

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 μ on \mathbb{R}^n . Integrating with respect to μ 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 μ , 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 μ is uniform or empirical and when the support of μ satisfies specific compactness conditions, the sequence of level sets of the Christoffel function associated to μ 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 μ 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

2010 *Mathematics Subject Classification.* 90C22; 47N10; 13J10.

Key words and phrases. noncommutative polynomials; Christoffel-Darboux kernels; semialgebraic set; Bernstein-Markov property; semidefinite programming; trace optimization; GNS construction.

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 μ 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, WM20] allow one to reduce the associated computational burden. The case of more general trace polynomials has been investigated in [KMV20]. 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 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

Given $k, n \in \mathbb{N}$, let us denote by $\mathbb{M}_k(\mathbb{C})$ (resp. \mathbb{S}_k) the space of all complex (resp. hermitian) matrices of order k , and by \mathbb{S}_k^n the set of n -tuples $\underline{A} = (A_1, \dots, A_n)$ of hermitian matrices A_i of order k . Let I_k stands for the identity matrix of order k . For a fixed $n \in \mathbb{N}$, we consider a finite alphabet X_1, \dots, X_n of symmetric letters and generate all possible words of finite length in these letters. The empty word is denoted by 1. For $d \in \mathbb{N}$, $\langle \underline{X} \rangle_d$ is the subset of all resulting words of length at most d . The length of a word w is denoted by $|w|$. We use the expressions $<_{\text{gl}}$ and \leq_{gl} to denote the graded lexicographic order over n -variate words. Let $\mathbb{C}\langle \underline{X} \rangle$ be the set of polynomials in noncommutative variables $\underline{X} = (X_1, \dots, X_n)$. Let us denote by $\mathbf{W}_d(\underline{X})$ the vector of all words of $\langle \underline{X} \rangle_d$ ordered w.r.t. \leq_{gl} . Note that the dimension of $\mathbb{C}\langle \underline{X} \rangle_d$ equals the length of $\mathbf{W}_d(\underline{X})$, which is $\sigma(n, d) := \sum_{i=0}^d n^i = \frac{n^{d+1}-1}{n-1}$. We equip the set $\mathbb{C}\langle \underline{X} \rangle$ with the involution \star that conjugates the elements of \mathbb{C} , fixes $\{X_1, \dots, X_n\}$ point-wise and reverses words, in such a way that $\mathbb{C}\langle \underline{X} \rangle$ is the \star -algebra freely generated by n self-adjoint letters X_1, \dots, X_n . (It will be sometimes convenient to identify the set of words in n letters and the set of monomials in n selfadjoint indeterminates. We shall systematically use this identification in notations as well. Thus, if the word w has p letters, $w = \ell_1 \ell_2 \ell_3 \cdots \ell_p$, then we identify it with the degree p monomial $\underline{X}^w = X_{\ell_1} X_{\ell_2} X_{\ell_3} \cdots X_{\ell_p}$. Sometimes it will be convenient to denote a monomial simply by the word that it is indexed by, or a word by the monomial it indexes.) The set of all *self-adjoint elements* is defined as $\text{Sym}\mathbb{C}\langle \underline{X} \rangle := \{f \in \mathbb{C}\langle \underline{X} \rangle : f = f^\star\}$. Let $\tau : \mathbb{C}\langle \underline{X} \rangle \rightarrow \mathbb{C}$ be a positive tracial functional. Let $M_d(\tau)$ be the moment matrix of τ , i.e., the symmetric matrix indexed by words of $\langle \underline{X} \rangle_d$ ordered by \leq_{gl} , which is defined by $M_d(\tau)_{u,v} := \tau(u^\star v)$. Since τ is positive, the moment matrix $M_d(\tau)$ is positive semi-definite. If, in addition, τ 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, we consider a special subcategory of the set of such traces, which we shall call *bounded traces*. These are traces τ that satisfy the additional condition that there exists a real number $M > 0$ such that $|\tau(w)| < M^d$ for all $d \in \mathbb{N}$, $w \in \langle \underline{X} \rangle_d$. If M is known/given, we say that τ is bounded by M .

3. NONCOMMUTATIVE CHRISTOFFEL-DARBOUX KERNELS

3.1. Orthonormal polynomials. Assume τ is positive and faithful. For all $d \in \mathbb{N}$, let us define the family of orthonormal polynomials, $\{P_w\}_{w \in \langle \underline{X} \rangle_d}$, ordered according to the lexicographic order, and satisfying for all $v, w \in \langle \underline{X} \rangle_d$:

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

Of course, $\text{span}\{w : w <_{\text{gl}} v\} = \text{span}\{P_w : w <_{\text{gl}} v\}$ for all $v \in \langle \underline{X} \rangle$. Such a family can be constructed easily by using the classical Gram-Schmidt orthonormalization

process. Written in the language of the τ -induced inner product, it is given recursively by the initial condition $P_1(\underline{X}) = 1 \in \mathbb{C}$ and the general expression

$$P_w(\underline{X}) = \frac{w - \sum_{v <_{\text{gl}} w} \tau(P_v^*(\underline{X})w)P_v(\underline{X})}{\left(\tau(w^*w) - \sum_{v <_{\text{gl}} w} |\tau(P_v^*(\underline{X})w)|^2 \right)^{\frac{1}{2}}}, \quad w \in \langle \underline{X} \rangle_d.$$

It is immediately clear from the construction that $\tau(P_w^*P_w) = 1$, $\tau(P_v^*P_w) = 0$ if $v <_{\text{gl}} w$, and $\tau(w^*P_w) = \frac{\tau(w^*w) - \sum_{v <_{\text{gl}} w} \tau(P_v^*(\underline{X})w)\tau(w^*P_v(\underline{X}))}{\left(\tau(w^*w) - \sum_{v <_{\text{gl}} w} |\tau(P_v^*(\underline{X})w)|^2 \right)^{\frac{1}{2}}} > 0$. Of course, since

$\tau(P_w^*w) = \overline{\tau(w^*P_w)}$, we also have $\tau(P_w^*w) > 0$. Finally, if $v_0 <_{\text{gl}} w$, then v_0 is a linear combination of elements P_v , $v \leq_{\text{gl}} v_0$, so that $\tau(P_w v_0^*) = \sum_{v \leq_{\text{gl}} v_0} \bar{e}_v \tau(P_w P_v^*) = 0$. In the context, we should emphasize that faithfulness of τ automatically implies the denominator in the expression of $P_w(\underline{X})$ above is nonzero. If τ is not faithful, it may well happen that $\tau(w^*w) = \sum_{v <_{\text{gl}} w} |\tau(P_v^*(\underline{X})w)|^2$; 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]. The index set $\langle \underline{X} \rangle, \langle \underline{X} \rangle_d$ usually become smaller, but the choice of which monomials to “drop” may not be unique.

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

The fact that both $\{P_w\}_{w \in \langle \underline{X} \rangle_d}$ and $\{P_w^*\}_{w \in \langle \underline{X} \rangle_d}$ are orthonormal bases implies that any correspondence sending one to the other is a unitary transformation of $L^2(\mathbb{C}\langle \underline{X} \rangle_d, \tau)$. Thus, there exists a unitary matrix $U = (U_{v,w})_{v,w \in \langle \underline{X} \rangle_d}$ such that $P_{v_0}^*(\underline{X}) = \sum_{v \in \langle \underline{X} \rangle} U_{v_0,v} P_v(\underline{X})$. (As seen above, U leaves $\text{span}\{P_w : w \in \langle \underline{X} \rangle_d \setminus \langle \underline{X} \rangle_{d-1}\}$ 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, $P_{v_0}(\underline{X}) = \sum_{v \in \langle \underline{X} \rangle} (U^*)_{v_0,v} P_v^*(\underline{X}) = \sum_{v \in \langle \underline{X} \rangle} \overline{U_{v,v_0}} P_v^*(\underline{X})$, and $\delta_{v=v_0} = \sum_{w \in \langle \underline{X} \rangle_d} U_{v,w} \overline{U_{v_0,w}}$. Moreover, the orthonormality relations yield $\tau(P_{v_0}^*(\underline{X})P_{v_1}(\underline{X})) = \sum_{v \in \langle \underline{X} \rangle} U_{v_0,v} \tau(P_v(\underline{X})P_{v_1}^*(\underline{X})) = U_{v_0,v_1}$. By traciality of τ , we obtain that $U_{v_0,v_1} = U_{v_1,v_0}$ for all $v_0, v_1 \in \langle \underline{X} \rangle_d$, 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.

It should be noted that the polynomials P_w are not usually selfadjoint, and the construction offered above cannot be expected to provide selfadjoint orthonormal bases form a non-selfadjoint set of monomials. However, it is quite easy to create a family of selfadjoint orthogonal polynomials out of the hermitization of our monomials. Let us outline one of the several possible variations on the Gram-Schmidt process starting from, and yielding, selfadjoint elements. As it is well-known (and easy to see), if \underline{X}^w is not selfadjoint, then $\text{Span}_{\mathbb{C}}\{\underline{X}^w, \underline{X}^{w^*}\} = \text{Span}_{\mathbb{C}}\left\{ \frac{\underline{X}^w + \underline{X}^{w^*}}{2}, \frac{\underline{X}^w - \underline{X}^{w^*}}{2i} \right\}$, and $\frac{\underline{X}^{w^*} + \underline{X}^w}{2}, \frac{\underline{X}^w - \underline{X}^{w^*}}{2i}$ are both selfadjoint. Thus,

for any $d \in \mathbb{N}$, we take the set $\langle \underline{X} \rangle_d \setminus \langle \underline{X} \rangle_{d-1}$ and, to make a choice, we proceed through the set of monomials according to the lexicographic order. We proceed the following way: when we see a selfadjoint element, we leave it alone; when we see a non-selfadjoint monomial \underline{X}^w , we replace it with its *real part* $\Re \underline{X}^w = \frac{\underline{X}^w + \underline{X}^{w^*}}{2}$. At the same time, we replace \underline{X}^{w^*} (which, according to this procedure, must necessarily be greater than \underline{X}^w) with the *imaginary part* of \underline{X}^w , namely $\Im \underline{X}^w = \frac{\underline{X}^w - \underline{X}^{w^*}}{2i}$ (of course, the first element in the list is selfadjoint, namely X_1^d). This process yields a basis of $\text{Span}_{\mathbb{C}}(\langle \underline{X} \rangle_d \setminus \langle \underline{X} \rangle_{d-1})$ 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 selfadjoints X_1, \dots, X_n). For the sake of simplicity, we leave it ordered according to the same lexicographic order, where the position of selfadjoint monomials remains unchanged, the position of $\Re \underline{X}^w$ is at w , and at position $w^*_{\text{gl}} > w$ we find $\Im \underline{X}^w$. We apply the Gram-Schmidt procedure to this basis which, we claim, yields a family $\{S_w\}_{w \in \langle \underline{X} \rangle} \subset \text{Sym} \mathbb{C} \langle \underline{X} \rangle \subset \mathbb{C} \langle \underline{X} \rangle$ of selfadjoint orthogonal polynomials. Indeed, this is easily argued by induction after the ordered set of words. For 1, X_1 this is obvious ($S_1(\underline{X}) = P_1(\underline{X}) = 1$, $S_{X_1}(\underline{X}) = P_{X_1}(\underline{X}) = \frac{X_1 - \tau(X_1)}{\tau(X_1^2) - \tau(X_1)^2}$). Just in this paragraph, we agree to denote by $w \in \langle \underline{X} \rangle$ not the monomial \underline{X}^w , but the selfadjoint basis element it indexes according to the above procedure (thus, if $w = w^*$, then $w = \underline{X}^w$, if $w <_{\text{gl}} w^*$, then $w = \Re \underline{X}^w$, if $w^* <_{\text{gl}} w$, then $w = \Im \underline{X}^w$). If we assume that all $S_w(\underline{X})$, $w <_{\text{gl}} w_0$, were constructed, and are selfadjoints, then, as before, we define

$$S_{w_0}(\underline{X}) = \frac{w_0 - \sum_{v <_{\text{gl}} w_0} \tau(S_v(\underline{X})w_0)S_v(\underline{X})}{\left(\tau(w_0^*w_0) - \sum_{v <_{\text{gl}} w_0} |\tau(S_v(\underline{X})w_0)|^2 \right)^{\frac{1}{2}}}, \quad w_0 \in \langle \underline{X} \rangle_d.$$

By hypothesis, $w_0 = w_0^*$, $S_v(\underline{X}) = S_v^*(\underline{X})$. Moreover, $\overline{\tau(S_v(\underline{X})w_0)} = \tau((S_v(\underline{X})w_0)^*) = \tau(w_0^*S_v^*(\underline{X})) = \tau(w_0S_v(\underline{X})) = \tau(S_v(\underline{X})w_0)$, so that $\tau(S_v(\underline{X})w_0) \in \mathbb{R}$. Thus, $S_{w_0}(\underline{X}) = S_{w_0}^*(\underline{X})$. This completes our argument.

One may generally perform the above procedure also if τ is not faithful, but with more difficulty.

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 3.1.

- (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 $\mathbb{C} \langle X \rangle = \mathbb{C}[X]$. The state τ corresponding to it is simply the integration of the polynomial with respect to the given probability measure. τ 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 $\mathbb{C}[\underline{X}] = \mathbb{C}\langle \underline{X} \rangle / \langle X_i X_j = X_j X_i : 1 \leq i, j, \leq n \rangle$ with respect to the trace τ 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 $\mathbb{C}[\underline{X}]$. However, it is quite clear that such a trace does *not* satisfy our condition of faithfulness: since $-(X_1 X_2 - X_2 X_1)^2 = (X_1 X_2 - X_2 X_1)(X_2 X_1 - X_1 X_2) = (X_2 X_1 - X_1 X_2)^*(X_2 X_1 - X_1 X_2)$ is a positive polynomial in $\mathbb{C}\langle \underline{X} \rangle \setminus \{0\}$ but it is the zero polynomial in $\mathbb{C}[\underline{X}]$, we have $\tau(-(X_1 X_2 - X_2 X_1)^2) = \tau(0) = 0$, so τ is far from being faithful on $\mathbb{C}\langle \underline{X} \rangle$ 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 τ is bounded and the components X_1, \dots, X_n of \underline{X} are free [Voi85] with respect to τ . Assume moreover that τ is faithful, which, in this case, is equivalent to requiring that the - classical - distribution of X_j with respect to τ 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 τ . Given $w \in \langle \underline{X} \rangle_d$, we write it as $w = (w_1, w_2, \dots, w_d)$. We partition w in intervals given by consecutively repeating indices: $w = (\pi_1, \pi_2, \dots, \pi_\ell)$ for some $\ell \in \{1, \dots, d\}$. For example, if $n = 3, d = 10$, and $w = (2, 2, 2, 1, 3, 3, 1, 1, 1, 1)$, 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 π_j the orthonormal polynomial $P_{|\pi_j|}^{(\iota_j)}(X_{\iota_j})$, where $|\pi_j|$ is the cardinality of the block π_j , and $\iota_j \in \{1, \dots, n\}$ is the index that is contained by π_j . Then $P_w(\underline{X}) = P_{|\pi_1|}^{(\iota_1)}(X_{\iota_1}) P_{|\pi_2|}^{(\iota_2)}(X_{\iota_2}) \dots P_{|\pi_\ell|}^{(\iota_\ell)}(X_{\iota_\ell})$, $w \in \langle \underline{X} \rangle$, is a system of orthonormal polynomials for τ . (For example, with $w = (2, 2, 2, 1, 3, 3, 1, 1, 1, 1)$, 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, $\mathbb{C}\langle X_1, \dots, X_p \rangle$ and $\mathbb{C}\langle X_{p+1}, \dots, X_n \rangle$ are free with respect to τ is treated conceptually precisely the same way, and the orthonormal polynomials associated to τ on $\mathbb{C}\langle \underline{X} \rangle$ are again alternating products of orthonormal polynomials corresponding to the restrictions of τ to $\mathbb{C}\langle X_1, \dots, X_p \rangle$ and $\mathbb{C}\langle X_{p+1}, \dots, X_n \rangle$, respectively, but the notation becomes more involved. Faithfulness of τ on either of $\mathbb{C}\langle X_1, \dots, X_p \rangle$ or $\mathbb{C}\langle X_{p+1}, \dots, X_n \rangle$ 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 \mathcal{V}, \mathcal{W} over \mathbb{C} , the algebraic tensor product vector space over \mathbb{C} , $\mathcal{V} \otimes \mathcal{W}$, is well-defined. If \mathcal{V}, \mathcal{W} are complex algebras, this vector space has itself two algebra structures: $\mathcal{V} \otimes \mathcal{W}$ with the multiplication $(\sum_i v_i \otimes w_i)(\sum_j \tilde{v}_j \otimes \tilde{w}_j) = \sum_{i,j} v_i \tilde{v}_j \otimes w_i \tilde{w}_j$ and $\mathcal{V} \otimes \mathcal{W}^{\text{op}}$ with the multiplication $(\sum_i v_i \otimes w_i)(\sum_j \tilde{v}_j \otimes \tilde{w}_j) = \sum_{i,j} v_i \tilde{v}_j \otimes \tilde{w}_j w_i$. This second multiplication is important because it allows in some

circumstances the identification $\mathcal{V} \otimes \mathcal{V}^{\text{op}} \simeq \mathcal{L}(\mathcal{V}, \mathcal{V})$, the space of linear operators from \mathcal{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 $\mathbb{M}_k(\mathbb{C})$ and, when \mathcal{V} is a von Neumann algebra, there are completions of the range of $\mathcal{V} \otimes \mathcal{V}^{\text{op}}$ in $\mathcal{L}(\mathcal{V}, \mathcal{V})$ that are quite important in the theory of operator spaces.

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

$$\begin{aligned} (P_{ij}(\underline{X}))_{i,j=1}^k &= \left(\sum_w \alpha_{ij}^{(w)} P_w(\underline{X}) \right)_{i,j=1}^k &= \sum_{i,j=1}^k \sum_w \alpha_{ij}^{(w)} e_{ij} \otimes P_w(\underline{X}) \\ &= \sum_w \left(\sum_{i,j=1}^k \alpha_{ij}^{(w)} e_{ij} \right) \otimes P_w(\underline{X}) \\ &= \sum_w C_w((P_{ij}(\underline{X}))_{i,j=1}^k) \otimes P_w(\underline{X}), \end{aligned}$$

where $C_w((P_{ij}(\underline{X}))_{i,j=1}^k) \in \mathbb{M}_k(\mathbb{C})$, $C_w((P_{ij}(\underline{X}))_{i,j=1}^k) = \sum_{i,j=1}^k \alpha_{ij}^{(w)} e_{ij}$. Written under this form, the star operation is simply $(\sum_w C_w((P_{ij}(\underline{X}))_{i,j=1}^k) \otimes P_w(\underline{X}))^* = \sum_w C_w((P_{ij}(\underline{X}))_{i,j=1}^k)^* \otimes P_w^*(\underline{X})$. We define an $\mathbb{M}_k(\mathbb{C})$ -valued sesquilinear form $\langle (P_{ij}(\underline{X}))_{i,j=1}^k, (Q_{ij}(\underline{X}))_{i,j=1}^k \rangle = (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau)((P_{ij}(\underline{X}))_{i,j=1}^k [(Q_{ij}(\underline{X}))_{i,j=1}^k]^*) = \sum_w C_w((P_{ij}(\underline{X}))_{i,j=1}^k) C_w((Q_{ij}(\underline{X}))_{i,j=1}^k)^*$. With these notations, if $A \in \mathbb{M}_k(\mathbb{C})$, $\mathbf{P}, \mathbf{Q} \in \mathbb{M}_k(\mathbb{C}\langle \underline{X} \rangle)$, then $\langle \mathbf{P}A, \mathbf{Q} \rangle = \langle \mathbf{P}, \mathbf{Q}A^* \rangle$, $\langle \mathbf{A}\mathbf{P}, \mathbf{Q} \rangle = A\langle \mathbf{P}, \mathbf{Q} \rangle$, $\langle \mathbf{P}, \mathbf{A}\mathbf{Q} \rangle = \langle \mathbf{P}, \mathbf{Q} \rangle A^*$. This makes $\mathbb{M}_k(\mathbb{C}\langle \underline{X} \rangle)$ into a $\mathbb{M}_k(\mathbb{C})$ -Hilbert module (see [Pau02, Chapter 14]). From now on, we denote matrix-valued polynomials by boldface letters.

Remark 3.2. For fixed $0 \neq k \in \mathbb{N}$, a polynomial $\mathbf{P} \in \mathbb{M}_k(\mathbb{C}\langle \underline{X} \rangle)$, $\mathbf{P}(\underline{X}) = \sum_{w \in \langle \underline{X} \rangle} a_w \otimes P_w(\underline{X})$ is an nc function. Indeed, consider an arbitrary unital C^* -algebra \mathcal{A} (we will mostly need the case $\mathcal{A} = \mathbb{M}_l(\mathbb{C})$ for some l , possibly different from k) and an $\mathbb{M}_k(\mathbb{C})$ - \mathcal{A} -bimodule \mathcal{B} (usually $\mathbb{M}_{k \times l}(\mathbb{C})$). One defines the function

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

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

$$\mathbf{P}(\underline{a}, b) = \sum_{w \in \langle \underline{X} \rangle} (a_w \otimes I_l) b P_w(\underline{a}) = \sum_{w \in \langle \underline{X} \rangle} [\text{diag}(\underbrace{a_w, \dots, a_w}_{\iota \text{ times}})] b P_w(\underline{a}).$$

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

evaluations on possibly non-selfadjoint tuples, $\mathbf{P}(\underline{A})^* = \left[\sum_{w \in \langle \underline{X} \rangle} c_w \otimes \underline{A}^w \right]^* = \sum_{w \in \langle \underline{X} \rangle} c_w^* \otimes (\underline{A}^w)^* = \sum_{w \in \langle \underline{X} \rangle} c_w^* \otimes (\underline{A}^*)^{w^*} = \sum_{w \in \langle \underline{X} \rangle} c_{w^*}^* \otimes (\underline{A}^*)^w = \mathbf{P}^*(\underline{A}^*)$. It is useful to record this equality, together with the one corresponding to scalar polynomials:

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

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

$$\begin{aligned} \langle C, \mathbf{P}(\underline{A})(D) \rangle_{\text{HS}} &= \text{Tr}_k \left(\left[\sum_{w \in \langle \underline{X} \rangle} c_w D \underline{A}^w \right]^* C \right) = \text{Tr}_k \left(\sum_{w \in \langle \underline{X} \rangle} (\underline{A}^w)^* D^* c_w^* C \right) \\ &= \text{Tr}_k \left(D^* \sum_{w \in \langle \underline{X} \rangle} c_w^* C (\underline{A}^*)^w \right) = \text{Tr}_k (D^* \mathbf{P}^*(\underline{A}^*)(C)) \\ &= \langle \mathbf{P}^*(\underline{A}^*)(C), D \rangle_{\text{HS}}, \quad C, D \in \mathbb{M}_k(\mathbb{C}), \underline{A} \in \mathbb{M}_k(\mathbb{C})^n, k \in \mathbb{N}. \end{aligned}$$

It should be emphasized that re-normalizing Tr changes nothing in the above. As a consequence, for a II_1 factor \mathcal{M} with normal faithful trace-state tr , if $\mathbf{P}(\underline{X}) \in \mathcal{M} \otimes \mathbb{C}\langle \underline{X} \rangle$, then $[\mathbf{P}(\underline{a})]^*(c) = \mathbf{P}^*(\underline{a}^*)(c)$ for any $\underline{a} \in \mathcal{M}^n, c \in L^2(\mathcal{M}, \text{tr})$.

