y)^* \kappa_{\tau,d}(X, Y)) = \sum_{v, w \in \langle X \rangle_d} (\tau \otimes \tau)(P_v^*(X) P_w(X) \otimes P_w^*(Y) P_v(Y)) = \sum_{v, w \in \langle X \rangle_d} \tau(P_v^*(X) P_w(X)) \tau(P_w^*(Y) P_v(Y)) = \sum_{w \in \langle X \rangle_d} \tau(P_w^*(X) P_w(X)) \tau(P_w(Y) P_w^*(Y)) = \sum_{w \in \langle X \rangle_d} 1 \cdot 1 = \sigma(n, d),
\]

by orthonormality of the \( \{ P_w \}_{w \in \langle X \rangle_d} \).

The bi-variate polynomial \( \kappa_{\tau,d} \) is in fact a globally defined completely positive noncommutative kernel in the sense of [BMV16, Sections 2.3 and 2.4] (for exact definitions, which are not relevant for our current paper, we refer to conditions (2.3), (2.4), and (2.10) in this reference); indeed, for any strictly positive \( k, k' \in \mathbb{N} \), \( \kappa_{\tau,d} \) acts on \( M_{k \times k'}(\mathbb{C}) \) via \( M_{k}(\mathbb{C})^n \times M_{k'}(\mathbb{C})^n \to \mathcal{L}(M_{k \times k'}(\mathbb{C}), M_{k \times k'}(\mathbb{C})), (A, B) \mapsto X \mapsto \sum_{w \in \langle X \rangle_d} P_w(A) X P_w(B) \) and thus \( \kappa_{\tau,d}(A, A^*) \geq 0 \), i.e. it is completely positive, for all \( A \in M_k(\mathbb{C})^n \). Note that we obtain a different action on \( M_{k \times k'}(\mathbb{C}) \) for each choice of \( A \) and \( B \). We also note for future use that, since \( P_1(X) = 1 \) (see (3.1)), we have

\[
\kappa_{\tau,d}(A, A^*)(C) = \sum_{w \in \langle X \rangle_d} P_w(A) C P_w(A)^* \geq P_1(A) C P_1(A)^* = C
\]

whenever evaluated on elements of a unital \( C^* \)-algebra in which the coordinates of \( A \) are selfadjoint and \( C \geq 0 \).

For each \( A = (A_1, \ldots, A_n) \in M_k(\mathbb{C})^n \), let us define \( \Lambda_{\tau,d}(A) := \kappa_{\tau,d}(A, A^*)(I_k) \), as an operator in \( M_k(\mathbb{C}) \). The function \( \Lambda_{\tau,d} \) is called the \textit{noncommutative Christoffel function} associated to \( \tau \) and \( d \). We see from the above that \( \Lambda_{\tau,d} \) is well-defined on \( A^n \) for any unital \( C^* \)-algebra \( A \), and in particular it is well defined on all of \( M_k(\mathbb{C})^n \), \( k \in \mathbb{N} \), and \( \Lambda_{\tau,d}(A) \leq I_k, A \in M_k(\mathbb{C})^n \).
We now prove the noncommutative analog of [LP19, Theorem 3.1], which can be found in e.g. [DX14] and [Nev86]. We emphasize that the min appearing in (3.8) is a minimum of a matrix-valued quantity.

**Theorem 3.4.** Let \( \Lambda \in \mathbb{M}_k(\mathbb{C})^n \) be fixed, arbitrary. Then

\[
\Lambda_{\tau,d}(\Lambda) = \min \{ (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau)(PP^*) : P \in \mathbb{M}_k(\mathbb{C}(\mathcal{A}))_d, P(\Lambda)(I_k) = I_k \}.
\]

It is remarkable that the set in the right-hand side of (3.8) has a minimum, given that the set of positive operators is not a lattice. Moreover, under the assumption that \( \tau \) is bounded, it can be realized as the distribution of a tuple \( \mathcal{A}_w \) of selfadjoints in a von Neumann algebra with respect to a normal faithful trace \( \text{tr} \), and then \((\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau)(PP^*) = (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \text{tr})(P(\mathcal{A}_w)P(\mathcal{A}_w)\dagger), \) the conditional expectation of a positive element in a von Neumann algebra.

**Proof of Theorem 3.4.** Let us define \( c_w(\Lambda) = \Lambda_{\tau,d}(\Lambda)P_w(\Lambda) \in \mathbb{M}_k(\mathbb{C}) \), for all \( w \in (\mathcal{X})_d \) and \( P_w(\mathcal{X}) := \sum_{w \in (\mathcal{X})} c_w(\Lambda) \otimes P_w(\mathcal{X}) \in \mathbb{M}_k(\mathbb{C}(\mathcal{X}))_d \). (If \( \tau \) is not faithful, we again drop monomials from the set \( (\mathcal{X})_d \) so that it becomes a basis for \( L^2(\tau, C(\mathcal{X}_d)) \).) This polynomial \( P_w(\mathcal{X}) \) is feasible for (3.8) since one has

\[
P_w(\mathcal{X})(I_k) = \sum_{w \in (\mathcal{X})_d} c_w(\Lambda)I_kP_w(\mathcal{A})^\dagger = \Lambda_{\tau,d}(\Lambda)\sum_{w \in (\mathcal{X})_d} P_w(\mathcal{A})I_kP_w(\mathcal{A})^\dagger = \Lambda_{\tau,d}(\Lambda)\kappa_{\tau,d}(\Lambda, \mathcal{A}^\dagger)(I_k) = I_k.
\]

In addition, one has (we use the op algebra structure on the second tensor)

\[
P_w(\mathcal{X})P_w(\mathcal{X})^\dagger = \sum_{v,w \in (\mathcal{X})_d} c_w(\Lambda)c_v(\Lambda)^\dagger \otimes P_v(\mathcal{X})P_w(\mathcal{X}),
\]

so

\[
(\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau)(P_w(\mathcal{X})P_w(\mathcal{X})^\dagger) = \sum_{w \in (\mathcal{X})_d} c_w(\Lambda)c_w(\Lambda)^\dagger
\]

\[
= \sum_{w \in (\mathcal{X})_d} \Lambda_{\tau,d}(\Lambda)P_w(\mathcal{A})P_w(\mathcal{A})^\dagger \Lambda_{\tau,d}(\Lambda)
\]

\[
= \Lambda_{\tau,d}(\Lambda)\kappa_{\tau,d}(\Lambda, \mathcal{A}^\dagger)(I_k)\Lambda_{\tau,d}(\Lambda) = \Lambda_{\tau,d}(\Lambda),
\]

proving that the set in the right-hand side of (3.8) contains \( \Lambda_{\tau,d}(\Lambda) \).

To investigate the minimality of the element \( \Lambda_{\tau,d}(\Lambda) \), recall the Hilbert module structure on \( \mathbb{M}_k(\mathbb{C})^{(\mathcal{X}_d)\tau(n,d)} : \langle (a_w)_w, (b_w)_w \rangle_d = \sum_{w \in (\mathcal{X})_d} a_w b_w^* \).

Consider an arbitrary polynomial \( P(\mathcal{X}) = \sum_{w \in (\mathcal{X})_d} a_w \otimes P_w(\mathcal{X}) \) such that \( P(\Lambda)(I_k) = I_k \). By the definition of the \( \mathbb{M}_k(\mathbb{C}) \)-valued inner product, this last expression is equivalent to stating \( \langle (a_w)_w, (P_w(\mathcal{A})^\dagger)_w \rangle_d = I_k \). With the Hilbert module structure (see [Pau02, Chapter 14]), we are guaranteed the positivity

\[
\sum_{w \in (\mathcal{X})_d} a_w \otimes P_w(\mathcal{X})^\dagger \otimes P_w(\mathcal{X})^\dagger a_w^* \geq 0.
\]

As \( \langle (P_w(\mathcal{A})^\dagger)_w, (P_w(\mathcal{A})^\dagger)_w \rangle_d^{-1} = |\sum P_w(\mathcal{A})^\dagger P_w(\mathcal{A})|^{-1} = \Lambda_{\tau,d}(\Lambda) \) (see (3.6)) is known to exist, the positivity of the above matrix is equivalent, by considering
the Schur complement [HJ85, Theorem 7.7.6], or [Pau02, Lemma 3.1], to the relation
\[ I_k((P_w(A^*w))(P_w(A^*)w)^{-1}I_k \preceq \langle (a_w)w,(a_w)w \rangle_d = \sum_{w \in \mathcal{X}^d} a_w^*a_w^* = (\text{Id}_{M_k(C)} \otimes \tau)(PP^*), \]
as claimed.

We mention below some re-interpretations of Theorem 3.4; one remarkable point is that the normalization \( P(A)(I_k) = I_k \) is in fact quite arbitrary, even though useful.

Remark 3.5. (1) The positivity of the matrix in (3.9) imposes automatically the domination relation \( \langle (a_w)w,(P_w(A^*)w) \rangle_d \Lambda_{\tau,d}(A)((P_w(A^*))w,(a_w)w) \rangle_d \leq \langle (a_w)w,(a_w)w \rangle_d \)
for all vectors \((a_w)w \in M_k(C)^{\mathcal{X}^d}\). Identifying the optimal solution \((a_w)w \in \mathcal{X}^d = (\Lambda_{\tau,d}(A)P_w^*(A))w \in \mathcal{X}^d\)
comes simply to identifying the element which transforms \( \preceq \)
into \( = \) while not creating a null space through conjugation with \((a_w)w,(P_w(A^*)w) \rangle_d\), i.e. while imposing that this element stays invertible (of course, the most convenient choice to impose is \((a_w)w,(P_w^*(A))w \rangle_d = I_k\).

(2) Under the assumption that \((a_w)w,(a_w)w) \rangle_d \) is invertible, positivity of the left-most matrix in relation (3.9) is equivalent, via the Schur complement, to
\[ \langle (P_w(A^*)w)w,(a_w)w \rangle_d (\langle (a_w)w,(a_w)w \rangle_d)^{-1} \langle (a_w)w,(P_w(A^*)w) \rangle_d \leq \langle (P_w(A^*)w),(P_w(A^*)w) \rangle_d \]
After substituting \((P_w(A^*)w),(P_w(A^*)w) \rangle_d = \kappa_{\tau,d}(A,A^*)(I_k)\), \((a_w)w,(a_w)w) \rangle_d = (\text{Id}_{M_k(C)} \otimes \tau)(P(X)P^*(X))\), and \((a_w)w,(P_w(A^*)w) \rangle_d = P(A)(I_k)\) we obtain
\[ P(A)(I_k)^*(\text{Id}_{M_k(C)} \otimes \tau)(P(X)P^*(X))^{-1}P(A)(I_k) \preceq \kappa_{\tau,d}(A,A^*)(I_k)\]
By employing an arbitrarily small perturbation, the above holds also if \((a_w)w,(a_w)w) \rangle_d \) is not invertible, in the sense that the left-hand side stays bounded by \( \kappa_{\tau,d}(A,A^*)(I_k)\) when the perturbation tends to zero. This relation can be viewed as the noncommutative equivalent of the inequality \( |P(x)|^2 \leq \kappa_{\mu,d}(x,x)\int |P(t)|^2 d\mu(t)\), \( x \in \mathbb{R}^n\).
Moreover, Equation (3.8) is then rewritten as
\[ \kappa_{\tau,d}(A,A^*)(I_k) = \max\{P(A)(I_k)^*P(A)(I_k) : P \in M_k(C(X)d), (\text{Id}_{M_k(C)} \otimes \tau)(PP^*) = I_k\}\]
Indeed, the Schur complement in the left-hand side of the previous domination relation remains invariant under multiplication of \( P \) to the left with \( G \otimes 1 \) for any \( G \in GL_k(C)\).

(3) One may of course extend the definition of \( \Lambda_{\tau,d}(A,B)(C) = (\kappa_{\tau,d}(A,B^*)(C))^{-1} \) for all \( A,B,C \) for which the inverse is well-defined. While that does not seem to be a useful extension, we would like to point out that \( \Lambda_{\tau,d}(A,B)(C) := (\kappa_{\tau,d}(A,B^*)(C))^{-1} \) is well defined for any \( C \succ 0 \), and moreover \( \Lambda_{\tau,d}(A,B)(C) = \min\{(\text{Id}_{M_k(C)} \otimes \tau)(PP^*) : P \in M_k(C(X)d), P(A)(C^{1/2}) = I_k\}\).

(4) A brief inspection of the proof of Theorem 3.4 shows that the result holds as well when one replaces \( M_k(C) \) with any finite factor \( A \) (and in particular, \( M_k(C(X)d) \)
with \( A \otimes C(X)\)). The essential element of the proof is that \( A^{\tau(n,d)} \) accepts the \( A \)-Hilbert module structure \((a_w)w,(b_w)w) \rangle_d = \sum_{w \in \mathcal{X}^d} a_w^*b_w^*\).

(5) We record for future use the expression of the minimizer in (3.8): for any operator algebra \( A \) on a Hilbert space,
\[ P_{d}(X) = (\Lambda_{\tau,d}(a) \otimes 1) \sum_{w \in \mathcal{X}^d} P_w(a) \otimes P_w^*(X) \in A \otimes C(X)d, \quad a \in (A^{sa})^n.\]
In particular, for $A = M_k(\mathbb{C})$, we obtain the formula in Theorem 3.4. The optimization problem stated in (3.8) is the noncommutative analog of [LP19, (3.2)].

6 The view of the evaluation $P(A)(I_k)$ as a matrix-valued inner product has been essential in the proof of Theorem 3.4. Another re-interpretation of this evaluation might be more intuitive, namely as a matrix product. Tautologically, for $P(X) = \sum_{w \in (X)_d} c_w \otimes X^w$, we have

$$P(A)(I_k) = cA = \begin{bmatrix} c_0 & c_{X_1} & \cdots & c_{X_{n-1}}^d & c_{X_n^d} \\ A_1 \\ \vdots \\ A_{n-1}^{d-1} A_{n-1} \\ A_n^d \end{bmatrix}.$$  

(3.11)

The matrix $c$ is the $k \times \sigma(n,d)k$ complex matrix formed by the row of $k \times k$ blocks $c_w, w \in (X)_d$. Similarly, $A \in M_{k\sigma(n,d)}(M_k(\mathbb{C}))$ is the column matrix having as entries the $k \times k$ blocks $A^w, w \in (X)_d$. The matrix entries of the row vector $c$ are ordered lexicographically from left to right, and the matrix entries $A^w$ of $A$ are ordered lexicographically from top to bottom. It is clear that one may amplify $c$ to a matrix $c \in M_{k\sigma(n,d)}(\mathbb{C})$ by adding rows of zeros under $c$ and similarly one may add columns of zeros to the right of $A$ in order to form a square matrix $A$ of the same size. Then

$$P(A)(I_k) = [cA]_{1,1}$$  

is the $(1,1)$ entry of the product $cA \in M_{k\sigma(n,d)}(M_k(\mathbb{C}))$, or as the upper left $k \times k$ block of $M_{k\sigma(n,d)}(\mathbb{C})$.

Let $D_d(\tau)$ be the lower triangular matrix whose rows are the coefficients of the polynomials $P_w$ ordered by $\leq_{\text{gl}}$. Then one can prove that the moment matrix of $\tau$ satisfies $M_d(\tau)^{-1} = D_d(\tau)^TD_d(\tau)$. As a consequence we obtain the following:

Proposition 3.6. Recall that $W_d(X) = (X^n)_d$ is the vector of all words of $(X)_d$. Let $\tau: \mathbb{C}(X) \to \mathbb{C}$ be a faithful tracial state. For any $n$-tuples of selfadjont operators $A, B$ and any bounded operator $C$ on a Hilbert space, one has

$$\kappa_{\tau,d}(A,B)(C) = W_d(A)^TD_d(\tau)^TCd(\tau)W_d^*(B).$$

In particular, $\kappa_{\tau,d}(A,B)(1) = W_d(A)^TM_d(\tau)^{-1}W_d^*(B)$.

3.3. Noncommutative Siciak extremal functions. Our next objective is to investigate the asymptotic properties of $\Lambda_{\tau,d}$ as $d \to \infty$ when evaluated on various domains. In order to perform this analysis, we are forced to place some restrictions both on the nature of $\tau$ and the domains on which we evaluate $\Lambda_{\tau,d}$. This necessity should probably not be too surprising, as even in the much more studied classical context, a full description of this asymptotic behavior is not known. We are unaware of any previous work in the noncommutative context.

For a fixed positive, bounded trace-state $\tau: \mathbb{C}(X) \to \mathbb{C}$, we construct via the GNS construction [Tak79, Chapter I, Definition 9.15] the finite von Neumann algebra generated by the $n$ selfadjoint elements (the images of $X_1, \ldots, X_n \in \mathbb{C}(X)$ via the GNS representation) whose joint law equals $\tau$ [AGZ10, Proposition 5.2.14(d)]. We denote it by $W^*(\tau)$. 
In the following, the faithfulness of \( \tau \) will often be assumed for convenience. We will however try to indicate what changes occur in various arguments if \( \tau \) is not assumed to be faithful.

Our methods parallel to a significant extent the classical ones [GZ17, Kli91, ST97]. The way to investigate the asymptotic behavior of the Christoffel-Darboux kernel in classical analysis involves the so-called Siciak extremal function. Let us start by proposing free noncommutative versions of this object. We consider \( \varrho_r \) to be a tuple of selfadjoint variables in a von Neumann algebra \( \mathcal{M} \) which is distributed according to \( \tau \) with respect to the normal faithful tracial state \( \text{tr} \) on \( \mathcal{M} \) (in particular, \( W^*(\tau) \) is isomorphic as a von Neumann algebra to the von Neumann subalgebra of \( \mathcal{M} \) generated by the \( n \)-tuple \( \varrho_r \)). For each \( d, k \in \mathbb{N}, A \in \mathbb{M}_k(\mathbb{C})^n \), we define

\[
\Phi^2_{\tau,d}(A) = \sup \{ \text{tr}(P(A)(I_k)^* P(A)(I_k)) : P \in \mathbb{M}_k(\mathbb{C}(X_d)), \|P(a_r)\| \leq 1 \},
\]

\[
\Phi^\infty_{\tau,d}(A) = \sup \{ \|P(A)(I_k)^2 : P \in \mathbb{M}_k(\mathbb{C}(X_d)), \|P(a_r)\| \leq 1 \}.
\]

Recall that \( P(a_r) \in \mathbb{M}_k(W^*(\tau)) \), so that the norm \( \|P(a_r)\| \) is the operator norm on \( \mathbb{M}_k(W^*(\tau)) \approx \mathbb{M}_k(\mathbb{C}) \otimes W^*(\tau) \subseteq \mathbb{M}_k(\mathbb{C}) \otimes \mathcal{M} \). Note that in (3.13) we have defined \( \Phi^2_{\tau,d} \) as the trace of the product of the two matrices obtained by applying the linear operator \( P(A) \) to \( I_k \). However, we would have obtained precisely the same result had we defined it as \( \text{tr}((P^*(A^*)P(A))(I_k)) \). Clearly \( \Phi^2_{\tau,d}(A) \leq \Phi^\infty_{\tau,d}(A) \) for all \( A \).

Let us list next various properties of these functions which will be of use to us later.

3.3.1. Plurisubharmonicity. Both \( \Phi^2_{\tau,d}(A) \) and \( \Phi^\infty_{\tau,d}(A) \) are suprema of norms of analytic polynomials, hence entire plurisubharmonic functions (see Section 2.3) in the classical sense when viewed as functions on the complex Euclidean space \( \mathbb{C}^{nk^2} \approx (\mathbb{M}_k(\mathbb{C}))^n \). Moreover, \( \|P(A)(I_k)\| = \lim_{m \to \infty} \text{tr}((P(A)(I_k)^*P(A)(I_k))^m)^{1/2m} \).

3.3.2. Monotonicity. If \( X' \subseteq X \) is a subset, then \( \Phi^\bullet_{\tau,|X'|,d}(A') \leq \Phi^\bullet_{\tau,d}(A \setminus A') \), for \( \bullet \in \{2, \infty\}, A = (A', A \setminus A') \in \mathbb{M}_k(\mathbb{C}(X_d))^n, d \in \mathbb{N} \). This follows easily by noting that a polynomial \( P \in \mathbb{C}(X_d) \) belongs tautologically to \( \mathbb{C}(X_d) \) as well, and that the embedding \( W^*(\tau|_{X'}) \to W^*(\tau) \) is completely isometric.

3.3.3. Compactness. The set of polynomials \( \{P \in \mathbb{M}_k(\mathbb{C}(X_d)) : \|P(a_r)\| \leq 1 \} \) is compact, in the sense that for any given basis \( \{b_1, \ldots, b_{\sigma(n,d)}\} \) of \( \mathbb{C}(X_d) \), the set \( \{(c_1, \ldots, c_{\sigma(n,d)}) \in \mathbb{M}_k(\mathbb{C})^{\sigma(n,d)} : \|\sum c_j \otimes b_j(a_r)\| \leq 1 \} \) is compact\(^1\) in the Euclidean space \( \mathbb{M}_k(\mathbb{C})^{\sigma(n,d)} \). Indeed, all bases are equivalent via an invertible, finite-dimensional linear transformation, so it is enough to show this for \( \{b_1, \ldots, b_{\sigma(n,d)}\} = \{P_w\}_{w \in \mathbb{C}(X_d)} \), the basis of orthonormal polynomials. In that case, \( \sum c_w r^2 \otimes \text{tr}(P_w^*(a_r)P_w(a_r)) = (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \text{tr})(\sum c_w r^2 \otimes P_w^*(a_r)P_w(a_r)) = (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \text{tr})(\sum c_w r^2 \otimes P_w(a_r)) = (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \text{tr})(\sum c_w r^2 \otimes P_w(a_r)) \)

\(^1\)It is important to remember that we work here under the hypothesis of faithfulness of \( \tau \). If, for instance, \( \varrho_r \) were to satisfy some algebraic relation – that is, if there were a \( P \in \mathbb{C}(X) \) such that \( P(\varrho) = 0 \) – the bound in (3.15) would easily fail. In that case, one would need to work on a space of dimension strictly smaller than \( \sigma(n,d) \) in order for this same result to hold – see [BPSS20]. The reader may keep in mind the most extreme case, when all coordinates of \( \varrho_r \) are complex multiples of the algebra’s unit.
for any $P(X) = \sum_w c_w \otimes P_w(X) \in \mathcal{M}_k(\mathbb{C}(X)_d)$. Thus, the condition $\|P(\alpha)\| \leq 1$ implies
\[(3.15)\] 
\[1 \geq \|P(\alpha)\|^2 = \|P(\alpha)P(\alpha)^*\| \geq \|\text{Id}_{\mathcal{M}_k(\mathbb{C}(X))} \otimes \text{tr}(P(\alpha)P(\alpha)^*)\| = \left\| \sum_w c_w c_w^* \right\|
\] 
so that $\|c_w\|^2 \leq 1$ for all $w \in (X)_d$. This, together with the continuity of the functions involved, guarantees that the suprema in both (3.13) and (3.14) are in fact maxima.