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

$$(3.3) \quad \sup_{\|C\|_{\text{HS}}=1} \|\mathbf{P}(\underline{A})(C)\|_{\text{HS}}^2 = \sup_{\|C\|_{\text{HS}}=1} \text{Tr}_k \left(\sum_{v, w \in \langle \underline{X} \rangle} (\underline{A}^v)^* C^* c_v^* c_w C \underline{A}^w \right),$$

$$(3.4) \quad \sup \|\mathbf{P}(\underline{A}) \sum_{i, j=1}^k \alpha_{i, j} e_i \otimes e_j\|_2^2 = \sup \sum_{v, w \in \langle \underline{X} \rangle} \sum_{i, j, p, q=1}^k \alpha_{i, j} \bar{\alpha}_{p, q} \langle c_v e_i, c_w e_p \rangle \langle \underline{A}^v e_j, \underline{A}^w e_q \rangle,$$

where the supremum in the second equality is taken after all arrays $\alpha_{i, j}, 1 \leq i, j \leq k$ 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 Tr_k or tr_k in defining the HS norm.) The first expression identifies with the second by simply splitting C into matrix units. This form identifies $\mathbf{P}(\underline{A})$ acting on $(\mathbb{M}_k(\mathbb{C}), \langle \cdot, \cdot \rangle_{\text{HS}})$ as an element in the tensor product of von Neumann algebras $\mathbb{M}_k(\mathbb{C}) \otimes \mathbb{M}_k(\mathbb{C})^{\text{op}}$ and $\mathbf{P}(\underline{A})$ acting on $(\mathbb{C}^k \otimes \mathbb{C}^k, \langle \cdot, \cdot \rangle)$ as an element of $\mathbb{M}_k(\mathbb{C}) \otimes \mathbb{M}_k(\mathbb{C})$. In this case it follows that $\mathbf{P}(\underline{A})$ has the same norm regardless of whether it is viewed in one or in the other. It should also be noted that while the spectral properties of $\mathbf{P}(\underline{A})$ do not change regardless of whether it is viewed as acting on $\mathbb{C}^k \otimes \mathbb{C}^k$, on $(\mathbb{M}_k(\mathbb{C}), \|\cdot\|_{\text{HS}})$, or on $(\mathbb{M}_k(\mathbb{C}), \|\cdot\|)$, positivity properties do change, in the sense that $\mathbf{P}(\underline{A})$ may be a positive operator on $(\mathbb{M}_k(\mathbb{C}), \|\cdot\|_{\text{HS}})$ but not (completely) positive on $(\mathbb{M}_k(\mathbb{C}), \|\cdot\|)$, or vice-versa.

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

$$(3.5) \quad \kappa_{\tau, d}(\underline{X}, \underline{Y}) := \sum_{w \in \langle \underline{X} \rangle_d} P_w(\underline{X}) \otimes P_w^*(\underline{Y}).$$

If τ is *not faithful*, then the above notation $w \in \langle \underline{X} \rangle_d$ should be understood as running only through a subset of *linearly independent* monomials in the completion $L^2(\tau) := L^2(\mathbb{C}\langle \underline{X} \rangle, \tau)$ of $\mathbb{C}\langle \underline{X} \rangle$ with respect to $\|P\|_2 = \tau(P^*P)^{\frac{1}{2}}$ which span $\mathbb{C}\langle \underline{X} \rangle_d$ – 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, $\{P_v^*\}_{v \in \langle \underline{X} \rangle_d}$ is an orthonormal basis as well, so that the correspondence between the elements of $\{P_w\}_{w \in \langle \underline{X} \rangle_d}$ and $\{P_v^*\}_{v \in \langle \underline{X} \rangle_d}$ is unitary. With the notations from Section 3.1,

$$\begin{aligned}
\sum_{w \in \langle \underline{X} \rangle_d} P_w(\underline{X}) \otimes P_w^*(\underline{Y}) &= \sum_{w \in \langle \underline{X} \rangle_d} \left(\sum_{v_0 \in \langle \underline{X} \rangle_d} (U^*)_{w, v_0} P_{v_0}^*(\underline{X}) \right) \otimes \left(\sum_{v \in \langle \underline{X} \rangle_d} U_{w, v} P_v(\underline{Y}) \right) \\
&= \sum_{v, v_0 \in \langle \underline{X} \rangle_d} \left(\sum_{w \in \langle \underline{X} \rangle_d} U_{w, v} (U^*)_{w, v_0} \right) P_{v_0}^*(\underline{X}) \otimes P_v(\underline{Y}) \\
(3.6) \quad &= \sum_{v, v_0 \in \langle \underline{X} \rangle_d} \delta_{v=v_0} P_{v_0}^*(\underline{X}) \otimes P_v(\underline{Y}) = \sum_{v \in \langle \underline{X} \rangle_d} P_v^*(\underline{X}) \otimes P_v(\underline{Y}).
\end{aligned}$$

(As shown in Section 3.1, U is symmetric, so that $U_{w, v} (U^*)_{w, v_0} = U_{v, w} (U^*)_{w, v_0}$.)

One may view $\kappa_{\tau, d}$ as a reproducing kernel for the finite-dimensional space $L^2(\mathbb{C}\langle \underline{X} \rangle_d, \tau)$: we tautologically have

$$\begin{aligned}
(\text{Id}_{\mathbb{C}\langle \underline{X} \rangle} \otimes \tau)(\kappa_{\tau, d}(\underline{X}, \underline{Y})(1 \otimes P(\underline{Y}))) &= \sum_{v \in \langle \underline{X} \rangle_d} (\text{Id}_{\mathbb{C}\langle \underline{X} \rangle} \otimes \tau)(P_v(\underline{X}) \otimes P_v^*(\underline{Y})P(\underline{Y})) \\
&= \sum_{v \in \langle \underline{X} \rangle_d} P_v(\underline{X}) \otimes \tau(P_v^*(\underline{Y})P(\underline{Y}))1 \\
&= \sum_{v \in \langle \underline{X} \rangle_d} \tau(P_v^*(\underline{Y})P(\underline{Y}))P_v(\underline{X}) \otimes 1 \\
&= P(\underline{X}) \otimes 1,
\end{aligned}$$

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

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

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

$$\begin{aligned}
(\tau \otimes \tau)(\kappa_{\tau, d}(\underline{X}, \underline{Y})^* \kappa_{\tau, d}(\underline{X}, \underline{Y})) &= \sum_{v, w \in \langle \underline{X} \rangle_d} (\tau \otimes \tau)(P_v^*(\underline{X})P_w(\underline{X}) \otimes P_w^*(\underline{Y})P_v(\underline{Y})) \\
&= \sum_{v, w \in \langle \underline{X} \rangle_d} \tau(P_v^*(\underline{X})P_w(\underline{X}))\tau(P_v(\underline{Y})P_w^*(\underline{Y})) \\
&= \sum_{w \in \langle \underline{X} \rangle_d} \tau(P_w^*(\underline{X})P_w(\underline{X}))\tau(P_w(\underline{Y})P_w^*(\underline{Y})) \\
&= \sum_{w \in \langle \underline{X} \rangle_d} 1 \cdot 1 = \sigma(n, d),
\end{aligned}$$

by orthonormality of the $\{P_w\}_{w \in \langle \underline{X} \rangle_d}$.

$\kappa_{\tau,d}$ is in fact a noncommutative kernel; indeed, it acts on $\mathbb{M}_{k \times k'}(\mathbb{C})$ via $\mathbb{M}_k(\mathbb{C})^n \times \mathbb{M}_{k'}(\mathbb{C})^n \mapsto \mathcal{L}(\mathbb{M}_{k \times k'}(\mathbb{C}), \mathbb{M}_{k \times k'}(\mathbb{C}))$, $(\underline{A}, \underline{B}) \mapsto \left[X \mapsto \sum_{w \in \langle \underline{X} \rangle_d} P_w(\underline{A}) X P_w^*(\underline{B}) \right]$ and thus $\kappa_{\tau,d}(\underline{A}, \underline{A}^*) \geq 0$, i.e. it is completely positive, for all $\underline{A} \in \mathbb{M}_k(\mathbb{C})^n$. We note for future use that, since $P_1(\underline{X}) = 1$ (see (3.1)), we have $\kappa_{\tau,d}(\underline{A}, \underline{A})(C) = \sum_{w \in \langle \underline{X} \rangle_d} P_w(\underline{A}) C P_w^*(\underline{A}) \succeq P_1(\underline{A}) C P_1^*(\underline{A}) = C$ whenever evaluated on elements of a unital C^* -algebra in which the coordinates of \underline{A} are selfadjoint and $C \succeq 0$.

For each $\underline{A} = (A_1, \dots, A_n) \in \mathbb{M}_k(\mathbb{C})^n$, let us define $\Lambda_{\tau,d}(\underline{A}) := \kappa_{\tau,d}(\underline{A}, \underline{A}^*)(I_k)^{-1}$, as an operator in $\mathbb{M}_k(\mathbb{C})$. The function $\Lambda_{\tau,d}$ is called the *noncommutative Christoffel function* associated to τ and d . We see from the above that $\Lambda_{\tau,d}$ is well-defined on \mathcal{A}^n for any unital C^* -algebra \mathcal{A} , and in particular it is well defined on all of $\mathbb{M}_k(\mathbb{C})^n$, $k \in \mathbb{N}$, and $\Lambda_{\tau,d}(\underline{A}) \preceq I_k$, $\underline{A} \in \mathbb{M}_k(\mathbb{C})^n$.

We now prove the noncommutative analog of [LP19, Theorem 3.1], which can be found in e.g. [DX14] and [Nev86].

Theorem 3.3. *Let $\underline{A} \in \mathbb{M}_k(\mathbb{C})^n$ be fixed, arbitrary. Then*

$$(3.8) \quad \Lambda_{\tau,d}(\underline{A}) = \min \left\{ (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau)(\mathbf{P}\mathbf{P}^*) : \mathbf{P} \in \mathbb{M}_k(\mathbb{C}\langle \underline{X} \rangle_d), \mathbf{P}(\underline{A})(I_k) = I_k \right\}.$$

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

$$\mathbf{P}_d(\underline{A})(I_k) = \sum_{w \in \langle \underline{X} \rangle_d} c_w(\underline{A}) I_k P_w(\underline{A})^* = \Lambda_{\tau,d}(\underline{A}) \sum_{w \in \langle \underline{X} \rangle_d} P_w(\underline{A}) I_k P_w(\underline{A})^* = \Lambda_{\tau,d}(\underline{A}) \kappa_{\tau,d}(\underline{A}, \underline{A}^*)(I_k) = I_k.$$

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

$$\mathbf{P}_d(\underline{X}) \mathbf{P}_d^*(\underline{X}) = \sum_{v,w \in \langle \underline{X} \rangle_d} c_w(\underline{A}) c_v(\underline{A})^* \otimes P_v(\underline{X}) P_w^*(\underline{X}),$$

so

$$\begin{aligned} (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau)(\mathbf{P}_d(\underline{X}) \mathbf{P}_d^*(\underline{X})) &= \sum_{w \in \langle \underline{X} \rangle_d} c_w(\underline{A}) c_w(\underline{A})^* \\ &= \sum_{w \in \langle \underline{X} \rangle_d} \Lambda_{\tau,d}(\underline{A}) P_w(\underline{A}) P_w(\underline{A})^* \Lambda_{\tau,d}(\underline{A}) \\ &= \Lambda_{\tau,d}(\underline{A}) \kappa_{\tau,d}(\underline{A}, \underline{A}^*)(I_k) \Lambda_{\tau,d}(\underline{A}) = \Lambda_{\tau,d}(\underline{A}), \end{aligned}$$

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

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

Consider an arbitrary polynomial $\mathbf{P}(\underline{X}) = \sum_{w \in \langle \underline{X} \rangle_d} a_w \otimes P_w(\underline{X})$ such that $\mathbf{P}(\underline{A})(I_k) =$

I_k . By the definition of the $\mathbb{M}_k(\mathbb{C})$ -valued inner product, this last expression is equivalent to stating $\langle (a_w)_w, (P_w(\underline{A})^*)_w \rangle_d = I_k$. With the Hilbert module structure (see [Pau02, Chapter 14]), we are guaranteed the positivity

$$(3.9) \quad \begin{aligned} &\begin{bmatrix} \langle (a_w)_w, (a_w)_w \rangle_d & \langle (a_w)_w, (P_w(\underline{A})^*)_w \rangle_d \\ \langle (P_w(\underline{A})^*)_w, (a_w)_w \rangle_d & \langle (P_w(\underline{A})^*)_w, (P_w(\underline{A})^*)_w \rangle_d \end{bmatrix} \\ &= \begin{bmatrix} \langle (a_w)_w, (a_w)_w \rangle_d & I_k \\ I_k & \langle (P_w(\underline{A})^*)_w, (P_w(\underline{A})^*)_w \rangle_d \end{bmatrix} \succeq 0. \end{aligned}$$

As $\langle (P_w(\underline{A})^*)_w, (P_w(\underline{A})^*)_w \rangle_d^{-1} = [\sum P_w(\underline{A})^* P_w(\underline{A})]^{-1} = \Lambda_{\tau,d}(\underline{A})$ (see (3.6)) is known to exist, the positivity of the above matrix is equivalent to the relation

$$I_k \langle (P_w(\underline{A})^*)_w, (P_w(\underline{A})^*)_w \rangle_d^{-1} I_k \preceq \langle (a_w)_w, (a_w)_w \rangle_d = \sum_{w \in \langle \underline{X} \rangle_d} a_w a_w^* = (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau)(\mathbf{P}\mathbf{P}^*),$$

as claimed. \square

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 τ is bounded, it can be realized as the distribution of a tuple \underline{a}_τ of selfadjoints in a von Neumann algebra with respect to a normal faithful trace tr , and then $(\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau)(\mathbf{P}\mathbf{P}^*) = (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \text{tr})(\mathbf{P}(\underline{a}_\tau)\mathbf{P}(\underline{a}_\tau)^*)$, the conditional expectation of a positive element in a von Neumann algebra. We mention below some re-interpretations of Theorem 3.3; one remarkable point is that the normalization $\mathbf{P}(\underline{A})(I_k) = I_k$ is in fact quite arbitrary, even though useful.

Remark 3.4.(1) *The positivity of the matrix in (3.9) imposes automatically the domination relation $\langle (a_w)_w, (P_w(\underline{A})^*)_w \rangle_d \Lambda_{\tau,d}(\underline{A}) \langle (P_w(\underline{A})^*)_w, (a_w)_w \rangle_d \preceq \langle (a_w)_w, (a_w)_w \rangle_d$ for all vectors $(a_w)_w \in \mathbb{M}_k(\mathbb{C})^{|\langle \underline{X} \rangle_d|}$. Identifying the optimal solution $(a_w)_{w \in \langle \underline{X} \rangle_d} = (\Lambda_{\tau,d}(\underline{A}) P_w^*(\underline{A}))_{w \in \langle \underline{X} \rangle_d}$ comes simply to identifying the element which transforms \preceq into $=$ while not creating a null space through conjugation with $\langle (a_w)_w, (P_w(\underline{A})^*)_w \rangle_d$, i.e. while imposing that this element stays invertible (of course, the most convenient choice to impose is $\langle (a_w)_w, (P_w^*(\underline{A}))_w \rangle_d = I_k$).*
(2) *Under the assumption that $\langle (a_w)_w, (a_w)_w \rangle_d$ is invertible, positivity of the left-most matrix in relation (3.9) is equivalent to*

$$\mathbf{P}(\underline{A})(I_k)^* (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau)(\mathbf{P}(\underline{X})\mathbf{P}^*(\underline{X}))^{-1} \mathbf{P}(\underline{A})(I_k) \preceq \kappa_{\tau,d}(\underline{A}, \underline{A}^*)(I_k).$$

By employing an arbitrarily small perturbation, the above holds also if $\langle (a_w)_w, (a_w)_w \rangle_d$ is not invertible, in the sense that the left-hand side stays bounded by $\kappa_{\tau,d}(\underline{A}, \underline{A}^)(I_k)$ when the perturbation tends to zero. This relation can be viewed as the noncommutative equivalent of the inequality $|P(\underline{x})|^2 \leq \kappa_{\mu,d}(\underline{x}, \underline{x}) \int |P(t)|^2 d\mu(t)$, $\underline{x} \in \mathbb{R}^n$. Moreover, Equation (3.8) is then rewritten as*

$$\begin{aligned} & \kappa_{\tau,d}(\underline{A}, \underline{A}^*)(I_k) \\ &= \max\{\mathbf{P}(\underline{A})(I_k)^* \mathbf{P}(\underline{A})(I_k) : \mathbf{P} \in \mathbb{M}_k(\mathbb{C}\langle \underline{X} \rangle_d), (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau)(\mathbf{P}\mathbf{P}^*) = I_k\}. \end{aligned}$$

Indeed, the Schur complement in the left-hand side of the previous domination relation remains invariant under multiplication of \mathbf{P} to the left with $G \otimes 1$ for any $G \in GL_k(\mathbb{C})$.

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

(4) *A brief inspection of the proof of Theorem 3.3 shows that the result holds as well when one replaces $\mathbb{M}_k(\mathbb{C})$ with any finite factor \mathcal{A} (and in particular, $\mathbb{M}_k(\mathbb{C}\langle \underline{X} \rangle)$ with $\mathcal{A} \otimes \mathbb{C}\langle \underline{X} \rangle$). The essential element of the proof is that $\mathcal{A}^{\sigma^{(n,d)}}$ accepts the \mathcal{A} -Hilbert module structure $\langle (a_w)_w, (b_w)_w \rangle_d = \sum a_w b_w^*$.*

We record for future use the expression of the minimizer in (3.8): for any operator algebra \mathcal{A} on a Hilbert space,

$$(3.10) \quad \mathbf{P}_d(\underline{X}) = (\Lambda_{\tau,d}(\underline{a}) \otimes 1) \sum_{w \in \langle \underline{X} \rangle_d} P_w(\underline{a}) \otimes P_w^*(\underline{X}) \in \mathcal{A} \otimes \mathbb{C}\langle \underline{X} \rangle_d, \quad \underline{a} \in (\mathcal{A}^{\text{sa}})^n.$$

In particular, for $\mathcal{A} = \mathbb{M}_k(\mathbb{C})$, we obtain the formula in Theorem 3.3. The optimization problem stated in (3.8) is the noncommutative analog of [LP19, (3.2)].

The view of the evaluation $\mathbf{P}(\underline{A})(I_k)$ as a matrix-valued inner product has been essential in the proof of Theorem 3.3. Another re-interpretation of this evaluation might be more intuitive, namely as a matrix product. Tautologically, for $\mathbf{P}(\underline{X}) = \sum_{w \in \langle \underline{X} \rangle_d} c_w \otimes \underline{X}^w$, we have

$$(3.11) \quad \mathbf{P}(\underline{A})(I_k) = \mathbf{c}\mathfrak{A} = \begin{bmatrix} c_0 & c_{X_1} & \cdots & c_{X_n^{d-1}X_{n-1}} & c_{X_n^d} \end{bmatrix} \begin{bmatrix} 1 \\ A_1 \\ \vdots \\ A_n^{d-1}A_{n-1} \\ A_n^d \end{bmatrix}.$$

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

$$(3.12) \quad \mathbf{P}(\underline{A})(I_k) = [\mathbf{c}\mathbf{A}]_{1,1}$$

is the $(1, 1)$ entry of the product $\mathbf{c}\mathbf{A} \in \mathbb{M}_{\sigma(n,d)}(\mathbb{M}_k(\mathbb{C}))$, or as the upper left $k \times k$ block of $\mathbb{M}_{k\sigma(n,d)}(\mathbb{C})$.

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

Proposition 3.5. *Let $\tau: \mathbb{C}\langle \underline{X} \rangle \rightarrow \mathbb{C}$ be a faithful tracial state. For any n -tuples of selfadjoint operators $\underline{A}, \underline{B}$ and any bounded operator C on a Hilbert space, one has $\kappa_{\tau,d}(\underline{A}, \underline{B})(C) = \mathbf{W}_d(\underline{A})^T D_d(\tau)^T C D_d(\tau) \mathbf{W}_d^*(\underline{B})$; in particular, $\kappa_{\tau,d}(\underline{A}, \underline{B})(1) = \mathbf{W}_d(\underline{A})^T M_d(\tau)^{-1} \mathbf{W}_d^*(\underline{B})$.*

3.3. Nc plurisubharmonic functions. Our next objective is to investigate the asymptotic properties of $\Lambda_{\tau,d}$ as $d \rightarrow \infty$ when evaluated on various domains. In order to be able to perform this analysis, we are forced to place some restrictions both on the nature of τ 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 $\tau: \mathbb{C}\langle \underline{X} \rangle \rightarrow \mathbb{C}$, we construct via the GNS construction the finite von Neumann algebra generated by the n selfadjoint elements whose joint law equals τ . We denote it by $W^*(\tau)$. Since it is a finite von Neumann algebra, it is endowed with at least one normal faithful tracial state

tr (in brief, for any finite von Neumann algebra \mathcal{M} , there exists a unique tracial conditional expectation onto its center, $\mathcal{C}: \mathcal{M} \rightarrow \mathcal{Z}(\mathcal{M})$; any normal faithful tracial state on \mathcal{M} is obtained by composing \mathcal{C} with a state on $\mathcal{Z}(\mathcal{M})$; since $\mathcal{Z}(\mathcal{M})$ is an Abelian C^* -algebra, any state on it is known via the Gelfand-Naimark and the Riesz representation theorems; in particular, if \mathcal{M} is a factor, then $\mathcal{Z}(\mathcal{M}) = \mathbb{C} \cdot 1$ and \mathcal{C} is the unique normal faithful tracial state on \mathcal{M}). Denote by $\underline{x}_\tau = (x_1, \dots, x_n) \in W^*(\tau)$ the tuple of selfadjoint elements generating $W^*(\tau)$ whose joint law is τ ; that is, we choose tr such that the star-algebra homomorphism $(\mathbb{C}\langle \underline{X} \rangle, \tau) \mapsto (W^*(\tau), \text{tr})$ given by $X_j \mapsto x_j$ is injective and trace-preserving, i.e. $\tau(X_{i_1} \cdots X_{i_l}) = \text{tr}(x_{i_1} \cdots x_{i_l}), l \in \mathbb{N}, i_1, \dots, i_l \in \{1, \dots, n\}$.

In the following, the faithfulness of τ will often be assumed for convenience. We will however try to indicate what changes occur in various arguments if τ 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 \underline{a}_τ to be a tuple of selfadjoint variables in a von Neumann algebra \mathcal{M} which is distributed according to τ with respect to the normal faithful tracial state tr on \mathcal{M} . For each $d, k \in \mathbb{N}$, $\underline{A} \in \mathbb{M}_k(\mathbb{C})^n$, we define

$$(3.13) \quad \Phi_{\tau,d}^2(\underline{A}) = \sup \{ \text{tr}_k(\mathbf{P}(\underline{A})(I_k)^* \mathbf{P}(\underline{A})(I_k)) : \mathbf{P} \in \mathbb{M}_k(\mathbb{C}\langle \underline{X} \rangle_d), \|\mathbf{P}(\underline{a}_\tau)\| \leq 1 \},$$

$$(3.14) \quad \Phi_{\tau,d}^\infty(\underline{A}) = \sup \{ \|\mathbf{P}(\underline{A})(I_k)\|^2 : \mathbf{P} \in \mathbb{M}_k(\mathbb{C}\langle \underline{X} \rangle_d), \|\mathbf{P}(\underline{a}_\tau)\| \leq 1 \}.$$

Recall that $\mathbf{P}(\underline{a}_\tau) \in \mathbb{M}_k(W^*(\tau))$, so that the norm $\|\mathbf{P}(\underline{a}_\tau)\|$ is the operator norm on $\mathbb{M}_k(W^*(\tau)) \simeq \mathbb{M}_k(\mathbb{C}) \otimes W^*(\tau) \subseteq \mathbb{M}_k(\mathbb{C}) \otimes \mathcal{M}$. Note that in (3.13) we have defined $\Phi_{\tau,d}^2$ as the trace of the product of the two matrices obtained by applying the linear operator $\mathbf{P}(\underline{A})$ to I_k . However, we would have obtained precisely the same result had we defined it as $\text{tr}_k([\mathbf{P}^*(\underline{A}^*)\mathbf{P}(\underline{A})](I_k))$. Clearly $\Phi_{\tau,d}^2(\underline{A}) \leq \Phi_{\tau,d}^\infty(\underline{A})$ for all \underline{A} . Let us list next various properties of these functions which will be of use to us later.

3.3.1. *Plurisubharmonicity.* Both $\Phi_{\tau,d}^2(\underline{A})$ and $\Phi_{\tau,d}^\infty(\underline{A})$ are supremums of norms of analytic polynomials, hence entire plurisubharmonic functions in the classical sense when viewed as functions on the complex Euclidean space $\mathbb{C}^{nk^2} \simeq (\mathbb{M}_k(\mathbb{C}))^n$. Moreover, $\|\mathbf{P}(\underline{A})(I_k)\| = \lim_{m \rightarrow \infty} \text{tr}_k([\mathbf{P}(\underline{A})(I_k)^* \mathbf{P}(\underline{A})(I_k)]^m)^{1/2m}$.

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

3.3.3. *Compactness.* The set of polynomials $\{\mathbf{P} \in \mathbb{M}_k(\mathbb{C}\langle \underline{X} \rangle_d) : \|\mathbf{P}(\underline{a}_\tau)\| \leq 1\}$ is compact, in the sense that for any given basis $\{b_1, \dots, b_{\sigma(n,d)}\}$ of $\mathbb{C}\langle \underline{X} \rangle_d$, the set

$\{(a_1, \dots, a_{\sigma(n,d)}) \in \mathbb{M}_k(\mathbb{C})^{\sigma(n,d)} : \|\sum a_j \otimes b_j(\underline{a}_\tau)\| \leq 1\}$ is compact¹ 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 $\{b_1, \dots, b_{\sigma(n,d)}\} = \{P_w\}_{w \in \langle \underline{X} \rangle_d}$, the basis of orthonormal polynomials. In that case, $\sum_w a_w a_w^* = \sum_{v,w} a_w a_v^* \otimes \text{tr}(P_v^*(\underline{a}_\tau) P_w(\underline{a}_\tau)) = (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \text{tr})(\sum_{v,w} a_w a_v^* \otimes P_v^*(\underline{a}_\tau) P_w(\underline{a}_\tau)) = (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \text{tr})((\sum_w a_w \otimes P_w(\underline{a}_\tau))(\sum_w a_w \otimes P_w(\underline{a}_\tau))^*) = (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \text{tr})(\mathbf{P}(\underline{a}_\tau) \mathbf{P}(\underline{a}_\tau)^*)$ for any $\mathbf{P}(\underline{X}) = \sum_w a_w \otimes P_w(\underline{X}) \in \mathbb{M}_k(\mathbb{C}\langle \underline{X} \rangle_d)$. Thus, the condition $\|\mathbf{P}(\underline{a}_\tau)\| \leq 1$ implies

(3.15)

$$1 \geq \|\mathbf{P}(\underline{a}_\tau)\|^2 = \|\mathbf{P}(\underline{a}_\tau) \mathbf{P}(\underline{a}_\tau)^*\| \geq \|(\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \text{tr})(\mathbf{P}(\underline{a}_\tau) \mathbf{P}(\underline{a}_\tau)^*)\| = \left\| \sum_{w \in \langle \underline{X} \rangle_d} a_w a_w^* \right\|$$

so that $\|a_w\|^2 \leq 1$ for all $w \in \langle \underline{X} \rangle_d$. 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. The set $\{\mathbf{P} \in \mathbb{M}_k(\mathbb{C}\langle \underline{X} \rangle_d) : \|\mathbf{P}(\underline{a}_\tau)\| \leq 1\}$ is also convex. In fact it is matrix convex, and the proof is quite trivial: if $\mathbf{P} \in \mathbb{M}_k(\mathbb{C}\langle \underline{X} \rangle)$, $\mathbf{P}' \in \mathbb{M}_{k'}(\mathbb{C}\langle \underline{X} \rangle)$ satisfy the desired norm inequality, then $\|\mathbf{P}(\underline{a}_\tau) \oplus \mathbf{P}'(\underline{a}_\tau)\| \leq 1$ as well in $\mathbb{M}_{k+k'}(\mathbb{C}) \otimes \mathcal{M}$ (to be clear, if $\mathbf{P}(\underline{X}) = \sum_w c_w \otimes P_w(\underline{X})$, $\mathbf{P}'(\underline{X}) = \sum_w c'_w \otimes P_w(\underline{X})$, we view $\mathbf{P}(\underline{X}) \oplus \mathbf{P}'(\underline{X}) = \sum_w (c_w \oplus c'_w) \otimes P_w(\underline{X}) \in \mathbb{M}_{k+k'}(\mathbb{C}\langle \underline{X} \rangle)$). If $\alpha \in \mathbb{M}_{k \times k'}(\mathbb{C})$ satisfies $\alpha^* \alpha = I_{k'}$ (so implicitly $k' \leq k$), then

$$\left\| \sum_w (\alpha^* c_w \alpha) \otimes P_w(\underline{a}_\tau) \right\| = \|(\alpha^* \otimes 1) \mathbf{P}(\underline{a}_\tau) (\alpha \otimes 1)\| \leq \|\alpha\|^2 \|\mathbf{P}(\underline{a}_\tau)\| \leq 1.$$

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 $\mathbf{P} \in \mathbb{M}_k(\mathbb{C}\langle \underline{X} \rangle_d)$ is a polynomial such that one of the equalities $\Phi_{\tau,d}^2(\underline{A}) = \text{tr}_k(\mathbf{P}(\underline{A})(I_k)^* \mathbf{P}(\underline{A})(I_k))$, $\Phi_{\tau,d}^\infty(\underline{A}) = \|\mathbf{P}(\underline{A})(I_k)\|^2$ takes place, then the corresponding equality still holds for the polynomial $(U \otimes 1) \mathbf{P}$ for any $U \in \mathbb{M}_k(\mathbb{C})$ satisfying $UU^* = I_k$, i.e. for any unitary matrix U . Indeed, this is straightforward since $\text{tr}_k([(U \otimes 1) \mathbf{P}](\underline{A})(I_k)^* [(U \otimes 1) \mathbf{P}](\underline{A})(I_k)) = \text{tr}_k((\sum_w (U c_w) \underline{A}^w)^* (\sum_v (U c_v) \underline{A}^v)) = \text{tr}_k((\sum_w (\underline{A}^w)^* c_w^*) U^* U (\sum_v c_v \underline{A}^v)) = \text{tr}_k(\mathbf{P}(\underline{A})(I_k)^* \mathbf{P}(\underline{A})(I_k))$, $\|[(U \otimes 1) \mathbf{P}](\underline{A})(I_k)\| = \|\sum_w U c_w \underline{A}^w\| = \|U(\sum_w c_w \underline{A}^w)\| = \|\sum_w c_w \underline{A}^w\| = \|\mathbf{P}(\underline{A})(I_k)\|$. Moreover, we have as well that if $\|\mathbf{P}(\underline{a}_\tau)\| \leq 1$, then $\|[(U \otimes 1) \mathbf{P}](\underline{a}_\tau)\| \leq 1$ as well. This is simply because $U \otimes 1$ is unitary in $\mathbb{M}_k(\mathbb{C}) \otimes W^*(\tau)$ whenever U is in $\mathbb{M}_k(\mathbb{C})$ and $\|[(U \otimes 1) \mathbf{P}](\underline{a}_\tau)\| = \|(U \otimes 1)(\mathbf{P}(\underline{a}_\tau))\| = \|\mathbf{P}(\underline{a}_\tau)\|$. 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 \mathbf{P} satisfies the condition

¹It is important to remember that we work here under the hypothesis of faithfulness for τ . If, for instance, \underline{a}_τ were to satisfy some algebraic relation – that is, there is a $P \in \mathbb{C}\langle \underline{X} \rangle$ such that $P(\underline{a}_\tau) = 0$ – 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 \underline{a}_τ are complex multiples of the algebra's unit.

$\mathbf{P}(\underline{A})(I_k) \succeq 0$: indeed, if $\mathbf{P}(\underline{A})(I_k) = U|\mathbf{P}(\underline{A})(I_k)|$, then we replace \mathbf{P} by $(U^* \otimes 1)\mathbf{P}$. Unfortunately that does not mean $\mathbf{P} \geq 0$.

With the same methods, one shows that $\Phi_{\tau,d}^\bullet(\underline{A}) \leq \Phi_{\tau,d}^\bullet(U\underline{A}U^*)$, $\bullet \in \{2, \infty\}$, for all unitary $k \times k$ complex matrices U , so that $\Phi_{\tau,d}^\bullet$ is constant on unitary orbits. Indeed, picking $\mathbf{P}(\underline{X}) = \sum_{w \in \langle \underline{X} \rangle_d} c_w \otimes \underline{X}^w$ for which the maximum $\Phi_{\tau,d}^\bullet(\underline{A}) = \|\mathbf{P}(\underline{A})(I_k)\|_\bullet^2$ is reached, one has

$$\begin{aligned} \mathbf{Q}(U\underline{A}U^*)(I_k)\mathbf{Q}(U\underline{A}U^*)(I_k)^* &= \sum_{v \in \langle \underline{X} \rangle_d} (c_v U^*) U \underline{A}^v U^* \left(\sum_{w \in \langle \underline{X} \rangle_d} (c_w U^*) U \underline{A}^w U^* \right)^* \\ &= \sum_{v,w \in \langle \underline{X} \rangle_d} c_v (U^* U) \underline{A}^v (\underline{A}^w)^* (U^* U) c_w^* \\ &= \mathbf{P}(\underline{A})(I_k) \mathbf{P}(\underline{A})(I_k)^*, \end{aligned}$$

where $\mathbf{Q}(\underline{X}) = \mathbf{P}(\underline{X})[U^* \otimes 1]$. As $1 = \|\mathbf{P}(\underline{a}_\tau)\| = \|\mathbf{P}(\underline{a}_\tau)[U^* \otimes 1]\| = \|\mathbf{Q}(\underline{a}_\tau)\|$, it follows that $\Phi_{\tau,d}^\bullet(U\underline{A}U^*) \geq \|\mathbf{Q}(U\underline{A}U^*)(I_k)\|_\bullet^2 = \|\mathbf{P}(\underline{A})(I_k)\|_\bullet^2 = \Phi_{\tau,d}^\bullet(\underline{A})$, 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 $\mathbf{P}(\underline{X}) = \sum_{w \in \langle \underline{X} \rangle_d} c_w \otimes \underline{X}^w$ and V, W are $k \times k$ unitaries, then

$$V \sum_{w \in \langle \underline{X} \rangle_d} c_w \underline{A}^w W = \sum_{w \in \langle \underline{X} \rangle_d} V c_w W (W^* \underline{A} W)^w = [(V \otimes 1) \mathbf{P}(W \otimes 1)] (W^* \underline{A} W)(I_k);$$

since $\|[(V \otimes 1) \mathbf{P}(W \otimes 1)](\underline{a}_\tau)\| = \|\mathbf{P}(\underline{a}_\tau)\|$, it follows that if $\Phi_{\tau,d}^\bullet(\underline{A}) = \|\mathbf{P}(\underline{A})(I_k)\|_\bullet^2$ is achieved on \mathbf{P} , then $\Phi_{\tau,d}^\bullet(\underline{A}) = \Phi_{\tau,d}^\bullet(W^* \underline{A} W) \geq \|[(V \otimes 1) \mathbf{P}(W \otimes 1)](W^* \underline{A} W)(I_k)\|_\bullet^2 = \|\mathbf{P}(\underline{A})(I_k)\|_\bullet^2 = \Phi_{\tau,d}^\bullet(\underline{A})$, that is,

$$(3.16) \quad \Phi_{\tau,d}^\bullet(\underline{A}) = \left\| V \sum_{w \in \langle \underline{X} \rangle_d} c_w \underline{A}^w W \right\|_\bullet^2 = \|[(V \otimes 1) \mathbf{P}(W \otimes 1)](W^* \underline{A} W)(I_k)\|_\bullet^2.$$

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

3.3.6. Domination. The condition $\|\mathbf{P}(\underline{a}_\tau)\| \leq 1$ in relations (3.13) – (3.14) is equivalent to either of $\mathbf{P}(\underline{a}_\tau)^* \mathbf{P}(\underline{a}_\tau) \preceq I_k \otimes 1$, $\mathbf{P}(\underline{a}_\tau) \mathbf{P}(\underline{a}_\tau)^* \preceq I_k \otimes 1$ in $\mathbb{M}_k(\mathbb{C}) \otimes W^*(\tau)$. If there exists a $0 \prec C \preceq I_k$ such that, say, $\mathbf{P}(\underline{a}_\tau) \mathbf{P}(\underline{a}_\tau)^* \preceq C \otimes 1$, then $(C^{-1/2} \otimes 1) \mathbf{P}(\underline{a}_\tau) \mathbf{P}(\underline{a}_\tau)^* (C^{-1/2} \otimes 1) \preceq I_k \otimes 1$, ensuring that the polynomial $(C^{-1/2} \otimes 1) \mathbf{P}(\underline{X})$ belongs to the acceptable set in the right hand side of (3.13) and (3.14). Then $((C^{-1/2} \otimes 1) \mathbf{P}(\underline{A}))(I_k) = C^{-1/2} \mathbf{P}(\underline{A})(I_k)$, so that $\text{tr}_k(\mathbf{P}(\underline{A})(I_k)^* C^{-1} \mathbf{P}(\underline{A})(I_k)) \geq \text{tr}_k(\mathbf{P}(\underline{A})(I_k)^* \mathbf{P}(\underline{A})(I_k))$, $\|\mathbf{P}(\underline{A})(I_k)^* C^{-1} \mathbf{P}(\underline{A})(I_k)\| \geq \|\mathbf{P}(\underline{A})(I_k)^* \mathbf{P}(\underline{A})(I_k)\|$. (Since $C^{-1} \not\preceq I_k$, equality in the previous inequality of traces above can only take place if $\ker \mathbf{P}(\underline{A})(I_k)^* \mathbf{P}(\underline{A})(I_k) \neq \{0\}$.) This makes it clear that without loss of generality we may assume in both (3.13) and (3.14) that there is no element $C \not\preceq I_k$ which is not a projection such that $\mathbf{P}(\underline{a}_\tau) \mathbf{P}(\underline{a}_\tau)^* \preceq C \otimes 1$.

However, we note that one may obtain a “minimal maximizer”: assume that $\Phi_{\tau,d}^\infty(\underline{A})$ is achieved at $\mathbf{P}_\underline{A}(\underline{X}) = \sum_{w \in \langle \underline{X} \rangle_d} c_w \otimes \underline{X}^w$. Then there exists a rank-one projection $p \in \mathbb{M}_k(\mathbb{C})$ such that $\|\mathbf{P}_\underline{A}(\underline{A})(I_k)\|_\bullet^2 = \|p \mathbf{P}_\underline{A}(\underline{A})(I_k)\|_\bullet^2$ (one simply picks

p to be the projection onto a norm-one vector $\xi \in \mathbb{C}^k$ such that $\|\mathbf{P}_{\underline{A}}(\underline{A})(I_k)\|^2 = \langle \mathbf{P}_{\underline{A}}(\underline{A})(I_k) \mathbf{P}_{\underline{A}}(\underline{A})(I_k)^* \xi, \xi \rangle$. Since $\mathbf{P}_{\underline{A}}(\underline{A})(I_k) = \sum_{w \in \langle \underline{X} \rangle_d} c_w \underline{A}^w$ it follows that

$$\begin{aligned} \Phi_{\tau,d}^\infty(\underline{A}) &= \|p \mathbf{P}_{\underline{A}}(\underline{A})(I_k)\|^2 = \left\| p \left(\sum_{w \in \langle \underline{X} \rangle_d} c_w \underline{A}^w \right) \left(\sum_{w \in \langle \underline{X} \rangle_d} c_w \underline{A}^w \right)^* p \right\| \\ &= \left\| \left(\sum_{w \in \langle \underline{X} \rangle_d} p c_w \underline{A}^w \right) \left(\sum_{w \in \langle \underline{X} \rangle_d} p c_w \underline{A}^w \right)^* \right\| = \|[(p \otimes 1) \mathbf{P}_{\underline{A}}](\underline{A})(I_k)\|^2 \end{aligned}$$

and of course $[(p \otimes 1) \mathbf{P}_{\underline{A}}](\underline{X}) = \sum_{w \in \langle \underline{X} \rangle_d} (p c_w) \otimes \underline{X}^w$, so $\|[(p \otimes 1) \mathbf{P}_{\underline{A}}](\underline{a}_\tau)\| = 1$, with $[(p \otimes 1) \mathbf{P}_{\underline{A}}](\underline{a}_\tau) [(p \otimes 1) \mathbf{P}_{\underline{A}}](\underline{a}_\tau)^* \preceq p \otimes 1$. On the other hand, for $\Phi_{\tau,d}^2$ it is clear that the orthogonal complement of the kernel of $\mathbf{P}_{\underline{A}}(\underline{A})(I_k)^*$ is the biggest projection with which we can multiply $\mathbf{P}_{\underline{A}}$ on the left.

3.3.7. Linear change of coordinates. Assume that $\mathbf{u} \in GL_n(\mathbb{R})$. We consider the n -tuple $\underline{\mathfrak{J}}_\theta = \mathbf{u} \underline{a}_\tau$. 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 \mathbf{f} defined on an nc neighborhood of \underline{a}_τ so that $\mathbf{f}(\underline{a}_\tau)$ is a tuple of selfadjoints). One can identify $\mathbb{C}\langle \underline{X} \rangle$ with the Fock space of the n -dimensional Hilbert space \mathbb{C}^n via the identification of 1 with the vacuum vector Ω and X_j with e_j , the j^{th} canonical base vector of \mathbb{C}^n . The product $X_i X_j$ corresponds to $e_i \otimes e_j$ etc. With this view, one sees immediately that the effect of \mathbf{u} on \underline{a}_τ extends to all of $\mathbb{C}\langle \underline{a}_\tau \rangle$ by acting on $(\mathbb{C}^n)^{\otimes p}$ as $\mathbf{u}^{\otimes p}$. Thus, action on the subspace $\mathbb{C}\Omega \oplus \bigoplus_{j=1}^d (\mathbb{C}^n)^{\otimes j}$ is done via $\text{Id}_{\mathbb{C}\Omega} \oplus \bigoplus_{j=1}^d \mathbf{u}^{\otimes j}$. Obviously, this assigns (easily computable but long) coefficients $\mathbf{u}_{w,v} \in \mathbb{R}$ in the block-matrix expression of the operator: $\mathbf{u}^{\otimes j} e_w = \sum_{|v|=j} (\mathbf{u}^{\otimes j})_{w,v} e_v$. Given a matrix-valued polynomial $\mathbf{P}(\underline{X}) = \sum_{w \in \langle \underline{X} \rangle_d} c_w \otimes \underline{X}^w$, the change of variable $\underline{\mathfrak{J}}_\theta = \mathbf{u} \underline{a}_\tau$ translates into

$$\begin{aligned} \mathbf{P}(\underline{\mathfrak{J}}_\theta) &= \mathbf{P}(\mathbf{u} \underline{a}_\tau) = \sum_{w \in \langle \underline{X} \rangle_d} c_w \otimes (\mathbf{u} \underline{a}_\tau)^w \\ &= c_0 \otimes 1 + \sum_{j=1}^d [I_k \otimes \mathbf{u}^{\otimes j}] \sum_{w \in \langle \underline{X} \rangle_j \setminus \langle \underline{X} \rangle_{j-1}} c_w \otimes \underline{a}_\tau^w \\ &= c_0 \otimes 1 + \sum_{j=1}^d \sum_{w \in \langle \underline{X} \rangle_j \setminus \langle \underline{X} \rangle_{j-1}} c_w \otimes \left(\sum_{v \in \langle \underline{X} \rangle_j \setminus \langle \underline{X} \rangle_{j-1}} (\mathbf{u}^{\otimes j})_{w,v} \underline{a}_\tau^v \right) \\ &= c_0 \otimes 1 + \sum_{j=1}^d \sum_{v \in \langle \underline{X} \rangle_j \setminus \langle \underline{X} \rangle_{j-1}} \underbrace{\left[\sum_{w \in \langle \underline{X} \rangle_j \setminus \langle \underline{X} \rangle_{j-1}} (\mathbf{u}^{\otimes j})_{w,v} c_w \right]}_{=d_v} \otimes \underline{a}_\tau^v = \mathbf{Q}(\underline{a}_\tau), \end{aligned}$$

where $\mathbf{Q}(\underline{X}) = \sum_{v \in \langle \underline{X} \rangle_d} d_v \otimes \underline{X}^v$. Tautologically $\|\mathbf{P}(\underline{\mathfrak{J}}_\theta)\| = 1 \iff \|\mathbf{Q}(\underline{a}_\tau)\| = 1$. For given $\underline{A} \in \mathbb{M}_k(\mathbb{C})^n$, assume \mathbf{Q} is a polynomial on which the maximum in (3.13) (resp. (3.14)) is achieved for \underline{a}_τ . Thus, from the above,

$$\Phi_{\tau,d}^\bullet(\underline{A}) = \|\mathbf{Q}(\underline{A})(I_k)\|_\bullet^2 = \left\| \sum_{v \in \langle \underline{X} \rangle_d} d_v \underline{A}^v \right\|_\bullet^2.$$

$$\begin{aligned}
&= \left\| c_0 + \sum_{j=1}^d \sum_{v \in \langle \underline{X} \rangle_j \setminus \langle \underline{X} \rangle_{j-1}} \left[\sum_{w \in \langle \underline{X} \rangle_j \setminus \langle \underline{X} \rangle_{j-1}} (\mathbf{u}^{\otimes j})_{w,v} c_w \right] \underline{A}^v \right\|_{\bullet}^2 \\
&= \left\| c_0 + \sum_{j=1}^d \sum_{w \in \langle \underline{X} \rangle_j \setminus \langle \underline{X} \rangle_{j-1}} c_w \left[\sum_{v \in \langle \underline{X} \rangle_j \setminus \langle \underline{X} \rangle_{j-1}} (\mathbf{u}^{\otimes j})_{w,v} \underline{A}^v \right] \right\|_{\bullet}^2 \\
&= \left\| \sum_{w \in \langle \underline{X} \rangle_d} c_w (\mathbf{u} \underline{A})^w \right\|_{\bullet}^2 = \|\mathbf{P}(\mathbf{u} \underline{A})(I_k)\|_{\bullet}^2 \leq \Phi_{\theta, d}^{\bullet}(\mathbf{u} \underline{A}).
\end{aligned}$$

Since the transformation \mathbf{u} is invertible, it follows that strict inequality in the last relation would lead to the existence of a $\tilde{\mathbf{P}}$ on which the supremum is realized, and then the corresponding $\tilde{\mathbf{Q}}$ would achieve the obvious contradiction to the above $\|\tilde{\mathbf{Q}}(\underline{A})(I_k)\|_{\bullet}^2 > \Phi_{\tau, d}^{\bullet}(\underline{A})$. Thus,

$$(3.17) \quad \Phi_{\tau, d}^{\bullet}(\underline{A}) = \Phi_{\theta, d}^{\bullet}(\mathbf{u} \underline{A}), \quad \underline{A} \in \mathbb{M}_k(\mathbb{C})^n.$$

Clearly, faithfulness of τ does not play a role in this argument; moreover, faithfulness of τ is equivalent to faithfulness of θ . The case of invertible, analytic, but nonlinear \mathbf{u} is more involved and leads to an acceptably strong result only for the objects defined in (3.18)–(3.19) below. We are not concerned with it at this moment.