3.3.4. Matrix convexity. Recall that a set $K$ of matrices over $\mathbb{C}$ consists of subsets $K_k \subset \mathcal{M}_k(\mathbb{C})$, for all $k \in \mathbb{N}$. Such a set $K$ is called matrix convex [EW97, Section 3] if for all $A \in K_k$ and $B \in K_{k'}$, one has $A \oplus B \in K_{k+k'}$, and for all $A \in K_k$ and $\alpha \in \mathcal{M}_k(\mathbb{C})$ with $\alpha^*\alpha = I_k$, one has $\alpha^*A\alpha \in K_k$. The set $\{P \in \mathcal{M}_k(\mathbb{C}(X)) : \|P(\alpha)\| \leq 1\}$ is convex, and in fact it is also matrix convex, which is verified as follows: if $P \in \mathcal{M}_k(\mathbb{C}(X))$, $P' \in \mathcal{M}_k(\mathbb{C}(X))$ satisfy the desired norm inequality, then $\|P(\alpha) \oplus P'(\alpha)\| \leq 1$ as well in $\mathcal{M}_{k+k'}(\mathbb{C}) \otimes \mathcal{M}$ (to be clear, if $P(X) = \sum_w c_w \otimes P_w(X)$, $P'(X) = \sum_w c_w' \otimes P_w(X)$, we view $P(X) \oplus P'(X) = \sum_w (c_w \otimes c_w') \otimes P_w(X) \in \mathcal{M}_{k+k'}(\mathbb{C}(X))$). If $\alpha \in \mathcal{M}_{k \times k'}(\mathbb{C})$ satisfies $\alpha^*\alpha = I_{k'}$ (so implicitly $k' \leq k$), then
\[\left\| \sum_w (\alpha^*c_w\alpha) \otimes P_w(\alpha) \right\| = \|(\alpha^* \otimes 1)P(\alpha)(\alpha \otimes 1)\| \leq \|\alpha\|^2 \|P(\alpha)\| \leq 1.
\] Obviously, this implies that the maximizers providing equality in (3.13) and (3.14) are extremal points in these (classically) convex sets.

3.3.5. Invariance under left unitary action of the set of maximality. Equally relevant, let us observe that if $P \in \mathcal{M}_k(\mathbb{C}(X)_d)$ is a polynomial such that one of the equalities $\Phi^*_k \Phi_k(\alpha) = \text{tr}_k(P(\alpha)(\alpha)P(\alpha)(\alpha))$ or $\Phi^* \Phi(\alpha) = \|P(\alpha)(\alpha)|\|$ takes place, then the corresponding equality still holds for the polynomial $(U \otimes 1)P$ for any $U \in \mathcal{M}_k(\mathbb{C})$ satisfying $UU^* = I_k$, i.e. for any unitary matrix $U$. Indeed, this is straightforward since $\text{tr}_k((U \otimes 1)P(\alpha)(\alpha)(U \otimes 1)P(\alpha)(\alpha)) = \text{tr}_k\left(\left(\sum_w (Uc_w)^w\right)^* \left(\sum_w (Uc_w)^w\right)\right) = \text{tr}_k\left(\left(\sum_w (A^w)^*\right) (U^*U) \left(\sum_w c_w A^w\right)\right) = \|\sum_w Uc_w A^w\|^2 = \|P(\alpha)(\alpha)\|$ whenever $U$ is in $\mathcal{M}_k(\mathbb{C})$ and $\|(U \otimes 1)P(\alpha)\| = \|P(\alpha)(\alpha)\|$. Moreover, we have that if $\|P(\alpha)\| \leq 1$, then $\|(U \otimes 1)P(\alpha)\| \leq \|P(\alpha)(\alpha)\|$. Thus, the two sets on which maxima are achieved in (3.13), (3.14) are invariant under the left action of the unitary group of the $k$ by $k$ complex matrices. As an immediate consequence of the polar decomposition of operators, we may assume without loss of generality that the maximizer $P$ satisfies the condition $\Phi_k(\alpha) \geq 0$: indeed, if $\Phi_k(\alpha) = U \Phi_k(\alpha)(U^*\alpha)$, then we replace $P$ by $U^* \otimes 1)P$. Unfortunately, that does not mean $P \geq 0$.

With the same methods, one shows that $\Phi^*_k \Phi_k(\alpha) \leq \Phi^* \Phi(\alpha)(U^*\alpha)$, $\alpha \in \{2, \infty\}$, for all unitary $k \times k$ complex matrices $U$, so that $\Phi^*_k$ is constant on unitary orbits. Indeed, picking $P(X) = \sum_{w \in (X)_d} c_w \otimes X^w$ for which the maximum $\Phi^*_k(X) = \...$
\[ ||P(A)(I_k)||^2 \] is reached, one has
\[ Q(U A^* U)(I_k) Q(U A^* U)(I_k)^* = \sum_{w \in \mathcal{X}_d} (c_w U^*) U A^* U^* \left( \sum_{w \in \mathcal{X}_d} (c_w U^*) U A^* U^* \right)^* = \sum_{v, w \in \mathcal{X}_d} c_v (U^* U) A^w \langle (A^w)^* (U^* U) c_w^* \right. \]
\[ = P(A)(I_k) P(A)(I_k)^*, \]
where \( Q(\mathcal{X}) = P(\mathcal{X})[U^* \otimes 1] \). As \( 1 = ||P(\mathcal{X})|| = ||P(\mathcal{X})[U^* \otimes 1]|| = ||Q(\mathcal{X})|| \), it follows that \( \Phi_{\tau, d}(U A^* U) \geq ||Q(U A^* U)(I_k)||^2 = ||P(A)(I_k)||^2 = \Phi_{\tau, d}(A) \), for any \( k \times k \) unitary matrix \( U \).

The proofs of the two statements above can be put together to yield a slightly stronger technical result: if \( P(\mathcal{X}) = \sum_{w \in \mathcal{X}_d} c_w \otimes X^w \) and \( V, W \) are \( k \times k \) unitaries, then
\[ V \sum_{w \in \mathcal{X}_d} c_w A^w W = \sum_{w \in \mathcal{X}_d} V c_w W(W^* A W)^w = [(V \otimes 1) P(W \otimes 1)](W^* A W)(I_k); \]

since \( ||[(V \otimes 1) P(W \otimes 1)](\mathcal{X})|| = ||P(\mathcal{X})|| \), it follows that if \( \Phi_{\tau, d}(A) = ||P(\mathcal{X})(I_k)||^2 \) is achieved on \( P \), then \( \Phi_{\tau, d}(A) = \Phi_{\tau, d}(W^* A W) \geq ||[(V \otimes 1) P(W \otimes 1)](W^* A W)(I_k)||^2 = ||P(\mathcal{X})(I_k)||^2 = \Phi_{\tau, d}(A) \), that is,
\[ \Phi_{\tau, d}(A) = \left( V \sum_{w \in \mathcal{X}_d} c_w A^w W \right)^2 = ||[(V \otimes 1) P(W \otimes 1)](W^* A W)(I_k)||^2. \]

In particular, \( \Phi_{\tau, d}(A) = \Phi_{\tau, d}(W^* A W) \) can be achieved at an element that is diagonal in whichever basis it is desired.

**3.3.6. Domination.** The condition \( ||P(\mathcal{X})|| \leq 1 \) in relations (3.13) – (3.14) is equivalent to either of \( P(\mathcal{X})^* P(\mathcal{X}) \preceq I_k \otimes 1 \) or \( P(\mathcal{X})^* P(\mathcal{X}) \preceq I_k \otimes 1 \) in \( M_k(\mathbb{C}) \otimes W^*(\tau) \).

If there exists a \( 0 \prec C \preceq I_k \) such that, say, \( P(\mathcal{X})^* P(\mathcal{X}) \preceq C = I_k \otimes 1 \), then \( (C^{-1} \otimes 1) P(\mathcal{X})^* P(\mathcal{X}) \preceq C^{-1} \otimes 1 \) \( \preceq I_k \otimes 1 \) ensuring that the polynomial \( (C^{-1} \otimes 1) P(\mathcal{X}) \) belongs to the acceptable set in the right hand side of (3.13) and (3.14). Then
\[ ||(C^{-1} \otimes 1) P(\mathcal{X})(I_k) = C^{-1} \otimes 1 P(\mathcal{X})(I_k), \quad \text{so that} \quad \text{tr}_{\mathcal{X}}(P(\mathcal{X})(I_k)^* C^{-1} P(\mathcal{X})(I_k)) \geq \text{tr}_{\mathcal{X}}(P(\mathcal{X})(I_k) P(\mathcal{X})(I_k)), \quad ||P(\mathcal{X})(I_k) C^{-1} P(\mathcal{X})(I_k) || \geq ||P(\mathcal{X})(I_k)||. \]

If \( C^{-1} \preceq I_k \), equality in the previous inequality of traces above can only take place if \( \ker P(\mathcal{X})(I_k) = \{ 0 \} \). This makes it clear that without loss of generality we may assume in both (3.13) and (3.14) that there is no element \( C \npreceq I_k \) which is not a projection such that \( P(\mathcal{X})^* P(\mathcal{X}) \preceq C = I_k \otimes 1 \).

However, we note that one may obtain a “minimal maximizer”: assume that \( \Phi_{\tau, d}(A) \) is achieved at \( P(\mathcal{X}) = \sum_{w \in \mathcal{X}_d} c_w \otimes X^w \). Then there exists a rank-one projection \( p \in M_k(\mathbb{C}) \) such that \( ||P(\mathcal{X})(I_k)||^2 = ||p P(\mathcal{X})(I_k)||^2 \) (one simply picks \( p \) to be the projection onto a norm-one vector \( \xi \in \mathbb{C}^N \) such that \( ||P(\mathcal{X})(I_k)||^2 = \langle P(\mathcal{X})(I_k) P(\mathcal{X})(I_k)^* \xi, \xi \rangle \)). Since \( P(\mathcal{X})(I_k) = \sum_{w \in \mathcal{X}_d} c_w X^w \) it follows that
\[ \Phi_{\tau, d}(A) = ||p P(\mathcal{X})(I_k)||^2 \]
\[ = p \left( \sum_{w \in \mathcal{X}_d} c_w X^w \right)^* \left( \sum_{w \in \mathcal{X}_d} c_w X^w \right) p ||p \| 

\]
3.14

and of course \((p \otimes 1)P_\Delta (A) = \sum_{w \in (X)_d} (p \otimes \Lambda w) \otimes X_w\), so \(\|(p \otimes 1)P_\Delta (a_r)\| = 1\), with \((p \otimes 1)P_\Delta (a_r) (p \otimes 1)P_\Delta (a_r)^* \leq p \otimes 1\). On the other hand, for \(\Phi^*_r, d\) it is clear that the orthogonal complement of the kernel of \(P_\Delta (A) (I_k)^*\) is the biggest projection with which we can multiply \(P_\Delta\) on the left.

3.3.7. Linear change of coordinates. Assume that \(u \in GL_n(\mathbb{R})\). We consider the \(n\)-tuple of selfadjoint operators \(\mathbb{A}_u = u_{\mathbb{A}}\). That is, \(\mathbb{A}_u = (A_1, A_2, \ldots, A_n)\) is a tuple of selfadjoint variables such that \(A_j = \sum_{k=1}^u u_{jk} a_k\), so that the law \(\theta: \mathbb{C}(X) \to \mathbb{C}\) of the \(n\)-tuple \(\mathbb{A}_u\) is given by \(\theta(P(X_1, X_2, \ldots, X_n)) = \tau(P(\sum_{k=1}^u u_{1k} X_k, \ldots, \sum_{k=1}^u u_{uk} X_k))\). It is obvious that \(W^*(\theta) \simeq W^*(\tau)\) as von Neumann algebras (in fact it is equally obvious the case for any invertible – under composition – nc map \(f\) defined on an nc neighborhood of \(a_r\) so that \(f(a_r)\) is a tuple of selfadjoints).

One can identify \(\mathbb{C}(X)\) with the Fock space of the \(n\)-dimensional Hilbert space \(\mathbb{C}^n\) via the identification of 1 with the vacuum vector \(\Omega\) and \(X_j\) with \(e_j\), the \(j\)th canonical basis vector of \(\mathbb{C}^n\). The product \(X_i X_j\) corresponds to \(e_i \otimes e_j\). With this view, one sees immediately that the effect of \(u\) on \(a_r\) extends to all of \(\mathbb{C}(a_r)\) by acting on \((\mathbb{C}^n)^{\otimes \rho}\) as \(u^{\otimes \rho}\). Thus, action on the subspace \(\mathbb{C} \otimes \bigoplus_{j=1}^d (\mathbb{C}^n)^{\otimes \rho}\) is done via \(Id_{\mathbb{C}^n} \otimes \bigoplus_{j=1}^d u^{\otimes j}\). Obviously, this assigns (easily computable but long) coefficients \(u_{w,v} \in \mathbb{R}\) in the block-matrix expression of the operator: \(u^{\otimes j} e_w = \sum_{v \in (\mathbb{V})_w} (u^{\otimes j})_{w,v} e_v\). Given a matrix-valued polynomial \(P(X) = \sum_{w \in (X)_d} c_w \otimes X_w\), the change of variable \(\mathbb{A}_u = u_{\mathbb{A}}\) translates into

\[
P(\mathbb{A}_u) = P(\mathbb{A}_u) = \sum_{w \in (X)_d} c_w \otimes (u_{\mathbb{A}})^w
\]

\[
= c_0 \otimes 1 + \sum_{j=1}^d [I_k \otimes u^{\otimes j}] \sum_{w \in (X)_j \setminus (X)_{j-1}} c_w \otimes a_v^w
\]

\[
= c_0 \otimes 1 + \sum_{j=1}^d \sum_{w \in (X)_j \setminus (X)_{j-1}} c_w \otimes \left( \sum_{v \in (\mathbb{V})_w} (u^{\otimes j})_{w,v} a_v^w \right)
\]

\[
= c_0 \otimes 1 + \sum_{j=1}^d \sum_{v \in (\mathbb{V})_w} \left( \sum_{w \in (X)_j \setminus (X)_{j-1}} (u^{\otimes j})_{w,v} c_w \right) a_v^w = Q(\mathbb{A}_u),
\]

where \(Q(X) = \sum_{v \in (\mathbb{V})_w} d_v \otimes X_v\). Tautologically \(\|P(\mathbb{A}_u)\| = 1 \iff \|P(a_r)\| = 1\). For given \(\mathbb{A} \in M_k(\mathbb{C})^n\), assume \(Q\) is a polynomial on which the maximum in (3.13) (resp. (3.14)) is achieved for \(a_r\). Thus, from the above,

\[
\Phi^*_r, d(\mathbb{A}) = \|Q(\mathbb{A} (I_k))\| = \left\| \sum_{w \in (X)_d} d_v A_v^w \right\|^2
\]

\[
= \left\| c_0 + \sum_{j=1}^d \sum_{v \in (\mathbb{V})_w} \left( \sum_{w \in (X)_j \setminus (X)_{j-1}} (u^{\otimes j})_{w,v} c_w \right) A_v^w \right\|^2.
\]
\[ \frac{1}{2d} \log \Phi_{\tau,d}(A) \leq \log \| A \|_d + c_{d,\tau}. \]

(As it is usual, \( \log^+ t = \max\{\log t, 0\} \).) It is remarkably convenient that from this point of view, the two norms we use are the same. Indeed, \( \| A \|_{2d}^2 = \text{tr}_k(A_1^* A_1 + \cdots + A_n^* A_n) \leq n \max_{1 \leq i \leq n} \text{tr}(A_i^* A_i) \leq n \max_{1 \leq i \leq n} \| A_i \|^2 = n \| A \|_d^2 \) and \( \| A \|_d^2 = \max_{1 \leq i \leq n} \| A_i \|^2 \leq k \max_{1 \leq i \leq n} \text{tr}(A_i^* A_i) \leq k \text{tr}(A_1^* A_1 + \cdots + A_n^* A_n) = k \| A \|_d^2 \), so that \( \frac{1}{\sqrt{n}} \| A \|_2 \leq \| A \|_d \leq \sqrt{k} \| A \|_2 \). By taking log, we obtain \( \log \| A \|_2 - \log \sqrt{n} \leq \log \| A \|_d \leq \log \| A \|_2 + \log \sqrt{k} \). Thus, we will compare everything to \( \log \| A \|_2 \). Now, for any given polynomial \( P \in M_k(\mathbb{C}(X)_d \setminus M_k(\mathbb{C}(X)_{d-1}) \), it follows from the classical theory that

\[ \limsup_{\| A \|_2 \to \infty} \frac{\log \text{tr}_k(P(A))(I_k)}{2d \log \| A \|_2} \leq 1. \]

Indeed, the numerator is the logarithm of a classical polynomial of degree \( 2d \) on \( \mathbb{C}^{n^2} \). According to Section 3.3.3, \( \{ \log \text{tr}_k(P(A))(I_k)^* P(A)(I_k) : P \in M_k(\mathbb{C}(X)_d) \}, \| P(a_r) \|_2 \leq 1 \) is a locally bounded family, so [ST97, Theorem 1.6, Appendix B] guarantees that \( \frac{1}{2d} \log \Phi_{\tau,d}(A) \) satisfies the same growth condition. By the same theorem, together with the description of the operator norm from Section 3.3.1, so does \( \frac{1}{2d} \log \Phi_{\tau,d}(A) \).

3.3.9. Single variable bounds. As noted before, our hypothesis that \( \tau \) is positive guarantees that \( a_r \) can be chosen as a tuple of self-adjoint random variables, \( a_r^* = a_r \), which generate a tracial \( C^* \)-algebra \( (C^*(\tau), \text{tr}) \). For each coordinate of \( a_r \), we consider, via functional calculus, the identification of the \( C^* \)-subalgebra it generates with the space of continuous functions on its spectrum, endowed with the restriction of \( \text{tr} \). Without loss of generality, pick the first coordinate, \( a_{r,1} \), for our analysis. Denote by \( \mu_1 \) its distribution with respect to \( \tau \), which is a classical, compactly supported Borel probability measure on \( \mathbb{R} \). Assume that \( \text{supp}(\mu_1) \) is an infinite set in
the real line, so that \( C^* (\tau|_{\mathcal{C}(X,d)}) = C(\sigma(\alpha_{r,1}), \mu_1) \) is an infinite-dimensional Abelian \( C^* \)-algebra. For each \( k \in \mathbb{N} \), we may choose \( k \) distinct points in \( \sigma(\alpha_{r,1}) \); the evaluation of an element from \( C(\sigma(\alpha_{r,1}), \mu_1) \) at those points yields an algebra morphism onto \( \mathbb{C}^k \), viewed as the diagonal of \( M_k(\mathbb{C}) \). In particular, for given matrix size \( k \), one may pick pairwise distinct points \( \mathcal{P} = \{ p_1, \ldots, p_k \} \subset \sigma(\alpha_{r,1}) \), and define the linear map \( \mathcal{C}_\mathcal{P} : C(\sigma(\alpha_{r,1}), \mu_1) \to \mathbb{C}^k \) by \( f \mapsto \mathcal{C}_\mathcal{P}(f) = (f(p_1), \ldots, f(p_k)) \). This is tautologically a surjective \( C^* \)-algebra morphism, and in particular it is completely contractive, completely positive, and unit preserving. Thus, for any \( N \in \mathbb{N} \), the map \( \text{id}_N \otimes \mathcal{C}_\mathcal{P} \) is a positive contraction. In particular, if \( \mathbf{P}(X) = \sum_{w \in (X)_d} c_w \otimes X_w \in M_k(\mathbb{C}(X)_d) \), then
\[
\left\| \sum_{w \in (X)_d} c_w \otimes \mathcal{C}_\mathcal{P}(a_{r,1}^w) \right\| = \left\| (\text{id}_N \otimes \mathcal{C}_\mathcal{P})(\mathbf{P}(a_{r,1})) \right\| \leq \left\| \mathbf{P}(a_{r,1}) \right\|.
\]
However, \( \mathcal{C}_\mathcal{P} \) is an algebra morphism, so that \( \mathcal{C}_\mathcal{P}(a_{r,1}^w) = \mathcal{C}_\mathcal{P}(a_{r,1})^w \) for all \( w \in \mathbb{N} \) (or \( w \in \mathbb{Z} \), if \( a_{r,1} \) is invertible). As noted after equations (3.3)–(3.4), this, together with the fact that \( \text{tr}_k(I_k) = 1 \), guarantees that \( \text{tr}_k(\mathbf{P}(\mathcal{C}_\mathcal{P}(a_{r,1}))(I_k))^* \mathbf{P}(\mathcal{C}_\mathcal{P}(a_{r,1}))(I_k) \leq \| \mathbf{P}(a_{r,1}) \|^2 \) for any collection \( \mathcal{P} \subset \sigma(\alpha_{r,1}) \) of \( k \) mutually disjoint points. In particular, we have established that for any selfadjoint matrix \( A \in M_k(\mathbb{C}) \) whose spectrum is included in the spectrum of \( a_{r,1} \), the condition \( \| \mathbf{P}(a_{r,1}) \| \leq 1 \) implies automatically \( \Phi_{\tau|_{\mathcal{C}(X,d)}}^2(A) \leq 1 \).

We define
\[
\Phi_{\tau,d}^2(A) = \left( \limsup_{d \to \infty} \left( \Phi_{\tau,d}^2(A) \right)^{1/d} \right)^*,
\]
\[
\Phi_{\tau}^\infty(A) = \left( \limsup_{d \to \infty} \left( \Phi_{\tau,d}^\infty(A) \right)^{1/d} \right)^*.
\]
(The asterisk denotes the upper semicontinuous regularization: the upper semicontinuous regularization \( f^* \) for a function \( f \) is defined by \( f^*(\zeta) = \limsup_{z \to \zeta} f(z) \), and is the smallest upper semicontinuous function with the property that \( f^* \geq f \).)

We also define
\[
\Sigma_{\tau,d}^2(A) = \left[ \sup_{d \in \mathbb{N}} \left( \Phi_{\tau,d}^2(A) \right)^{1/d} \right]^*;
\]
\[
\Sigma_{\tau}^\infty(A) = \left[ \sup_{d \in \mathbb{N}} \left( \Phi_{\tau,d}^\infty(A) \right)^{1/d} \right]^*.
\]
Clearly, \( \Sigma_{\tau,d}^2(A) \geq \Phi_{\tau,d}^*(A), \bullet \in \{ 2, \infty \}, A \in M_k(\mathbb{C})^n, k \in \mathbb{N} \). It is known (see, for instance, [ST97, Equation (2.17) in Appendix B2]) that \( \Sigma \) and \( \Phi \) coincide in the classical case. That might not happen for us, as it can be easily seen from (3.25) below and the comments following it. Before enumerating below the consequences of the above-listed properties of \( \Phi_{\tau,d}^* \) on these functions, we recall (see Section 2.3) that a set is pluripolar if it is contained in the \(-\infty\)-level set of a non-constant plurisubharmonic function. A property holds quasi everywhere (q.e.) in a set \( S \) if it holds on \( S \setminus E \) where \( E \) is a pluripolar set.