3.3.8. *Growth.* The functions $\underline{A} \mapsto \frac{1}{2d} \log \Phi_{\tau, d}^{\bullet}(\underline{A})$ have logarithmic growth at infinity, in the sense that there exists a real constant $c_{d, \bullet, \tau}$ such that

$$\frac{1}{2d} \log \Phi_{\tau, d}^{\bullet}(\underline{A}) \leq \log^+ \|\underline{A}\|_{\bullet} + c_{d, \bullet, \tau}.$$

(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, $\|\underline{A}\|_2^2 = \text{tr}_k(A_1^* A_1 + \cdots + A_n^* A_n) \leq n \max_{1 \leq i \leq n} \text{tr}_k(A_i^* A_i) \leq n \max_{1 \leq i \leq n} \|A_i\|^2 = n \|\underline{A}\|_2^2$ and $\|\underline{A}\|_{\bullet}^2 = \max_{1 \leq i \leq n} \|A_i\|^2 \leq k \max_{1 \leq i \leq n} \text{tr}_k(A_i^* A_i) \leq k \text{tr}_k(A_1^* A_1 + \cdots + A_n^* A_n) = k \|\underline{A}\|_2^2$, so that $\frac{1}{\sqrt{n}} \|\underline{A}\|_2 \leq \|\underline{A}\|_{\bullet} \leq \sqrt{k} \|\underline{A}\|_2$. By taking log, we obtain $\log \|\underline{A}\|_2 - \log \sqrt{n} \leq \log \|\underline{A}\|_{\bullet} \leq \log \|\underline{A}\|_2 + \log \sqrt{k}$. Thus, we will compare everything to $\log \|\underline{A}\|_2$. Now, for any given polynomial $\mathbf{P} \in \mathbb{M}_k(\mathbb{C}\langle \underline{X} \rangle_d) \setminus \mathbb{M}_k(\mathbb{C}\langle \underline{X} \rangle_{d-1})$, it follows from the classical theory that

$$\limsup_{\|\underline{A}\|_2 \rightarrow \infty} \frac{\log \text{tr}_k(\mathbf{P}(\underline{A})(I_k)^* \mathbf{P}(\underline{A})(I_k))}{2d \log \|\underline{A}\|_2} \leq 1.$$

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

3.3.9. *Single variable bounds.* As noted before, our hypothesis that τ is positive guarantees that \underline{a}_{τ} can be chosen as a tuple of selfadjoint random variables, $\underline{a}_{\tau}^* = \underline{a}_{\tau}$, which generate a tracial C^* -algebra $(C^*(\tau), \text{tr})$. For each coordinate of \underline{a}_{τ} , 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 tr . Without loss of generality, pick the first coordinate, $a_{\tau,1}$, for our analysis. Denote by μ_1 its distribution with respect to τ , 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 $C^*(\tau|_{\mathbb{C}\langle X_1 \rangle}) = C(\sigma(a_{\tau,1}), \mu_1)$ is an infinite-dimensional Abelian C^* -algebra. For each $k \in \mathbb{N}$, we may choose k distinct points in $\sigma(a_{\tau,1})$; the evaluation of an element from $C(\sigma(a_{\tau,1}), \mu_1)$ at those points yields an algebra morphism onto \mathbb{C}^k , viewed as the diagonal of $\mathbb{M}_k(\mathbb{C})$. In particular, for given matrix size k , one may pick pairwise distinct points $\mathcal{P} = \{p_1, \dots, p_k\} \subset \sigma(a_{\tau,1})$, and define the linear map $\mathcal{C}_{\mathcal{P}}: C(\sigma(a_{\tau,1}), \mu_1) \rightarrow \mathbb{C}^k$ by $f \mapsto \mathcal{C}_{\mathcal{P}}(f) = (f(p_1), \dots, f(p_k))$. 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 $\mathbf{P}(X) = \sum_{w \in \langle X \rangle_d} c_w \otimes X^w \in \mathbb{M}_k(\mathbb{C}\langle X \rangle_d)$, then $\left\| \sum_{w \in \langle X \rangle_d} c_w \otimes \mathcal{C}_{\mathcal{P}}(a_{\tau,1}^w) \right\| = \|(\text{id}_k \otimes \mathcal{C}_{\mathcal{P}})(\mathbf{P}(a_{\tau,1}))\| \leq \|\mathbf{P}(a_{\tau,1})\|$. However, $\mathcal{C}_{\mathcal{P}}$ is an algebra morphism, so that $\mathcal{C}_{\mathcal{P}}(a_{\tau,1}^w) = \mathcal{C}_{\mathcal{P}}(a_{\tau,1})^w$ for all $w \in \mathbb{N}$ (or $w \in \mathbb{Z}$, if $a_{\tau,1}$ is invertible). As noted after equations (3.3)–(3.4), this, together with the fact that $\text{tr}_k(I_k) = 1$, guarantees that $\text{tr}_k(\mathbf{P}(\mathcal{C}_{\mathcal{P}}(a_{\tau,1}))(I_k)^* \mathbf{P}(\mathcal{C}_{\mathcal{P}}(a_{\tau,1}))(I_k)) \leq \|\mathbf{P}(a_{\tau,1})\|^2$ for any collection $\mathcal{P} \subset \sigma(a_{\tau,1})$ of k mutually disjoint points. In particular, we have established that for any selfadjoint matrix $A \in \mathbb{M}_k(\mathbb{C})$ whose spectrum is included in the spectrum of $a_{\tau,1}$, the condition $\|\mathbf{P}(a_{\tau,1})\| \leq 1$ implies automatically $\Phi_{\tau|_{\mathbb{C}\langle X_1 \rangle}, d}^2(A) \leq 1$.

We define

$$(3.18) \quad \Phi_{\tau}^2(\underline{A}) = \left[\limsup_{d \rightarrow \infty} (\Phi_{\tau, d}^2(\underline{A}))^{1/d} \right]^*,$$

$$(3.19) \quad \Phi_{\tau}^{\infty}(\underline{A}) = \left[\limsup_{d \rightarrow \infty} (\Phi_{\tau, d}^{\infty}(\underline{A}))^{1/d} \right]^*.$$

(The asterisk denotes the upper semicontinuous regularization: the upper semicontinuous regularization f^* for a function f is defined by $f^*(\zeta) = \limsup_{z \rightarrow \zeta} f(z)$, and is the smallest upper semicontinuous function with the property that $f^* \geq f$.)

We also define

$$(3.20) \quad \Sigma_{\tau}^2(\underline{A}) = \left[\sup_{d \in \mathbb{N}} (\Phi_{\tau, d}^2(\underline{A}))^{1/d} \right]^*,$$

$$(3.21) \quad \Sigma_{\tau}^{\infty}(\underline{A}) = \left[\sup_{d \in \mathbb{N}} (\Phi_{\tau, d}^{\infty}(\underline{A}))^{1/d} \right]^*.$$

Clearly, $\Sigma_{\tau}^{\bullet}(\underline{A}) \geq \Phi_{\tau}^{\bullet}(\underline{A})$, $\bullet \in \{2, \infty\}$, $\underline{A} \in \mathbb{M}_k(\mathbb{C})^n$, $k \in \mathbb{N}$. It is known (see, for instance, [ST97, Equation (2.17) in Appendix B2]) that Σ and Φ 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. We enumerate below the consequences of the above-listed properties of $\Phi_{\tau, d}^{\bullet}$ on these functions.

$$(1) \quad (3.22) \quad \Phi_{\tau}^2(\underline{A}) \leq \Phi_{\tau}^{\infty}(\underline{A}), \quad \Sigma_{\tau}^2(\underline{A}) \leq \Sigma_{\tau}^{\infty}(\underline{A}) \quad \text{quasi-everywhere,} \quad \underline{A} \in \mathbb{M}_k(\mathbb{C})^n, k \in \mathbb{N};$$

- (2) As limits of plurisubharmonic functions (see 3.3.1), according to [GZ17, Proposition 1.40 and Corollary 1.47] or [Kli91, Theorem 2.9.14 and Proposition 2.9.17], $\Sigma_\tau^2(\underline{A})$, $\Sigma_\tau^\infty(\underline{A})$, $\Phi_\tau^2(\underline{A})$, and $\Phi_\tau^\infty(\underline{A})$ are plurisubharmonic as functions on the complex Euclidean space \mathbb{C}^{nk^2} . More precisely, if the family $\underline{A} \mapsto \left(\Phi_{\tau,d}^\infty(\underline{A})\right)^{1/d}$, $d \in \mathbb{N}$, is locally bounded from the above, then all the above functions are plurisubharmonic and the set where they differ from their upper semicontinuous regularization has Lebesgue measure zero in \mathbb{C}^{nk^2} (in fact it is of zero inner/outer capacity [GZ17, Corollary 4.43]). In the following we assume implicitly that $\underline{A} \mapsto \left(\Phi_{\tau,d}^\infty(\underline{A})\right)^{1/d}$ is locally bounded from the above in order for most of our statements to be non-vacuous, and will show in Example 3.10 that this happens in numerous cases;
- (3) If we denote $\tau_1 = \tau|_{\mathbb{C}\langle X_1, \dots, X_n \rangle}$, $\tau_2 = \tau|_{\mathbb{C}\langle X_{n+1}, \dots, X_{n+m} \rangle}$, then $\Phi_\tau^\bullet(\underline{A}, \underline{B}) \geq \max\{\Phi_{\tau_1}^\bullet(\underline{A}), \Phi_{\tau_2}^\bullet(\underline{B})\}$, $\Sigma_\tau^\bullet(\underline{A}, \underline{B}) \geq \max\{\Sigma_{\tau_1}^\bullet(\underline{A}), \Sigma_{\tau_2}^\bullet(\underline{B})\}$ for any $(\underline{A}, \underline{B}) \in \mathbb{M}_k(\mathbb{C})^n \times \mathbb{M}_k(\mathbb{C})^m$, $\bullet \in \{2, \infty\}$. 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

$$(3.23) \quad \Phi_\tau^\bullet(\underline{A}) = \Phi_\tau^\bullet(U\underline{A}U^*), \quad \Sigma_\tau^\bullet(\underline{A}) = \Sigma_\tau^\bullet(U\underline{A}U^*), \quad \underline{A} \in \mathbb{M}_k(\mathbb{C})^n, U \in \mathbb{M}_k(\mathbb{C}), UU^* = I_k.$$

As an immediate consequence, we record

$$\Phi_\tau^\bullet \left(\begin{bmatrix} \underline{A} & \underline{B} \\ \underline{B} & \underline{A}^* \end{bmatrix} \right) = \Phi_\tau^\bullet \left(\begin{bmatrix} \underline{A}^* & \underline{B} \\ \underline{B} & \underline{A} \end{bmatrix} \right), \quad \Sigma_\tau^\bullet \left(\begin{bmatrix} \underline{A} & \underline{B} \\ \underline{B} & \underline{A}^* \end{bmatrix} \right) = \Sigma_\tau^\bullet \left(\begin{bmatrix} \underline{A}^* & \underline{B} \\ \underline{B} & \underline{A} \end{bmatrix} \right),$$

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

- (5) If $\mathbf{u} \in GL_n(\mathbb{R})$, τ is the distribution of \underline{a}_τ , and θ is the distribution of $\mathbf{u}\underline{a}_\tau$, then by (3.17),

$$(3.24) \quad \Phi_\tau^\bullet(\underline{A}) = \Phi_\theta^\bullet(\mathbf{u}\underline{A}), \quad \Sigma_\tau^\bullet(\underline{A}) = \Sigma_\theta^\bullet(\mathbf{u}\underline{A}) \quad \underline{A} \in \mathbb{M}_k(\mathbb{C})^n, k \in \mathbb{N}, \bullet \in \{2, \infty\}.$$

Let us record some obvious properties of Φ_τ^\bullet , $\bullet \in \{2, \infty\}$.

- (1) $\Phi_\tau^\bullet(\underline{A}) \geq 1$ for all $\bullet \in \{2, \infty\}$, $\underline{A} \in \mathbb{M}_k(\mathbb{C})^n$, $k \in \mathbb{N}$. This follows easily by fixing a polynomial \mathbf{P} in the right hand side of either of (3.13), (3.14) and noting that $\Phi_{\tau,d}^\bullet(\underline{A}) \geq \|\mathbf{P}(\underline{A})(I_k)\|_\bullet^2$; we conclude $\Phi_\tau^\bullet(\underline{A}) \geq \lim_{d \rightarrow \infty} \|\mathbf{P}(\underline{A})(I_k)\|_\bullet^{2/d} = 1$. The conclusion for Σ is obvious;
- (2) $\{\underline{A} \in \mathbb{M}_k(\mathbb{C})^n : \limsup_{d \rightarrow \infty} \Phi_{\tau,d}^\bullet(\underline{A}) = 1\} \neq \emptyset$ for all $k > 1$. Indeed, on the set of jointly nilpotent n -tuples of matrices, evaluation of terms of sufficiently large degree yields zero, so that $\sup_{d \in \mathbb{N}} \Phi_{\tau,d}^\bullet(\underline{A}) < \infty$ whenever \underline{A} is jointly nilpotent. This example also shows that, unlike in the case of the classical pluripotential theory, it is possible that $\Phi_{\tau,d_1+d_2}^\bullet(\underline{A}) \not\geq \Phi_{\tau,d_1}^\bullet(\underline{A})\Phi_{\tau,d_2}^\bullet(\underline{A})$, and thus

$$(3.25) \quad \left\{ \Phi_{\tau,d}^\bullet(\underline{A})^{1/d} \right\}_{d \in \mathbb{N}} \text{ might not be an increasing sequence.}$$

In particular, 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 to get a different result before performing the upper semicontinuous regularization).

- (3) $\{\underline{A} \in \mathbb{M}_k(\mathbb{C})^n : \Phi_\tau^\bullet(\underline{A}) > 1\} \neq \emptyset$ for all $k > 1$. To find such an example, one considers $d > 1$ and a selfadjoint polynomial $P \in \mathbb{C}\langle \underline{X} \rangle_d$ (identified with $I_k \otimes P \in \mathbb{M}_k(\mathbb{C}\langle \underline{X} \rangle_d)$) such that $\|P(\underline{a}_\tau)\| = 1$, and one picks a selfadjoint tuple \underline{A} such that $\|P(\underline{A})\|_\bullet > 1$. Then $\Phi_\tau^\bullet(\underline{A}) \geq \|(P^*P)^{2^n}(\underline{A})\|_\bullet^{\frac{1}{2^n d}} \geq \|(P^*P)(\underline{A})\|_\bullet^{\frac{1}{d}} > 1$ for all sufficiently large n .
- (4) $\Phi_\tau^2(\underline{A} \oplus \underline{B}) \geq \max\{\Phi_\tau^2(\underline{A}), \Phi_\tau^2(\underline{B})\}$, $\underline{A} \in \mathbb{M}_{k_1}(\mathbb{C})^n$, $\underline{B} \in \mathbb{M}_{k_2}(\mathbb{C})^n$. Indeed, this follows easily from the fact that if $\mathbf{P}_j \in \mathbb{M}_{k_j}(\mathbb{C}\langle \underline{X} \rangle_d)$, $j = 1, 2$, are maximizing elements in (3.13) corresponding to \underline{A} and \underline{B} , respectively, then $\mathbf{P}_1 \oplus \mathbf{P}_2$ still satisfies $\|(\mathbf{P}_1 \oplus \mathbf{P}_2)(\underline{a}_\tau)\| \leq 1$ and thus $\Phi_{\tau,d}^2(\underline{A} \oplus \underline{B}) \geq \text{tr}_{k_1+k_2}(|\mathbf{P}_1 \oplus \mathbf{P}_2)(\underline{A} \oplus \underline{B})(I_{k_1+k_2})|^2) = \frac{k_1}{k_1+k_2} \text{tr}_{k_1}(|\mathbf{P}_1(\underline{A})(I_{k_1})|^2) + \frac{k_2}{k_1+k_2} \text{tr}_{k_2}(|\mathbf{P}_2(\underline{B})(I_{k_2})|^2) = \frac{k_1}{k_1+k_2} \Phi_{\tau,d}^2(\underline{A}) + \frac{k_2}{k_1+k_2} \Phi_{\tau,d}^2(\underline{B})$. If one of $\Phi_\tau^2(\underline{A})$, $\Phi_\tau^2(\underline{B})$ is infinite, then so is $\Phi_\tau^2(\underline{A} \oplus \underline{B})$. Assume that $\Phi_\tau^2(\underline{A}) > \Phi_\tau^2(\underline{B})$. Pick a subsequence $\{d_\ell\}_{\ell \in \mathbb{N}}$ on which $[\Phi_{\tau,d_\ell}^2(\underline{A})]^{\frac{1}{d_\ell}} \geq [\Phi_{\tau,d_\ell}^2(\underline{B})]^{\frac{1}{d_\ell}}$ and $\Phi_\tau^2(\underline{A})$ is achieved. Then $\limsup_{d \rightarrow \infty} (\Phi_{\tau,d}^2(\underline{A} \oplus \underline{B}))^{\frac{1}{d}} \geq \limsup_{\ell \rightarrow \infty} (\Phi_{\tau,d_\ell}^2(\underline{A} \oplus \underline{B}))^{\frac{1}{d_\ell}} \geq \lim_{\ell \rightarrow \infty} (\Phi_{\tau,d_\ell}^2(\underline{A}))^{\frac{1}{d_\ell}} \left(\frac{k_1}{k_1+k_2} + \frac{k_2}{k_1+k_2} \frac{\Phi_{\tau,d_\ell}^2(\underline{B})}{\Phi_{\tau,d_\ell}^2(\underline{A})} \right)^{\frac{1}{d_\ell}} = \Phi_\tau^2(\underline{A})$, since the term under the parenthesis is bounded from above by 1. The case $\Phi_\tau^2(\underline{A}) = \Phi_\tau^2(\underline{B})$ is equally easy (one picks a subsequence as above, but on which one of the two is no less than the other). The same may not hold for Σ^2 (sup might be achieved at a finite d).
- (5) $\Phi_\tau^\infty(\underline{A} \oplus \underline{B}) \geq \max\{\Phi_\tau^\infty(\underline{A}), \Phi_\tau^\infty(\underline{B})\}$. As in the previous item, for each given d we find maximizers $\mathbf{P}_j \in \mathbb{M}_{k_j}(\mathbb{C}\langle \underline{X} \rangle_d)$, $j = 1, 2$, and then $\Phi_{\tau,d}^\infty(\underline{A} \oplus \underline{B}) \geq \|(\mathbf{P}_1 \oplus \mathbf{P}_2)(\underline{A} \oplus \underline{B})(I_{k_1+k_2})\| = \max\{\|\mathbf{P}_1(\underline{A})(I_{k_1})\|, \|\mathbf{P}_2(\underline{B})(I_{k_2})\|\}$. The rest of the argument is identical to the one above.

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

Definition 3.6. Consider a unit-preserving positive bounded linear map $\varphi: \mathbb{C}\langle \underline{X} \rangle \rightarrow \mathbb{C}$ and a tuple of noncommutative selfadjoint random variables \underline{a}_φ in the von Neumann algebra generated by φ via the GNS construction such that the law of \underline{a}_φ is φ . We say that φ 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

$$\|\mathbf{P}(\underline{a}_\varphi)\| \leq M_d (\text{tr}_k \otimes \varphi)(\mathbf{P}(\underline{a}_\varphi)^* \mathbf{P}(\underline{a}_\varphi))^{\frac{1}{2}} = M_d \|\mathbf{P}(\underline{X})\|_{L^2(\text{tr}_k \otimes \varphi)}$$

for all $\mathbf{P} \in \mathbb{M}_k(\mathbb{C}\langle \underline{X} \rangle_d)$ of degree d , and

$$\limsup_{d \rightarrow \infty} M_d^{\frac{1}{d}} = 1 \text{ for each } k.$$

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 φ 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 φ . 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 $\underline{a}_\varphi = \underline{a}_\tau$ 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 $\{\mathbf{P}_d\}_{d \in \mathbb{N}}$, $\deg \mathbf{P}_d \leq d$,

$$(3.26) \quad \limsup_{d \rightarrow \infty} \left(\frac{\|\mathbf{P}_d(\underline{a}_\tau)\|}{\|\mathbf{P}_d(\underline{a}_\tau)\|_{L^2(\text{tr}_k \otimes \tau)}} \right)^{\frac{1}{d}} \leq 1;$$

(3.27)

$$\forall \varepsilon > 0 \exists C(\varepsilon, \tau) > 0 \text{ such that } \|\mathbf{P}_d(\underline{a}_\tau)\| \leq C(\varepsilon, \tau)(1 + \varepsilon)^d \|\mathbf{P}_d(\underline{a}_\tau)\|_{L^2(\text{tr}_k \otimes \tau)}.$$

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 3.7. *Consider a bounded positive unital trace $\tau: \mathbb{C}\langle \underline{X} \rangle \rightarrow \mathbb{C}$ and a tuple of noncommutative selfadjoint random variables \underline{a}_τ in the von Neumann algebra $W^*(\tau)$ generated by τ via the GNS construction such that the law of \underline{a}_τ is τ . We say that τ satisfies the \mathfrak{C} -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*

$$\|\mathbf{P}(\underline{a}_\tau)\| \leq M_d(\text{tr}_k \otimes \text{tr})(\mathbf{P}(\underline{a}_\tau)^* \mathbf{P}(\underline{a}_\tau))^{\frac{1}{2}} = M_d \|\mathbf{P}(\underline{X})\|_{L^2(\text{tr}_k \otimes \tau)}$$

for all $\mathbf{P} \in \mathbb{M}_k(\mathbb{C}\langle \underline{X} \rangle_d)$ of degree d , and

$$\limsup_{d \rightarrow \infty} M_d^{\frac{1}{d}} = \mathfrak{C} \text{ for each } k.$$

Observe first that in order to verify the \mathfrak{C} -Bernstein-Markov condition, one only needs to check it in the case $k = 1$. Indeed, if $\mathbf{P}(\underline{a}_\tau) = (P_{i,j}(\underline{a}_\tau))_{i,j=1}^k$, then $\mathbf{P}^*(\underline{a}_\tau) = (P_{j,i}^*(\underline{a}_\tau))_{i,j=1}^k$, so that $\mathbf{P}(\underline{a}_\tau) \mathbf{P}^*(\underline{a}_\tau) = \left(\sum_{l=1}^k P_{i,l}(\underline{a}_\tau) P_{j,l}^*(\underline{a}_\tau) \right)_{i,j=1}^k$. Then

$$\|\mathbf{P}(\underline{X})\|_{L^2(\text{tr}_k \otimes \tau)}^2 = \frac{1}{k} \sum_{i,l=1}^k \text{tr}(P_{i,l}(\underline{a}_\tau) P_{i,l}^*(\underline{a}_\tau)) = \frac{1}{k} \sum_{i,l=1}^k \|P_{i,l}(\underline{X})\|_{L^2(\tau)}^2.$$

We write

$$\frac{\|\mathbf{P}(\underline{a}_\tau)\|}{\|\mathbf{P}(\underline{X})\|_{L^2(\text{tr}_k \otimes \tau)}} = \frac{\sqrt{k} \|\mathbf{P}(\underline{a}_\tau)\|}{\left[\sum_{i,l=1}^k \|P_{i,l}(\underline{X})\|_{L^2(\tau)}^2 \right]^{\frac{1}{2}}} \leq \frac{\sqrt{k} \sum_{i,l=1}^k \|P_{i,l}(\underline{a}_\tau)\|}{\left[\sum_{i,l=1}^k \|P_{i,l}(\underline{X})\|_{L^2(\tau)}^2 \right]^{\frac{1}{2}}}.$$

Assume $\mathfrak{C} = \mathfrak{C}_k$ depends on k . By considering diagonal polynomials, it is obvious that \mathfrak{C}_k can only be non-decreasing. Consider a sequence of polynomials $\mathbf{P} = \mathbf{P}_d$, $d \in \mathbb{N}$ being the degree of \mathbf{P} , so that $\lim_{d \rightarrow \infty} \left(\frac{\|\mathbf{P}(\underline{a}_\tau)\|}{\|\mathbf{P}(\underline{X})\|_{L^2(\text{tr}_k \otimes \tau)}} \right)^{\frac{1}{d}} = \mathfrak{C}_k$. That

means,

$$\begin{aligned}
\mathfrak{C}_k &= \lim_{d \rightarrow \infty} \left(\frac{\|\mathbf{P}(\underline{a}_\tau)\|}{\|\mathbf{P}(\underline{X})\|_{L^2(\text{tr}_k \otimes \tau)}} \right)^{\frac{1}{d}} \leq \lim_{d \rightarrow \infty} \left(\frac{\sqrt{k} \sum_{i,l=1}^k \|P_{i,l}(\underline{a}_\tau)\|}{\left[\sum_{i,l=1}^k \|P_{i,l}(\underline{X})\|_{L^2(\tau)}^2 \right]^{\frac{1}{2}}} \right)^{\frac{1}{d}} \\
(3.28) &= \lim_{d \rightarrow \infty} \left(\frac{\|P_{i_0, l_0}(\underline{a}_\tau)\|}{\|P_{i_0, l_0}(\underline{X})\|_{L^2(\tau)}} \right)^{\frac{1}{d}} \left(\frac{1 + \sum_{(i,l) \neq (i_0, l_0)} \frac{\|P_{i,l}(\underline{a}_\tau)\|}{\|P_{i_0, l_0}(\underline{a}_\tau)\|}}{\left[1 + \sum_{(i,l) \neq (i_0, l_0)} \frac{\|P_{i,l}(\underline{X})\|_{L^2(\tau)}^2}{\|P_{i_0, l_0}(\underline{X})\|_{L^2(\tau)}^2} \right]^{\frac{1}{2}}} \right)^{\frac{1}{d}}.
\end{aligned}$$

This holds for any pair of indices $(i_0, l_0) \in \{1, \dots, k\}^2$. Obviously, for each such pair, the sequence $\left(\frac{\|P_{i_0, l_0}(\underline{a}_\tau)\|}{\|P_{i_0, l_0}(\underline{X})\|_{L^2(\tau)}} \right)^{\frac{1}{d}}$ indexed by the degrees d of the polynomials \mathbf{P} has an uppermost limit point. We choose the pair $(i_0, l_0) \in \{1, \dots, k\}^2$ 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 $\frac{\|P_{i_0, l_0}(\underline{a}_\tau)\|}{\|P_{i_0, l_0}(\underline{X})\|_{L^2(\tau)}}$, at least on a subsequence on which the upper limit is reached, grows no slower than $\frac{\|P_{i,l}(\underline{a}_\tau)\|}{\|P_{i,l}(\underline{X})\|_{L^2(\tau)}}$ 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, $\lim_{d \rightarrow \infty} \left(\frac{\|P_{i_0, l_0}(\underline{a}_\tau)\|}{\|P_{i_0, l_0}(\underline{X})\|_{L^2(\tau)}} \right)^{\frac{1}{d}}$, for the limit along this subsequence. We claim next

that $\left(\frac{1 + \sum_{(i,l) \neq (i_0, l_0)} \frac{\|P_{i,l}(\underline{a}_\tau)\|}{\|P_{i_0, l_0}(\underline{a}_\tau)\|}}{\left[1 + \sum_{(i,l) \neq (i_0, l_0)} \frac{\|P_{i,l}(\underline{X})\|_{L^2(\tau)}^2}{\|P_{i_0, l_0}(\underline{X})\|_{L^2(\tau)}^2} \right]^{\frac{1}{2}}} \right)^{\frac{1}{d}}$ is bounded from above by one along this

same subsequence. Indeed, the only way for this to fail is if at least one $\frac{\|P_{i_1, l_1}(\underline{a}_\tau)\|}{\|P_{i_0, l_0}(\underline{a}_\tau)\|}$ tends very fast to infinity while all of $\frac{\|P_{i,l}(\underline{X})\|_{L^2(\tau)}}{\|P_{i_0, l_0}(\underline{X})\|_{L^2(\tau)}}$ either stay bounded or tend much slower to infinity. However,

$$\frac{\|P_{i_1, l_1}(\underline{a}_\tau)\|}{\|P_{i_0, l_0}(\underline{a}_\tau)\|} = \frac{\|P_{i_1, l_1}(\underline{a}_\tau)\|}{\|P_{i_1, l_1}(\underline{X})\|_{L^2(\tau)}} \cdot \frac{\|P_{i_1, l_1}(\underline{X})\|_{L^2(\tau)}}{\|P_{i_0, l_0}(\underline{X})\|_{L^2(\tau)}} \cdot \frac{\|P_{i_0, l_0}(\underline{X})\|_{L^2(\tau)}}{\|P_{i_0, l_0}(\underline{a}_\tau)\|},$$

and, according to our choice of (i_0, l_0) and of the subsequence, for d large enough, $\frac{\|P_{i_1, l_1}(\underline{a}_\tau)\|}{\|P_{i_1, l_1}(\underline{X})\|_{L^2(\tau)}} \cdot \frac{\|P_{i_0, l_0}(\underline{X})\|_{L^2(\tau)}}{\|P_{i_0, l_0}(\underline{a}_\tau)\|} < 2$. Thus, $\frac{\|P_{i_1, l_1}(\underline{a}_\tau)\|}{\|P_{i_0, l_0}(\underline{a}_\tau)\|}$ cannot grow more than twice

as fast as $\frac{\|P_{i_1, l_1}(\underline{X})\|_{L^2(\tau)}}{\|P_{i_0, l_0}(\underline{X})\|_{L^2(\tau)}}$. We conclude that $\left(\frac{1 + \sum_{(i,l) \neq (i_0, l_0)} \frac{\|P_{i,l}(\underline{a}_\tau)\|}{\|P_{i_0, l_0}(\underline{a}_\tau)\|}}{\left[1 + \sum_{(i,l) \neq (i_0, l_0)} \frac{\|P_{i,l}(\underline{X})\|_{L^2(\tau)}^2}{\|P_{i_0, l_0}(\underline{X})\|_{L^2(\tau)}^2} \right]^{\frac{1}{2}}} \right)^{\frac{1}{d}}$

is indeed bounded from above by 1. By (3.28), we conclude that $\mathfrak{C}_k \leq \mathfrak{C}_1$ for all $k \in \mathbb{N}$, so $\mathfrak{C}_k = \mathfrak{C}$ indeed does not depend on k . It should also be noted that if τ satisfies the \mathfrak{C} -Bernstein-Markov property, then $\mathfrak{C} \leq \sup_d M_d^{\frac{1}{d}} < \infty$, and the first inequality may be strict.

Second, in order to verify the \mathfrak{C} -Bernstein-Markov property, it is tautologically enough to verify it on positive polynomials: $Q_d(\underline{X}) \in \mathbb{C}\langle \underline{X} \rangle_d$, $d \in \mathbb{N}$, $Q_d \geq 0$ in

$\mathfrak{C}\langle \underline{X} \rangle$,

$$(3.29) \quad \limsup_{d \rightarrow \infty} \left(\frac{\|Q_d(\underline{a}_\tau)\|}{\text{tr}(Q_d(\underline{a}_\tau))} \right)^{\frac{1}{2d}} = \mathfrak{C},$$

with no value larger than \mathfrak{C} being attainable. For the Bernstein-Markov property ($\mathfrak{C} = 1$), it is enough to simply show that for any sequence Q_d as above,

$$\limsup_{d \rightarrow \infty} \left(\frac{\|Q_d(\underline{a}_\tau)\|}{\text{tr}(Q_d(\underline{a}_\tau))} \right)^{\frac{1}{2d}} = 1.$$

(As an aside, we note that for any $d \in \mathbb{N}$ the optimization problem $\max\{\|Q_d(\underline{a}_\tau)\| : \text{tr}(Q_d(\underline{a}_\tau)) \leq 1, Q_d \succeq 0, Q_d(\underline{X}) \in \mathfrak{C}\langle \underline{X} \rangle_d\}$ is linear, on a convex finite-dimensional set, hence solvable in principle.)

These facts should probably arouse no astonishment in view of Theorem 3.3. Indeed, the connection of $\kappa_{\tau,d}$ to the Bernstein-Markov property (respectively the \mathfrak{C} -Bernstein-Markov property) becomes obvious in view of this theorem and Remark 3.4 following it: recall that

$$\kappa_{\tau,d}(\underline{A}, \underline{A}^*)(I_k) = \max_{\mathbf{P} \in \mathbb{M}_k(\mathfrak{C}\langle \underline{X} \rangle_d)} \mathbf{P}(\underline{A})(I_k)^* (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau)(\mathbf{P}\mathbf{P}^*)^{-1} \mathbf{P}(\underline{A})(I_k),$$

where we assume the obvious conditions of invertibility. As noted in Remark 3.4(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 \mathbf{P} so that $\|\mathbf{P}(\underline{a}_\tau)\| = 1$ and $(\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau)(\mathbf{P}\mathbf{P}^*) = \|\mathbf{P}(\underline{X})\|_{L^2(\text{tr}_k \otimes \tau)}^2 I_k$ (i.e. the $k \times k$ complex matrix $(\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau)(\mathbf{P}\mathbf{P}^*)$ is a multiple of the identity). Now consider a polynomial $\tilde{\mathbf{P}}$ such that $\|\tilde{\mathbf{P}}(\underline{a}_\tau)\| = 1$ and $\frac{1}{\|\tilde{\mathbf{P}}(\underline{a}_\tau)\|_{L^2(\text{tr}_k \otimes \tau)}} = \sup_{\mathbf{P} \in \mathbb{M}_k(\mathfrak{C}\langle \underline{X} \rangle_d)} \frac{\|\mathbf{P}(\underline{a}_\tau)\|}{\|\mathbf{P}(\underline{X})\|_{L^2(\text{tr}_k \otimes \tau)}} := M_d$. Of course, $\kappa_{\tau,d}(\underline{A}, \underline{A}^*)(I_k) \succeq \tilde{\mathbf{P}}(\underline{A})(I_k)^* (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau)(\tilde{\mathbf{P}}\tilde{\mathbf{P}}^*)^{-1} \tilde{\mathbf{P}}(\underline{A})(I_k)$ if $(\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau)(\tilde{\mathbf{P}}\tilde{\mathbf{P}}^*)$ is invertible. If not, we pick an $\epsilon > 0$ arbitrarily small and choose $\tilde{\mathbf{P}}_\epsilon$ approximating $\tilde{\mathbf{P}}$ such that $\|\tilde{\mathbf{P}}_\epsilon(\underline{a}_\tau)\| = 1$, $(\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau)(\tilde{\mathbf{P}}_\epsilon \tilde{\mathbf{P}}_\epsilon^*)$ is invertible, and $\frac{1}{\|\tilde{\mathbf{P}}_\epsilon(\underline{X})\|_{L^2(\text{tr}_k \otimes \tau)}} > M_d - \epsilon$. Then the above majorization by $\kappa_{\tau,d}(\underline{A}, \underline{A}^*)(I_k)$ holds for $\tilde{\mathbf{P}}_\epsilon$ as well. This allows us to write

$$\kappa_{\tau,d}(\underline{A}, \underline{A}^*)(I_k) = \max_{\mathbf{P} \in \mathbb{M}_k(\mathfrak{C}\langle \underline{X} \rangle_d)} \frac{\mathbf{P}(\underline{A})(I_k)^* \mathbf{P}(\underline{A})(I_k)}{\|\mathbf{P}(\underline{X})\|_{L^2(\text{tr}_k \otimes \tau)}^2},$$

imposing the same invertibility condition as for the previously displayed relation.

With this definition in hand, we assume that τ does satisfy a \mathfrak{C} -Bernstein-Markov condition. Then $\left\| \frac{1}{M_d} \mathbf{P}_A(\underline{a}_\tau) \right\| \leq 1$, so that $\Phi_{\tau,d}^\infty(\underline{A}) \geq \left\| \frac{1}{M_d} \mathbf{P}_A(\underline{A})(I_k) \right\|^2$, which, according to Remark 3.4(2), implies

$$\Phi_\tau^\infty(\underline{A}) \geq \limsup_{d \rightarrow \infty} \left\| \frac{1}{M_d} \mathbf{P}_A(\underline{A})(I_k) \right\|^{\frac{2}{d}} \geq \frac{1}{\limsup_{d \rightarrow \infty} M_d^{\frac{2}{d}}} \limsup_{d \rightarrow \infty} \|\kappa_{\tau,d}(\underline{A}, \underline{A}^*)(I_k)\|^{\frac{1}{d}}.$$

Applying the hypothesis that τ satisfies the \mathfrak{C} -Bernstein-Markov property yields

$$(3.30) \quad \mathfrak{C}^2 \Phi_\tau^\infty(\underline{A}) \geq \limsup_{d \rightarrow \infty} \|\kappa_{\tau,d}(\underline{A}, \underline{A}^*)(I_k)\|^{\frac{1}{d}}.$$

Conversely, if we consider the embodiment of $\kappa_{\tau,d}(\underline{A}, \underline{A}^*)(I_k)$ as the maximum after \mathbf{P} of $\mathbf{P}(\underline{A})(I_k)^* (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau)(\mathbf{P}\mathbf{P}^*)^{-1} \mathbf{P}(\underline{A})(I_k)$ as in Theorem 3.3, then we may renormalize \mathbf{P} with a scalar constant so that $\|\mathbf{P}(\underline{a}_\tau)\| = 1$. It follows that $(\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau)$

$\tau(\mathbf{P}\mathbf{P}^*)^{-1} \succeq I_k$, so $\|\kappa_{\tau,d}(\underline{A}, \underline{A}^*)(I_k)\| = \|\mathbf{P}(\underline{A})(I_k)^*(\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau)(\mathbf{P}\mathbf{P}^*)^{-1}\mathbf{P}(\underline{A})(I_k)\| \geq \|\mathbf{P}(\underline{A})(I_k)^*\mathbf{P}(\underline{A})(I_k)\|$ for all \mathbf{P} , so $\|\kappa_{\tau,d}(\underline{A}, \underline{A}^*)(I_k)\| \geq \Phi_{\tau,d}^\infty(\underline{A})$. This proves that

$$(3.31) \quad \limsup_{d \rightarrow \infty} \|\kappa_{\tau,d}(\underline{A}, \underline{A}^*)(I_k)\|^{\frac{1}{d}} \geq \Phi_\tau^\infty(\underline{A}), \quad \underline{A} \in \mathbb{M}_k(\mathbb{C})^n, k \in \mathbb{N}.$$

Under the assumption of \mathfrak{C} -Bernstein-Markov property for τ , equation (3.30) guarantees that $\mathfrak{C}^2 \Phi_\tau^\infty(\underline{A}) \geq \limsup_{d \rightarrow \infty} \|\kappa_{\tau,d}(\underline{A}, \underline{A}^*)(I_k)\|^{\frac{1}{d}} \geq \Phi_\tau^\infty(\underline{A})$, which, for $\mathfrak{C} = 1$ (i.e. the Bernstein-Markov property), yields

$$(3.32) \quad \Phi_\tau^\infty(\underline{A}) = \limsup_{d \rightarrow \infty} \|\kappa_{\tau,d}(\underline{A}, \underline{A}^*)(I_k)\|^{\frac{1}{d}}, \quad \underline{A} \in \mathbb{M}_k(\mathbb{C})^n, k \in \mathbb{N}.$$

Let us consider next the other noncommutative version of Siciak's extremal function, namely Φ_τ^2 . For fixed d , we have seen above that we may renormalize any polynomial \mathbf{P} in Remark 3.4(2) so that $\|\mathbf{P}(\underline{a}_\tau)\| = 1$, which yields $\kappa_{\tau,d}(\underline{A}, \underline{A}^*)(I_k) \succeq \mathbf{P}(\underline{A})(I_k)^*(\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau)(\mathbf{P}\mathbf{P}^*)^{-1}\mathbf{P}(\underline{A})(I_k)$. As $(\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau)(\mathbf{P}\mathbf{P}^*) = (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \text{tr})(\mathbf{P}(\underline{a}_\tau)\mathbf{P}(\underline{a}_\tau)^*)$, and $(\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \text{tr})$ is a conditional expectation and hence (completely) positive, $\|\mathbf{P}(\underline{a}_\tau)\| \leq 1 \iff \mathbf{P}(\underline{a}_\tau)\mathbf{P}(\underline{a}_\tau)^* \preceq I_k \otimes 1 \implies (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \text{tr})(\mathbf{P}(\underline{a}_\tau)\mathbf{P}(\underline{a}_\tau)^*) \preceq I_k \implies (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \text{tr})(\mathbf{P}(\underline{a}_\tau)\mathbf{P}(\underline{a}_\tau)^*)^{-1} \succeq I_k$, which yields the conclusion $\kappa_{\tau,d}(\underline{A}, \underline{A}^*)(I_k) \succeq \mathbf{P}(\underline{A})(I_k)^*\mathbf{P}(\underline{A})(I_k)$. Thus, $\text{tr}_k(\kappa_{\tau,d}(\underline{A}, \underline{A}^*)(I_k)) \geq \text{tr}_k(\mathbf{P}(\underline{A})(I_k)^*\mathbf{P}(\underline{A})(I_k))$ for all \mathbf{P} as above, so that $\text{tr}_k(\kappa_{\tau,d}(\underline{A}, \underline{A}^*)(I_k)) \geq \Phi_{\tau,d}^2(\underline{A})$. As \underline{A} is arbitrary, we obtain

$$\limsup_{d \rightarrow \infty} \text{tr}_k(\kappa_{\tau,d}(\underline{A}, \underline{A}^*)(I_k))^{\frac{1}{d}} \geq \Phi_\tau^2(\underline{A}).$$

Under the assumption of the \mathfrak{C} -Bernstein-Markov property for τ , Remark 3.4(2) allows us to consider a polynomial $\mathbf{P}_{\underline{A}} \in \mathbb{M}_k(\mathbb{C}\langle \underline{X} \rangle_d)$ such that $\kappa_{\tau,d}(\underline{A}, \underline{A}^*)(I_k) = \mathbf{P}_{\underline{A}}(\underline{A})(I_k)^*\mathbf{P}_{\underline{A}}(\underline{A})(I_k)$ and $(\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \text{tr})(\mathbf{P}_{\underline{A}}(\underline{a}_\tau)\mathbf{P}_{\underline{A}}(\underline{a}_\tau)^*) = I_k$. This naturally yields $\Phi_{\tau,d}^2(\underline{A})^{1/d} \geq \text{tr}_k\left(\frac{\mathbf{P}_{\underline{A}}(\underline{A})(I_k)^*\mathbf{P}_{\underline{A}}(\underline{A})(I_k)}{\|\mathbf{P}_{\underline{A}}(\underline{a}_\tau)\|^2}\right)^{\frac{1}{d}} = \frac{1}{\|\mathbf{P}_{\underline{A}}(\underline{a}_\tau)\|^{2/d}} \text{tr}_k(\kappa_{\tau,d}(\underline{A}, \underline{A}^*)(I_k))^{\frac{1}{d}}$. Definition 3.6 guarantees that $\frac{1}{\|\mathbf{P}_{\underline{A}}(\underline{a}_\tau)\|^{2/d}} \geq \frac{1}{M_d^{2/d}(\text{tr}_k \otimes \text{tr})(\mathbf{P}_{\underline{A}}(\underline{a}_\tau)^*\mathbf{P}_{\underline{A}}(\underline{a}_\tau))^{1/d}} = \frac{1}{M_d^{2/d} \text{tr}_k(I_k)^{1/d}} = \frac{1}{M_d^{2/d}} \rightarrow \frac{1}{\mathfrak{C}^2}$ as $d \rightarrow \infty$. Thus, $\Phi_\tau^2(\underline{A}) = \limsup_d \Phi_{\tau,d}^2(\underline{A})^{1/d} \geq \frac{1}{\mathfrak{C}^2} \limsup_d \text{tr}_k(\kappa_{\tau,d}(\underline{A}, \underline{A}^*)(I_k))^{\frac{1}{d}}$. Together with the above-displayed relation, it yields the following estimate for distributions τ that satisfy the \mathfrak{C} -Bernstein-Markov inequality

$$(3.33) \quad \mathfrak{C}^2 \Phi_\tau^2(\underline{A}) \geq \limsup_{d \rightarrow \infty} \text{tr}_k(\kappa_{\tau,d}(\underline{A}, \underline{A}^*)(I_k))^{\frac{1}{d}} \geq \Phi_\tau^2(\underline{A})$$

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

$$(3.34) \quad \Phi_\tau^2(\underline{A}) = \limsup_{d \rightarrow \infty} \text{tr}_k(\kappa_{\tau,d}(\underline{A}, \underline{A}^*)(I_k))^{\frac{1}{d}}, \quad \underline{A} \in \mathbb{M}_k(\mathbb{C})^n, k \in \mathbb{N}.$$

Finally, let us establish the relation between the Christoffel-Darboux kernel of τ and the functions Σ_τ^\bullet , $\bullet \in \{2, \infty\}$, introduced in (3.20)–(3.21). First observe that if $\sup_{d \in \mathbb{N}} (\Phi_{\tau,d}^\bullet(\underline{A}))^{1/d}$ is *not* achieved at a finite d , then $\Sigma_\tau^\bullet = \Phi_\tau^\bullet$, and relations (3.30), (3.31), and/or (3.33) hold. Thus, assume that $\Sigma_\tau^\bullet(\underline{A}) = \sup_{d \in \mathbb{N}} (\Phi_{\tau,d}^\bullet(\underline{A}))^{\frac{1}{d}} = \max_{d \in \mathbb{N}} (\Phi_{\tau,d}^\bullet(\underline{A}))^{\frac{1}{d}} = \Phi_{\tau,d_0}^\bullet(\underline{A})^{1/d_0}$. As before, we pick $\mathbf{P}_0 \in \mathbb{M}_k(\mathbb{C}\langle \underline{X} \rangle_{d_0})$ such that $\|\mathbf{P}_0(\underline{a}_\tau)\| = 1$ and $\Sigma_\tau^\bullet(\underline{A}) = \Phi_{\tau,d_0}^\bullet(\underline{A})^{\frac{1}{d_0}} = \|\mathbf{P}_0(\underline{A})(I_k)\|_2^{\frac{1}{d_0}}$. Then $\kappa_{\tau,d_0}(\underline{A}, \underline{A}^*)(I_k) \succeq \mathbf{P}_0(\underline{A})(I_k)^*(\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \text{tr})(\mathbf{P}_0^*(\underline{a}_\tau)\mathbf{P}_0(\underline{a}_\tau))^{-1}\mathbf{P}_0(\underline{A})(I_k) \succeq \mathbf{P}_0(\underline{A})(I_k)^*\mathbf{P}_0(\underline{A})(I_k)$, which immediately implies $\text{tr}_k(\kappa_{\tau,d_0}(\underline{A}, \underline{A}^*)(I_k))^{\frac{1}{d_0}} \geq \Sigma_\tau^2(\underline{A})$,

$\|\kappa_{\tau,d_0}(\underline{A}, \underline{A}^*)(I_k)\|_{d_0}^{\frac{1}{d_0}} \geq \Sigma_{\tau}^{\infty}(\underline{A})$. For the opposite inequality, we again take a polynomial $\mathbf{P}_0 \in \mathbb{M}_k(\mathbb{C}\langle \underline{X} \rangle_{d_0})$, but this time chosen so to achieve κ_{τ,d_0} : $(\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \text{tr})(\mathbf{P}_0^*(\underline{a}_{\tau})\mathbf{P}_0(\underline{a}_{\tau})) = \|\mathbf{P}_0\|_{L^2(\text{tr}_k \otimes \tau)}^2 I_k$ and $\kappa_{\tau,d_0}(\underline{A}, \underline{A}^*)(I_k) = \mathbf{P}_0(\underline{A})(I_k)^*(\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \text{tr})(\mathbf{P}_0^*(\underline{a}_{\tau})\mathbf{P}_0(\underline{a}_{\tau}))^{-1}\mathbf{P}_0(\underline{A})(I_k) = \frac{\mathbf{P}_0(\underline{A})(I_k)^*\mathbf{P}_0(\underline{A})(I_k)}{\|\mathbf{P}_0\|_{L^2(\text{tr}_k \otimes \tau)}^2}$. We assume that τ satisfies the \mathfrak{C} -Bernstein-Markov property for some $\mathfrak{C} \in [1, +\infty)$. Then, as noted before,

$$\mathfrak{C} \leq \sup_{d \in \mathbb{N}} M_d^{\frac{1}{d}} < \infty$$

depends only on τ . We thus have $\|\kappa_{\tau,d_0}(\underline{A}, \underline{A}^*)(I_k)\|_{d_0}^{\frac{1}{d_0}} = \frac{\|\mathbf{P}_0(\underline{A})(I_k)^*\mathbf{P}_0(\underline{A})(I_k)\|_{d_0}^{\frac{1}{d_0}}}{\|\mathbf{P}_0\|_{L^2(\text{tr}_k \otimes \tau)}^{\frac{2}{d_0}}}$, $\text{tr}_k(\kappa_{\tau,d_0}(\underline{A}, \underline{A}^*)(I_k))^{1/d_0} = \frac{\text{tr}_k(\mathbf{P}_0(\underline{A})(I_k)^*\mathbf{P}_0(\underline{A})(I_k))^{1/d_0}}{\|\mathbf{P}_0\|_{L^2(\text{tr}_k \otimes \tau)}^{2/d_0}}$. Since by definition $M_{d_0} \geq \frac{\|\mathbf{P}_0(\underline{a}_{\tau})\|}{\|\mathbf{P}_0\|_{L^2(\text{tr}_k \otimes \tau)}}$, one has $\|\kappa_{\tau,d_0}(\underline{A}, \underline{A}^*)(I_k)\|_{d_0}^{\frac{1}{d_0}} \leq \sup_{d \in \mathbb{N}} M_d^{\frac{2}{d}} \left[\frac{\|\mathbf{P}_0(\underline{A})(I_k)^*\mathbf{P}_0(\underline{A})(I_k)\|}{\|\mathbf{P}_0(\underline{a}_{\tau})\|^2} \right]^{\frac{1}{d_0}} \leq \sup_{d \in \mathbb{N}} M_d^{\frac{2}{d}} \Sigma_{\tau}^{\infty}(\underline{A})$, $\text{tr}_k(\kappa_{\tau,d_0}(\underline{A}, \underline{A}^*)(I_k))^{1/d_0} \leq \sup_{d \in \mathbb{N}} M_d^{\frac{2}{d}} \Sigma_{\tau}^2(\underline{A})$.