**Remark 3.7.** The following statements hold:

1. \( \Phi_{\tau,d}^2(A) \leq \Phi_{\tau}^\infty(A), \Sigma_{\tau,d}^2(A) \leq \Sigma_{\tau}^\infty(A) \quad \text{q.e., } A \in M_k(\mathbb{C})^n, k \in \mathbb{N}; \)
In the following we assume implicitly that \( A \mapsto \left( \Phi_{r,d}^{\infty}(A) \right)^{1/d} \) is locally bounded from the above in order for most of our statements to be non-vacuous (this happens in numerous cases as shown later on in Example 3.12).

(3) If we denote \( \tau_1 = \tau|_{\mathbb{C}(X_1, \ldots, X_n)} \), \( \tau_2 = \tau|_{\mathbb{C}(X_{n+1}, \ldots, X_{n+m})} \), then \( \Phi_\tau(A, B) \geq \max\{\Phi_{\tau_1}(A), \Phi_{\tau_2}(B)\} \), \( \Sigma_\tau(A, B) \geq \max\{\Sigma_{\tau_1}(A), \Sigma_{\tau_2}(B)\} \) for any \((A, B) \in M_k(\mathbb{C})^n \times M_k(\mathbb{C})^m, \bullet \in \{2, \infty\}\). This follows immediately from Section 3.3.2.

(4) By 3.3.5, for any \( k \in \mathbb{N}, \bullet \in \{2, \infty\} \), we have

\[
\Phi_\tau^*(A) = \Phi_\tau^*(U A U^*), \quad \Sigma_\tau(A) = \Sigma_\tau(U A U^*), \quad A \in M_k(\mathbb{C})^n, U \in M_k(\mathbb{C}), U U^* = I_k.
\]

As an immediate consequence, we record

\[
\Phi_\tau^*\left( \begin{bmatrix} A & B \\ \overline{B} & \overline{A} \end{bmatrix} \right) = \Phi_\tau^*\left( \begin{bmatrix} A^* & B^* \\ \overline{B} & \overline{A} \end{bmatrix} \right), \quad \Sigma_\tau\left( \begin{bmatrix} A & B \\ \overline{B} & \overline{A} \end{bmatrix} \right) = \Sigma_\tau\left( \begin{bmatrix} A^* & B^* \\ \overline{B} & \overline{A} \end{bmatrix} \right),
\]

for all \( A, B \in M_k(\mathbb{C})^n \), \( k \in \mathbb{N}, \bullet \in \{2, \infty\} \).

(5) If \( u \in GL_n(\mathbb{R}) \), \( \tau \) is the distribution of \( a_{\tau} \), and \( \theta \) is the distribution of \( u a_{\tau} u^* \), then by (3.17),

\[
\Phi_\tau^*(A) = \Phi_\theta^*(u A), \quad \Sigma_\tau(A) = \Sigma_\theta(u A), \quad A \in M_k(\mathbb{C})^n, k \in \mathbb{N}, \bullet \in \{2, \infty\}.
\]

Let us record some obvious properties of \( \Phi_\tau^*, \bullet \in \{2, \infty\} \).

1. \( \Phi_\tau^*(A) \geq 1 \) for all \( \bullet \in \{2, \infty\}, A \in M_k(\mathbb{C})^n, k \in \mathbb{N} \). This follows easily by fixing a polynomial \( P \) in the right hand side of either of (3.13), (3.14) and noting that \( \Phi_{r,d}(A) \geq \|P(A)(I_k)\|^2 \); we conclude \( \Phi_\tau^*(A) \geq \lim_{d \to \infty} \|P(A)(I_k)\|^2/d = 1 \). The conclusion for \( \Sigma \) is obvious.

2. In the classical theory of several complex variables, it is known (see [ST97, Appendix B.2, (2.17)]) that the sequence \( \{\Phi_{r,d}(A)^{1/d}\}_{d \in \mathbb{N}} \) is increasing for any \( A \in M_1(\mathbb{C})^n \simeq \mathbb{C}^n \). When the matrix size \( k > 1 \), this monotonicity statement may turn out to be false. This fact is essentially caused by the existence of nilpotent elements in \( M_k(\mathbb{C}) \) whenever \( k \geq 2 \). Surprisingly, a counterexample can be found even when \( n = 1 \). Let \( P(X) = \sum_{j=0}^d c_j \otimes X^j = \begin{bmatrix} P_{1,1}(X) & P_{1,2}(X) \\ P_{2,1}(X) & P_{2,2}(X) \end{bmatrix} \in M_2(\mathbb{C}(X)) \) and \( A = \begin{bmatrix} 0 & t \\ 0 & 0 \end{bmatrix} \). Then

\[
P(A)(I_2) = P\left( \begin{bmatrix} 0 & t \\ 0 & 0 \end{bmatrix} \right) \left( \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) = \sum_{j=0}^d c_j \begin{bmatrix} 0 & t^j \\ 0 & 0 \end{bmatrix}^j
\]
\[
= \begin{bmatrix}
(c_0)_{1,1} & (c_0)_{1,2} + t(c_1)_{1,1} \\
(c_0)_{2,1} & (c_0)_{2,2} + t(c_1)_{2,1}
\end{bmatrix} = \begin{bmatrix}
P(1,1)(0) & tP(1,1)(0) + P(1,2)(0) \\
P(2,1)(0) & tP(2,1)(0) + P(2,2)(0)
\end{bmatrix}.
\]

Pick \( \tau \) to be Wigner’s semicircle law, \( \tau(P(X)) = \frac{2}{\pi} \int_{-1}^{1} P(x) \sqrt{1 - x^2} \, dx \) (in fact, for our example, any distribution given by a probability measure with infinite and compact support containing zero would do). By definition, \( \Phi^2_{\tau,d}(A) \) is the maximum of \( \text{tr}_d(P(A)(I_d)^*P(A)(I_d)) \) on the set of polynomials \( P(X) \in M_d(\mathbb{C}(X)) \) satisfying \( \|P(a_\tau)\| \leq 1 \). With the above notation, \( \text{tr}_2(P(A)(I_2)^*P(A)(I_2)) = \frac{1}{2} (|P(1,0)(0)|^2 + |tP(1,1)(0) + P(1,2)(0)|^2 + |P(2,1)(0)|^2 + |tP(2,1)(0) + P(2,2)(0)|^2) \). Now, \( P_{k,j} \) are classical polynomials in one variable, to which Bernstein’s estimate [Tim94, (33), Section 4.8.7] applies: \( |P_{k,j}(0)| \leq \deg P_{k,j} \leq d \|P_{k,j}\|, \) where \( \| \cdot \| \) is the sup, or \( C^* \)-algebra, norm on \([-1,1]\) (the estimate is optimal: an example can be found in the Chebyshev orthogonal polynomials of the first kind, given by \( T_d(x) = \cos(d \arccos(x)) \), or \( T_d(\cos \theta) = \cos(d \theta), \theta \in [0,\pi], \) so that \( \|T_d\| = 1, \|T_d(0)\| = d \sin(d\pi/2) = d \) whenever \( d \) is odd). The requirement \( \|P(a_\tau)\| \leq 1 \) imposes \( |P_{k,j}(0)| \leq 1, 1 \leq i,j \leq 2, \) so that \( \text{tr}_2(P(A)(I_2)^*P(A)(I_2)) < 1 + (1 + td)^2 \) holds for all degrees \( d \in \mathbb{N} \) and parameters \( t \in \mathbb{C} \). On the other hand, by choosing the other entries to be zero, one sees that the estimate \( \text{tr}_2(P(A)(I_2)^*P(A)(I_2)) \geq \frac{|tP(1,0)(0)|^2}{(t^2d^2 - 1)/2} \) holds at least on a subsequence of degrees \( d \). Since both sequences \( (1 + (1 + td)^2)^{1/d} \) and \( (t^2d^2 - 1)/2 \) are both eventually decreasing and tend to the same limit as \( d \to \infty \), one obtains that

\[
(3.25)\quad \left\{ \Phi^2_{\tau,d}(A) \right\} \quad (d \in \mathbb{N}) \quad \text{might not be an increasing sequence.}
\]

Thus, if one replaces limsup in (3.18)-(3.19) by sup as in (3.20)-(3.21), one may get a different function whenever \( k > 1 \) (and in fact one is guaranteed the existence of a context in which a different result occurs before performing the upper semicontinuous regularization).

(3) \( \{A \in M_k(\mathbb{C})^n : \Phi^*_k(A) > 1 \} \neq \emptyset \) for all \( k > 1 \). To find such an example, one considers \( k > 1 \) and a selfadjoint polynomial \( P \in \mathbb{C}(X)_d \) (identified with \( I_k \otimes P \in M_k(\mathbb{C}(X)_d) \)) such that \( \|P(a_\tau)\| = 1 \) (for instance, \( P(X) = (X_{\tau,k}^{1,1} I) \)) and one picks a selfadjoint tuple \( A \) such that \( \|P(A)\|_\star > 1 \) (if \( P(X) = (X_{\tau,k}^{1,1} I) \), then any tuple \( A \) whose first coordinate is a positive matrix whose spectrum is included in \( (\|a_\tau\|, +\infty) \) will do). Then \( \Phi^*_k(A) \geq \|P^*(P)^*_k(A)\|_\star \geq \|P^*(P)^*_k(A)\|_\star \geq 1 \) for all sufficiently large \( n \).

(4) \( \Phi^2_{\tau,d}(A + B) \geq \max(\Phi^2_{\tau,d}(A), \Phi^2_{\tau,d}(B)), A \in M_{k_1}(\mathbb{C})^n, B \in M_{k_2}(\mathbb{C})^n \). Indeed, this follows easily from the fact that if \( P_j \in M_{k_j}(\mathbb{C}(X)_d), j = 1, 2, \) are maximizing elements in (3.13) corresponding to \( A \) and \( B \), respectively, then \( P_1 \oplus P_2 \) still satisfies \( \| (P_1 \oplus P_2)(a_\tau) \| \leq 1 \) and thus \( \Phi^2_{\tau,d}(A + B) \geq k_1k_2 \text{tr}_{k_1+k_2}((P_1 \oplus P_2)(I_{k_1+k_2})^2) = k_1k_2 \text{tr}_{k_1+k_2}(P(1)(I_{k_1})^2) + k_2 \text{tr}_{k_1+k_2}(P(2)(I_{k_2})^2) = k_1k_2 \Phi^2_{\tau,d}(A) + k_2 \Phi^2_{\tau,d}(B) \). If one of \( \Phi^2_{\tau,d}(A) \), \( \Phi^2_{\tau,d}(B) \) is infinite, then so is \( \Phi^2_{\tau,d}(A + B) \). Assume that \( \Phi^2_{\tau,d}(A) > \Phi^2_{\tau,d}(B) \). Pick a subsequence \( \{d_{\tau}\} \in \mathbb{N} \) on which \( \Phi^2_{\tau,d_{\tau}}(A) \geq \Phi^2_{\tau,d_{\tau}}(B) \) and \( \Phi^2_{\tau,d_{\tau}}(A) \) is
exists an \( \{ \phi \} \) such that the expression of the Bernstein-Markov property in terms of the (real) random variables \( \phi \) is a state, it has norm equal to one, so that \( (\text{tr} \otimes \phi)(Y) \) satisfies this equality, which satisfies this equality, and \( (\text{tr} \otimes \phi)(Y) \) is the classical definition of the Bernstein-Markov property: if \( \kappa = 1 \) and \( \phi \) is a probability measure on a Euclidean space, then the above is the expression of the Bernstein-Markov property in terms of the (real) random variables whose joint distribution is \( \phi \). For numerous details on this matter, we refer to [BLPW15, Blo97]. In this paper we will only consider the case when \( \phi = \tau \) is a positive bounded tracial state (usually, but not always, faithful), and thus \( \phi = \tau \) is a tuple of selfadjoint operators in a finite von Neumann algebra.

The conditions in the definition have some obvious reformulations: given any sequence of polynomials \( \{ P_d \}_{d \in \mathbb{N}}, \deg P_d \leq d \),

\[
\limsup_{d \to \infty} \left( \frac{\| P_d(\bar{a}_\phi) \|}{\| P_d(\bar{a}_\phi) \|_{L^2(\text{tr} \otimes \phi)}} \right)^{\frac{1}{d}} \leq 1;
\]

3.4. Level sets of the noncommutative Christoffel-Darboux polynomial: comparison of \( \Phi^*_\tau, \kappa_\tau \) and \( \kappa_\tau \). According to the second item of Remark 3.5, we find \( \kappa_{\tau,d}(\bar{a}, \bar{a}^*) = \max \{ \Phi^*_\tau(\bar{a}), \Phi^*_\tau(\bar{a}) \} \) as the maximum of \( \Phi(\bar{a})(I_k)I(\bar{a})^*(I_k)I(\bar{a}) + (PP^*) \) under the condition that \( (\text{Id}_\mathbf{M}_k(\mathbb{C}) \otimes \tau)(PP^*) = I_k \). For the element, call it \( \bar{P}_\bar{a} \), which satisfies this equality, we have tautologically \( \| \bar{P}_\bar{a}(\bar{a}) \|_{L^2(\text{tr} \otimes \phi)} = \text{tr}_k(I_k) = 1 \). In order to productively connect this to our versions of Siciak's extremal function, we need

**Definition 3.8.** Consider a unit-preserving positive bounded linear map \( \varphi : \mathbb{C}(\mathbb{X}) \to \mathbb{C} \) and a tuple of noncommutative selfadjoint random variables \( \bar{a}_\varphi \) in the von Neumann algebra generated by \( \varphi \) via the GNS construction such that the law of \( \bar{a}_\varphi \) is \( \varphi \). We say that \( \varphi \) satisfies the Bernstein-Markov property if for all \( d \in \mathbb{N} \) there exists an \( M_d = M(\tau)_d \geq 0 \) (possibly depending on \( k \)) such that

\[
\| P(\bar{a}_\varphi) \| \leq M_d(\text{tr}_k \otimes \varphi)(P(\bar{a}_\varphi)'P(\bar{a}_\varphi))^{\frac{1}{d}} = M_d\| P(\mathbb{X}) \|_{L^2(\text{tr}_k \otimes \varphi)}
\]

for all \( \mathbb{P} \in \mathbb{M}_k(\mathbb{C}(\mathbb{X})_d) \) of degree \( d \), and

\[
\limsup_{d \to \infty} M_d^{\frac{1}{d}} = 1 \text{ for each } k.
\]

As \( \text{tr}_k \otimes \varphi \) is a state, it has norm equal to one, so that \( (\text{tr}_k \otimes \varphi)(Y^*Y) \leq \| Y^*Y \| = \| Y \|^2 \), i.e. \( \| Y \|_{L^2(\text{tr}_k \otimes \varphi)} \leq \| Y \| \). Thus, automatically \( M_d \geq 1 \) for all \( d, k \).

This actually is the classical definition of the Bernstein-Markov property: if \( k = 1 \) and \( \varphi \) is a probability measure on a Euclidean space, then the above is the expression of the Bernstein-Markov property in terms of the (real) random variables whose joint distribution is \( \varphi \). For numerous details on this matter, we refer to [BLPW15, Blo97]. In this paper we will only consider the case when \( \varphi = \tau \) is a positive bounded tracial state (usually, but not always, faithful), and thus \( \bar{a}_\tau = \bar{a}_\tau \) is a tuple of selfadjoint operators in a finite von Neumann algebra.

The conditions in the definition have some obvious reformulations: given any sequence of polynomials \( \{ P_d \}_{d \in \mathbb{N}}, \deg P_d \leq d \),

\[
\limsup_{d \to \infty} \left( \frac{\| P_d(\bar{a}_\tau) \|}{\| P_d(\bar{a}_\tau) \|_{L^2(\text{tr} \otimes \tau)}} \right)^{\frac{1}{d}} \leq 1;
\]
(3.27) 
\[ \forall \varepsilon > 0 \exists C(\varepsilon, \tau) > 0 \text{ such that } \| P_d(\varrho_\tau) \| \leq C(\varepsilon, \tau)(1 + \varepsilon)^d \| P_d(\varrho_\tau) \|_{L^2(\tr_k \otimes \tau)}. \]

As of now, we are not aware of this notion having previously appeared in free probability or free analysis. In particular, it seems far from obvious the precise way it is related to established notions from free probability such as free entropy, conjugate variables, or even free independence. However, unlike in classical pluripotential theory, the Bernstein-Markov property appears to be somewhat less ubiquitous in the (highly) noncommutative context. We shall thus introduce an obvious modification:

**Definition 3.9.** Consider a bounded positive unital trace \( \tau : \mathcal{C}(\mathcal{X}) \to \mathbb{C} \) and a tuple of noncommutative selfadjoint random variables \( \varrho_\tau \) in the von Neumann algebra \( W^*(\tau) \) generated by \( \tau \) via the GNS construction such that the law of \( \varrho_\tau \) is \( \tau \). Given a real constant \( \epsilon \in [1, \infty) \), we say that \( \tau \) satisfies the \( \epsilon \)-Bernstein-Markov property if for all \( d \in \mathbb{N} \) there exists an \( M_d = M(\tau)_d \geq 0 \) (possibly depending on \( k \)) such that

\[ \| P_d(\varrho_\tau) \| \leq M_d(\tr_k \otimes \tr)(P_d(\varrho_\tau)^*P_d(\varrho_\tau))^{1/2} = M_d\| P(\mathcal{X}) \|_{L^2(\tr_k \otimes \tau)} \]

for all \( P \in \mathcal{M}_k(\mathcal{C}(\mathcal{X})_d) \) of degree \( d \), and

\[ \limsup_{d \to \infty} M_d^2 = \epsilon \text{ for each } k. \]

The next proposition states the main results of this section (compare with [BPSS20, Lemma 2.9] and its proof):

**Proposition 3.10.** For any given bounded positive trace \( \tau \) on \( \mathcal{C}(\mathcal{X}) \), we have

(3.28) \[ \limsup_{d \to \infty} \| \kappa_{\tau,d}(A, A^*) (I_k)^{1/2} \| \geq \Phi_\tau^\infty(A), \quad \sup_{d \in \mathbb{N}} \| \kappa_{\tau,d}(A, A^*) (I_k)^{1/2} \| \geq \Sigma_\tau^\infty(A), \]

(3.29) \[ \limsup_{d \to \infty} \tr_k(\kappa_{\tau,d}(A, A^*) (I_k)^{1/2}) \geq \Phi_\tau^2(A), \quad \sup_{d \in \mathbb{N}} \tr_k(\kappa_{\tau,d}(A, A^*) (I_k)^{1/2}) \geq \Sigma_\tau^2(A), \]

for all \( A \in \mathcal{M}_k(\mathcal{C})^n, k \in \mathbb{N} \). If \( \tau \) satisfies the \( \epsilon \)-Bernstein-Markov property 3.9, then

(3.30) \[ \mathcal{C}^2\Phi_\tau^\infty(A) \geq \limsup_{d \to \infty} \| \kappa_{\tau,d}(A, A^*) (I_k)^{1/2} \| \geq \Phi_\tau^\infty(A), \]

(3.31) \[ \mathcal{C}^2\Phi_\tau^2(A) \geq \limsup_{d \to \infty} \tr_k(\kappa_{\tau,d}(A, A^*) (I_k)^{1/2}) \geq \Phi_\tau^2(A), \]

(3.32) \[ \Sigma_\tau^\infty(A) \sup_{d \in \mathbb{N}} M_d^2 \geq \sup_{d \in \mathbb{N}} \| \kappa_{\tau,d}(A, A^*) (I_k)^{1/2} \| \geq \Sigma_\tau^\infty(A), \]

(3.33) \[ \Sigma_\tau^2(A) \sup_{d \in \mathbb{N}} M_d^2 \geq \sup_{d \in \mathbb{N}} \tr_k(\kappa_{\tau,d}(A, A^*) (I_k)^{1/2}) \geq \Sigma_\tau^2(A), \quad A \in \mathcal{M}_k(\mathcal{C})^n, k \in \mathbb{N}. \]

Before proving the proposition, we give some intuition for the Bernstein-Markov property by explaining how to verify it in practice. Observe first that in order to verify the \( \epsilon \)-Bernstein-Markov condition, one only needs to check it in the case \( k = 1 \). Indeed, if \( P_d(\varrho_\tau) = (P_{i,j}(\varrho_\tau))_{i,j=1}^k \), then \( P^*(\varrho_\tau) = (P^*_{j,i}(\varrho_\tau))_{i,j=1}^k \), so that

\[ P_d(\varrho_\tau)P^*(\varrho_\tau) = \left( \sum_{l=1}^k P_{i,l}(\varrho_\tau)P^*_{j,l}(\varrho_\tau) \right)_{i,j=1}^k. \]

\[ \| P(\mathcal{X}) \|_{L^2(\tr_k \otimes \tau)}^2 = \frac{1}{k} \sum_{i,l=1}^k \tr \left( P_{i,l}(\varrho_\tau)P^*_{i,l}(\varrho_\tau) \right) = \frac{1}{k} \sum_{i,l=1}^k \| P_{i,l}(\mathcal{X}) \|_{L^2(\tau)}^2, \]

as desired.
We write
\[
\|P(a_r)\|_{L^2(L^2(\tau) \otimes \tau)} = \frac{\sqrt{\mathcal{C}} \|P(a_r)\|}{\sum_{i,l=1}^{k} \|P_{i,l}(a_r)\|_{L^2(\tau)}^{2}} \leq \frac{\sqrt{\mathcal{C}} \sum_{i,l=1}^{k} \|P_{i,l}(a_r)\|_{L^2(\tau)}^{2}}{\sum_{i,l=1}^{k} \|P_{i,l}(a_r)\|_{L^2(\tau)}^{2}}^{\frac{1}{2}}.
\]

Assume \(\mathcal{C} = \mathcal{C}_k\) depends on \(k\). By considering diagonal polynomials, it is obvious that \(\{\mathcal{C}_k\}_k\) can only be non-decreasing. Consider a subsequence of polynomials \(P = P_d, d \in \mathbb{N}\) being the degree of \(P\), so that \(\lim_{d \to \infty} \left(\frac{\|P(a_r)\|}{\|P(a_r)\|_{L^2(L^2(\tau) \otimes \tau)}}\right)^{\frac{1}{d}} = \mathcal{C}_k\). That means,

\[
\mathcal{C}_k = \lim_{d \to \infty} \left(\frac{\|P(a_r)\|}{\|P(a_r)\|_{L^2(L^2(\tau) \otimes \tau)}}\right)^{\frac{1}{d}} \leq \lim_{d \to \infty} \left(\frac{\sqrt{\mathcal{C}} \sum_{i,l=1}^{k} \|P_{i,l}(a_r)\|_{L^2(\tau)}^{2}}{\sum_{i,l=1}^{k} \|P_{i,l}(a_r)\|_{L^2(\tau)}^{2}}^{\frac{1}{2}}\right)^{\frac{1}{d}}.
\]