Let us put these results together in a proposition (compare with [BPSS20, Lemma 2.9] and its proof):

Proposition 3.8. *For any given bounded positive trace τ on $\mathbb{C}\langle \underline{X} \rangle$, we have*

$$(3.35) \quad \limsup_{d \rightarrow \infty} \|\kappa_{\tau,d}(\underline{A}, \underline{A}^*)(I_k)\|_{d}^{\frac{1}{d}} \geq \Phi_{\tau}^{\infty}(\underline{A}), \quad \sup_{d \in \mathbb{N}} \|\kappa_{\tau,d}(\underline{A}, \underline{A}^*)(I_k)\|_{d}^{\frac{1}{d}} \geq \Sigma_{\tau}^{\infty}(\underline{A}),$$

$$(3.36)$$

$$\limsup_{d \rightarrow \infty} \text{tr}_k(\kappa_{\tau,d}(\underline{A}, \underline{A}^*)(I_k))^{1/d} \geq \Phi_{\tau}^2(\underline{A}), \quad \sup_{d \in \mathbb{N}} \text{tr}_k(\kappa_{\tau,d}(\underline{A}, \underline{A}^*)(I_k))^{1/d} \geq \Sigma_{\tau}^2(\underline{A}),$$

for all $\underline{A} \in \mathbb{M}_k(\mathbb{C})^n$, $k \in \mathbb{N}$. If τ satisfies the \mathfrak{C} -Bernstein-Markov property 3.7, then

$$(3.37) \quad \mathfrak{C}^2 \Phi_{\tau}^{\infty}(\underline{A}) \geq \limsup_{d \rightarrow \infty} \|\kappa_{\tau,d}(\underline{A}, \underline{A}^*)(I_k)\|_{d}^{\frac{1}{d}} \geq \Phi_{\tau}^{\infty}(\underline{A}),$$

$$(3.38) \quad \mathfrak{C}^2 \Phi_{\tau}^2(\underline{A}) \geq \limsup_{d \rightarrow \infty} \text{tr}_k(\kappa_{\tau,d}(\underline{A}, \underline{A}^*)(I_k))^{1/d} \geq \Phi_{\tau}^2(\underline{A}),$$

$$(3.39) \quad \Sigma_{\tau}^{\infty}(\underline{A}) \sup_{d \in \mathbb{N}} M_d^{\frac{2}{d}} \geq \sup_{d \in \mathbb{N}} \|\kappa_{\tau,d}(\underline{A}, \underline{A}^*)(I_k)\|_{d}^{\frac{1}{d}} \geq \Sigma_{\tau}^{\infty}(\underline{A}),$$

$$(3.40) \quad \Sigma_{\tau}^2(\underline{A}) \sup_{d \in \mathbb{N}} M_d^{\frac{2}{d}} \geq \sup_{d \in \mathbb{N}} \text{tr}_k(\kappa_{\tau,d}(\underline{A}, \underline{A}^*)(I_k))^{1/d} \geq \Sigma_{\tau}^2(\underline{A}), \quad \underline{A} \in \mathbb{M}_k(\mathbb{C})^n, k \in \mathbb{N}.$$

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

We have noted that Definition 3.6 actually generalizes the classical Bernstein-Markov property. Let us put this in context through an example:

Example 3.9. Let τ be the semicircular (Wigner) distribution on \mathbb{R} : $\tau: \mathbb{C}\langle X \rangle \rightarrow \mathbb{C}$, $\tau(X^d) = \frac{2}{\pi} \int_{-1}^1 t^d \sqrt{1-t^2} dt$. It is well-known (see for instance [BLPW15]), and easy to prove, that τ has the Bernstein-Markov property in the classical sense: for instance, the orthonormal polynomials of τ 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 $P \in \mathbb{C}\langle X \rangle_d$ as $P(X) = \sum_{v=0}^d c_v U_v(X)$, then $\|P(a_\tau)\| \leq \sum_{v=0}^d |c_v| \|U_v(a_\tau)\| \leq \sum_{v=0}^d |c_v| (v+1) \leq (d+1) \sum_{v=0}^d |c_v|$, while $\|P(a_\tau)\|_2 = \sqrt{\sum_{v=0}^d |c_v|^2}$. By the Schwarz-Cauchy inequality, $\sum_{v=0}^d |c_v| \leq \sqrt{d+1} \sqrt{\sum_{v=0}^d |c_v|^2}$, so that $\|P(a_\tau)\| \leq (d+1)^{\frac{3}{2}} \|P(a_\tau)\|_2$ for all $P \in \mathbb{C}\langle X \rangle_d$. Since $\lim_{d \rightarrow \infty} (d+1)^{\frac{3}{2d}} = 1$, τ satisfies the classical Bernstein-Markov property. As shown above, this means τ satisfies the Bernstein-Markov property for matrix-valued polynomials as well. Now, Proposition 3.8 shows that $\Phi_\tau^\infty(A) = \limsup_{d \rightarrow \infty} \|\kappa_{\tau,d}(A, A^*)(I_k)^{\frac{1}{d}}\|$ and $\Phi_\tau^2(A) = \limsup_{d \rightarrow \infty} \text{tr}_k(\kappa_{\tau,d}(A, A^*)(I_k))^{\frac{1}{d}}$, $A \in \mathbb{M}_k(\mathbb{C})$. 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 , and thus the existence of $\limsup_{d \rightarrow \infty} \|\kappa_{\tau,d}(A, A^*)(I_k)^{\frac{1}{d}}\|$ is simply guaranteed by the convergence of $\kappa_{\tau,d}(z, \bar{z})(1)^{\frac{1}{d}}$ 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 $A \in \mathbb{M}_k(\mathbb{C})$ is in upper triangular form $A = D + T$, where D is diagonal and T strictly upper triangular. With the notation $D = \text{diag}(z_1, \dots, z_k)$, $T = (t_{i,i+j})_{1 \leq i < k, 1 \leq j \leq k-i}$, we make the assumptions that $z_i \neq z_l$ if $l \neq i$ (which excludes a set of Lebesgue measure zero in $\mathbb{M}_k(\mathbb{C})$), and that $|z_1| \leq \dots \leq |z_k|$ (which implies no loss of generality, as any upper triangular matrix is unitarily equivalent to one whose eigenvalues are ordered increasingly). We then have the following formula for $U_d(A)$:

$$(U_d(A))_{i,i+j} = \left[\frac{U_d(z_i)}{z_i - z_{i+j}} + \frac{U_d(z_{i+j})}{z_{i+j} - z_i} \right] t_{i,i+j} + \sum_{l=1}^{j-1} \left[\sum_{i < i_1 < \dots < i_l < i+j} \left[\sum_{s=i, i_1, \dots, i_l, i+j} \frac{U_d(z_s)}{\prod_{r=i, i_1, \dots, i_l, i+j, r \neq s} [z_s - z_r]} \right] t_{i, i_1} t_{i_1, i_2} \dots t_{i_l, i+j} \right],$$

$(U_d(A))_{i,i} = U_d(z_i)$, while $(U_d(A))_{i,k} = 0$ if $k < i$. 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_d 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,

$$U_d \left(\begin{bmatrix} z_1 & t_{1,2} & t_{1,3} \\ 0 & z_2 & t_{2,3} \\ 0 & 0 & z_3 \end{bmatrix} \right) =$$

$$\begin{bmatrix} U_d(z_1) & \left[\frac{U_d(z_1)}{z_1-z_2} + \frac{U_d(z_2)}{z_2-z_1} \right] t_{1,2} & \left[\frac{U_d(z_1)}{z_1-z_3} + \frac{U_d(z_3)}{z_3-z_1} \right] t_{1,3} + \left[\frac{U_d(z_1)}{[z_1-z_2][z_1-z_3]} + \frac{U_d(z_2)}{[z_2-z_1][z_2-z_3]} + \frac{U_d(z_3)}{[z_3-z_1][z_3-z_2]} \right] t_{1,2}t_{2,3} \\ 0 & U_d(z_2) & \left[\frac{U_d(z_2)}{z_2-z_3} + \frac{U_d(z_3)}{z_3-z_2} \right] t_{2,3} \\ 0 & 0 & U_d(z_3) \end{bmatrix}$$

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: $U_d(A) = (2\pi i)^{-1} \int_\gamma U_d(\zeta)(\zeta I_k - A)^{-1} d\zeta$ for any smooth curve γ surrounding $\sigma(A)$ once.)

More general examples can be derived from the estimates obtained in Section 3.3.9: specifically, assuming that the spectrum $\sigma(a_\tau)$ of our generating variable is sufficiently large so that the compact subset $S_{a_\tau, k} := \{A = A^* \in \mathbb{M}_k(\mathbb{C}) : \sigma(A) \subset \sigma(a_\tau)\}$ of \mathbb{S}_k is not pluripolar in $\mathbb{M}_k(\mathbb{C}) = \mathbb{C}^{k^2}$ (which is the case, for instance, if $\sigma(a_\tau)$ has nonempty interior², which happens trivially for the semicircular) we have seen that $\Phi_{\tau, d}^2(A) \leq 1$ for all $A \in S_{a_\tau, k}$. This, together with the estimate in Section 3.3.8, implies that both Σ_τ^2 and Φ_τ^2 are finite q.e. on $\mathbb{M}_k(\mathbb{C})$ (see, for instance, [ST97, Theorem 1.6, Appendix B]). Moreover, they are also no greater than one on $S_{a_\tau, k}$. Since $\|\cdot\|_2$ and $\|\cdot\|$ are comparable on $\mathbb{M}_k(\mathbb{C})$, it follows that Σ_τ^∞ and Φ_τ^∞ are finite q.e. on $\mathbb{M}_k(\mathbb{C})$ as well.

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

Example 3.10. Consider two copies τ_1, τ_2 of the semicircular (Wigner) distribution from Example 3.9 above, and take their *free product* $\tau = \tau_1 * \tau_2 : \mathbb{C}\langle X_1, X_2 \rangle \rightarrow \mathbb{C}$: τ is completely specified by the following evaluation rules: τ is \mathbb{C} -linear, $\tau(P(X_j)) = \tau_j(P(X_j))$, $j = 1, 2$, and $\tau(P_{i_1}(X_{i_1})P_{i_2}(X_{i_2}) \cdots P_{i_r}(X_{i_r})) = 0$ whenever $r \in \mathbb{N}$, $r \geq 1$, $i_1, i_2, \dots, i_r \in \{1, 2\}$ are such that $i_1 \neq i_2 \neq \cdots \neq i_r$ and $\tau(P_{i_j}(X_{i_j})) = 0, 1 \leq j \leq r$. It is known that τ is a faithful, bounded, tracial state (see [Voi85] for details). As mentioned in Example 3.1(3), a family of orthonormal polynomials for τ is $\{U_w\}_{w \in \langle X_1, X_2 \rangle_d}$, $U_w(X_1, X_2) = U_{d_{i_1}}(X_{i_1})U_{d_{i_2}}(X_{i_2}) \cdots U_{d_{i_r}}(X_{i_r})$ whenever $w = X_{i_1}^{d_{i_1}} X_{i_2}^{d_{i_2}} \cdots X_{i_r}^{d_{i_r}}$, $r \in \mathbb{N}$, $i_1, i_2, \dots, i_r \in \{1, 2\}$, $i_1 \neq i_2 \neq \cdots \neq i_r$. Here, as before, $U_{d_{i_j}}(X_{i_j})$ are the Chebyshev polynomials of the second kind introduced in Example 3.9 above.

Let us show that Σ_τ^\bullet , and hence $\Phi_\tau^\bullet, \bullet \in \{2, \infty\}$, is finite q.e. First, it follows from Section 3.3.2 (and has been argued in item (3) in the list following Section 3.3.7) that $\Phi_\tau^\bullet(A_1, A_2) \geq \max\{\Phi_{\tau_1}^\bullet(A_1), \Phi_{\tau_2}^\bullet(A_2)\}$, $A_1, A_2 \in \mathbb{M}_k(\mathbb{C})$ (with a similar statement for Σ), so that $\Phi_\tau^\bullet, \Sigma_\tau^\bullet$ are not constantly 1.

For an arbitrary $\mathbf{f} \in \mathbb{M}_k(\mathbb{C}\langle X_1, X_2 \rangle_d)$, the following Fourier-like expansion holds:

$$\begin{aligned} \mathbf{f}(X_1, X_2) &= \sum_{w \in \langle X_1, X_2 \rangle_d} \langle \mathbf{f}, I_k \otimes U_w \rangle \otimes U_w(X_1, X_2) \\ &= \sum_{w \in \langle X_1, X_2 \rangle_d} (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau)((I_k \otimes U_w^*)\mathbf{f}) \otimes U_w(X_1, X_2). \end{aligned}$$

²This immediately implies that the set $S_{a_\tau, k}$ has nonempty interior in \mathbb{S}_k , so that it cannot be pluripolar.

Then

$$\begin{aligned} & \mathbf{f}(A_1, A_2)(I_k)^* \mathbf{f}(A_1, A_2)(I_k) \\ &= \sum_{v, w \in \langle X_1, X_2 \rangle_d} U_v(A_1, A_2)^* (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau) (\mathbf{f}^* I_k \otimes U_v) (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau) (I_k \otimes U_w^* \mathbf{f}) U_w(A_1, A_2). \end{aligned}$$

For $v = w$, then

$$\begin{aligned} & U_w(A_1, A_2)^* (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau) (\mathbf{f}^* I_k \otimes U_w) (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau) (I_k \otimes U_w^* \mathbf{f}) U_w(A_1, A_2) \\ & \leq \| \mathbf{f}(a_{\tau_1}, a_{\tau_2}) \|^2 U_w(A_1, A_2)^* [\tau(U_w U_w^*) I_k] U_w(A_1, A_2) \\ & = \| \mathbf{f}(a_{\tau_1}, a_{\tau_2}) \|^2 U_w(A_1, A_2)^* U_w(A_1, A_2), \end{aligned}$$

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: $U_v(A_1, A_2)^* (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau) (\mathbf{f}^* I_k \otimes U_v) (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau) (I_k \otimes U_w^* \mathbf{f}) U_w(A_1, A_2) + U_w(A_1, A_2)^* (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau) (\mathbf{f}^* I_k \otimes U_w) (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau) (I_k \otimes U_v^* \mathbf{f}) U_v(A_1, A_2)$. The positivity of $(V \pm W)^*(V \pm W)$ guarantees that $-V^*V - W^*W \leq V^*W + W^*V \leq V^*V + W^*W$. Applying this observation to our sums yields

$$\begin{aligned} & \pm [U_v(A_1, A_2)^* (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau) (\mathbf{f}^* I_k \otimes U_v) (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau) (I_k \otimes U_w^* \mathbf{f}) U_w(A_1, A_2) \\ & \quad + U_w(A_1, A_2)^* (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau) (\mathbf{f}^* I_k \otimes U_w) (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau) (I_k \otimes U_v^* \mathbf{f}) U_v(A_1, A_2)] \\ & \leq \| \mathbf{f}(a_{\tau_1}, a_{\tau_2}) \|^2 (U_w(A_1, A_2)^* U_w(A_1, A_2) + U_v(A_1, A_2)^* U_v(A_1, A_2)). \end{aligned}$$

We obtain the (rather brutal) majorization

$$\begin{aligned} & \mathbf{f}(A_1, A_2)(I_k)^* \mathbf{f}(A_1, A_2)(I_k) \\ & \leq \| \mathbf{f}(\underline{a}_{\tau_1}, \underline{a}_{\tau_2}) \|^2 \left[\sum_{w \in \langle X_1, X_2 \rangle_d} U_w(A_1, A_2)^* U_w(A_1, A_2) \right. \\ (3.41) \quad & \left. + \sum_{\substack{v, w \in \langle X_1, X_2 \rangle_d \\ v <_{g^1} w}} U_w(A_1, A_2)^* U_w(A_1, A_2) + U_v(A_1, A_2)^* U_v(A_1, A_2) \right]. \end{aligned}$$

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 $\left\| \frac{\mathbf{P}(\underline{A})(I_k)}{\|\mathbf{P}(a_{\tau_1})\|} \right\|_{\bullet}^2 \leq \Phi_{\tau_1, d}^{\bullet}(\underline{A})$ according to the definition of $\Phi_{\tau_1, d}^{\bullet}$, by employing the embedding $U_d \mapsto I_k \otimes U_d$ we have the estimate $\|U_d(A_j)\|_{\bullet}^2 \leq \|U_d(a_{\tau_j})\|^2 \Phi_{\tau_j, d}^{\bullet}(A_j)$. According to (3.20) and (3.21), by taking power $1/d$, we obtain $\|U_d(A_j)\|_{\bullet}^{2/d} \leq \|U_d(a_{\tau_j})\|^{2/d} \Sigma_{\tau_j}^{\bullet}(A_j)$, $A_j \in \mathbb{M}_k(\mathbb{C})$, $k \in \mathbb{N}$.

Consider the case $\bullet = \infty$. Given $w = X_{i_1}^{d_{i_1}} X_{i_2}^{d_{i_2}} \dots X_{i_r}^{d_{i_r}} \in \langle X_1, X_2 \rangle_d$, we have

$$\begin{aligned} & \|U_w(A_1, A_2)^* U_w(A_1, A_2)\| \\ & \leq \|U_{d_{i_1}}(A_{i_1})\|^2 \dots \|U_{d_{i_r}}(A_{i_r})\|^2 \\ & \leq \|U_{d_{i_1}}(a_{\tau_{i_1}})\|^2 \dots \|U_{d_{i_r}}(a_{\tau_{i_r}})\|^2 \Sigma_{\tau_{i_1}}^{\infty}(A_{i_1})^{d_{i_1}} \dots \Sigma_{\tau_{i_r}}^{\infty}(A_{i_r})^{d_{i_r}} \\ & \leq (d_{i_1} + 1)^3 \dots (d_{i_r} + 1)^3 \Sigma_{\tau_{i_1}}^{\infty}(A_{i_1})^{d_{i_1}} \dots \Sigma_{\tau_{i_r}}^{\infty}(A_{i_r})^{d_{i_r}}. \end{aligned}$$

We choose the w for which $\max_{v \in \langle X_1, X_2 \rangle_d} \|U_v(A_1, A_2)\|$ is achieved. By a most brutal estimate in (3.41),

$$\begin{aligned} \|\mathbf{f}(A_1, A_2)(I_k)\|^2 &\leq 2^{2d} \|\mathbf{f}(a_{\tau_1}, a_{\tau_2})\|^2 \max_{v \in \langle X_1, X_2 \rangle_d} \|U_v(A_1, A_2)\| \\ &\leq 4^d \|\mathbf{f}(a_{\tau_1}, a_{\tau_2})\|^2 (d_{i_1} + 1)^3 \cdots (d_{i_r} + 1)^3 \Sigma_{\tau_{i_1}}^\infty(A_{i_1})^{d_{i_1}} \cdots \Sigma_{\tau_{i_r}}^\infty(A_{i_r})^{d_{i_r}}, \end{aligned}$$

so that, by dividing with $\|\mathbf{f}(a_{\tau_1}, a_{\tau_2})\|^2$ and taking power $1/d$,

$$(3.42) \quad \begin{aligned} \Sigma_\tau^\infty(A_1, A_2) &\leq 4[(d_{i_1} + 1) \cdots (d_{i_r} + 1)]^{3/d} \Sigma_{\tau_{i_1}}^\infty(A_{i_1})^{d_{i_1}/d} \cdots \Sigma_{\tau_{i_r}}^\infty(A_{i_r})^{d_{i_r}/d} \\ &< 32 \max\{\Sigma_{\tau_1}^\infty(A_1), \Sigma_{\tau_2}^\infty(A_2)\}. \end{aligned}$$

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 3.4(4), the characterization of $\kappa_{\tau,d}$ as minimum after certain sets of polynomials as in Theorem 3.3 (and parts (2) and (3) of the same Remark 3.4) hold also when $\kappa_{\tau,d}$ is evaluated on tuples of elements in a finite von Neumann algebra. As the entire proof of Proposition 3.8 is based on this characterization, it follows that it holds also for \underline{A} replaced by $\underline{a}_\tau \in W^*(\tau)^n$. 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 $\Phi_{\tau,d}^2((a_{\tau_1}, a_{\tau_2})) = \sup\{\text{tr}(\mathbf{P}((a_{\tau_1}, a_{\tau_2}))(1)^* \mathbf{P}((a_{\tau_1}, a_{\tau_2}))(1)) : \|\mathbf{P}((a_{\tau_1}, a_{\tau_2}))\| \leq 1, \mathbf{P} \in W^*(\tau_1 * \tau_2) \otimes \mathbb{C}\langle \underline{X} \rangle_d\} \leq 1$ because $\text{tr}(\mathbf{P}((a_{\tau_1}, a_{\tau_2}))(1)^* \mathbf{P}((a_{\tau_1}, a_{\tau_2}))(1)) \leq \|\mathbf{P}((a_{\tau_1}, a_{\tau_2}))\|^2$. Thus, $\Phi_\tau^2((a_{\tau_1}, a_{\tau_2})) = 1$, while $\limsup_{d \rightarrow \infty} \text{tr}(\kappa_{\tau,d}((a_{\tau_1}, a_{\tau_2}), (a_{\tau_1}, a_{\tau_2}))(1)) = 2$. Thus, the free product $\tau = \tau_1 * \tau_2$ does *not* satisfy the Bernstein-Markov property, even though it does satisfy a \mathfrak{C} -Bernstein-Markov property for some $2 \leq \mathfrak{C} \leq 8\sqrt{2}$, as it will be seen in Remark 3.12.

The reader might legitimately hope that some of the inequalities in (3.37)–(3.38) can be eventually shown to be equalities even if $\mathfrak{C} > 1$. It will follow from Theorem 3.11 below, the remark following it, and Examples 3.10 and 3.9 that this is unlikely to happen in many cases.

Some of the methods from Example 3.10 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 3.11. *Let $\underline{X}_1 = (X_1, \dots, X_n)$ and $\underline{X}_2 = (X_{n+1}, \dots, X_{n+m})$ be selfadjoint non-commuting indeterminates. Assume $\tau_1 : \mathbb{C}\langle \underline{X}_1 \rangle \rightarrow \mathbb{C}$ and $\tau_2 : \mathbb{C}\langle \underline{X}_2 \rangle \rightarrow \mathbb{C}$ are two faithful positive bounded traces, and denote $\tau = \tau_1 * \tau_2 : \mathbb{C}\langle \underline{X}_1, \underline{X}_2 \rangle \rightarrow \mathbb{C}$ their free product. Suppose that τ_j satisfies the \mathfrak{C}_j -Bernstein-Markov property for some $\mathfrak{C}_j \in [1, +\infty)$, $j = 1, 2$. Then for each $k \in \mathbb{N}$, one has*

$$(3.43) \quad \begin{aligned} \max\{\Sigma_{\tau_1}^\infty(\underline{A}_1), \Sigma_{\tau_2}^\infty(\underline{A}_2)\} &\leq \Sigma_\tau^\infty(\underline{A}_1, \underline{A}_2) \\ &\leq (m+n)^2 \max\left\{\Sigma_{\tau_1}^\infty(\underline{A}_1) \sup_{q \in \mathbb{N}} M(\tau_1)_{q,q}^{\frac{2}{q}}, \Sigma_{\tau_2}^\infty(\underline{A}_2) \sup_{q \in \mathbb{N}} M(\tau_2)_{q,q}^{\frac{2}{q}}\right\}, \end{aligned}$$

for all $(\underline{A}_1, \underline{A}_2) \in \mathbb{M}_k(\mathbb{C})^n \times \mathbb{M}_k(\mathbb{C})^m$, and

$$\max\{\Phi_{\tau_1}^\infty(\underline{A}_1), \Phi_{\tau_2}^\infty(\underline{A}_2)\} \leq \Phi_\tau^\infty(\underline{A}_1, \underline{A}_2)$$

$$(3.44) \quad \leq (m+n)^2 \max \left\{ \Phi_{\tau_1}^\infty(\underline{A}_1) \sup_{q \in \mathbb{N}} M(\tau_1)_{\frac{2}{q}}, \Phi_{\tau_2}^\infty(\underline{A}_2) \sup_{q \in \mathbb{N}} M(\tau_2)_{\frac{2}{q}} \right\},$$

for all $(\underline{A}_1, \underline{A}_2) = (\underline{A}_1, \underline{A}_2)^* \in \mathbb{M}_k(\mathbb{C})^n \times \mathbb{M}_k(\mathbb{C})^m$. In particular, under the hypotheses of the theorem, the functions Σ_τ^∞ , Σ_τ^2 , Φ_τ^2 , and Φ_τ^∞ are finite q.e. on the Euclidean space $\mathbb{M}_k(\mathbb{C})^n \times \mathbb{M}_k(\mathbb{C})^m$ and nontrivial.

Proof. As in Example 3.10, our proof parallels to some extent the proof provided in [Kli91] for Theorem 5.1.8. We start by recalling that $\Phi_{\tau_1 * \tau_2}^\bullet(\underline{A}_1, \underline{A}_2) \geq \max\{\Phi_{\tau_1}^\bullet(\underline{A}_1), \Phi_{\tau_2}^\bullet(\underline{A}_2)\}$, $\Sigma_{\tau_1 * \tau_2}^\bullet(\underline{A}_1, \underline{A}_2) \geq \max\{\Sigma_{\tau_1}^\bullet(\underline{A}_1), \Sigma_{\tau_2}^\bullet(\underline{A}_2)\}$. 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 3.1 (as shown in [Ans10, Section 3]) provides an explicit formula for orthonormal polynomials corresponding to free products of distributions. If $\{P_{w, \tau_1}\}_{w \in \langle \underline{X}_1 \rangle}$ and $\{P_{w, \tau_2}\}_{w \in \langle \underline{X}_2 \rangle}$ are the orthonormal polynomials associated to noncommutative distributions τ_1 and τ_2 , respectively, then the orthonormal polynomials $\{P_{w, \tau}\}_{w \in \langle \underline{X}_1, \underline{X}_2 \rangle}$ corresponding to $\tau = \tau_1 * \tau_2$ are indexed by words in letters from \underline{X}_1 and \underline{X}_2 , and if $w = u_1 u_2 u_3 \cdots u_r \in \langle \underline{X}_1, \underline{X}_2 \rangle$, $u_j \in \langle \underline{X}_{i_j} \rangle$, $1 \leq j \leq r$, $r \in \mathbb{N}$, where $i_1 \neq i_2 \neq i_3 \neq \cdots \neq i_r$, then

$$P_{w, \tau}(\underline{X}_1, \underline{X}_2) = P_{u_1, \tau_{i_1}}(\underline{X}_{i_1}) P_{u_2, \tau_{i_2}}(\underline{X}_{i_2}) P_{u_3, \tau_{i_3}}(\underline{X}_{i_3}) \cdots P_{u_r, \tau_{i_r}}(\underline{X}_{i_r}).$$

Pick $d \in \mathbb{N}$ and $\mathbf{f} \in \mathbb{M}_k(\mathbb{C}\langle \underline{X}_1, \underline{X}_2 \rangle_d)$ such that $\|\mathbf{f}(\underline{a}_{\tau_1}, \underline{a}_{\tau_2})\| = 1$. Then

$$\begin{aligned} \mathbf{f}(\underline{X}_1, \underline{X}_2) &= \sum_{w \in \langle \underline{X}_1, \underline{X}_2 \rangle_d} \langle \mathbf{f}, I_k \otimes P_{w, \tau} \rangle \otimes P_{w, \tau}(\underline{X}_1, \underline{X}_2) \\ &= \sum_{w \in \langle \underline{X}_1, \underline{X}_2 \rangle_d} (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau)((I_k \otimes P_{w, \tau}^* \mathbf{f}) \otimes P_{w, \tau}(\underline{X}_1, \underline{X}_2)). \end{aligned}$$

It then follows that

$$\begin{aligned} &\mathbf{f}(\underline{A}_1, \underline{A}_2)(I_k)^* \mathbf{f}(\underline{A}_1, \underline{A}_2)(I_k) \\ &= \left[\sum_{w \in \langle \underline{X}_1, \underline{X}_2 \rangle_d} (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau)(I_k \otimes P_{w, \tau}^* \mathbf{f}) P_w(\underline{A}_1, \underline{A}_2) \right]^* \\ &\quad \times \left[\sum_{w \in \langle \underline{X}_1, \underline{X}_2 \rangle_d} (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau)(I_k \otimes P_{w, \tau}^* \mathbf{f}) P_w(\underline{A}_1, \underline{A}_2) \right] \\ &= \sum_{v, w \in \langle \underline{X}_1, \underline{X}_2 \rangle_d} P_{v, \tau}(\underline{A}_1, \underline{A}_2)^* (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau)(\mathbf{f}^* I_k \otimes P_{v, \tau}) (\text{Id}_{\mathbb{M}_k(\mathbb{C})} \otimes \tau)(I_k \otimes P_{w, \tau}^* \mathbf{f}) P_{w, \tau}(\underline{A}_1, \underline{A}_2). \end{aligned}$$

As in Example 3.10, we obtain

$$(3.45) \quad \begin{aligned} &\mathbf{f}(\underline{A}_1, \underline{A}_2)(I_k)^* \mathbf{f}(\underline{A}_1, \underline{A}_2)(I_k) \\ &\leq \|\mathbf{f}(\underline{a}_{\tau_1}, \underline{a}_{\tau_2})\|^2 \left[\sum_{w \in \langle \underline{X}_1, \underline{X}_2 \rangle_d} P_{w, \tau}(\underline{A}_1, \underline{A}_2)^* P_{w, \tau}(\underline{A}_1, \underline{A}_2) \right. \\ &\quad \left. + \sum_{\substack{v, w \in \langle \underline{X}_1, \underline{X}_2 \rangle_d \\ v <_{g_1} w}} P_{w, \tau}(\underline{A}_1, \underline{A}_2)^* P_{w, \tau}(\underline{A}_1, \underline{A}_2) + P_{v, \tau}(\underline{A}_1, \underline{A}_2)^* P_v(\underline{A}_1, \underline{A}_2) \right]. \end{aligned}$$

This holds for any polynomial $\mathbf{f} \in \mathbb{M}_k(\mathbb{C}) \otimes \mathbb{C}\langle \underline{X}_1, \underline{X}_2 \rangle_d$ of degree at most d which has norm one on $(\underline{a}_{\tau_1}, \underline{a}_{\tau_2})$.

Again as in Example 3.10,

$$\begin{aligned} & P_{w,\tau}(\underline{A}_1, \underline{A}_2)^* P_{w,\tau}(\underline{A}_1, \underline{A}_2) \\ &= P_{u_1, \tau_{i_1}}(\underline{A}_{i_1}) P_{u_2, \tau_{i_2}}(\underline{A}_{i_2}) P_{u_3, \tau_{i_3}}(\underline{A}_{i_3}) \cdots P_{u_r, \tau_{i_r}}(\underline{A}_{i_r}) \\ &\quad \times P_{u_r, \tau_{i_r}}(\underline{A}_{i_r})^* \cdots P_{u_3, \tau_{i_3}}(\underline{A}_{i_3})^* P_{u_2, \tau_{i_2}}(\underline{A}_{i_2})^* P_{u_1, \tau_{i_1}}(\underline{A}_{i_1})^*, \end{aligned}$$

where we remind the reader that $w = u_1 u_2 u_3 \cdots u_r \in \langle \underline{X}_1, \underline{X}_2 \rangle$, $u_j \in \langle \underline{X}_{i_j} \rangle$, $1 \leq j \leq r$, $r \in \mathbb{N}$, 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

$$\begin{aligned} & P_{w,\tau}(\underline{A}_1, \underline{A}_2)^* P_{w,\tau}(\underline{A}_1, \underline{A}_2) \\ &\leq \|P_{w,\tau}(\underline{A}_1, \underline{A}_2)^* P_{w,\tau}(\underline{A}_1, \underline{A}_2)\| \\ &\leq \|P_{u_1, \tau_{i_1}}(\underline{A}_{i_1})\|^2 \|P_{u_2, \tau_{i_2}}(\underline{A}_{i_2})\|^2 \|P_{u_3, \tau_{i_3}}(\underline{A}_{i_3})\|^2 \cdots \|P_{u_r, \tau_{i_r}}(\underline{A}_{i_r})\|^2, \end{aligned}$$

and in particular $\|P_{w,\tau}(\underline{A}_1, \underline{A}_2)\|_{\bullet} \leq \|P_{u_1, \tau_{i_1}}(\underline{A}_{i_1})\| \|P_{u_2, \tau_{i_2}}(\underline{A}_{i_2})\| \|P_{u_3, \tau_{i_3}}(\underline{A}_{i_3})\| \cdots \|P_{u_r, \tau_{i_r}}(\underline{A}_{i_r})\|$.

By definition, $\left\| \frac{\mathbf{P}(\underline{A}_j)(I_k)}{\|\mathbf{P}(\underline{a}_{\tau_j})\|} \right\|_{\bullet}^2 \leq \Phi_{\tau_j, d}^{\bullet}(\underline{A}_j)$ for any $\mathbf{P}(\underline{X}_j) \in \mathbb{M}_k(\mathbb{C}\langle \underline{X}_j \rangle_d)$, that is, for any polynomial \mathbf{P} with $k \times k$ scalar matrix coefficients of degree at most d . We embed the orthonormal polynomials in $\mathbb{M}_k(\mathbb{C}\langle \underline{X}_j \rangle_d)$ via the obvious identification $P(\underline{X}_j) \mapsto I_k \otimes P(\underline{X}_j)$, which is isometric on operators. Thus, $\|P_{u_j, \tau_{i_j}}(\underline{A}_{i_j})\|_{\bullet}^2 = \|P_{u_j, \tau_{i_j}}(\underline{A}_{i_j})^*\|_{\bullet}^2 \leq \|P_{u_j, \tau_{i_j}}(\underline{a}_{\tau_{i_j}})\|^2 \Phi_{\tau_{i_j}, |u_j|}^{\bullet}(\underline{A}_{i_j})$, which implies $\|P_{u_j, \tau_{i_j}}(\underline{A}_{i_j})\|_{\bullet}^2 \leq \|P_{u_j, \tau_{i_j}}(\underline{a}_{\tau_{i_j}})\|^2 \Sigma_{\tau_{i_j}}^{\infty}(\underline{A}_{i_j})^{|u_j|}$. The cardinality of $\langle \underline{X}_1, \underline{X}_2 \rangle_d$ is $(m+n)^d$. The number of summands in (3.45) is thus $\frac{(m+n)^d [(m+n)^d + 1]}{2}$, 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 $\|P_{w,\tau}(\underline{A}_1, \underline{A}_2)\|_{\bullet}^2$. We have

$$\begin{aligned} (3.46) \quad & \|P_{w,\tau}(\underline{A}_1, \underline{A}_2)\|_{\bullet}^2 \leq \|P_{u_1, \tau_{i_1}}(\underline{A}_{i_1})\|^2 \|P_{u_2, \tau_{i_2}}(\underline{A}_{i_2})\|^2 \|P_{u_3, \tau_{i_3}}(\underline{A}_{i_3})\|^2 \cdots \|P_{u_r, \tau_{i_r}}(\underline{A}_{i_r})\|^2 \\ & \leq \|P_{u_1, \tau_{i_1}}(\underline{a}_{\tau_{i_1}})\|^2 \|P_{u_2, \tau_{i_2}}(\underline{a}_{\tau_{i_2}})\|^2 \|P_{u_3, \tau_{i_3}}(\underline{a}_{\tau_{i_3}})\|^2 \cdots \|P_{u_r, \tau_{i_r}}(\underline{a}_{\tau_{i_r}})\|^2 \\ & \quad \times \Sigma_{\tau_{i_1}}^{\infty}(\underline{A}_{i_1})^{|u_1|} \Sigma_{\tau_{i_2}}^{\infty}(\underline{A}_{i_2})^{|u_2|} \Sigma_{\tau_{i_3}}^{\infty}(\underline{A}_{i_3})^{|u_3|} \cdots \Sigma_{\tau_{i_r}}^{\infty}(\underline{A}_{i_r})^{|u_r|}. \end{aligned}$$

Since our traces τ_i , $i = 1, 2$ satisfy a \mathfrak{C}_i -Bernstein-Markov property, the norms of the individual orthonormal polynomials are bounded from above:

$$(3.47) \quad \|P_{u_j, \tau_{i_j}}(\underline{a}_{\tau_{i_j}})\|^2 / |u_j| \leq M(\tau_{i_j})_{|u_j|}^{2/|u_j|} \leq \sup_{q \in \mathbb{N}} M(\tau_{i_j})_q^{2/q} < \infty.$$

(As in Definitions 3.6, 3.7, it is necessary to specify here the dependence of the parameter M_q on the tracial state τ_{i_j} .) Together with the above estimate for $\|P_{w,\tau}(\underline{A}_1, \underline{A}_2)\|_{\bullet}$, this yields

$$\begin{aligned} & \|P_{w,\tau}(\underline{A}_1, \underline{A}_2)\|_{\bullet}^2 \leq \|P_{u_1, \tau_{i_1}}(\underline{a}_{\tau_{i_1}})\|_{\bullet}^2 \|P_{u_2, \tau_{i_2}}(\underline{a}_{\tau_{i_2}})\|_{\bullet}^2 \cdots \|P_{u_r, \tau_{i_r}}(\underline{a}_{\tau_{i_r}})\|_{\bullet}^2 \\ & \quad \times \Sigma_{\tau_{i_1}}^{\infty}(\underline{A}_{i_1})^{\frac{|u_1|}{d}} \Sigma_{\tau_{i_2}}^{\infty}(\underline{A}_{i_2})^{\frac{|u_2|}{d}} \cdots \Sigma_{\tau_{i_r}}^{\infty}(\underline{A}_{i_r})^{\frac{|u_r|}{d}} \\ & \leq \left[\sup_{q \in \mathbb{N}} M(\tau_{i_1})_q^{\frac{2}{q}} \right]^{\frac{|u_1|}{d}} \left[\sup_{q \in \mathbb{N}} M(\tau_{i_2})_q^{\frac{2}{q}} \right]^{\frac{|u_2|}{d}} \cdots \left[\sup_{q \in \mathbb{N}} M(\tau_{i_r})_q^{\frac{2}{q}} \right]^{\frac{|u_r|}{d}} \\ & \quad \times \Sigma_{\tau_{i_1}}^{\infty}(\underline{A}_{i_1})^{\frac{|u_1|}{d}} \Sigma_{\tau_{i_2}}^{\infty}(\underline{A}_{i_2})^{\frac{|u_2|}{d}} \cdots \Sigma_{\tau_{i_r}}^{\infty}(\underline{A}_{i_r})^{\frac{|u_r|}{d}} \end{aligned}$$

$$(3.48) \quad \leq \max \left\{ \Sigma_{\tau_1}^\infty(\underline{A}_1) \sup_{q \in \mathbb{N}} M(\tau_1) \frac{2}{q}, \Sigma_{\tau_2}^\infty(\underline{A}_2) \sup_{q \in \mathbb{N}} M(\tau_2) \frac{2}{q} \right\}.$$

Combining this estimate with (3.45) yields

$$\begin{aligned} & \|\mathbf{f}(\underline{A}_1, \underline{A}_2)(I_k)\|_{\frac{2}{d}}^2 \\ & \leq \|\mathbf{f}(\underline{a}_{\tau_1}, \underline{a}_{\tau_2})\|_{\frac{2}{d}}^2 (m+n)^2 \max \left\{ \Sigma_{\tau_1}^\infty(\underline{A}_1) \sup_{q \in \mathbb{N}} M(\tau_1) \frac{2}{q}, \Sigma_{\tau_2}^\infty(\underline{A}_2) \sup_{q \in \mathbb{N}} M(\tau_2) \frac{2}{q} \right\} \end{aligned}$$

which, by the definition of Σ^∞ , provides the desired estimate

$$(3.49) \quad \Sigma_\tau^\infty(\underline{A}_1, \underline{A}_2) \leq (m+n)^2 \max \left\{ \Sigma_{\tau_1}^\infty(\underline{A}_1) \sup_{q \in \mathbb{N}} M(\tau_1) \frac{2}{q}, \Sigma_{\tau_2}^\infty(\underline{A}_2) \sup_{q \in \mathbb{N}} M(\tau_2) \frac{2}{q} \right\},$$

for all $(\underline{A}_1, \underline{A}_2) \in \mathbb{M}_k(\mathbb{C})^{m+n}$.

Note that $\Phi_\theta^\infty(\underline{A})$ is finite if and only if $\Sigma_\theta^\infty(\underline{A})$ is. Thus, the above estimate guarantees that $\Sigma_\tau^\infty(\underline{A}_1, \underline{A}_2) < +\infty$, $\Sigma_\tau^2(\underline{A}_1, \underline{A}_2) < +\infty$, $\Phi_\tau^2(\underline{A}_1, \underline{A}_2) < +\infty$, $\Phi_\tau^\infty(\underline{A}_1, \underline{A}_2) < +\infty$ whenever $\Phi_{\tau_j}^\infty(\underline{A}_j) < \infty$ and τ_j satisfy a \mathfrak{C}_j -Bernstein-Markov property, $j = 1, 2$. In particular, if $\Phi_{\tau_j}^\infty$ are finite q.e., then so are all four functions associated to τ .

Assume next that the tuples \underline{A}_1 and \underline{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 $P \in \mathbb{C}\langle \underline{X}_{i_j} \rangle$, one has $P(\underline{A}_{i_j})^* = P^*(\underline{A}_{i_j})$, so that $P(\underline{A}_{i_j})P(\underline{A}_{i_j})^* = (PP^*)(\underline{A}_{i_j})$. In particular, if $P \in \mathbb{C}\langle \underline{X}_{i_j} \rangle_d$, then $(PP^*)^P \in \mathbb{C}\langle \underline{X}_{i_j} \rangle_{2dp}$. In light of this observation, relation $\|P_{u_j, \tau_{i_j}}(\underline{A}_{i_j})\|_{\bullet}^2 \leq \|P_{u_j, \tau_{i_j}}(\underline{a}_{\tau_{i_j}})\|_{\bullet}^2 \Phi_{\tau_{i_j}, |u_j|}^\bullet(\underline{A}_{i_j})$ immediately implies

$$\begin{aligned} & \|P_{u_j, \tau_{i_j}}(\underline{A}_{i_j})\|_{\bullet}^{2^{N+1}} \\ & \leq \|P_{u_j, \tau_{i_j}}(\underline{A}_{i_j})P_{u_j, \tau_{i_j}}(\underline{A}_{i_j})^*\|_{\bullet}^{2^N} = \|(P_{u_j, \tau_{i_j}}P_{u_j, \tau_{i_j}})^*(\underline{A}_{i_j})\|_{\bullet}^{2^N} \\ & \leq \left\| \left[(P_{u_j, \tau_{i_j}}P_{u_j, \tau_{i_j}})^* \right]^{2^{N-1}}(\underline{A}_{i_j}) \right\|_{\bullet}^2 \\ & \leq \left\| \left[(P_{u_j, \tau_{i_j}}P_{u_j, \tau_{i_j}})^* \right]^{2^{N-1}}(\underline{a}_{\tau_{i_j}}) \right\|_{\bullet}^2 \Phi_{\tau_{i_j}, 2^N |u_j|}^\bullet(\underline{A}_{i_j}) \\ & = \left\| (P_{u_j, \tau_{i_j}}P_{u_j, \tau_{i_j}})^*(\underline{a}_{\tau_{i_j}}) \right\|_{\bullet}^{2^N} \Phi_{\tau_{i_j}, 2^N |u_j|}^\bullet(\underline{A}_{i_j}) \\ & = \left\| P_{u_j, \tau_{i_j}}(\underline{a}_{\tau_{i_j}}) \right\|_{\bullet}^{2^{N+1}} \Phi_{\tau_{i_j}, 2^N |u_j|}^\bullet(\underline{A}_{i_j}); \end{aligned}$$

for $\bullet = \infty$, we have used the equality $\|YY^*\| = \|Y\|^2$, and for $\bullet = 2$, the inequality $\text{tr}_k(Y)\text{tr}_k(Y^*) \leq \text{tr}_k(YY^*)$. It follows that

$$\begin{aligned} \|P_{u_j, \tau_{i_j}}(\underline{A}_{i_j})\|_{\bullet}^2 & \leq \left\| P_{u_j, \tau_{i_j}}(\underline{a}_{\tau_{i_j}}) \right\|_{\bullet}^2 \limsup_{N \rightarrow \infty} \Phi_{\tau_{i_j}, 2^N |u_j|}^\bullet(\underline{A}_{i_j}) \frac{1}{2^N} \\ & \leq \left\| P_{u_j, \tau_{i_j}}(\underline{a}_{\tau_{i_j}}) \right\|_{\bullet}^2 \limsup_{q \rightarrow \infty} \Phi_{\tau_{i_j}, q}^\bullet(\underline{A}_{i_j}) \frac{|u_j|}{q} \\ & \leq \left\| P_{u_j, \tau_{i_j}}(\underline{a}_{\tau_{i_j}}) \right\|_{\bullet}^2 \Phi_{\tau_{i_j}}^\bullet(\underline{A}_{i_j})^{|u_j|}, \end{aligned}$$

for all $1 \leq j \leq r$. Relation (3.46) becomes

$$\|P_{w, \tau}(\underline{A}_1, \underline{A}_2)\|_{\bullet}^2 \leq \|P_{u_1, \tau_{i_1}}(\underline{A}_{i_1})\|_{\bullet}^2 \|P_{u_2, \tau_{i_2}}(\underline{A}_{i_2})\|_{\bullet}^2 \|P_{u_3, \tau_{i_3}}(\underline{A}_{i_3})\|_{\bullet}^2 \cdots \|P_{u_r, \tau_{i_r}}(\underline{A}_{i_r})\|_{\bullet}^2$$

$$\begin{aligned} &\leq \|P_{u_1, \tau_{i_1}}(\underline{a}_{\tau_{i_1}})\|^2 \|P_{u_2, \tau_{i_2}}(\underline{a}_{\tau_{i_2}})\|^2 \|P_{u_3, \tau_{i_3}}(\underline{a}_{\tau_{i_3}})\|^2 \cdots \|P_{u_r, \tau_{i_r}}(\underline{a}_{\tau_{i_r}})\|^2 \\ &\quad \times \Phi_{\tau_{i_1}}^\infty(\underline{A}_{i_1})^{|u_1|} \Phi_{\tau_{i_2}}^\infty(\underline{A}_{i_2})^{|u_2|} \Phi_{\tau_{i_3}}^\infty(\underline{A}_{i_3})^{|u_3|} \cdots \Phi_{\tau_{i_r}}^\infty(\underline{A}_{i_r})^{|u_r|}. \end{aligned}$$

As for Σ , by taking limsup as d tends to infinity and picking maximizing elements $\mathbf{f} \in \mathbb{M}_k(\mathbb{C}\langle \underline{X}_1, \underline{X}_2 \rangle_d)$ for each d , we obtain

$$\Phi_\tau^\infty(\underline{A}_1, \underline{A}_2) \leq (m+n)^2 \max \left\{ \Phi_{\tau_1}^\infty(\underline{A}_1) \sup_{q \in \mathbb{N}} M(\tau_1)_{q,q}^{\frac{2}{q}}, \Phi_{\tau_2}^\infty(\underline{A}_2) \sup_{q \in \mathbb{N}} M(\tau_2)_{q,q}^{\frac{2}{q}} \right\},$$

for all *selfadjoint* tuples $(\underline{A}_1, \underline{A}_2) \in \mathbb{M}_k(\mathbb{C})^{m+n}$. \square

Remark 3.12. *Under the hypotheses of Theorem 3.11, τ satisfies the \mathfrak{C} -Bernstein-Markov property for some constant \mathfrak{C} satisfying the inequalities $\min\{\mathfrak{C}_1, \mathfrak{C}_2\} \leq \mathfrak{C} \leq \sqrt{m+n} \max \left\{ \sup_{q \in \mathbb{N}} M(\tau_1)_{q,q}^{\frac{1}{q}}, \sup_{q \in \mathbb{N}} M(\tau_2)_{q,q}^{\frac{1}{q}} \right\}$. This is indeed a very rough estimate: if $P \in \mathbb{C}\langle \underline{X}_1, \underline{X}_2 \rangle_d$, then*

$$\begin{aligned} \|P(\underline{a}_{\tau_1}, \underline{a}_{\tau_2})\| &= \left\| \sum_{w \in \langle \underline{X}_1, \underline{X}_2 \rangle_d} c_w P_{w, \tau}(\underline{a}_{\tau_1}, \underline{a}_{\tau_2}) \right\| \\ &\leq \sum_{w \in \langle \underline{X}_1, \underline{X}_2 \rangle_d} |c_w| \|P_{w, \tau}(\underline{a}_{\tau_1}, \underline{a}_{\tau_2})\| \\ &\leq \left(\max_{w \in \langle \underline{X}_1, \underline{X}_2 \rangle_d} \|P_{w, \tau}(\underline{a}_{\tau_1}, \underline{a}_{\tau_2})\| \right) \sum_{w \in \langle \underline{X}_1, \underline{X}_2 \rangle_d} |c_w|. \end{aligned}$$

By Cauchy-Schwarz, $\sum_{w \in \langle \underline{X}_1, \underline{X}_2 \rangle_d} |c_w| \leq \left(\sum_{w \in \langle \underline{X}_1, \underline{X}_2 \rangle_d} |c_w|^2 \right)^{\frac{1}{2}} (m+n)^{\frac{d}{2}} = (m+n)^{\frac{d}{2}} \|P\|_{L^2(\tau)}$ (where $(m+n)^{\frac{d}{2}}$ is the two-norm of the vector of ones of length equal to the cardinality of $\langle \underline{X}_1, \underline{X}_2 \rangle_d$). The operator norm $\|P_{w, \tau}(\underline{a}_{\tau_1}, \underline{a}_{\tau_2})\|$ of orthonormal polynomials associated to $\tau = \tau_1 * \tau_2$ has been estimated in the proof of Theorem 3.11: if $w = u_1 u_2 \cdots u_r \in \langle \underline{X}_1, \underline{X}_2 \rangle$, $u_j \in \langle \underline{X}_{i_j} \rangle$, $1 \leq j \leq r$, $r \leq d$, where $i_1 \neq i_2 \neq i_3 \neq \cdots \neq i_r$ and $|u_1| + |u_2| + \cdots + |u_r| = d$, then

$$\begin{aligned} \|P_{w, \tau}(\underline{a}_{\tau_1}, \underline{a}_{\tau_2})\|^{\frac{2}{d}} &\leq \left[\|P_{u_1, \tau_{i_1}}(\underline{a}_{\tau_{i_1}})\|^2 \|P_{u_2, \tau_{i_2}}(\underline{a}_{\tau_{i_2}})\|^2 \cdots \|P_{u_r, \tau_{i_r}}(\underline{a}_{\tau_{i_r}})\|^2 \right]^{1/d} \\ &\leq \max \left\{ \sup_{q \in \mathbb{N}} M(\tau_1)_{q,q}^{\frac{2}{q}}, \sup_{q \in \mathbb{N}} M(\tau_2)_{q,q}^{\frac{2}{q}} \right\}. \end{aligned}$$

Combining the last two estimates yields

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

The lower bound is obvious.