(3.34) \(= \lim_{d \to \infty} \left(\frac{\|P_{i_0,l_0}(a_r)\|}{\|P_{i_0,l_0}(a_r)\|_{L^2(L^2(\tau) \otimes \tau)}}\right)^{\frac{1}{d}} \leq \lim_{d \to \infty} \left(\frac{\sqrt{\mathcal{C}} \sum_{i,l=1}^{k} \|P_{i,l}(a_r)\|_{L^2(\tau)}^{2}}{\sum_{i,l=1}^{k} \|P_{i,l}(a_r)\|_{L^2(\tau)}^{2}}^{\frac{1}{2}}\right)^{\frac{1}{d}}.
\]

This holds for any pair of indices \((i_0, l_0) \in \{1, \ldots, k\}^2\) corresponding to a non-zero entry. Obviously, for each such pair, the sequence \(\frac{\|P_{i_0,l_0}(a_r)\|}{\|P_{i_0,l_0}(a_r)\|_{L^2(L^2(\tau) \otimes \tau)}}\) indexed by the degrees \(d\) of the polynomials \(P\) has an uppermost limit point. We choose the pair \((i_0, l_0) \in \{1, \ldots, k\}^2\) for which this uppermost limit point is the largest; if there are two or more which reach this largest upper limit, we choose \((i_0, l_0)\) such that the proportion \(\frac{\|P_{i_0,l_0}(a_r)\|}{\|P_{i_0,l_0}(a_r)\|_{L^2(L^2(\tau) \otimes \tau)}}\) grows no slower than \(\frac{\|P_{i,l}(a_r)\|}{\|P_{i,l}(a_r)\|_{L^2(L^2(\tau) \otimes \tau)}}\) for any other pair \((i, l)\). We pass to a subsequence of \(d\)'s such that this uppermost limit is reached along a sequence which is fastest growing among all pairs \((i, l)\). By an abuse of notation, we still write limit after \(d\), that is, \(\lim_{d \to \infty} \left(\frac{\|P_{i_0,l_0}(a_r)\|}{\|P_{i_0,l_0}(a_r)\|_{L^2(L^2(\tau) \otimes \tau)}}\right)^{\frac{1}{d}}\), for the limit along this subsequence. We claim that the limit points of

\[
\left(\frac{1 + \sum_{i,l \neq (i_0,l_0)} \|P_{i,l}(a_r)\|}{1 + \sum_{i,l \neq (i_0,l_0)} \|P_{i,l}(a_r)\|^{\frac{2}{d}}}ight)^{\frac{1}{d}}
\]

along this same subsequence are bounded from above by one. Indeed, the only way for this to fail is if at least one \(\frac{\|P_{i_0,l_0}(a_r)\|}{\|P_{i_0,l_0}(a_r)\|_{L^2(L^2(\tau) \otimes \tau)}}\) tends very fast to infinity while all of \(\frac{\|P_{i,l}(a_r)\|}{\|P_{i_0,l_0}(a_r)\|_{L^2(L^2(\tau) \otimes \tau)}}\) either stay bounded or tend much slower to infinity. However,

\[
\frac{\|P_{i_0,l_0}(a_r)\|}{\|P_{i_0,l_0}(a_r)\|_{L^2(L^2(\tau) \otimes \tau)}} = \frac{\|P_{i_0,l_0}(a_r)\|}{\|P_{i_0,l_0}(a_r)\|_{L^2(\tau)}} \cdot \frac{\|P_{i_0,l_0}(a_r)\|_{L^2(\tau)}}{\|P_{i_0,l_0}(a_r)\|_{L^2(L^2(\tau) \otimes \tau)}}.
\]

and, according to our choice of \((i_0, l_0)\) and of the subsequence, for \(d\) large enough, \(\|P_{i_0,l_0}(a_r)\|_{L^2(L^2(\tau) \otimes \tau)}\) cannot grow more than twice
as fast as \( \frac{\|P_{(1,1)}(X)\|_{L^2(L^2)}}{\|P_{(0,0)}(\Delta)\|_{L^2(L^2)}} \). We conclude that

\[
\left( 1 + \sum_{(i,j) \neq (0,0)} \frac{\|P_{(i,j)}(\bar{\tau})\|}{\|P_{(i,j)}(\bar{\tau})\|_{L^2(L^2)}} \right)^{\frac{1}{2}} \leq \left( 1 + \sum_{(i,j) \neq (0,0)} \frac{\|P_{(i,j)}(X)\|_{L^2(L^2)}}{\|P_{(i,j)}(\bar{\tau})\|_{L^2(L^2)}} \right)^{\frac{1}{2}}
\]

is indeed bounded from above by 1. By (3.34), we conclude that \( C_k \leq C_1 \) for all \( k \in \mathbb{N} \), so \( \mathcal{C}_k = \mathcal{C} \) indeed does not depend on \( k \). It should also be noted that if \( \tau \) satisfies the \( \mathcal{C} \)-Bernstein-Markov property, then \( \mathcal{C} \leq \sup_d M_d^2 < \infty \), and the first inequality may be strict.

Second, in order to verify the \( \mathcal{C} \)-Bernstein-Markov property, it is enough to verify it on positive polynomials: there exists \( Q_d(X) \in \mathbb{C}(X)_{2d}, d \in \mathbb{N} \), with \( Q_d \geq 0 \) in \( \mathbb{C}(X) \),

\[
\limsup_{d \to \infty} \left( \frac{\|Q_d(\bar{\tau})\|}{\text{tr}(Q_d(\bar{\tau}))} \right)^{\frac{1}{2}} = \mathcal{C},
\]

with no value larger than \( \mathcal{C} \) being attainable. (Here we performed the substitution \( Q_d = P_d P_d^* \). For the Bernstein-Markov property (\( \mathcal{C} = 1 \)), it is enough to simply show that for any sequence \( Q_d \) as above, \( \limsup_{d \to \infty} \left( \frac{\|Q_d(\bar{\tau})\|}{\text{tr}(Q_d(\bar{\tau}))} \right)^{\frac{1}{2}} = 1 \). As an aside, we note that for any \( d \in \mathbb{N} \) the optimization problem

\[
\max\{\|Q_d(\bar{\tau})\|: \text{tr}(Q_d(\bar{\tau})) \leq 1, Q_d \geq 0, Q_d(X) \in \mathbb{C}(X)_d\}
\]

is convex, on a convex finite-dimensional set, hence solvable in principle.

These facts should probably arouse no astonishment in view of Theorem 3.4. Indeed, the connection of \( \kappa_{\tau,d} \) to the Bernstein-Markov property (respectively the \( \mathcal{C} \)-Bernstein-Markov property) becomes obvious in view of this theorem and Remark 3.5 following it: recall that

\[
\kappa_{\tau,d}(A^*)(I_k) = \max_{P \in M_k(\mathbb{C}(X)_d)} \frac{\text{P}(A)(I_k)^*(\text{Id}_{M_k(\mathbb{C})} \otimes \tau)(\text{PP}^*)^{-1}\text{P}(A)(I_k)}{1 + \frac{1}{\|P_{(0,0)}(\Delta)\|_{L^2(L^2)}}} = \max_{P \in M_k(\mathbb{C}(X)_d)} \frac{\|P(\bar{\tau})\|}{\|P_{(0,0)}(\Delta)\|_{L^2(L^2)}} = \kappa_{\tau,d}(A^*)(I_k),
\]

where we assume the obvious conditions of invertibility. As noted in Remark 3.5(2), the set of polynomials on which the above maximum is achieved is invariant under multiplication to the left with a scalar invertible \( k \times k \) complex matrix, so in the above we may pick a polynomial \( P \) so that \( \|P(\bar{\tau})\| = 1 \) and \( \|P(X)\|_{L^2(L^2)} = \|P(\bar{\tau})\|_{L^2(L^2)} \). Of course, \( \kappa_{\tau,d}(A^*)(I_k) \geq \frac{\|P(A)(I_k)^*(\text{Id}_{M_k(\mathbb{C})} \otimes \tau)(\text{PP}^*)^{-1}\text{P}(A)(I_k)}{1 + \frac{1}{\|P_{(0,0)}(\Delta)\|_{L^2(L^2)}}} = M_d \). Of course, \( \kappa_{\tau,d}(A^*)(I_k) \geq \frac{\|P(A)(I_k)^*(\text{Id}_{M_k(\mathbb{C})} \otimes \tau)(\text{PP}^*)^{-1}\text{P}(A)(I_k)}{1 + \frac{1}{\|P_{(0,0)}(\Delta)\|_{L^2(L^2)}}} = M_d \). Of course, \( \kappa_{\tau,d}(A^*)(I_k) \geq \frac{\|P(A)(I_k)^*(\text{Id}_{M_k(\mathbb{C})} \otimes \tau)(\text{PP}^*)^{-1}\text{P}(A)(I_k)}{1 + \frac{1}{\|P_{(0,0)}(\Delta)\|_{L^2(L^2)}}} = M_d \). Of course, \( \kappa_{\tau,d}(A^*)(I_k) \geq \frac{\|P(A)(I_k)^*(\text{Id}_{M_k(\mathbb{C})} \otimes \tau)(\text{PP}^*)^{-1}\text{P}(A)(I_k)}{1 + \frac{1}{\|P_{(0,0)}(\Delta)\|_{L^2(L^2)}}} = M_d \). Then the above majorization by \( \kappa_{\tau,d}(A^*)(I_k) \) holds for \( P_{\epsilon} \) as well. This allows us to write

\[
\kappa_{\tau,d}(A^*)(I_k) = \max_{P \in M_k(\mathbb{C}(X)_d)} \frac{\|P(X)\|_{L^2(L^2)}\|P(A)(I_k)^*(\text{Id}_{M_k(\mathbb{C})} \otimes \tau)(\text{PP}^*)^{-1}\text{P}(A)(I_k)}{1 + \frac{1}{\|P_{(0,0)}(\Delta)\|_{L^2(L^2)}}},
\]

imposing the same invertibility condition as for the previously displayed relation.
Proof of Proposition 3.10. With Definition 3.9 in hand, we assume that \( \tau \) does satisfy a \( \mathcal{C} \)-Bernstein-Markov condition. Then \( \frac{1}{M_d} \| P_A(\alpha_r) \| \leq 1 \) (recall the meaning of \( P_A \) from the very beginning of Section 3.4), so that \( \Phi_{\tau,d}^\infty(A) \geq \frac{1}{M_d} \| P_A(A)(I_k) \|^2 \), which, according to Remark 3.5(2), implies

\[
\Phi_{\tau}^\infty(A) \geq \limsup_{d \to \infty} \frac{1}{M_d} \| P_A(A)(I_k) \|^2 \geq \limsup_{d \to \infty} \frac{1}{M_d^{\frac{d}{2}}} \| \kappa_{\tau,d}(A, A^*)(I_k) \|^2.
\]

Applying the hypothesis that \( \tau \) satisfies the \( \mathcal{C} \)-Bernstein-Markov property yields

\[
\mathcal{C}^2 \Phi_{\tau}^\infty(A) \geq \limsup_{d \to \infty} \| \kappa_{\tau,d}(A, A^*)(I_k) \|^2.
\]

Conversely, if we consider the embodiment of \( \kappa_{\tau,d}(A, A^*)(I_k) \) as the maximum after \( P \) of \( P(A)(I_k)^* (Id_{M_d(C)} \otimes \tau)(PP^*)^{-1}P(A)(I_k) \) as in Theorem 3.4, then we may renormalize \( P \) with a scalar constant so that \( \| P(\alpha_r) \| = 1 \). It follows that

\[
\| \kappa_{\tau,d}(A, A^*)(I_k) \| \geq \| P(A)(I_k)^* (Id_{M_d(C)} \otimes \tau)(PP^*)^{-1}P(A)(I_k) \|
\]

for all \( P \), so \( \| \kappa_{\tau,d}(A, A^*)(I_k) \| \geq \Phi_{\tau,d}^\infty(A) \). This proves that

\[
\limsup_{d \to \infty} \| \kappa_{\tau,d}(A, A^*)(I_k) \|^2 \geq \Phi_{\tau,d}^\infty(A), \quad A \in M_k(C)^n, k \in \mathbb{N}.
\]

Under the assumption of \( \mathcal{C} \)-Bernstein-Markov property for \( \tau \), equation (3.36) guarantees that \( \mathcal{C}^2 \Phi_{\tau}^\infty(A) \geq \limsup_{d \to \infty} \| \kappa_{\tau,d}(A, A^*)(I_k) \|^2 \geq \Phi_{\tau,d}^\infty(A) \), which, for \( \mathcal{C} = 1 \) (i.e. the Bernstein-Markov property), yields

\[
\Phi_{\tau}^\infty(A) = \limsup_{d \to \infty} \| \kappa_{\tau,d}(A, A^*)(I_k) \|^2, \quad A \in M_k(C)^n, k \in \mathbb{N}.
\]

Let us consider next the other noncommutative version of Siciak’s extremal function, namely \( \Phi^2 \). For fixed \( d \), we have seen above that we may renormalize any polynomial \( P \) in Remark 3.5(2) so that \( \| P(\alpha_r) \| = 1 \), which yields \( \kappa_{\tau,d}(A, A^*)(I_k) \geq P(A)(I_k)^* (Id_{M_d(C)} \otimes \tau)(PP^*)^{-1}P(A)(I_k) \). As \( (Id_{M_d(C)} \otimes \tau)(PP^*) = (Id_{M_d(C)} \otimes tr)(P(\alpha_r)P(\alpha_r)^*) \), and \( (Id_{M_d(C)} \otimes tr) \) is a conditional expectation and hence (completely) positive, \( \| P(\alpha_r) \| \leq 1 \iff P(\alpha_r)P(\alpha_r)^* \leq I_k \iff (Id_{M_d(C)} \otimes tr)(P(\alpha_r)P(\alpha_r)^*) \leq I_k \), which yields the conclusion \( \kappa_{\tau,d}(A, A^*)(I_k) \geq P(A)(I_k)^* P(A)(I_k) \). Thus, \( \text{tr}_k(\kappa_{\tau,d}(A, A^*)(I_k)) \geq \text{tr}_k(P(A)(I_k)^* P(A)(I_k)) \) for all \( P \) as above, so that \( \text{tr}_k(\kappa_{\tau,d}(A, A^*)(I_k)) \geq \Phi_{\tau,d}^2(A) \).

As \( A \) is arbitrary, we obtain

\[
\limsup_{d \to \infty} \text{tr}_k(\kappa_{\tau,d}(A, A^*)(I_k)) \geq \Phi_{\tau}^2(A).
\]

Under the assumption of the \( \mathcal{C} \)-Bernstein-Markov property for \( \tau \), Remark 3.5(2) allows us to consider a polynomial \( P_A \in M_k(C(X)_d) \) such that \( \kappa_{\tau,d}(A, A^*)(I_k) = P_A(A)(I_k)^* P_A(A)(I_k) \) and \( (Id_{M_d(C)} \otimes tr)(P_A(\alpha_r)P_A(\alpha_r)^*) = I_k \). This naturally yields \( \Phi_{\tau,d}^2(A) \|^{1/d} \geq \text{tr}_k \left( \frac{P_A(A)(I_k)^* P_A(A)(I_k)}{\| P_A(A)(I_k) \|^2} \right) \|^{1/d} = \frac{1}{\| P_A(A)(I_k) \|^2} \text{tr}_k(\kappa_{\tau,d}(A, A^*)(I_k)) \|^{1/d}. \)

Definition 3.9 guarantees that

\[
\frac{1}{M_d^{1/d} \text{tr}_k(I_k)^{1/d}} = \frac{1}{M_d^{1/d}}. \quad \text{As } \limsup_{d \to \infty} M_d^{1/d} = \frac{1}{\mathcal{C}}, \text{ one has } \Phi_{\tau,d}^2(A) = \limsup_{d \to \infty} \Phi_{\tau,d}^2(A) \|^{1/d} \geq
\]
Together with the above-displayed relation, it yields the following estimate for distributions $\tau$ that satisfy the $C$-Bernstein-Markov inequality

$$\frac{1}{n} \limsup_{d\to\infty} \text{tr}_d (\kappa_{\tau,d}(A, A^*)(I_k)) \geq \Phi^2(A)$$

and the following equality for distributions that satisfy the Bernstein-Markov property:

$$\Phi^2(A) = \limsup_{d\to\infty} \text{tr}_d (\kappa_{\tau,d}(A, A^*)(I_k)) \geq \Phi^2(A)$$

Finally, let us establish the relation between the Christoffel-Darboux kernel of $\tau$ and the functions $\Sigma^\tau_1, \ldots, \Sigma^\tau_n \in \{1, \ldots, n\}$, introduced in (3.20)–(3.21). First observe that if $\sup_{d\in \mathbb{N}} (\Phi^\tau_{d,A}(A))^{1/d}$ is not achieved at a finite $d$, then $\Sigma^\tau_1 = \Phi^\tau_{d,A}(A)$, and relations (3.36), (3.37), and/or (3.39) hold. Thus, assume that $\Sigma^\tau_1(A) = \sup_{d\in \mathbb{N}} (\Phi^\tau_{d,A}(A))^{1/d}$ is achieved at a finite $d$. We thus have

$$\limsup_{d\to\infty} \text{tr}_d (\kappa_{\tau,d}(A, A^*)(I_k)) \geq \Phi^2(A)$$

where $\Phi^\tau_{d,A}(A) = \sup_{d\in \mathbb{N}} (\Phi^\tau_{d,A}(A))^{1/d}$.

As before, we pick $P_0 \in M_k(\mathbb{C})$ such that $\|P_0(a_r)\| = 1$ and $\Sigma^\tau_1(A) = \Phi^\tau_{d,A}(A)$, and relations (3.36), (3.37), and/or (3.39) hold. Then $\kappa_{\tau,d_0}(A, A^*)(I_k) = \Phi^\tau_{d_0}(A)(I_k) = \limsup_{d\to\infty} (\Phi^\tau_{d,A}(A))^{1/d} \geq \Phi^\tau_{d_0}(A)(I_k)$, which immediately implies $\text{tr}_d (\kappa_{\tau,d_0}(A, A^*)(I_k)) \geq \Phi^\tau_{d_0}(A)(I_k)$.

For the opposite inequality, we again take a polynomial $P_0 \in M_k(\mathbb{C})$, but this time chosen to achieve $\kappa_{\tau,d_0}(A, A^*)(I_k) = \Phi^\tau_{d_0}(A)(I_k)$, and $\kappa_{\tau,d_0}(A, A^*)(I_k) = \Phi^\tau_{d_0}(A)(I_k) = \limsup_{d\to\infty} (\Phi^\tau_{d,A}(A))^{1/d} \geq \Phi^\tau_{d_0}(A)(I_k) = \limsup_{d\to\infty} (\Phi^\tau_{d,A}(A))^{1/d} \geq \Phi^\tau_{d_0}(A)(I_k)$.

We assume that $\tau$ satisfies the $C$-Bernstein-Markov property for some $C \in (1, +\infty)$. Then, as noted before,

$$C \leq \sup_{d\in \mathbb{N}} M^\tau_d < \infty$$

depends only on $\tau$. We thus have $\|\kappa_{\tau,d_0}(A, A^*)(I_k)\|^{1/d_0} \leq \|P_0(a_r)\|^{1/d_0} = \|P_0(a_r)\|^{1/d_0}$.

tr$_d (\kappa_{\tau,d_0}(A, A^*)(I_k))^{1/d_0} \leq \|P_0(a_r)\|^{1/d_0}$. Since by definition $M_{d_0} \geq \|P_0(a_r)\|^{1/d_0}$, one has $\|\kappa_{\tau,d_0}(A, A^*)(I_k)\|^{1/d_0} \leq \sup_{d\in \mathbb{N}} M^\tau_d \left[ \|P_0(a_r)\|^{1/d_0} \right]^{1/d_0}$.

It should be noted that the above proposition does not state that the quantities involved are finite: neither of $\Phi^\tau_\infty, \Phi^\tau_\geq 2, \Sigma^\tau_1, \Sigma^\tau_\geq 2$ is guaranteed to be finite anywhere. The usefulness of this proposition in the study of the noncommutative Christoffel-Darboux kernel $\kappa_{\tau}$ depends on how well we understand the Siciak functions $\Phi$. In the classical pluripotential theory of several complex variables, the importance of $\Phi$ springs from the fact that it usually equals Green’s function associated to the support of the classical distribution $\tau$. In the noncommutative context, we do not even have a notion of support for $\tau$. We will go around this inconvenient fact (and also indirectly define a notion of support - or, in a certain sense, closure of support - for $\tau$) by making use of the classical theory of plurisubharmonic functions, which, as noted above, applies to the functions $\Phi$. 


We have noted that Definition 3.8 actually generalizes the classical Bernstein-Markov property. Let us put this in context through an example:

**Example 3.11.** Let \( \tau \) be the semicircular (Wigner) distribution on \( \mathbb{R} \): \( \tau : \mathbb{C}(X) \rightarrow \mathbb{C}, \tau(X^d) = \frac{2}{d+1} \int_1^d t^d \sqrt{t^2 - 1} \, dt \). It is well-known (see for instance [BLPW15]), and easy to prove, that \( \tau \) has the Bernstein-Markov property in the classical sense: for instance, the orthonormal polynomials of \( \tau \) are known to be the Chebyshev polynomials of the second kind, denoted \( \{U_d\}_{d \in \mathbb{N}} \), and \( U_d(\cos \theta) = \frac{\sin((d+1)\theta)}{\sin \theta} \), \( d \in \mathbb{N} \). The rate of growth of their infinity norm is of order \( d \) as the degree \( d \) tends to infinity (more precisely, the norm is reached at the boundary of the support, i.e. when \( \theta = 0 \) or \( \theta = \pi \), and then \( |U_d(\pm 1)| = d + 1 \)). If one writes an arbitrary polynomial \( P \in \mathbb{C}(X)_d \) as \( P(X) = \sum_{v=0}^d c_v U_v(X) \), then \( \|P(a_r)\| \leq \sum_{v=0}^d |c_v||U_v(a_r)| \leq \sum_{v=0}^d |c_v|(v + 1) \leq (d + 1) \sum_{v=0}^d |c_v| \), while \( \|P(a_r)\|_2 = \sqrt{\sum_{v=0}^d |c_v|^2} \). By the Schwarz-Cauchy inequality, \( \sum_{v=0}^d |c_v| \leq \sqrt{d + 1} \sum_{v=0}^d |c_v|^2 \), so that \( \|P(a_r)\| \leq (d + 1)^{\frac{3}{2}} \|P(a_r)\|_2 \) for all \( P \in \mathbb{C}(X)_d \). Since \( \lim_{d \rightarrow \infty} (d + 1)^{\frac{3}{2}} = 1 \), \( \tau \) satisfies the classical Bernstein-Markov property. As shown above, this means \( \tau \) satisfies the Bernstein-Markov property for matrix-valued polynomials as well.

Now, Proposition 3.10 shows that \( \Phi_{\tau}(A) = \lim sup_{d \rightarrow \infty} \|\kappa_{\tau,d}(A, A^*)(I_k)\| \) and \( \Phi_{\tau}^2(A) = \lim sup_{d \rightarrow \infty} \text{tr}_1(\kappa_{\tau,d}(A, A^*)(I_k))^{\frac{1}{2}} \), \( A \in M_k(\mathbb{C}) \).

Functional calculus shows that for fixed \( k \) the growth of \( \|\kappa_{\tau,d}(A, A^*)(I_k)\| \) as \( d \rightarrow \infty \) is governed by the eigenvalues of \( A \), so the existence of \( \lim_{d \rightarrow \infty} \|\kappa_{\tau,d}(A, A^*)(I_k)\| \) is simply guaranteed by the convergence of \( \kappa_{\tau,d}(z, \overline{z})(1)^{\frac{1}{2}} \) when \( z \in \mathbb{C} \), which is known from classical, one-variable potential theory. Specifically, by using analytic functional calculus, one obtains a manageable expression for \( U_d(A) \) when \( A \in M_k(\mathbb{C}) \) is in upper triangular form \( A = D + T \), where \( D \) is diagonal and \( T \) strictly upper triangular. With the notation \( D = \text{diag}(z_1, \ldots, z_k), T = (t_{i,i+j})_{1 \leq i, j, k \leq k-i} \), we make the assumptions that \( z_i \neq z_j \) if \( i \neq j \) (which excludes a set of Lebesgue measure zero in \( M_k(\mathbb{C}) \)), and that \( |z_1| \leq \cdots \leq |z_k| \) (which implies no loss of generality, as any upper triangular matrix is unitarily equivalent to one whose eigenvalues are ordered increasing)y). We then have the following formula for \( U_d(A) \):

\[
(U_d(A))_{i,i+j} = \left[ \frac{U_d(z_i)}{z_i - z_{i+j}} + \frac{U_d(z_{i+j})}{z_{i+j} - z_i} \right] t_{i,i+j} + \sum_{l=1}^{j-1} \sum_{1 < i_1 < \cdots < i_l < i_l+i+j} \sum_{s=i_1, \ldots, i_l, i+l, j} \frac{U_d(z_s)}{\prod_{r=i_1, \ldots, i_l, i+l, j, r \neq s} (z_s - z_r)} t_{i_{l+1}, i_{l+2}, \ldots, i_l+j},
\]

\[
(U_d(A))_{i,i} = U_d(z_i), \quad \text{while} \quad (U_d(A))_{i,k} = 0 \quad \text{if} \quad k < i.
\]

The formula does not make sense if \( z_r = z_s \) for some \( r \neq s \), but the expression is Lipschitz in all its diagonal entries, with derivatives of \( U_q \) occurring if two or more diagonal entries of \( A \) happen to be equal (specifically, if \( q \) entries are equal, then the derivatives of order up to and including \( q - 1 \) appear). The terms in the above sum can be easily seen to be indexed by the paths in the upper right corner between the diagonal entry \((i, i)\) and the diagonal entry \((i + j, i + j)\) which are formed by segments leaving from and
returning to diagonal entries. For instance,
\[
U_d \left( \begin{bmatrix}
  z_1 & t_{1,2} & t_{1,3} \\
  0 & z_2 & t_{2,3} \\
  0 & 0 & z_3
\end{bmatrix} \right) =
\begin{bmatrix}
  U_d(z_1) & U_d(z_2) & U_d(z_3) \\
  0 & U_d(z_2) & U_d(z_3) \\
  0 & 0 & U_d(z_3)
\end{bmatrix}
\]
for any \( z_1 \neq z_2 \neq z_3 \neq z_1 \in \mathbb{C} \). If any two of the three diagonal entries happen to be equal to each other, the first derivative of \( U_d \) evaluated in that entry is involved in the off-diagonal entries, and if all three are equal, so is the second derivative of \( U_d \). (Probably the easiest way to see these facts is by applying analytic functional calculus and the formulae for resolvents of upper triangular matrices: \( U_d(A) = (2\pi i)^{-1} \int_{\gamma} U_d(\zeta) (\zeta I_k - A)^{-2} d\zeta \) for any smooth curve \( \gamma \) surrounding \( \sigma(A) \) once.)

More general examples can be derived from the estimates obtained in Section 3.3.9: specifically, assuming that the spectrum \( \sigma(a_r) \) of our generating variable is sufficiently large so that the compact subset \( S_{n_r,k} := \{ A = A^* \in \mathcal{M}_k(\mathbb{C}): \sigma(A) \subset \sigma(a_r) \} \) of \( S_k \) is not pluripolar in \( \mathcal{M}_k(\mathbb{C}) = \mathbb{C}^{k^2} \) (which is the case, for instance, if \( \sigma(a_r) \) has nonempty interior\(^{2}\)), which happens trivially for the semicircular) we have seen that \( \Phi^2_{d,\tau}(A) \leq 1 \) for all \( A \in S_{n_r,k} \). This, together with the estimate in Section 3.3.8, implies that both \( \Sigma^2 \) and \( \Phi^2 \) are finite q.e. on \( \mathcal{M}_k(\mathbb{C}) \) (see, for instance, [ST97, Theorem 1.6, Appendix B]). Moreover, they are also no greater than one on \( S_{n_r,k} \). Since \( \| \cdot \|_2 \) and \( \| \cdot \| \) are comparable on \( \mathcal{M}_k(\mathbb{C}) \), it follows that \( \Sigma^\infty \) and \( \Phi^\infty \) are finite q.e. on \( \mathcal{M}_k(\mathbb{C}) \) as well.

A typical noncommutative context is provided by the case of free products. Let us consider next such an example.

**Example 3.12.** Consider two copies \( \tau_1, \tau_2 \) of the semicircular (Wigner) distribution from Example 3.11 above, and take their free product \( \tau = \tau_1 * \tau_2 \colon \mathbb{C}(X_1, X_2) \to \mathbb{C} \): \( \tau \) is completely specified by the following evaluation rules: \( \tau = \tau_j(P(X)), j = 1, 2 \), and \( \tau(P_{i_1}(X_{i_1}) P_{i_2}(X_{i_2}) \cdots P_{i_r}(X_{i_r})) = 0 \) whenever \( r \in \mathbb{N}, r \geq 1, i_1, i_2, \ldots, i_r \in \{1, 2\} \) are such that \( i_1 \neq i_2 \neq \cdots \neq i_r \) and \( \tau(P_{i_k}(X_{i_k})) = 0, 1 \leq j \leq r \). It is known that \( \tau \) is a faithful, bounded, tracial state (see [Voi85] for details). As mentioned in Example 3.2(3), a family of orthonormal polynomials for \( \tau \) is \( \{ U_w \}_{w \in \{X_1, X_2\}} \), \( U_w(X_1, X_2) = U_{d_1}(X_{i_1}) U_{d_2}(X_{i_2}) \cdots U_{d_r}(X_{i_r}) \) whenever \( w = i_1^{d_1}, i_2^{d_2}, \ldots, i_r^{d_r}, r \in \mathbb{N}, i_1, i_2, \ldots, i_r \in \{1, 2\}, i_1 \neq i_2 \neq \cdots \neq i_r \). Here, as before, \( U_{d_k}(X_{i_k}) \) are the Chebyshev polynomials of the second kind introduced in Example 3.11 above.

First, it follows from Section 3.3.2 (and has been argued in item (3) in the list following Section 3.3.7) that \( \Phi_{\tau_1}(A_1, A_2) \geq \max \{ \Phi_{\tau_2}(A_1), \Phi_{\tau_2}(A_2) \} \) for \( A_1, A_2 \in \mathcal{M}_k(\mathbb{C}) \) (with a similar statement for \( \Sigma \)), so that \( \Phi_{\tau_1}, \Sigma_{\tau_1} \) are not constantly 1.

Let us show that \( \Sigma_{\tau_2} \), and hence \( \Phi_{\tau_2} \in \{2, \infty\} \), is finite q.e. For an arbitrary \( f \in \mathcal{M}_k(\mathbb{C}(X_1, X_2)) \), the following Fourier-like expansion holds:
\[
f(X_1, X_2) = \sum_{w \in \{X_1, X_2\}\_d} \langle f, I_k \otimes U_w \rangle \otimes U_w(X_1, X_2)
\]

\(^{2}\)This immediately implies that the set \( S_{n_r,k} \) has nonempty interior in \( S_k \), so that it cannot be pluripolar.
\[
    = \sum_{w \in \langle X_1, X_2 \rangle_d} (\text{Id}_{M_k(C)} \otimes \tau)((I_k \otimes U^*_w)f) \otimes U_w(X_1, X_2).
\]

Then
\[
f(A_1, A_2)(I_k)^*f(A_1, A_2)(I_k)
= \sum_{v,w \in \langle X_1, X_2 \rangle_d} U_v(A_1, A_2)^*(\text{Id}_{M_k(C)} \otimes \tau)(f^*I_k \otimes U_v)(\text{Id}_{M_k(C)} \otimes \tau)(I_k \otimes U^*_w)f)U_w(A_1, A_2).
\]

For \(v = w\), then
\[
U_w(A_1, A_2)^*(\text{Id}_{M_k(C)} \otimes \tau)(f^*I_k \otimes U_w)(\text{Id}_{M_k(C)} \otimes \tau)(I_k \otimes U^*_w)f)U_w(A_1, A_2)
\leq \|f(a_{\tau_1}, a_{\tau_2})\|^2 U_w(A_1, A_2)^*|\tau(U_wU^*_w)I_k|U_w(A_1, A_2)
= \|f(a_{\tau_1}, a_{\tau_2})\|^2 U_w(A_1, A_2)^*U_w(A_1, A_2),
\]

for all \(w \in \langle X_1, X_2 \rangle_d\). Elements \(v \neq w \in \langle X_1, X_2 \rangle_d\) appear obviously in pairs, so we can group them accordingly: \(U_v(A_1, A_2)^*(\text{Id}_{M_k(C)} \otimes \tau)(f^*I_k \otimes U_v)(\text{Id}_{M_k(C)} \otimes \tau)(I_k \otimes U^*_w)f)U_w(A_1, A_2) + U_w(A_1, A_2)^*(\text{Id}_{M_k(C)} \otimes \tau)(f^*I_k \otimes U_w)(\text{Id}_{M_k(C)} \otimes \tau)(I_k \otimes U^*_w)f)U_v(A_1, A_2)\)
\[
\leq \|f(a_{\tau_1}, a_{\tau_2})\|^2 (U_w(A_1, A_2)^*U_w(A_1, A_2) + U_v(A_1, A_2)^*U_v(A_1, A_2)).
\]

We obtain the (rather brutal) majorization
\[
f(A_1, A_2)(I_k)^*f(A_1, A_2)(I_k)
\leq \|f(a_{\tau_1}, a_{\tau_2})\|^2 \left[ \sum_{w \in \langle X_1, X_2 \rangle_d} U_w(A_1, A_2)^*U_w(A_1, A_2) \right] + \sum_{v,w \in \langle X_1, X_2 \rangle_d \setminus \{x \}} U_w(A_1, A_2)^*U_w(A_1, A_2) + U_v(A_1, A_2)^*U_v(A_1, A_2).
\]

We estimate next the product \(U_v(A_1, A_2)^*U_v(A_1, A_2)\). Given the specific form of \(U_v(X_1, X_2)\), let us first focus on \(U_d(X)\). As \(\left\| \frac{P(A)(I_k)}{P(A)(a_{\tau_1})} \right\|^2 \leq \Phi_{\tau_1}^d(A)\) according to the definition of \(\Phi_{\tau_1}^d\), by employing the embedding \(U_d \rightarrow I_k \otimes U_d\) we have the estimate \(\|U_d(A_j)\|^2 \leq \|U_d^d(a_{\tau_1})\|^2 \Phi_{\tau_1}^d(A_j)\). According to (3.20) and (3.21), by taking power \(1/d\), we obtain \(\|U_d(A_j)\|^2 \leq \|U_d(a_{\tau_1})\|^2 \Sigma_{\tau_1}^d(A_j), A_j \in M_k(C), k \in \mathbb{N}\).

Consider the case \(\bullet = \infty\). Given \(w = X^{d_1}X^{d_2} \cdots X^{d_r} \in \langle X_1, X_2 \rangle_d\), we have
\[
\|U_w(A_1, A_2)^*U_w(A_1, A_2)\|
\leq \|U_{d_1}(A_i)\|^2 \cdots \|U_{d_r}(A_i)\|^2
\leq \|U_{d_1}(a_{\tau_1})\|^2 \cdots \|U_{d_r}(a_{\tau_r})\|^2 \Sigma_{\tau_1}^d(A_i)^{d_1} \cdots \Sigma_{\tau_r}^d(A_i)^{d_r}
\leq (d_1 + 1)^3 \cdots (d_r + 1)^3 \Sigma_{\tau_1}^d(A_i)^{d_1} \cdots \Sigma_{\tau_r}^d(A_i)^{d_r}.
\]
We choose the $w$ for which $\max_{v \in (X_1, X_2)_{d}} \|U_v(A_1, A_2)\|$ is achieved. By a most brutal estimate in (3.41),
\[
\|f(A_1, A_2)(I_k)\|^2 \leq 2^{2d}\|f(\tau_1, \tau_2)\|^2 \max_{v \in (X_1, X_2)_{d}} \|U_v(A_1, A_2)\|
\]
\[
\leq 4^d\|f(\tau_1, \tau_2)\|^2(d_i + 1)^3 \cdots (d_i + 1)^3 \Sigma_{\tau_1}^{\infty}(A_i)^{d_i} \cdots \Sigma_{\tau_r}^{\infty}(A_r)^{d_r},
\]
so that, by dividing with $\|f(\tau_1, \tau_2)\|^2$ and taking power $1/d$,
\[
\Sigma_{\tau}^{\infty}(A_1, A_2) \leq 4([d_i + 1] \cdots [d_i + 1])^{3/d}\Sigma_{\tau_1}^{\infty}(A_i)^{d_i} \cdots \Sigma_{\tau_r}^{\infty}(A_r)^{d_r}/d
\]
(3.42)
\[
< 32\max\{\Sigma_{\tau_1}^{\infty}(A_1), \Sigma_{\tau_2}^{\infty}(A_2)\}.
\]
This shows that all of $\Sigma_{\tau}^{\infty}(A_1, A_2), \Sigma_{\tau}^{2}(A_1, A_2), \Phi_{\tau}^{\infty}(A_1, A_2), \Phi_{\tau}^{2}(A_1, A_2)$ are non-trivial and finite q.e.

By Remark 3.5(4), the characterization of $\kappa_{\tau,d}$ as minimum over certain sets of polynomials as in Theorem 3.4 (and parts (2) and (3) of the same Remark 3.5) holds also when $\kappa_{\tau,d}$ is evaluated on tuples of elements in a finite von Neumann algebra. As the entire proof of Proposition 3.10 is based on this characterization, it follows that it holds also for $\mathcal{A}_{\tau}$ replaced by $\mathcal{A}_{\tau'} \in W^*(\tau)^n$. In particular, in our case, $\text{tr}(\kappa_{\tau,d}((a_{\tau_1}, a_{\tau_2}),(a_{\tau_1}, a_{\tau_2}))(1)) = 2^{d}$ from the definition of $\kappa_{\tau,d}$ and of orthonormal polynomials. On the other hand, by definition $\Phi_{\tau,d}^{2}((a_{\tau_1}, a_{\tau_2})) = \sup\{\text{tr}(\mathcal{P}((a_{\tau_1}, a_{\tau_2}))(1))\mathcal{P}((a_{\tau_1}, a_{\tau_2}))(1): \|\mathcal{P}((a_{\tau_1}, a_{\tau_2}))(1)\| \leq 1, \mathcal{P} \in W^*(\tau_1 \ast \tau_2) \otimes \mathbb{C} (\mathcal{X}_{d})\} \leq 1$ because $\text{tr}(\mathcal{P}((a_{\tau_1}, a_{\tau_2}))(1))\mathcal{P}((a_{\tau_1}, a_{\tau_2}))(1) \leq \|\mathcal{P}((a_{\tau_1}, a_{\tau_2}))\|^2$. Thus, $\Phi_{\tau}^{2}((a_{\tau_1}, a_{\tau_2})) = 1$, while $\lim_{d \to \infty} \text{tr}(\kappa_{\tau,d}((a_{\tau_1}, a_{\tau_2}),(a_{\tau_1}, a_{\tau_2}))(1)) = 2$. Thus, the free product $\tau = \tau_1 \ast \tau_2$ does not satisfy the Bernstein-Markov property, even though it does satisfy a $\mathcal{C}$-Bernstein-Markov property for some $2 \leq \mathcal{C} \leq 8\sqrt{2}$, as it will be seen in Remark 3.14.

The reader might legitimately hope that some of the inequalities in (3.30)-(3.31) can be eventually shown to be equalities even if $\mathcal{C} > 1$. It will follow from Theorem 3.13 below, the remark following it, and Examples 3.12 and 3.11 that this is unlikely to happen in many cases.

Some of the methods from Example 3.12 can be applied in a more general context in order to provide some estimates the noncommutative Siciak function. In view of (3.23), the following theorem appears to be intimately related to Voiculescu’s asymptotic freeness result for independent unitarily invariant random matrices, and can be viewed as a partial free analogue of Siciak’s Theorem [Kli91, Theorem 5.1.8].

**Theorem 3.13.** Let $\mathcal{X}_1 = (X_1, \ldots, X_n)$ and $\mathcal{X}_2 = (X_{n+1}, \ldots, X_{n+m})$ be selfadjoint non-commuting indeterminates. Assume $\tau_1: \mathcal{C}(\mathcal{X}_1) \to \mathcal{C}$ and $\tau_2: \mathcal{C}(\mathcal{X}_2) \to \mathcal{C}$ are two faithful positive bounded traces, and denote $\tau = \tau_1 \ast \tau_2: \mathcal{C}(\mathcal{X}_1, \mathcal{X}_2) \to \mathcal{C}$ their free product. Suppose that $\tau_j$ satisfies the $\mathcal{C}_j$-Bernstein-Markov property for some $\mathcal{C}_j \in [1, +\infty)$, $j = 1, 2$. Then for each $k \in \mathbb{N}$, one has
\[
\max\{\Sigma_{\tau_1}^{\infty}(\mathcal{A}_1), \Sigma_{\tau_2}^{\infty}(\mathcal{A}_2)\} \leq \Sigma_{\tau}^{\infty}(\mathcal{A}_1, \mathcal{A}_2)
\]
(3.43)
\[
\leq (m + n)^2 \max\left\{\Sigma_{\tau_1}^{\infty}(\mathcal{A}_1)\sup_{q \in \mathbb{N}} M(\tau_1)^{\frac{2}{q}}, \Sigma_{\tau_2}^{\infty}(\mathcal{A}_2)\sup_{q \in \mathbb{N}} M(\tau_2)^{\frac{2}{q}}\right\},
\]
for all $(\mathcal{A}_1, \mathcal{A}_2) \in \mathbb{M}_k(\mathbb{C})^n \times \mathbb{M}_k(\mathbb{C})^m$, and
\[
\max\{\Phi_{\tau_1}^{\infty}(\mathcal{A}_1), \Phi_{\tau_2}^{\infty}(\mathcal{A}_2)\} \leq \Phi_{\tau}^{\infty}(\mathcal{A}_1, \mathcal{A}_2)
\]
\[(3.44) \quad \leq (m + n)^2 \max \left\{ \Phi_{i, \tau}^\infty(A) \sup_{q \in \mathbb{N}} M(\tau_1)^{2q}, \Phi_{\tau}^\infty(A) \sup_{q \in \mathbb{N}} M(\tau_2)^{2q} \right\}, \]

for all \((A_1, A_2) = (A_1, A_2)^* \in M_k(\mathbb{C})^n \times M_k(\mathbb{C})^m\). In particular, under the hypotheses of the theorem, the functions \(\Sigma_{\tau_1}^x, \Sigma_{\tau_2}^x, \Phi_{\tau_1}^x, \text{ and } \Phi_{\tau_2}^x\) are finite q.e. on the Euclidean space \(M_k(\mathbb{C})^n \times M_k(\mathbb{C})^m\) and nontrivial.

**Proof.** As in Example 3.12, our proof parallels to some extent the proof provided in [Kli91] for Theorem 5.1.8. We start by recalling that \(\Phi_{\tau_1 \times \tau_2}^x(A_1, A_2) \geq \max\{\Phi_{\tau_1}^x(A_1), \Phi_{\tau_2}^x(A_2)\}\), \(\Sigma_{\tau_1 \times \tau_2}^x(A_1, A_2) \geq \max\{\Sigma_{\tau_1}^x(A_1), \Sigma_{\tau_2}^x(A_2)\}\). Indeed, this has nothing to do with freeness, and is a conclusion of Section 3.3.2, as it has been argued in item (3) in the list following Section 3.3.9. The proof of the other inequality is more challenging.

The third item of Example 3.2 (as shown in [Ans10, Section 3]) provides an explicit formula for orthonormal polynomials corresponding to free products of distributions. If \(\{P_{w, \tau_1}\}_{w \in \langle X_1 \rangle}\) and \(\{P_{w, \tau_2}\}_{w \in \langle X_2 \rangle}\) are the orthonormal polynomials associated to noncommutative distributions \(\tau_1\) and \(\tau_2\), respectively, then the orthonormal polynomials \(\{P_{w, \tau}\}_{w \in \langle X_1, X_2 \rangle}\) corresponding to \(\tau = \tau_1 \times \tau_2\) are indexed by words in letters from \(X_1\) and \(X_2\), and if \(w = u_1 u_2 u_3 \cdots u_r \in \langle X_1, X_2 \rangle, u_j \in \langle X_j \rangle, 1 \leq j \leq r, r \in \mathbb{N}\), where \(i_1 \neq i_2 \neq \cdots \neq i_r\), then

\[P_{w, \tau}(X_1, X_2) = P_{u_1, \tau_1}(X_1) P_{u_2, \tau_2}(X_2) P_{u_3, \tau_3}(X_3) \cdots P_{u_r, \tau_r}(X_r).\]

Pick \(d \in \mathbb{N}\) and \(f \in M_k(\mathbb{C}(X_1, X_2)_d)\) such that \(\|f(A_1, A_2)\| = 1\). Then

\[f(X_1, X_2) = \sum_{w \in \langle X_1, X_2 \rangle_d} \langle f, I_k \otimes P_{w, \tau} \rangle \otimes P_{w, \tau}(X_1, X_2)\]

\[= \sum_{w \in \langle X_1, X_2 \rangle_d} (Id_{M_k(\mathbb{C})} \otimes \tau)(I_k \otimes P_{w, \tau}^\ast f) \otimes P_{w, \tau}(X_1, X_2).\]

It then follows that

\[f(A_1, A_2)(I_k)^\ast f(A_1, A_2)(I_k) = \left[ \sum_{w \in \langle X_1, X_2 \rangle_d} (Id_{M_k(\mathbb{C})} \otimes \tau)(I_k \otimes P_{w, \tau}^\ast f) P_{w, \tau}(A_1, A_2) \right]^\ast\]

\[\times \left[ \sum_{w \in \langle X_1, X_2 \rangle_d} (Id_{M_k(\mathbb{C})} \otimes \tau)(I_k \otimes P_{w, \tau}^\ast f) P_{w, \tau}(A_1, A_2) \right]\]

\[= \sum_{w \in \langle X_1, X_2 \rangle_d} P_{w, \tau}(A_1, A_2)^\ast (Id_{M_k(\mathbb{C})} \otimes \tau)(f^\ast I_k \otimes P_{w, \tau}^\ast f) (Id_{M_k(\mathbb{C})} \otimes \tau)(I_k \otimes P_{w, \tau}^\ast f) P_{w, \tau}(A_1, A_2).\]

As in Example 3.12, we obtain

\[f(A_1, A_2)(I_k)^\ast f(A_1, A_2)(I_k) \leq \|f(A_1, A_2)\|^2 \sum_{w \in \langle X_1, X_2 \rangle_d} P_{w, \tau}(A_1, A_2)^\ast P_{w, \tau}(A_1, A_2) + \sum_{w \in \langle X_1, X_2 \rangle_d} P_{w, \tau}(A_1, A_2)^\ast P_{w, \tau}(A_1, A_2)\]

\[(3.45) \quad + \sum_{w \in \langle X_1, X_2 \rangle_d} P_{w, \tau}(A_1, A_2)^\ast P_{w, \tau}(A_1, A_2) + P_{w, \tau}(A_1, A_2)^\ast P_{w, \tau}(A_1, A_2)\].\]
This holds for any polynomial \( f \in \mathcal{M}_k(\mathbb{C}) \otimes \mathbb{C}(\mathbf{X}_1, \mathbf{X}_2)_d \) of degree at most \( d \) which has norm one on \( (\mathbf{u}_i, \mathbf{u}_j) \).

Again as in Example 3.12,
\[
P_{w, \tau}(A_1, A_2)^* P_{w, \tau}(A_1, A_2)
\leq \|P_{w, \tau}(A_1, A_2)^* P_{w, \tau}(A_1, A_2)\|
\leq \|P_{w, \tau}(A_1, A_2)\|^2 \|P_{w, \tau}(A_1, A_2)\|^2 \|P_{w, \tau}(A_1, A_2)\|^2 \cdots \|P_{w, \tau}(A_1, A_2)\|^2,
\]
where we remind the reader that \( w = u_1 u_2 u_3 \cdots u_r \in (\mathbf{X}_1, \mathbf{X}_2)_1, u_j \in (\mathbf{X}_j), 1 \leq j \leq r, r \in \mathbb{N}, \) where \( i_1 \neq i_2 \neq i_3 \neq \cdots \neq i_r \) and \(|u_1| + |u_2| + \cdots + |u_r| \leq d (|u_j| \) denotes the length of the word \( u_j \). This implies
\[
P_{w, \tau}(A_1, A_2)^* P_{w, \tau}(A_1, A_2)
\leq \|P_{w, \tau}(A_1, A_2)\|^2 \|P_{w, \tau}(A_1, A_2)\|^2 \|P_{w, \tau}(A_1, A_2)\|^2 \cdots \|P_{w, \tau}(A_1, A_2)\|^2,
\]
and in particular \( \|P_{w, \tau}(A_1, A_2)\||^2 \leq \|P_{w, \tau}(A_1, A_2)\|^2 \|P_{w, \tau}(A_1, A_2)\|^2 \|P_{w, \tau}(A_1, A_2)\|^2 \cdots \|P_{w, \tau}(A_1, A_2)\|^2 \).

By definition, \( \frac{\|P_{j}(A_i)(I_k)\|^2}{\|P_{j}(I_k)\|^2} \leq \|P_{j}(A_i)\| \) for any \( P(\mathbf{X}_j) \in \mathcal{M}_k(\mathbb{C}(\mathbf{X}_j)_d), \) that is, for any polynomial \( P \) with \( k \times k \) scalar matrix coefficients of degree at most \( d \). We embed the orthonormal polynomials in \( \mathcal{M}_k(\mathbb{C}(\mathbf{X}_j)_d) \) via the obvious identification \( P(\mathbf{X}_j) \mapsto I_k \otimes P(\mathbf{X}_j), \) which is isometric on operators. Thus, \( \|P_{\tau_j, \tau_j}(A_i,j)\|^2 = \|P_{\tau_j, \tau_j}(A_i,j)\|^2 \leq \|P_{\tau_j, \tau_j}(A_i,j)\|^2 \|P_{\tau_j, \tau_j}(A_i,j)\|^2 \|P_{\tau_j, \tau_j}(A_i,j)\|^2 \cdots \|P_{\tau_j, \tau_j}(A_i,j)\|^2 \), which implies \( \|P_{\tau_j, \tau_j}(A_i,j)\|^2 \leq \|P_{\tau_j, \tau_j}(A_i,j)\|^2 \|P_{\tau_j, \tau_j}(A_i,j)\|^2 \|P_{\tau_j, \tau_j}(A_i,j)\|^2 \cdots \|P_{\tau_j, \tau_j}(A_i,j)\|^2 \).

The cardinality of \( (\mathbf{X}_1, \mathbf{X}_2)_d \) is \( (m + n)^d \). The number of summands in (3.45) is thus \( (m + n)^{d(m + n)^d + 1} \), which we agree to majorize by \( (m + n)^{2d} \). We majorize the norm of the right-hand side of (3.45) by \( (m + n)^{2d} \) times the maximum after \( w \) of \( \|P_{w, \tau}(A_1, A_2)\|^2 \). We have
\[
\|P_{w, \tau}(A_1, A_2)\|^2 \leq \|P_{w, \tau}(A_1, A_2)\|^2 \|P_{w, \tau}(A_1, A_2)\|^2 \|P_{w, \tau}(A_1, A_2)\|^2 \cdots \|P_{w, \tau}(A_1, A_2)\|^2 \leq \|P_{w, \tau}(A_1, A_2)\|^2 \|P_{w, \tau}(A_1, A_2)\|^2 \|P_{w, \tau}(A_1, A_2)\|^2 \cdots \|P_{w, \tau}(A_1, A_2)\|^2 \times \Sigma_{\tau_j, \tau_j}(A_i,j)^{2/|u_j|} \Sigma_{\tau_j, \tau_j}(A_i,j)^{2/|u_j|} \Sigma_{\tau_j, \tau_j}(A_i,j)^{2/|u_j|} \cdots \Sigma_{\tau_j, \tau_j}(A_i,j)^{2/|u_j|}.
\]
Since our traces \( \tau_i = 1, 2 \) satisfy a \( c^*-\)Bernstein-Markov property, the norms of the individual orthonormal polynomials are bounded from above:
\[
\|P_{\tau_j, \tau_j}(A_i,j)\|^2/|u_j| \leq M(t_j) |u_j|^{2/|u_j|} \leq \sup_{q \in \mathbb{N}} M(t_j) |u_j|^{2/|u_j|} < \infty.
\]
(As in Definitions 3.8, 3.9, it is necessary to specify here the dependence of the parameter \( M_q \) on the tracial state \( \tau_i, \).) Together with the above estimate for \( \|P_{w, \tau}(A_1, A_2)\|^2 \), this yields
\[
\|P_{w, \tau}(A_1, A_2)\|^2 \leq \|P_{w, \tau}(A_1, A_2)\|^2 \|P_{w, \tau}(A_1, A_2)\|^2 \|P_{w, \tau}(A_1, A_2)\|^2 \cdots \|P_{w, \tau}(A_1, A_2)\|^2 \times \Sigma_{\tau_j, \tau_j}(A_i,j)^{2/|u_j|} \Sigma_{\tau_j, \tau_j}(A_i,j)^{2/|u_j|} \Sigma_{\tau_j, \tau_j}(A_i,j)^{2/|u_j|} \cdots \Sigma_{\tau_j, \tau_j}(A_i,j)^{2/|u_j|} \leq \left[ \sup_{q \in \mathbb{N}} M(t_j) |u_j|^{2/|u_j|} \right] \left[ \sup_{q \in \mathbb{N}} M(t_j) |u_j|^{2/|u_j|} \right] \cdots \left[ \sup_{q \in \mathbb{N}} M(t_j) |u_j|^{2/|u_j|} \right] \times \Sigma_{\tau_j, \tau_j}(A_i,j)^{2/|u_j|} \Sigma_{\tau_j, \tau_j}(A_i,j)^{2/|u_j|} \Sigma_{\tau_j, \tau_j}(A_i,j)^{2/|u_j|} \cdots \Sigma_{\tau_j, \tau_j}(A_i,j)^{2/|u_j|} \leq \left[ \sup_{q \in \mathbb{N}} M(t_j) |u_j|^{2/|u_j|} \right] \left[ \sup_{q \in \mathbb{N}} M(t_j) |u_j|^{2/|u_j|} \right] \cdots \left[ \sup_{q \in \mathbb{N}} M(t_j) |u_j|^{2/|u_j|} \right] \times \Sigma_{\tau_j, \tau_j}(A_i,j)^{2/|u_j|} \Sigma_{\tau_j, \tau_j}(A_i,j)^{2/|u_j|} \Sigma_{\tau_j, \tau_j}(A_i,j)^{2/|u_j|} \cdots \Sigma_{\tau_j, \tau_j}(A_i,j)^{2/|u_j|}.
\]
Combining this estimate with (3.45) yields
\[
\|f(A_1, A_2)(I_k)\|^2 \leq \|f(\varphi_{ij}, \tau_{ij})\|^2 (m + n)^2 \max \left\{ \Sigma_{\tau_1}^\infty(A_1) \sup_{q \in \mathbb{N}} M(\tau_1)_{ij}^{\frac{2}{q}}, \Sigma_{\tau_2}^\infty(A_2) \sup_{q \in \mathbb{N}} M(\tau_2)_{ij}^{\frac{2}{q}} \right\},
\]
which, by the definition of \(\Sigma^\infty\), provides the desired estimate
\[
(3.49) \quad \Sigma^\infty_\tau(A_1, A_2) \leq (m + n)^2 \max \left\{ \Sigma_{\tau_1}^\infty(A_1) \sup_{q \in \mathbb{N}} M(\tau_1)_{ij}^{\frac{2}{q}}, \Sigma_{\tau_2}^\infty(A_2) \sup_{q \in \mathbb{N}} M(\tau_2)_{ij}^{\frac{2}{q}} \right\},
\]
for all \((A_1, A_2) \in \mathcal{M}_k(\mathbb{C})^{m+n}\).

Note that \(\Phi_{ij}^\infty(A)\) is finite if and only if \(\Sigma_\theta^\infty(A)\) is. Thus, the above estimate guarantees that \(\Sigma^\infty_{\tau_1}(A_1, A_1) < +\infty, \Sigma^\infty_{\tau_2}(A_2, A_2) < +\infty, \Phi^\infty_{\tau_1}(A_1, A_2) < +\infty, \Phi^\infty_{\tau_2}(A_1, A_2) < +\infty\) whenever \(\Sigma^\infty_{\tau_1}(A_1) < \infty\) and \(\tau_j\) satisfy a \(C_j\)-Bernstein-Markov property, \(j = 1, 2\). In particular, if \(\Phi^\infty_{\tau_j}\) are finite q.c., then so are all four functions associated to \(\tau\).

Assume next that the tuples \(A_1\) and \(A_2\) are of selfadjoint matrices (or, equivalently for our purposes, of skew-selfadjoint matrices, or of any scalar multiples of selfadjoint matrices). Then for any polynomial \(P \in \mathbb{C}(\mathbb{X}_N)\), one has \(P(A_1)^* = P^*(A_1)\), so that \(P(A_1)P(A_1)^* = (PP^*)(A_1)\). In particular, if \(P \in \mathbb{C}(\mathbb{X}_N)\), then \((PP^*)^p \in \mathbb{C}(\mathbb{X}_N)^{2dp}\). In light of this observation, relation \(\|P_{u_j, \tau_j}(A_i)\|^2 \leq \|P_{u_j, \tau_j}(\varphi_{ij})\|^2 \Phi_{\tau_j, 2^N|u_j}(A_i)\) immediately implies
\[
\|P_{u_j, \tau_j}(A_i)\|_{\tau_j, 2^N|u_j}(A_i) = \|P_{u_j, \tau_j}(A_i)\|_{\tau_j, 2^N|u_j}(A_i);
\]
for \(\bullet = \infty\), we have used the equality \(\|YY^*\| = \|Y\|^2\), and for \(\bullet = 2\), the inequality \(\text{tr}\{Y\} \text{tr}\{Y^*\} \leq \text{tr}\{YY^*\}\). It follows that
\[
\|P_{u_j, \tau_j}(A_i)\|_{\tau_j, 2^N|u_j}(A_i) \leq \|P_{u_j, \tau_j}(A_i)\| \limsup_{N \to \infty} \Phi_{\tau_j, 2^N|u_j}(A_i) \frac{1}{N}
\]
for all \(1 \leq j \leq r\). Relation (3.46) becomes
\[
\|P_{w, \tau}(A_1, A_2)\|^2 \leq \|P_{u_1, \tau_1}(A_1)\|^2 \|P_{u_2, \tau_2}(A_2)\|^2 \|P_{u_3, \tau_3}(A_3)\|^2 \cdots \|P_{u_r, \tau_r}(A_r)\|^2.
\]
As for $\Sigma$, by taking $\sup$ as $d$ tends to infinity and picking maximizing elements $f \in M_k(\mathbb{C}(X_1, X_2))$ for each $d$, we obtain

$$
\Phi^\infty_\tau(A_1, A_2) \leq (m + n)^2 \max \left\{ \frac{2}{\tau_1} \sup_{q \in \mathbb{N}} M(\tau_1)^{\frac{2}{q}}, \frac{2}{\tau_2} \sup_{q \in \mathbb{N}} M(\tau_2)^{\frac{2}{q}} \right\},
$$

for all self-adjoint tuples $(A_1, A_2) \in M_k(\mathbb{C})^{m+n}$.

**Remark 3.14.** Under the hypotheses of Theorem 3.13, $\tau$ satisfies the $\mathcal{C}$-Bernstein-Markov property for some constant $\mathcal{C}$ satisfying the inequalities $\min\{\mathcal{C}_1, \mathcal{C}_2\} \leq \mathcal{C} \leq \sqrt{m + n} \max \left\{ \sup_{q \in \mathbb{N}} M(\tau_1)^{\frac{2}{q}}, \sup_{q \in \mathbb{N}} M(\tau_2)^{\frac{2}{q}} \right\}$. This is indeed a very rough estimate: if $P \in C(\mathbb{X}_1, \mathbb{X}_2)_d$, then

$$
\|P(a_{\tau_1}, a_{\tau_2})\| = \left\| \sum_{w \in C(\mathbb{X}_1, \mathbb{X}_2)_d} c_w P_{\tau, w}(a_{\tau_1}, a_{\tau_2}) \right\|
$$

$$
\leq \sum_{w \in C(\mathbb{X}_1, \mathbb{X}_2)_d} |c_w| \|P_{\tau, w}(a_{\tau_1}, a_{\tau_2})\|
$$

$$
\leq \left( \max_{w \in C(\mathbb{X}_1, \mathbb{X}_2)_d} \|P_{\tau, w}(a_{\tau_1}, a_{\tau_2})\| \right) \sum_{w \in C(\mathbb{X}_1, \mathbb{X}_2)_d} |c_w|.
$$

By Cauchy-Schwarz, $\sum_{w \in C(\mathbb{X}_1, \mathbb{X}_2)_d} |c_w| \leq \left( \sum_{w \in C(\mathbb{X}_1, \mathbb{X}_2)_d} |c_w|^2 \right)^{\frac{1}{2}} (m + n)^{\frac{1}{2}} = (m + n)^{\frac{1}{2}} \|P\|_{L^2(\tau)}$ (where $(m + n)^{\frac{1}{2}}$ is the two-norm of the vector of ones of length equal to the cardinality of $C(\mathbb{X}_1, \mathbb{X}_2)_d$). The operator norm $\|P_{\tau, w}(a_{\tau_1}, a_{\tau_2})\|$ of orthonormal polynomials associated to $\tau = \tau_1 \ast \tau_2$ has been estimated in the proof of Theorem 3.13: if $w = u_1 u_2 \cdots u_r \in \{X_1, X_2\}^{\times t}, u_j \in \{X_1, X_2\}, 1 \leq j \leq r, r \leq d$, where $i_1 \neq i_2 \neq \cdots \neq i_r$ and $|u_1| + |u_2| + \cdots + |u_r| = d$, then

$$
\|P_{\tau, w}(a_{\tau_1}, a_{\tau_2})\|^2 \leq \left[ \|P_{\tau_1, u_1}(a_{\tau_1})\|^2 \|P_{\tau_2, \tau_1}(a_{\tau_2})\|^2 \cdots \|P_{\tau_2, \tau_r}(a_{\tau_r})\|^2 \right]^{1/d}
$$

$$
\leq \max \left\{ \sup_{q \in \mathbb{N}} M(\tau_1)^{\frac{2}{q}}, \sup_{q \in \mathbb{N}} M(\tau_2)^{\frac{2}{q}} \right\}.
$$

Combining the last two estimates yields

$$
\|P(a_{\tau_1}, a_{\tau_2})\|^{\frac{2}{d}} \leq \sqrt{m + n} \max \left\{ \sup_{q \in \mathbb{N}} M(\tau_1)^{\frac{2}{q}}, \sup_{q \in \mathbb{N}} M(\tau_2)^{\frac{2}{q}} \right\} \|P\|_{L^2(\tau)}^{\frac{1}{d}}.
$$

The lower bound is obvious.

This shows that Theorem 3.13 applies to any finite free products $\tau = \tau_1 \ast \tau_2 \ast \cdots \ast \tau_s$, $s \in \mathbb{N}$, of bounded tracial states $\tau_j$ which satisfy a $\mathcal{C}_j$-Bernstein-Markov property, $1 \leq j \leq s$.

We conclude this subsection by pointing out that $\frac{1}{2} \log \Phi^\bullet(A), \frac{1}{2} \log \Sigma^\bullet(A)$ have logarthmic growth as plurisubharmonic functions on the Euclidean space $\mathbb{C}^{nk'} \simeq M_k(\mathbb{C})^n$ whenever they are non-trivial. Indeed, it follows from Section 3.3.8 that if the sequence $\left\{ (\Phi^\infty_\tau, d(A))^{1/d} \right\}$ is locally bounded, then, as log is increasing,
\( \frac{1}{2} \log \Sigma^\infty_\tau(A) = \left[ \sup_{d \in \mathbb{N}} \frac{1}{2n} \log \left( \Phi^\infty_{\tau,d}(A) \right) \right] \) has logarithmic growth, according to [ST97, Theorem 1.6, Appendix B]. Since it dominates the other three functions, it follows that \( \frac{1}{2} \log \Phi^\infty_\tau(A), \frac{1}{2} \log \Sigma^\infty_\tau(A) \) have logarithmic growth as \( \|A\|_\bullet \to \infty \).

### 3.5. Derivatives.

One of the equivalent characterizations of plurisubharmonicity is via positivity: if \( \Omega \) is an open subset of \( C^n \), then for any plurisubharmonic function \( u \) on \( \Omega \) and \( \xi \in C^n \),

\[
\sum_{j,k=1}^m \xi_j \frac{\partial^2 u}{\partial z_j \partial \bar{z}_k}
\]

is a positive distribution in \( \Omega \) (that is, it is a positive Borel measure on \( \Omega \), possibly \( \equiv 0 \)) - see [GZ17, Proposition 1.43]. We have established in the previous section that \( \frac{1}{2} \log \Phi^\infty_\tau(A), \frac{1}{2} \log \Sigma^\infty_\tau(A), \bullet \in \{2, \infty \} \) are classical plurisubharmonic functions when viewed as functions on \( C^{nk^2} \), and that they have at most logarithmic growth at infinity. Thus, according to, for instance, [GZ17, Proposition 3.34],

\[
\int_{C^{nk^2}} (d^c d^\ast \log \Phi^\infty_\tau(A) \frac{1}{2})^{nk^2} \leq 1, \quad \int_{C^{nk^2}} (d^c d^\ast \log \Sigma^\infty_\tau(A) \frac{1}{2})^{nk^2} \leq 1,
\]

where \( \varphi \to (d^c d^\ast \varphi)^{nk^2} = (2\pi)^{-nk^2} \det \left( \left( \partial^2_j \varphi \right)_{1 \leq i, j \leq nk^2} \right) dV \) is the Monge-Ampère operator if \( \varphi \) is of class \( C^2 \) (\( V \) is the Lebesgue measure). Otherwise this should be understood in the weak sense. Thus, \( (d^c d^\ast \log \Phi^\infty_\tau(A) \frac{1}{2})^{nk^2}, (d^c d^\ast \log \Sigma^\infty_\tau(A) \frac{1}{2})^{nk^2} \) are nonnegative measures on \( C^{nk^2} \) of mass at most one. For the purposes of this article, we agree to call such a measure the Monge-Ampère measure of the corresponding function. In the same [GZ17, Proposition 3.34] it is shown that if \( \frac{1}{2} \log \Phi^\infty_\tau(A), \frac{1}{2} \log \Sigma^\infty_\tau(A) \) have exactly logarithmic growth at infinity, then these measures have total mass precisely one, i.e. the inequalities in the above displayed relations are equalities.

We show next that indeed \( A \mapsto \frac{1}{2} \log \Sigma^\infty_\tau(A) \) is exactly of logarithmic growth for a large class of distributions \( \tau \). In order to do this, we consider first the case of \( n = 1 \) and then use item (3) from Lemma 3.7 to conclude for arbitrary \( n \). We look at polynomials \( L_{A,d}(X) = \frac{1}{d!} \Phi^\tau(A)^d \otimes X^d \). We assume without loss of generality that \( \|a_{\tau} \| = 1 \) and, with loss of generality, that \( \sigma(a_{\tau}) \) has non-empty interior. We clearly have \( \|L_{A,d}(a_{\tau})\| = \frac{1}{d!} \|\Phi^\tau(A)^d\| \|a_{\tau}\|^d = 1 \) and

\[
\|L_{A,d}(A)(Ik)^*L_{A,d}(A)(Ik)\| = \frac{1}{d!} \|\Phi^\tau(A)^d\|^2 \|A\|^d = \|A\|^d.
\]

It follows that \( \Phi^\infty_{\tau,d}(A) \geq \|A\|^d \). By definition, \( \Sigma^\infty_\tau(A) = \left[ \sup_{d \in \mathbb{N}} \Phi^\infty_{\tau,d}(A) \frac{1}{2} \right] \geq \sup_{d \in \mathbb{N}} \|A\|^d \geq \|A\|^2 \).

Taking logarithms and recalling Example 3.11 shows that \( \frac{1}{2} \log \Sigma^\infty_\tau(A) \geq \log^+ \|A\| \) for \( \|A\| \) sufficiently large.

Recall that if \( \tau : C(X, X) \to C \) is a bounded positive trace, then \( \Sigma^\infty_\tau(A_1, A_2) \geq \max\{\Sigma^\infty_{\tau_{|\langle X \rangle}}(A_1), \Sigma^\infty_{\tau_{|\langle X \rangle}}(A_2)\} \), with a similar statement for \( \Phi^\infty_{\tau,\bullet} \). Thus, if \( \frac{1}{2} \log \Sigma^\infty_\tau(A_j) \geq \log^+ \|A_j\|_\bullet - K \) for some \( K \geq 0 \) not depending on \( A_j, j = 1, 2 \), then \( \frac{1}{2} \log \Sigma^\infty_\tau(A_1, A_2) \geq \max\{\log^+ \|A_1\|, \|A_2\|_\bullet - K, \log^+ \|A_2\|_\bullet - K\} \). Since \( \|A_1\|^2 + \|A_2\|^2 = \|A_1 + A_2\|^2 \), one has

\[
\log \|A_1 + A_2\|^2 = \log \max\{\|A_1\|^2, \|A_2\|^2\} + \log \left( 1 + \frac{\|A_1\|^2 \cdot \|A_2\|^2}{\min\{\|A_1\|^2, \|A_2\|^2\}} \right) \geq \max\{\log \|A_1\|^2, \log \|A_2\|^2\} + \log \sqrt{2}.
\]
Similarly, one has 
\[ \| (A_1, A_2) \| = \max\{ \| A_1 \|, \| A_2 \| \} \]
so that \( \log \| (A_1, A_2) \| = \max\{ \log \| A_1 \|, \log \| A_2 \| \} \). Putting these together, it follows that
\[
\frac{1}{2} \log \Sigma^\tau_\star (A_1, A_2) \geq \log \| (A_1, A_2) \|_\star - \sqrt{2} - K
\]
whenever \( \frac{1}{2} \log \Sigma^\tau_\star (A_{jk}) \geq \log^+ \| A_{jk} \|_\star - K \) for some \( K \geq 0 \), \( j, k = 1, 2 \).

We conclude that whenever \( \tau: \mathbb{C}(X) \to \mathbb{C} \) is a positive bounded tracial state such that \( \tau|_{\mathbb{C}(X)} \) is a probability distribution whose support contains a nonempty open set in \( \mathbb{R} \), the function \( \Sigma^\infty_\tau (A) \) has precisely logarithmic growth at infinity, i.e. belongs to the Lelong class \( L^+(\mathbb{C}^{nk^2}) = \{ u: \mathbb{C}^{nk^2} \to [-\infty, +\infty): u \text{ plurisubharmonic}, \log^+ \| z \| - K_u \leq u(z) \leq \log^+ \| z \| + K_u, z \in \mathbb{C}^{nk^2}, \text{for some } K_u > 0 \} \). Proposition 3.10 allows us to obtain the following result:

**Proposition 3.15.** Assume that \( \tau: \mathbb{C}(X) \to \mathbb{C} \) is a positive bounded tracial state such that \( \tau|_{\mathbb{C}(X)} \) is a probability distribution whose support contains a nonempty open set in \( \mathbb{R} \), \( 1 \leq j \leq n \). If any of \( \Sigma^\tau_\star, \Phi^\tau_\star, \bullet \in \{ 2, \infty \} \) is well-defined and finite q.e., then \( \Sigma^\infty_\tau \in L^+(\mathbb{C}^{nk^2}) \). If \( \tau \) satisfies in addition the \( \mathcal{C} \)-Bernstein-Markov property 3.9, then all of
\[
\left[ \sup_{d \in \mathbb{N}} \log \| \kappa_{\tau, d}(A^\star, A)(I_k) \|_\star \right]^*, \left[ \sup_{d \in \mathbb{N}} \text{tr}_{\mathcal{K}} \left( \kappa_{\tau, d}(A^\star, A)(I_k) \right) \right]^*_\star,
\]
\[
\frac{1}{2} \log \Sigma^\tau_\star \in L^+(\mathbb{C}^{nk^2}), \quad \bullet \in \{ 2, \infty \}.
\]
In particular, applying the Monge-Ampère operator to any of these plurisubharmonic functions yields a probability measure on \( \mathbb{C}^{nk^2} \) for any \( k \in \mathbb{N} \).

We remind the reader that, according to Theorem 3.13, \( \Sigma^2_\tau \) is finite q.e. whenever \( \tau|_{\mathbb{C}(X)} \) are as in the above proposition and \( \mathbb{C}(X) \), \( 1 \leq j \leq n \), are free with respect to \( \tau \). This means that a large class of distributions is covered by the result above.

It is also known that in many cases the plurisubharmonic functions of the kind described in Proposition 3.15 reach their minimum precisely at the points of the support of their Monge-Ampère measure. This, together with the considerations from Section 3.3, guarantees that the supports of these measures are compact and invariant under conjugation by a unitary \( k \times k \) matrix for each \( k \).

**Aside:** To conclude the discussion of plurisubharmonic functions in the context of noncommutative sets, polynomials, distributions, and functions, we make some remarks regarding connections between plurisubharmonicity and the difference-differential operators, which play the role of the derivatives from classical analysis. This has no direct bearing on the rest of our paper at this time. In the case of noncommutative polynomials (or, more generally, functions), one has the free difference quotient standing in for the derivative [Voi98, Voi00]. Specifically, let us consider now a polynomial \( F \in \mathbb{C}(\mathbb{Z}, \mathbb{Z}^*) \). We identify \( Z_j = X_j + iY_j, Z_j^* = X_j - iY_j \) where \( X_j, Y_j \) are selfadjoint indeterminates algebraically free from each other, \( j \in \{ 1, \ldots, n \} \). That is, we double the number of noncommuting indeterminates. The rule for “differentiation” is simple in the case of polynomials in noncommuting selfadjoint indeterminates: \( \partial_X X_j = \delta_{i1}1 \otimes 1, \partial_X Y_k = 0, \partial_X \) is linear from \( \mathbb{C}(\mathbb{Z}, \mathbb{Z}^*) \) to \( \mathbb{C}(\mathbb{Z}, \mathbb{Z}^*) \otimes \mathbb{C}(\mathbb{Z}, \mathbb{Z}^*) \), and \( \partial_X \) respects the Leibniz rule with respect to products of polynomials. For instance, \( \partial_X(7iX_1X_2Y_2X_1^2Y_2) = \ldots \).
7i(1 \otimes X_2Y_2X_2^2Y_2 + X_1X_2Y_2X_1 \otimes Y_2 + X_1X_2Y_2 \otimes Y_1Y_2). To parallel classical analysis, we work with the derivative with respect to \(Z\) and with respect to \(Z^*\). Specifically, \(\partial_{Z_j} = \frac{1}{2}(\partial X_j - i\partial Y_j), \partial_{Z_j^*} = \frac{1}{2}(\partial X_j + i\partial Y_j), \) a simple change of variable. Not surprisingly,

\[
\partial_Z Z_j = \frac{1}{2}(\partial X_j - i\partial Y_j)(X_j + iY_j) = \frac{1}{2}(1 \otimes 1 + i0 - i0 - i1 \otimes 1) = 1 \otimes 1,
\]

\[
\partial_Z Z_j^* = \frac{1}{2}(\partial X_j - i\partial Y_j)(X_j - iY_j) = \frac{1}{2}(1 \otimes 1 - i0 - i0 + i1 \otimes 1) = 0,
\]

\[
\partial_Z Z_j = \frac{1}{2}(\partial X_j + i\partial Y_j)(X_j + iY_j) = \frac{1}{2}(1 \otimes 1 + i0 + i0 + i1 \otimes 1) = 0,
\]

\[
\partial_Z Z_j^* = \frac{1}{2}(\partial X_j + i\partial Y_j)(X_j - iY_j) = \frac{1}{2}(1 \otimes 1 - i0 + i0 - i1 \otimes 1) = 1 \otimes 1.
\]

This clearly implies that \(\partial_{Z_j}(P(Z)^*) = 0 = \partial_{Z_j}P(Z)\) for all \(P \in \mathbb{C}(\mathcal{Z}), 1 \leq j \leq n\). That is, an analytic polynomial is “killed” by \(\partial_{Z_j}\) and a conjugate-analytic one is “killed” by \(\partial_{Z_j^*}\). We extend the definition of the derivative precisely the same way to polynomials with matrix coefficients \(P \in \mathbb{M}_k(\mathbb{C}(\mathcal{Z}, \mathcal{Z}^*))\).

When evaluated, the “derivative” becomes a proper derivative in the sense of Fréchet. Indeed, direct computation shows that if \(\mathcal{A}\) is a star-algebra and \(F \in \mathbb{C}(\mathcal{Z}, \mathcal{Z}^*)\), then \(\partial_{Z_j}F(a, a^*)(c)\) replaces in the above formula the tensor symbol \(\otimes\) with the variable \(c:\n\)

\[
\partial_{Z_j}F(a, a^*)(c) = (\partial_{Z_j}F)(a, a^*) \circ m_c,
\]

where \(m_c : \mathcal{A} \otimes \mathcal{A}^{op} \to \mathcal{A}, m_c(p \otimes q) = p \cdot q\). This makes the above equivalent definition of the classical plurisubharmonic functions “immitable” in the noncommutative context. Consider the easiest case, namely functions of the type \(\phi : \mathcal{A} \to \text{tr}_k(F(\mathcal{A}^*)F(\mathcal{A}))\) defined on \(\mathbb{M}_k(\mathbb{C})^n\) for each given polynomial \(F \in \mathbb{C}(\mathcal{X})\). We view \(F\) as living in \(\mathbb{C}(\mathcal{Z})\), that is, as an analytic polynomial. As mentioned before, when viewed as an element in \(\mathbb{C}(\mathcal{X})\), the evaluation \(F(\mathcal{A}^*) = F^*(\mathcal{A}^*)\), and when viewed as an element in \(\mathbb{C}(\mathcal{Z})\), we have \(F^*\) as an element in \(\mathbb{C}(\mathcal{Z}^*)\). The problem of differentiating becomes now very classical. For any given direction \(C \in \mathbb{M}_k(\mathbb{C}),\)

\[
\partial_{Z_j} \phi(A)(C) = \frac{\partial \phi(A + \delta_z C)}{\partial z} \bigg|_{z=0}, \quad \partial_{Z_j^*} \phi(A)(C) = \frac{\partial \phi(A + \delta_z C)}{\partial z} \bigg|_{z=0},
\]

where we have denoted \(A + \delta_z C = (A_1, \ldots, A_{j-1}, A_{j} + zC, A_{j+1}, \ldots, A_n)\). Moreover, for a vector \(C = (C_1, \ldots, C_n)^t \in \mathbb{M}_k(\mathbb{C})^n\), we write

\[
\sum_{i,j=1}^n \partial_{Z_i} \partial_{Z_j} \phi(A)(C_j)(C_i)
\]

for the Levi form of \(\phi\) at the point \(A\) evaluated in \(C\) [GZ17, (1.3.6)]. (It might seem normal to rather write \(\partial_{Z_j} \partial_{Z_j} \phi(A)(C_j)(C_j^*)\) in the formula above; however, with the convention we have introduced, we actually have \(\partial_{Z_j} \partial_{Z_j} \phi(A)(C_j)(C_j^*) = \partial_{Z_j} \partial_{Z_j} \phi(A)(C_j)(C_i)\) when \(\partial_{Z_j} \partial_{Z_j} \phi(A)\) is viewed as an operator. We shall keep in mind, and follow, this convention.) In terms of directional derivatives, one writes this relation naturally as \(\sum \partial_{Z_j} \partial_{Z_j} \phi(A)(C_j)(C_i) = \partial_{Z_i} \partial_{Z_i} \phi(A + zC)\). To avoid confusion as much as possible, we will keep the notation with iterated arguments.

Starting with Section 3.3, we introduced several examples of classically plurisubharmonic functions obtained from noncommutative functions. These functions,
while not being themselves noncommutative functions, have been seen to possess many properties reminiscent of those of noncommutative functions. Based on these properties and the machinery of the free difference quotient described above, we propose next some more general function spaces that accept our examples as members. For any noncommutative function $H$ defined on a noncommutative set $\Omega \subseteq \mathcal{A}^n$ in a $C^*$-algebra $\mathcal{A}$ endowed with a state $\varphi$, we may define $\mathcal{H} = \varphi \circ (H^*H)$—that is, $\mathcal{H}(a) = (\varphi \otimes \text{tr}_k)(H(a)^*H(a))$ for any $a \in \Omega^k$—and then the Levi form of $\mathcal{H}$ at the point $a$ evaluated in the direction $z \in$ is

$$\partial_2 \partial_z \mathcal{H}(a + z\omega)|_{z=0} = \sum_{i,j=1}^n \partial_{Z_i} \partial_{Z_j} \mathcal{H}(\omega)(c_j)(c_i).$$

The positivity of the above for all $(a, \omega) \in (\Omega \times \mathcal{A}^n)_{nc}$ could be taken as a definition of an nc plurisubharmonic function. Indeed, as in the classical case one can find in references [GZ17, ST97], the above means that the correspondence $z \mapsto \mathcal{H}(a + z\omega)$ subharmonic as a function from a complex neighborhood of zero into $[-\infty, +\infty)$ [GZ17, Definition 1.27], which seems to be a fair definition of plurisubharmonicity in our case too. However, as in the classical case, such a definition is likely to be less general than needed: here we have assumed $H$ to exist and be a free noncommutative function, hence analytic. In (3.50), derivatives exist in the strongest possible sense. It is undoubtedly necessary to consider at a minimum level-by-level weak closures of spaces of such functions (this is somewhat reminiscent of [GZ17, Example 1.41], but in our case analyticity should be replaced by the property of being a noncommutative function).

We should emphasize that the functions $\Phi^{\tau, d}_{\omega, \zeta}$ defined in (3.13) belong to the space of plurisubharmonic functions of the form $\varphi \circ (H^*H)$ for some nc function $H$, while the functions $\Phi^\infty_{\tau, d}$ from (3.14) generally do not, as they might not be classically differentiable at some levels. However, $\Phi^\infty_{\tau, d}$ is a well-defined (and monotone) limit of fractional powers of functions of the type $\Phi^p_{\tau, d}$ with $p$ tending to infinity along a properly chosen subsequence, hence plurisubharmonic in the sense that $z \mapsto \Phi^\infty_{\tau, d}(a + z\omega)$ is classically subharmonic on some neighborhood of zero in $\mathbb{C}$ for all $k \in \mathbb{N}, \omega, \zeta \in (M_k(\mathbb{C}))^n$ (see [GZ17, Theorem 1.46]). For these functions, (3.50) holds, but in a weak sense. This same statement remains true for $\Phi^2_{\tau, \Sigma}, \Sigma^2_{\tau}, \Sigma^\infty_{\tau}$, and $\Phi^\infty_{\tau}$ if they are less than $+\infty$. Thus,

$$\sum_{i,j=1}^n \partial_{Z_i} \partial_{Z_j} \Phi^2_{\tau, d}(\omega)(C_j)(C_i) \geq 0 \text{ pointwise, and}$$

$$\sum_{i,j=1}^n \partial_{Z_i} \partial_{Z_j} \Phi^\infty_{\tau, d}(\omega)(C_j)(C_i) \geq 0 \text{ in the sense of distributions,}$$

for all $k, d \in \mathbb{N}$, and all $\omega, \zeta \in (M_k(\mathbb{C}))^n$, with similar statements for $\Sigma$. By positivity in the sense of distributions we mean that $\sum_{i,j=1}^n \partial_{Z_i} \partial_{Z_j} \Phi^\infty_{\tau, d}(\omega)(C_j)(C_i)$ is a positive distribution on the Euclidean space $\mathbb{C}^{nk^2}$—the variables considered being $\omega$, while $\zeta$ should be viewed as parameters. We expect that eventually the correct use of the difference-differential operators will reveal a tighter connection between $\tau$ and the Monge-Ampère measures associated to $\frac{1}{2} \log \Phi^2_{\tau}, \frac{1}{2} \log \Sigma^2_{\tau}, \frac{1}{4} \log \Sigma^\infty_{\tau}$, and $\frac{1}{2} \log \Phi^\infty_{\tau}$, as well as to the logarithm of the norms of the Christoffel-Darboux kernels.
4. Conclusion, Perspectives, and Numerical Experiments

We have seen in the previous section that to a bounded, positive, tracial noncommutative distribution \( \tau : \mathcal{C}(\mathbf{X}) \to \mathbb{C} \) one associates, under some reasonable assumptions, sequences of Borel probability measures supported on \( \mathcal{M}_k(\mathbb{C})^n \simeq \mathbb{C}^{nk^2} \). We focus here on those associated to the logarithm of the Christoffel-Darboux kernel, which we denote by \( \{ \mu_{\tau,k} \}_{k \in \mathbb{N}} \). While we have investigated some properties of these measures, it is not yet clear how strongly they are related to \( \tau \) itself besides being determined by it. We conjecture that these measures might be good approximants for \( \tau \) in a very precise way, thus providing another means to find random matrix approximants to (some) noncommutative distributions.

Conjecture 4.1. Assume that \( \tau : \mathcal{C}(\mathbf{X}) \to \mathbb{C} \) is a bounded faithful tracial state. Suppose that \( \{ X_1, \ldots, X_n \} \) are free with respect to \( \tau \) and \( \tau|_{\mathcal{C}(X_j)} \) has a support with nonempty interior, \( 1 \leq j \leq n \). For all \( f \in \text{Sym}_k(\mathbf{X}) \), one has

\[
\lim_{d \to \infty} \lim_{k \to \infty} \sup_{\{ A \in \Omega_{\tau,d}^k \}} |\text{tr}(f(A)) - \tau(f(X))| = \lim_{d \to \infty} \lim_{k \to \infty} \sup_{\{ A \in \Omega_{\tau,d}^k \}} |\text{tr}(f(A)) - \text{tr}(f(A_\tau))| = 0,
\]

\[
\lim_{d \to \infty} \lim_{k \to \infty} \sup_{\{ A \in \Omega_{\tau,d}^k \}} \|f(A)\|_{\mathcal{M}_k(\mathbb{C})} = \|f(A_\tau)\|_{W^*(\tau)},
\]

where \( \Omega_{\tau,d}^k = \{ A : \|\sup_{d \to \infty} \text{tr}(\kappa_{\tau,d}(A, A^*) (I_k))^{1/d} \|_{*} \leq n \} \).

Let us note some immediate consequences in this context of Theorem 3.4. Observe that in the above conjecture, the matrix variables \( A \) are taken as approximants of \( A_\tau \): for this to be possible, it is necessary that the algebra \( W^*(\tau) \) satisfies Connes’ embedding property, such approximants exist. Viewed this way, our conjecture states that the approximation is good along the subspaces generated by the first \( \sigma(n, d) \) vectors in the Hilbert module generated by the orthogonal polynomials of \( \tau \) as \( d \) tends to infinity.

In commutative analysis, the level sets of the Christoffel polynomial are used to approximate the support of the distribution of the variables \( A_\tau \). From this perspective, Conjecture 4.1 can be viewed as specifying the notion of support for noncommutative distributions.

Here, we illustrate our theoretical framework for the computation of Christoffel-Darboux kernels for free semicircular and Poisson laws and attempt to provide some numerical data to support our Conjecture 4.1. Our experiments are performed with Mathematica 12, and we heavily rely on the NCAAlgebra package \([\text{HMS96}]\) to handle noncommutative polynomials. All results were obtained on an Intel Xeon(R) E-2176M CPU (2.70GHz × 6) with 32Gb of RAM. Our code is available online\(^3\).

Given \( \tau \) and \( d \), we first compute the Christoffel polynomial \( \kappa_{\tau,d} \) from the knowledge of the \( d \)-th order moment matrix associated to \( \tau \). Then, we consider approximations \( \tilde{\Phi}_{\tau,d}^{(k)}(A) = \text{tr}(\kappa_{\tau,d}(A, A^*) (I_k))^{1/2} \) of

\[
\left[ \lim_{d \to \infty} \sup_{\{ \kappa_{\tau,d}(A, A^*) (I_k) \}} \frac{d}{k} \right]^{*}.
\]

\(^3\)http://homepages.laas.fr/vmagron/files/NCCD.zip
For several values of the matrix size $k$, we sample $N$ matrices from the following set
\[
\tilde{\Omega}_{r,d}^{(k)} = \{ A : n - \varepsilon \leq \tilde{\Phi}_{r,d}^{(k)}(A) \leq n + \varepsilon \},
\]
with respect to a given matrix distribution. Overall, it boils down to sampling in an \(\varepsilon\)-perturbation of the \(n\)-level set of the approximate Siciak function \(\tilde{\Phi}_{r,d}^{(k)}\). Then, we choose an \(f \in \text{Sym}(\mathbb{R}(X))\) and empirically check if (4.1) from Conjecture 4.1 holds on average. This verification is performed by computing the expectation \(E(\text{tr}_k(f(A)))\) over the \(N\) selected samples \(A\) from \(\tilde{\Omega}_{r,d}^{(k)}\) and comparing it with \(\tau(f(A))\). For all experiments, we select \(N = 10^5\).

4.1. Single free semicircular. Here, we consider a single standard free semi- circular \(A\) of variance 1. By [MS17, Definition 4], the odd moments of \(A\) are 0 and the even moments are given by the Catalan numbers, i.e., \(\tau(A^{2d+1}) = 0\) and \(\tau(A^{2d}) = \frac{1}{d+1} \binom{2d}{d}\). The first four corresponding values of \(\kappa_{r,d}\) are:
\[
\begin{align*}
\kappa_{r,1}(A, A)(1) &= 1 + A^2, \\
\kappa_{r,2}(A, A)(1) &= 2 - A^2 + A^4, \\
\kappa_{r,3}(A, A)(1) &= 2 + 3A^2 - 3A^4 + A^6, \\
\kappa_{r,4}(A, A)(1) &= 3 - 3A^2 + 8A^4 - 5A^6 + A^8.
\end{align*}
\]

In Figure 1, we present the numerical experiments obtained after sampling \(\tilde{\Omega}_{r,d}^{(k)}\) with respect to the Gaussian orthogonal matrix distribution with unit scale parameter, for degree \(d \in \{1, \ldots, 15\}\) and matrix size \(k \in \{2, 3, 5, 10\}\). We choose a perturbation of \(\varepsilon = 0.7\) to sample \(\tilde{\Omega}_{r,d}^{(k)}\) and \(f(A) = A^2\). The figure confirms that the average \(E(\text{tr}_k(f(A)))\) over the samples of \(\tilde{\Omega}_{r,d}^{(k)}\) gets closer to \(\tau(f(A)) = 1\), when \(d\) and \(k\) increase.

4.2. Pair of free semicircular distributions. Here, we consider a pair \(A = (A_1, A_2)\) of standard free semicircular distributions of variance 1. The odd moments of \(A\) are 0 and the even moments are given by the number of non-crossing pairings
Figure 2. Averaging the normalized traces of $A_1A_2A_2A_1A_1$ over $N = 10^5$ samples of $\Omega^{(k)}_{\tau,d}$ for a pair of free semicirculars of variance 1.

which respect the indices, e.g., $\tau(A_1A_2A_1) = 0$ and $\tau(A_1A_2A_2A_1A_1) = 2$. The first three corresponding values of $\kappa_{\tau,d}$ are:

$$
\kappa_{\tau,1}(A, A)(1) = 1 + A_1^2 + A_2^2,
$$

$$
\kappa_{\tau,2}(A, A)(1) = 3 - A_1^2 - A_2^2 + A_1^4 + A_1 A_2^2 A_1 + A_2 A_1^2 A_2 + A_2^4,
$$

$$
\kappa_{\tau,3}(A, A)(1) = 3 + 5 A_1^2 + 5 A_2^2 - 3 A_1^4 - 2 A_1^2 A_2^2 - A_1 A_2^2 A_1 - A_2 A_1^2 A_2 - 2 A_2^2 A_1^2 - 3 A_1^6 + A_1^2 A_2^2 A_1^2 + A_1 A_2 A_1^2 A_2 A_1 + A_1 A_2^2 A_1 + A_2 A_1^2 A_2 + A_2 A_1 A_2^2 A_1 A_2 + A_2^6.
$$

In Figure 2, we present the numerical experiments obtained after sampling $\Omega^{(k)}_{\tau,d}$ with respect to the Gaussian orthogonal matrix distribution with unit scale parameter, for degree $d \in \{1, \ldots, 8\}$ and matrix size $k \in \{2, 3, 4\}$. We choose a perturbation of $\varepsilon = 0.7$ to sample $\Omega^{(k)}_{\tau,d}$ and $f(A) = A_1 A_1 A_2 A_2 A_1 A_1$. As for the single case, the average of normalized traces over the samples of $\Omega^{(k)}_{\tau,d}$ gets closer to $\tau(f(A)) = 2$, when $d$ and $k$ increase.

4.3. Pair of free Poisson distributions. Finally, we consider a pair $A = (A_1, A_2)$ of standard free Poisson distributions of rate $c$. Since the cumulants of each law are equal to $c$, one can get the moments thanks to the moment-cumulant formula, by computing the non-crossing partitions. For instance $\tau(A_1) = \tau(A_2) = c$, $\tau(A_1^2) = \tau(A_2^2) = c + c^2$, $\tau(A_1 A_2) = \tau(A_2 A_1) = c^2$ and $\tau(A_1^3) = c + 3c^2 + c^3$. The first and second values of $\kappa_{\tau,d}$ are:

$$
\kappa_{\tau,1}(A, A)(1) = 1 + 2c - 2A_1 - 2A_2 + \frac{A_1^2 + A_2^2}{c},
$$

$$
\kappa_{\tau,2}(A, A)(1) = 1 + 2c + 4c^2 - (4 + 8c)(A_1 + A_2) + \left(8 + \frac{5}{c} + \frac{1}{c^2}\right)(A_1^2 + A_2^2) + 4(A_1 A_2 + A_2 A_1) - \left(\frac{2}{c^2} + \frac{4}{c}\right)(A_1^3 + A_2^3)
$$
Recall that a Wishart matrix is of the form \( A = G G^* \), where \( G \) is a \( k \times M \) matrix with entries being standard complex Gaussian random variables with mean 0 and \( E(|G_{ij}|^2) = 1 \). According to [MS17, § 4.5.1], if one sends \( k \) and \( M \) to infinity so that the ratio \( M/k \) is kept fixed to the value \( c \), the limiting distribution is the free Poisson distribution of rate \( c \). For \( c = 5 \), the ratio \( f \) is associated to \( \tau_{d} \) with respect to the Wishart matrix distribution, for \( c = 5 \), with a perturbation of \( \epsilon = 0.1 \). For \( f(A) = A_1 + A_2 \), we obtain the successive averages of normalized traces: 10.42, 10.24, 10.16, 10.14 and 10.1 \( \approx 10 = \tau(A_1 + A_2) \), for \( d \in \{1, \ldots, 5\} \), respectively. For \( f(A) = A_1A_2 \), we obtain 26.78, 25.78, 25.7, 25.6 and 25.4, while \( \tau(A_1A_2) = \tau(A_1) + \tau(A_2) \). Eventually, for \( f(A) = A_1^3 \), we obtain 264, 255, 244, 242 and 239, while \( \tau(A_1^3) = 5 + 3 \times 5^2 + 5^3 = 205 \).

Overall, the empirical convergence behavior of the sequence of averaged normalized traces seems to match with the theoretical moment values. It is important to emphasize that due to the computational burden associated to the rapidly growing number of non-crossing partitions, our method is currently limited to problems of modest size. We plan to overcome these scalability issues by exploiting the sparsity properties of the noncommutative Christoffel-Darboux kernels, e.g., by relying on the framework derived in [HLP+08].

### 4.4. Application to trace polynomial optimization

We recall the tracial version of Lasserre’s hierarchy [BCKP13] to minimize the trace of a noncommutative polynomial on a noncommutative semialgebraic set. Given \( f \in \text{Sym} \mathbb{R}(\mathbf{X}) \), a positive integer \( m \) and \( G = \{g_1, \ldots, g_m\} \subseteq \text{Sym} \mathbb{R}(\mathbf{X}) \), the semialgebraic set \( D_G \) associated to \( G \) is defined as follows:

\[
D_G := \bigcup_{k \in \mathbb{N}} \{ A = (A_1, \ldots, A_n) \in S_k^n : g_j(A) \succeq 0, \quad j = 1, \ldots, m \}.
\]

Now fix a possibly infinite dimensional complex Hilbert space \( \mathcal{H} \) endowed with a scalar product \( \langle \cdot | \cdot \rangle \) and let \( B(\mathcal{H}) \) be the algebra of all bounded linear operators on \( \mathcal{H} \). As usual, positivity for an operator \( C \in B(\mathcal{H}) \) means that \( \langle C \xi | \xi \rangle \geq 0 \) for all \( \xi \in \mathcal{H} \), and, as for matrices, it is denoted \( C \succeq 0 \). Strict positivity means that in addition \( C \) is also invertible, and is denoted by \( C > 0 \). For any set \( \mathcal{M} \) of bounded operators on \( \mathcal{H} \), we define

\[
D_G^\mathcal{M} := \{ A = (A_1, \ldots, A_n) \in \mathcal{M}^n : A = A^*, g_j(A) \succeq 0 \text{ on } \mathcal{H}, \quad j = 1, \ldots, m \},
\]

and its noncommutative extension

\[
D_G^\mathcal{M,nc} := \bigcup_{k \in \mathbb{N}} \{ A = (A_1, \ldots, A_n) \in M_k^n(\mathcal{M})^n : g_j(A) \succeq 0 \text{ on } \mathcal{H}^k, \quad j = 1, \ldots, m \}.
\]

(We denote by \( M_k^n(\mathcal{M}) \) the set of selfadjoint matrices with entries from \( \mathcal{M} \).) The important case for us is when \( \mathcal{M} \) is a finite von Neumann algebra endowed with a normal faithful tracial state \( \text{tr} \) (see [Tak79, Chapter V.2]). In particular, the set \( D_G \) defined in (4.3) corresponds, with the notation from (4.4), to \( D_G^\mathcal{C} \).
Given an arbitrary algebra $A$ and $g, h \in A$, we denote by $[g, h] := gh - hg$ the
commutator of $g$ and $h$. In the particular case when $A = \mathbb{C} \langle X \rangle$, two nc polynomials
$g, h$ are called cyclically equivalent (denoted $g \mathcal{C} \sim h$) if $g - h$ is a sum of commutators.
In particular, $\tau(g) = \tau(h)$.

Given the trace $\tau$ and the polynomial $g_j$, the matrix $M_d(g_j \tau)$ is the localizing
matrix associated to the nc polynomial $g_j$. It is defined the following way: let $d_j = \lceil \deg g_j / 2 \rceil \in \mathbb{N}$. The localizing matrix $M_d(g_j \tau)$ is indexed on $\langle X \rangle_{d - d_j}$ with
entry $(v, w)$ being equal to $\tau(v^* g_j w)$. Note that the moment matrix $M_d(\tau)$ is the
localizing matrix associated to $g = 1$.

Let us define $\text{tr}_{\text{min}}(f, G)$ as follows:

$$\text{tr}_{\text{min}}(f, G) := \inf \{ \text{tr} f(\mathbf{A}) : \mathbf{A} \in D_G \}.$$  \hfill (4.5)

For an arbitrary finite von Neumann algebra $M$ on a separable Hilbert space $\mathcal{H}$, we define $\text{tr}_{\text{min}}(f, G)^M$ as the trace-minimum of $f$ on $D_G^M$:

$$\text{tr}_{\text{min}}(f, G)^M := \inf \{ \text{tr} f(\mathbf{A}) : \mathbf{A} \in D_G^M \}.$$  \hfill (4.6)

We approximate $\text{tr}_{\text{min}}(f, G)^M$ from below via the following hierarchy of semi-
definite programs, indexed by $d$:

$$\tau_d(f, K) := \inf_{\tau} \tau(f)$$

s.t. $M_d(\tau)_{u,v} = M_d(\tau)_{w,z}, \quad \text{for all } u^* v \mathcal{C} \sim w^* z$,

$$M_d(\tau)_{1,1} = 1,$$

$$M_d(\tau) \succeq 0, \quad M_d(\tau) \in S_{\sigma(n,d)}.$$

$$M_d(g_j \tau) \succeq 0, \quad M_d(g_j \tau) \in S_{\sigma(n,d - d_j)}, \quad j = 1, \ldots, m.$$  \hfill (4.7)

The optimization variables of the semidefinite program (4.7) are the entries of the
moment and localizing matrices. If at a given relaxation order $d$, the computed
tracial moment matrix $M_d(\tau)$ is flat, then finite convergence of the hierarchy is
guaranteed and one can extract the minimizers of the corresponding optimization
problem, see, e.g., [BKP16, Theorem 1.69]. One research investigation would be
to approximate such minimizers when the flatness condition does not hold. The
algorithmic scheme would be to build an approximation of the noncommutative
Christoffel-Darboux kernel associated to $\tau$ from the inverse moment matrix (as
stated in Proposition 3.6) and perform a sampling procedure as done in the three
former subsections.

Acknowledgements

STB would like to thank Ahmed Zeriahi and Norm Levenberg for very helpful
discussions regarding plurisubharmonic functions and the Bernstein-Markov property.
All authors are grateful to Jean-Bernard Lasserre for encouraging and promoting
the line of research that led to this paper. Part of this work was conducted in
November 2019 while STB was a visiting fellow supported by the Faculty of Natural
Sciences Distinguished Scientist Visitors Program of the Ben Gurion University
of the Negev, Beer-Sheva, Israel, and during the three authors’ visit at the MFO
workshop I.D.: 2010, “Real Algebraic Geometry with a View Toward Hyperbolic
Programming and Free Probability” (1–7 March 2020). The second-named author
was supported by the Tremplin ERC Stg Grant ANR-18-ERC2-0004-01 (T-COPS
project), the European Union’s Horizon 2020 research and innovation programme.
under the Marie Sklodowska-Curie Actions, grant agreement 813211 (POEMA), as well as the FMJH Program PGMO (EPICS project). We would also like to thank the anonymous referee for a careful reading and numerous helpful comments that allowed us to improve the paper.

References

[AGZ10] Greg W Anderson, Alice Guionnet, and Ofer Zeitouni. *An introduction to random matrices*. Number 118. Cambridge university press, 2010, 4, 14

[AL12] Miguel F. Anjos and Jean B. Lasserre, editors. *Handbook on semidefinite, conic and polynomial optimization*, volume 166 of *International Series in Operations Research & Management Science*. Springer, New York, 2012, 2

[Ans08a] Michael Anshelevich. Monic non-commutative orthogonal polynomials. *Proceedings of the American Mathematical Society*, 136(7):2395–2405, 2008, 2

[Ans08b] Michael Anshelevich. Orthogonal polynomials with a resolvent-type generating function. *Transactions of the American Mathematical Society*, 360(8):4125–4143, 2008, 2, 8

[Ans10] Michael Anshelevich. Product-type non-commutative polynomial states. In *Noncommutative harmonic analysis with applications to probability. II*, volume 89 of *Banach Center Publications*. Institute of Mathematics, Polish Academy of Sciences, WARSZAWA, 2010, 2, 7, 8, 33

[BC04] T. Banks and T. Constantinescu. Orthogonal polynomials in several non-commuting variables. II. *arXiv:math/0412528v1 [math.FA]*, 2004, 2

[BCKP13] Sabine Burgdorf, Kristijan Cafuta, Igor Klep, and Janez Povh. The tracial moment problem and trace-optimization of polynomials. *Math. Program.*, 137(1-2, Ser. A):557–578, 2013, 2, 44

[BKP16] Sabine Burgdorf, Igor Klep, and Janez Povh. *Optimization of polynomials in non-commuting variables*. SpringerBriefs in Mathematics. Springer, [Cham], 2016, 2, 45

[Blö97] Thomas Bloom. Orthogonal polynomials in $\mathbb{C}^n$. *Indiana Univ. Math. J.*, 46(2):427–452, 1997, 23

[BLPW15] Thomas Bloom, Norman Levenberg, Federico Piazzon, and Franck Wielonsky. Bernstein-Markov: a survey. *arXiv:1512.00739v1 [math.CV]*, 2015, 23, 29

[BMV16] Joseph A. Ball, Gregory Marx, and Victor Vinnikov. Noncommutative reproducing kernel Hilbert spaces. *Journal of Functional Analysis*, 271:1844–1920, 2016, 11

[BPSS20] Bernhard Beckermann, Mihai Putinar, Edward B Saff, and Nikos Stylianopoulos. Perturbations of Christoffel–Darboux Kernels: Detection of outliers. *Foundations of Computational Mathematics*, pages 1–54, 2020, 1, 5, 15, 24

[CF98] Raúl E. Curto and Lawrence A. Fialkow. Flat extensions of positive moment matrices: recursively generated relations. *Mem. Amer. Math. Soc.*, 136(648):x+56, 1998, 2

[CKP12] Kristijan Cafuta, Igor Klep, and Janez Povh. Constrained polynomial optimization problems with noncommuting variables. *SIAM J. Optim.*, 22(2):363–383, 2012, 2

[Con02] Tiberiu Constantinescu. Orthogonal polynomials in several variables. I. *arXiv:math/0005539v1 [math.FA]*, 2002, 2

[DX14] Charles F Dunkl and Yuan Xu. Orthogonal polynomials of several variables. Number 155. Cambridge University Press, 2014, 1, 7, 12

[ER00] Edward G. Effros and Zhong-Jin Ruan. *Operator Spaces*, volume 23 of *London Mathematical Society Monographs*. Oxford University Press, 2000, 8

[EW97] Edward G. Effros and Soren Winkler. Matrix Convexity: Operator Analogues of the Bipolar and Hahn-Banach Theorems. *Journal of Functional Analysis*, 144:117–152, 1997, 16

[GS14] Alice Guionnet and Dimitri Shlyakhtenko. Free monotone transport. *Inventiones Mathematicae*, pages 613–661, 2014, 8

[GZ17] Vincent Guedj and Ahmed Zeriahi. *Degenerate Complex Monge-Ampère Equations*. Number 26 in Tracts in Mathematics. European Mathematical Society, 2017, 3, 4, 15, 21, 37, 39, 40

[Hel02] J. William Helton. “Positive” noncommutative polynomials are sums of squares. *Ann. of Math. (2)*, 156(2):675–694, 2002, 2
[HJ85] Roger A Horn and Charles R Johnson. Matrix analysis. Cambridge University Press, 1985. 13

[HL05] Didier Henrion and Jean-Bernard Lasserre. Detecting global optimality and extracting solutions in gloptipoly. In Positive polynomials in control, pages 293–310. Springer, 2005. 2

[HP+08] J William Helton, Jean B Lasserre, Mihai Putinar, et al. Measures with zeros in the inverse of their moment matrix. The Annals of Probability, 36(4):1453–1471, 2008. 44

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

[HMS96] JW Helton, RL Miller, and M Stankus. Ncalgebra: A mathematica package for doing non commuting algebra. available from ncalg@ucsd. edu, 1996. 41

[Kli91] Maciej Klimek. Pluripotential theory. Number 6 in London Mathematical Society Monographs. Oxford: Clarendon Press, 1991. 4

[KMP21] Igor Klep, Victor Magron, and Janez Povh. Sparse noncommutative polynomial optimization. Mathematical Programming, pages 1–41, 2021. 2

[KMV21] Igor Klep, Victor Magron, and Jurij Volčič. Optimization over trace polynomials. In Annales Henri Poincaré, pages 1–34. Springer, 2021. 2

[KVV14] Dmitry S Kaliuzhnyi-Verbovetskyi and Victor Vinnikov. Foundations of free noncommutative function theory, volume 199. American Mathematical Soc., 2014. 9

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

[Lau09] Monique Laurent. Sums of squares, moment matrices and optimization over polynomials. In Emerging applications of algebraic geometry, volume 149 of IMA Vol. Math. Appl., pages 157–270, Springer, New York, 2009. 2

[LP19] Jean B. Lasserre and Edouard Pauwels. The empirical Christoffel function with applications in data analysis. Advances in Computational Mathematics, 45(3):1439–1468, 2019. 1, 12, 14

McC01 Scott McCullough. Factorization of operator-valued polynomials in several non-commuting variables. Linear Algebra Appl., 326(1-3):193–203, 2001. 2

[MFH19] V. Magron, M. Forets, and D. Henrion. Semidefinite approximations of invariant measures for polynomial systems. Discrete & Continuous Dynamical Systems - B, 24:6745, 2019. 1

[MN80] Attila Máté and Paul G Nevi. Bernstein’s inequality in Lp for and (c, 1) bounds for orthogonal polynomials. Annals of Mathematics, pages 145–154, 1980. 1

[MNT91] Attila Máté, Paul Nevi, and Vilmos Totik. Szego’s extremum problem on the unit circle. Annals of Mathematics, pages 433–453, 1991. 1

[MPW+19] Swann Marx, Edouard Pauwels, Tillmann Weisser, Didier Henrion, and Jean Lasserre. Tractable semi-algebraic approximation using Christoffel-Darboux kernel. arXiv preprint arXiv:1904.01833, 2019. 2

[MS17] James A Mingo and Roland Speicher. Free probability and random matrices, volume 35, Springer, 2017. 42, 44

[Nev86] Paul Nevi. Géza Freud, orthogonal polynomials and Christoffel functions. A case study. Journal of Approximation Theory, 48(1):3–167, 1986. 12

[NA08] Miguel Navascués, Stefano Pironio, and Antonio Acín. A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations. New J. Phys., 10(7):073013, 2008. 2

[Pau02] Vern Paulsen. Completely bounded maps and operator algebras., volume 78 of Cambridge studies in advanced mathematics. Cambridge University Press, 2002. 2, 8, 12, 13

[PBV18] Edouard Pauwels, Francis Bach, and Jean-Philippe Vert. Relating leverage scores and density using regularized Christoffel functions. In Advances in Neural Information Processing Systems, pages 1669–1672, 2018. 1

[Pis03] Gilles Pisier. Introduction to operator space theory. Cambridge University Press, 2003. 8

[PL16] Edouard Pauwels and Jean B Lasserre. Sorting out typicality with the inverse moment matrix SOS polynomial. In Advances in Neural Information Processing Systems, pages 190–198, 2016. 1
[PNA10] Stefano Pironio, Miguel Navascués, and Antonio Acín. Convergent relaxations of polynomial optimization problems with noncommuting variables. *SIAM J. Optim.*, 20(5):2157–2180, 2010.

[PPL20] Edouard Pauwels, Mihai Putinar, and Jean-Bernard Lasserre. Data analysis from empirical moments and the Christoffel function. *Foundations of Computational Mathematics*, pages 1–31, 2020.

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

[Sim08] Barry Simon. The Christoffel–Darboux kernel. In “Perspectives in PDE, Harmonic Analysis and Applications,” a volume in honor of VG Maz’ya’s 70th birthday, *Proceedings of Symposia in Pure Mathematics*, volume 79, pages 295–335, 2008.

[ST97] Edward B. Saff and Vilmos Totik. *Logarithmic Potentials with External Fields (with an appendix by Thomas Bloom)*. Number 316 in Grundlehren der mathematischen Wissenschaften. Springer-Verlag Berlin Heidelberg, 1997.

[Tak79] Masamichi Takesaki. *Theory of Operator Algebras I*. Springer-Verlag, 1979.

[Tim94] Aleksandr Filipovich Timan. *Theory of approximation of functions of a real variable*. Dover Publications, Inc., 1994.

[VBP19] Mai Trang Vu, François Bachoc, and Edouard Pauwels. Rate of convergence for geometric inference based on the empirical Christoffel function. *arXiv preprint arXiv:1910.14458*, 2019.

[Voi85] Dan Voiculescu. Symmetries of some reduced free product $c^*$-algebras. In *Operator Algebras and Their Connections with Topology and Ergodic Theory*, volume 1132 of *Lecture Notes in Math.* Springer, 1985.

[Voi98] Dan Voiculescu. The analogues of entropy and of Fisher’s information measure in free probability theory, V: Noncommutative Hilbert transforms. *Invent. Math.*, 132:189–227, 1998.

[Voi00] Dan Voiculescu. The Coalgebra of the Free Difference Quotient and Free Probability. *International Mathematics Research Notices*, 2000(2):79–106, 2000.

[WM21] Jie Wang and Victor Magron. Exploiting term sparsity in noncommutative polynomial optimization. *Computational Optimization and Applications*, 80(2):483–521, 2021.