This shows that Theorem 3.11 applies to any finite free products $\tau = \tau_1 * \tau_2 * \cdots * \tau_s$, $s \in \mathbb{N}$, of bounded tracial states τ_j which satisfy a \mathfrak{C}_j -Bernstein-Markov property, $1 \leq j \leq s$.

We conclude this subsection by pointing out that $\frac{1}{2} \log \Phi_\tau^\bullet(\underline{A})$, $\frac{1}{2} \log \Sigma_\tau^\bullet(\underline{A})$ have logarithmic growth as plurisubharmonic functions on the Euclidean space $\mathbb{C}^{nk^2} \simeq \mathbb{M}_k(\mathbb{C})^n$ whenever they are non-trivial. Indeed, this follows from Section 3.3.8 that if the sequence $\left\{ \left(\Phi_{\tau, d}^\infty(\underline{A}) \right)^{1/d} \right\}_{d \in \mathbb{N}}$ is locally bounded, then, as log is increasing,

$\frac{1}{2} \log \Sigma_\tau^\infty(\underline{A}) = \left[\sup_{d \in \mathbb{N}} \frac{1}{2d} \log \left(\Phi_{\tau,d}^\infty(\underline{A}) \right) \right]^*$ has, according to [ST97, Theorem 1.6, Appendix B], logarithmic growth. Since it dominates the other three functions, it follows that $\frac{1}{2} \log \Phi_\tau^\bullet(\underline{A}), \frac{1}{2} \log \Sigma_\tau^\bullet(\underline{A})$ have logarithmic growth as $\|\underline{A}\|_\bullet \rightarrow \infty$.

3.5. Derivatives. One of the equivalent characterizations of plurisubharmonicity is via positivity: if Ω is an open subset of \mathbb{C}^m , then for any plurisubharmonic

function u on Ω and $\xi \in \mathbb{C}^m$, $\sum_{j,k=1}^m \xi_j \bar{\xi}_k \frac{\partial^2 u}{\partial z_j \partial \bar{z}_k}$ is a positive distribution in Ω (that

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

$$\int_{\mathbb{C}^{nk^2}} (d d^c \log \Phi_\tau^\bullet(\underline{A})^{\frac{1}{2}})^{nk^2} \leq 1, \quad \int_{\mathbb{C}^{nk^2}} (d d^c \log \Sigma_\tau^\bullet(\underline{A})^{\frac{1}{2}})^{nk^2} \leq 1,$$

where $\varphi \mapsto (d d^c \varphi)^{nk^2} = (2\pi)^{-nk^2} \det [(\partial_i \bar{\partial}_j \varphi)_{1 \leq i, j \leq nk^2}] dV$ is the Monge-Ampère operator if φ is of class C^2 (V in the Lebesgue measure). Otherwise this should be understood in the weak sense. Thus, $(d d^c \log \Phi_\tau^\bullet(\underline{A})^{\frac{1}{2}})^{nk^2}, (d d^c \log \Sigma_\tau^\bullet(\underline{A})^{\frac{1}{2}})^{nk^2}$ are nonnegative measures on \mathbb{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 $\frac{1}{2} \log \Phi_\tau^\bullet(\underline{A}), \frac{1}{2} \log \Sigma_\tau^\bullet(\underline{A})$ 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 $\underline{A} \mapsto \frac{1}{2} \log \Sigma_\tau^\infty(\underline{A})$ is of logarithmic growth for a large class of distributions τ . In order to do this, we consider first the case of $n = 1$ and then use item (3) in the list following Section 3.3.9 to conclude for arbitrary n . We look at polynomials $\mathbf{L}_{A,d}(X) = \frac{1}{\|A^d\|} (A^*)^d \otimes X^d$. 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 $\|\mathbf{L}_{A,d}(a_\tau)\| = \frac{1}{\|A^d\|} \|(A^*)^d\| \|a_\tau\|^d = 1$, and $\|\mathbf{L}_{A,d}(A)(I_k)^* \mathbf{L}_{A,d}(A)(I_k)\| = \frac{1}{\|A^d\|^2} \|[(A^*)^d A^d]^2\| = \|A^d\|^2$. It follows that $\Phi_{\tau,d}^\infty(A) \geq \|A^d\|^2$. By definition, $\Sigma_\tau^\infty(A) = \left[\sup_{d \in \mathbb{N}} \Phi_{\tau,d}^\infty(A)^{\frac{1}{d}} \right]^* \geq \sup_{d \in \mathbb{N}} \|A^d\|^{\frac{2}{d}} \geq \|A\|^2$. Taking logarithms and recalling Example 3.9 shows that $\frac{1}{2} \log \Sigma_\tau^\infty(A) \geq \log^+ \|A\|$ for $\|A\|$ sufficiently large.

Recall that if $\tau: \mathbb{C}\langle \underline{X}_1, \underline{X}_2 \rangle \rightarrow \mathbb{C}$ is a bounded positive trace, then $\Sigma_\tau^\bullet(\underline{A}_1, \underline{A}_2) \geq \max\{\Sigma_{\tau|_{\mathbb{C}\langle \underline{X}_1 \rangle}}^\bullet(\underline{A}_1), \Sigma_{\tau|_{\mathbb{C}\langle \underline{X}_2 \rangle}}^\bullet(\underline{A}_2)\}$, with a similar statement for $\Phi_\tau^\bullet, \bullet \in \{2, \infty\}$. Thus, if $\frac{1}{2} \log \Sigma_{\tau|_{\mathbb{C}\langle \underline{X}_j \rangle}}^\bullet(\underline{A}_j) \geq \log^+ \|\underline{A}_j\|_\bullet - K$ for some $K \geq 0$ not depending on $\underline{A}_j, j = 1, 2$, then $\frac{1}{2} \log \Sigma_\tau^\bullet(\underline{A}_1, \underline{A}_2) \geq \max\{\log^+ \|\underline{A}_1\|_\bullet - K, \log^+ \|\underline{A}_2\|_\bullet - K\}$. Since $\|\underline{A}_1\|_2^2 + \|\underline{A}_2\|_2^2 = \|(\underline{A}_1, \underline{A}_2)\|_2^2$, one has

$$\begin{aligned} \log \|(\underline{A}_1, \underline{A}_2)\|_2 &= \log \max\{\|\underline{A}_1\|_2, \|\underline{A}_2\|_2\} + \log \sqrt{1 + \frac{\min\{\|\underline{A}_1\|_2^2, \|\underline{A}_2\|_2^2\}}{\max\{\|\underline{A}_1\|_2^2, \|\underline{A}_2\|_2^2\}}} \\ &\leq \max\{\log \|\underline{A}_1\|_2, \log \|\underline{A}_2\|_2\} + \log \sqrt{2}. \end{aligned}$$

Similarly, one has $\|(\underline{A}_1, \underline{A}_2)\| = \max\{\|\underline{A}_1\|, \|\underline{A}_2\|\}$, so that $\log \|(\underline{A}_1, \underline{A}_2)\| = \max\{\log \|\underline{A}_1\|, \log \|\underline{A}_2\|\}$. Putting these together, it follows that

$$\frac{1}{2} \log \Sigma_\tau^\bullet(\underline{A}_1, \underline{A}_2) \geq \log \|(\underline{A}_1, \underline{A}_2)\|_\bullet - \sqrt{2} - K$$

whenever $\frac{1}{2} \log \Sigma_{\tau|_{\mathbb{C}\langle X_j \rangle}}^\bullet(\underline{A}_j) \geq \log^+ \|\underline{A}_j\|_\bullet - K$ for some $K \geq 0$, $j = 1, 2$.

We conclude that whenever $\tau: \mathbb{C}\langle X \rangle \rightarrow \mathbb{C}$ is a positive bounded tracial state such that $\tau|_{\mathbb{C}\langle X_j \rangle}$ is a probability distribution whose support contains a nonempty open set in \mathbb{R} , the function $\Sigma_\tau^\infty(\underline{A})$ has precisely logarithmic growth at infinity, i.e. belongs to the Lelong class $\mathcal{L}^+(\mathbb{C}^{nk^2}) = \{u: \mathbb{C}^{nk^2} \rightarrow [-\infty, +\infty): u \text{ plurisubharmonic, } \log^+ |z| - K_u \leq u(z) \leq \log^+ |z| + K_u, z \in \mathbb{C}^{nk^2}, \text{ for some } K_u > 0\}$. Proposition 3.8 allows us to obtain the following result:

Proposition 3.13. *Assume that $\tau: \mathbb{C}\langle X \rangle \rightarrow \mathbb{C}$ is a positive bounded tracial state such that $\tau|_{\mathbb{C}\langle X_j \rangle}$ is a probability distribution whose support contains a nonempty open set in \mathbb{R} , $1 \leq j \leq n$. If any of $\Sigma_\tau^\bullet, \Phi_\tau^\bullet, \bullet \in \{2, \infty\}$ is well-defined and finite q.e., then $\Sigma_\tau^\infty \in \mathcal{L}^+(\mathbb{C}^{nk^2})$. If τ satisfies in addition the \mathfrak{C} -Bernstein-Markov property 3.7, then all of*

$$\left[\sup_{d \rightarrow \infty} \log \|\kappa_{\tau, d}(\underline{A}^*, \underline{A})(I_k)\|_{\frac{1}{2d}} \right]^*, \quad \left[\sup_{d \rightarrow \infty} \log \text{tr}_k(\kappa_{\tau, d}(\underline{A}^*, \underline{A})(I_k))_{\frac{1}{2d}} \right]^*,$$

$$\frac{1}{2} \log \Sigma_\tau^\bullet \in \mathcal{L}^+(\mathbb{C}^{nk^2}), \quad \bullet \in \{2, \infty\}.$$

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 the condition that Σ_τ^2 is finite q.e. if $\tau|_{\mathbb{C}\langle X_j \rangle}$ are as in the above proposition and $\mathbb{C}\langle X_j \rangle$, $1 \leq j \leq n$, are free with respect to τ , according to Theorem 3.11. This means there is a large class of distributions covered by the result above.

It is also known that in many cases the plurisubharmonic functions of the kind described in Proposition 3.13 reach their minimum precisely at the points of the support of their Monge-Ampère measure. This, together with the considerations from Section 3.3, guarantee that the supports of these measures are invariant under conjugation by a unitary $k \times k$ matrix for each k , and are compact.

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 $F \in \mathbb{C}\langle \underline{Z}, \underline{Z}^* \rangle$. 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, $j \in \{1, \dots, n\}$. That is, we double the number of noncommuting indeterminates. The rule for “differentiation” is simple in the case of polynomials in noncommuting selfadjoint indeterminates: $\partial_{X_i} X_j = \delta_{i,j} 1 \otimes 1$, $\partial_{X_i} Y_k = 0$, ∂_{X_i} is linear from $\mathbb{C}\langle \underline{Z}, \underline{Z}^* \rangle$ to $\mathbb{C}\langle \underline{Z}, \underline{Z}^* \rangle \otimes \mathbb{C}\langle \underline{Z}, \underline{Z}^* \rangle$, and ∂_{X_i} respects the Leibniz rule with respect to products of polynomials. For instance, $\partial_{X_1}(7iX_1X_2Y_2X_1^2Y_2) =$

$7i(1 \otimes X_2 Y_2 X_1^2 Y_2 + X_1 X_2 Y_2 X_1 \otimes Y_2 + X_1 X_2 Y_2 \otimes X_1 Y_2)$. 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,

$$\begin{aligned}\partial_{Z_j} Z_j &= \frac{1}{2}(\partial_{X_j} - i\partial_{Y_j})(X_j + iY_j) = \frac{1}{2}(1 \otimes 1 + i0 - i0 - i \cdot i1 \otimes 1) = 1 \otimes 1, \\ \partial_{Z_j} Z_j^* &= \frac{1}{2}(\partial_{X_j} - i\partial_{Y_j})(X_j - iY_j) = \frac{1}{2}(1 \otimes 1 - i0 - i0 + i \cdot i1 \otimes 1) = 0, \\ \partial_{Z_j^*} Z_j &= \frac{1}{2}(\partial_{X_j} + i\partial_{Y_j})(X_j + iY_j) = \frac{1}{2}(1 \otimes 1 + i0 + i0 + i \cdot i1 \otimes 1) = 0, \\ \partial_{Z_j^*} Z_j^* &= \frac{1}{2}(\partial_{X_j} + i\partial_{Y_j})(X_j - iY_j) = \frac{1}{2}(1 \otimes 1 - i0 + i0 - i \cdot i1 \otimes 1) = 1 \otimes 1.\end{aligned}$$

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

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 $F \in \mathbb{C}\langle \underline{Z}, \underline{Z}^* \rangle$, then $\partial_{Z_j} F(\underline{a}, \underline{a}^*)(c)$ replaces in the above formula the tensor symbol \otimes with the variable c :

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

where $m_c: \mathcal{A} \otimes \mathcal{A}^{\text{op}} \rightarrow \mathcal{A}$, $m_c(p \otimes q) = pcq$. 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 $\phi: \underline{A} \mapsto \text{tr}_k(F(\underline{A})^* F(\underline{A}))$ defined on $\mathbb{M}_k(\mathbb{C})^n$ for each given polynomial $F \in \mathbb{C}\langle \underline{X} \rangle$. We view F as living in $\mathbb{C}\langle \underline{Z} \rangle$, that is, as an analytic polynomial. As mentioned before, when viewed as an element in $\mathbb{C}\langle \underline{X} \rangle$, the evaluation $F(\underline{A})^* = F^*(\underline{A}^*)$, and when viewed as an element in $\mathbb{C}\langle \underline{Z} \rangle$, we have F^* as an element in $\mathbb{C}\langle \underline{Z}^* \rangle$. The problem of differentiating becomes now very classical. For any given direction $C \in \mathbb{M}_k(\mathbb{C})$,

$$\partial_{Z_i} \phi(\underline{A})(C) = \left. \frac{\partial \phi(\underline{A} + \delta_i z C)}{\partial z} \right|_{z=0}, \quad \partial_{Z_j^*} \phi(\underline{A})(C) = \left. \frac{\partial \phi(\underline{A} + \delta_j z C)}{\partial \bar{z}} \right|_{z=0},$$

where we have denoted $\underline{A} + \delta_i z C = (A_1, \dots, A_{i-1}, A_i + zC, A_{i+1}, \dots, A_n)$. Moreover, for a vector $\underline{C} = (C_1, \dots, C_n)^t \in \mathbb{M}_k(\mathbb{C})^n$, we write

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

for the Levi form of ϕ at the point \underline{A} evaluated in \underline{C} [GZ17, (1.3.6)]. (It might seem normal to rather write $\partial_{Z_i^*} \partial_{Z_j} \phi(\underline{A})(C_j)(C_i^*)$ in the formula above; however, with the convention we have introduced, we actually have $[\partial_{Z_i^*} \partial_{Z_j} \phi(\underline{A})](C_j)(C_i^*) = \partial_{Z_i^*} \partial_{Z_j} \phi(\underline{A})(C_j)(C_i)$ when $\partial_{Z_i^*} \partial_{Z_j} \phi(\underline{A})$ 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 $\sum \partial_{Z_i^*} \partial_{Z_j} \phi(\underline{A})(C_j)(C_i) = \partial_z \partial_{\bar{z}} \phi(\underline{A} + z\underline{C})$.

To avoid confusion as much as possible, we will keep the notation with iterated arguments. Generally, 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 φ , we may define

$\mathcal{H} = \varphi \circ (H^*H)$ – that is, $\mathcal{H}(\underline{a}) = (\varphi \otimes \text{tr}_k)(H(\underline{a})^*H(\underline{a}))$ for any $\underline{a} \in \Omega^{(k)}$ – and then the Levi form of \mathcal{H} at the point \underline{a} evaluated in the direction \underline{c} is

$$(3.50) \quad \partial_z \partial_{\bar{z}} \mathcal{H}(\underline{a} + z\underline{c})|_{z=0} = \sum_{i,j=1}^n \partial_{Z_i^*} \partial_{Z_j} \mathcal{H}(\underline{a})(c_j)(c_i).$$

The positivity of the above for all $(\underline{a}, \underline{c}) \in (\Omega \times \mathcal{A}^n)_{\text{nc}}$ can be taken as a definition of an nc plurisubharmonic function (indeed, as in the classical case one can find in references [GZ17, ST97], \mathcal{H} is plurisubharmonic because $z \mapsto \mathcal{H}(\underline{a} + z\underline{c})$ is subharmonic as function from a complex neighborhood of zero into $[-\infty, +\infty)$ [GZ17, Definition 1.27], which is a fair definition of plurisubharmonicity in our case too) except for the fact that it would not be complete, as we have assumed H to exist and be a free noncommutative function, hence analytic. In the above, derivatives exist in the strongest possible sense. We do need to consider weak closures of spaces of such functions. We should however emphasize that the functions $\Phi_{\tau,d}^2$ defined in (3.13) obviously belong to this space, while the functions $\Phi_{\tau,d}^\infty$ from (3.14) do not. However, $\Phi_{\tau,d}^\infty$ is a well-defined (and monotone) limit of fractional powers of functions of the type $\Phi_{\tau,p}^2$ with p tending to infinity along a properly chosen subsequence, hence plurisubharmonic in the sense agreed just above (see [GZ17, Theorem 1.46]). For these functions, (3.50) does hold, but in a weak sense. This same statement remains true for $\Phi_\tau^2, \Sigma_\tau^2, \Sigma_\tau^\infty$, and Φ_τ^∞ if they are less than $+\infty$. Thus,

$$(3.51) \quad \sum_{i,j=1}^n \partial_{Z_i^*} \partial_{Z_j} \Phi_{\tau,d}^2(\underline{A})(C_j)(C_i) \geq 0 \text{ pointwise, and}$$

$$(3.52) \quad \sum_{i,j=1}^n \partial_{Z_i^*} \partial_{Z_j} \Phi_{\tau,d}^\infty(\underline{A})(C_j)(C_i) \geq 0 \text{ in the sense of distributions,}$$

for all $k, d \in \mathbb{N}$, and all $\underline{A}, \underline{C} \in \mathbb{M}_k(\mathbb{C})^n$, with similar statements for Σ . By positivity in the sense of distributions we mean that $\sum_{i,j=1}^n \partial_{Z_i^*} \partial_{Z_j} \Phi_{\tau,d}^\infty(\underline{A})(C_i)(C_j)$ is a positive distribution on the Euclidean space \mathbb{C}^{nk^2} – the variables considered being \underline{A} , while \underline{C} should be viewed as parameters. We expect that eventually the correct use of the difference-differential operators will reveal a tighter connection between τ and the Monge-Ampère measures associated to $\frac{1}{2} \log \Phi_\tau^2, \frac{1}{2} \log \Sigma_\tau^2, \frac{1}{2} \log \Sigma_\tau^\infty$, and $\frac{1}{2} \log \Phi_\tau^\infty$, 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 $\tau: \mathbb{C}\langle X \rangle \rightarrow \mathbb{C}$ one associates, under some reasonable assumptions, sequences of Borel probability measures supported on $\mathbb{M}_k(\mathbb{C})^n \simeq \mathbb{C}^{nk^2}$. 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 τ itself besides being determined by it. We conjecture that these measures might be good approximants for τ in a very precise way, thus providing another means to find random matrix approximants to (some) noncommutative distributions.

Conjecture 4.1. *Assume that $\tau: \mathbb{C}\langle \underline{X} \rangle \rightarrow \mathbb{C}$ is a bounded faithful tracial state. Suppose that $\{X_1, \dots, X_n\}$ are free with respect to τ and $\tau|_{\mathbb{C}\langle X_j \rangle}$ has a support with nonempty interior, $1 \leq j \leq n$. For all $f \in \text{Sym } \mathbb{R}\langle \underline{X} \rangle$, one has*

$$(4.1) \quad \lim_{d \rightarrow \infty} \limsup_{k \rightarrow \infty} \sup_{\underline{A} \in \Omega_{\tau, d}^{(k)}} |\text{tr}_k(f(\underline{A})) - \tau(f(\underline{X}))| = \lim_{d \rightarrow \infty} \limsup_{k \rightarrow \infty} \sup_{\underline{A} \in \Omega_{\tau, d}^{(k)}} |\text{tr}_k(f(\underline{A})) - \text{tr}(f(\underline{a}_\tau))| = 0,$$

$$(4.2) \quad \lim_{d \rightarrow \infty} \limsup_{k \rightarrow \infty} \sup_{\underline{A} \in \Omega_{\tau, d}^{(k)}} \|f(\underline{A})\|_{\mathbb{M}_k(\mathbb{C})} = \|f(\underline{a}_\tau)\|_{W^*(\tau)},$$

where $\Omega_{\tau, d}^{(k)} = \{\underline{A} : [\limsup_{d \rightarrow \infty} \text{tr}_k(\kappa_{\tau, d}(\underline{A}, \underline{A}^*)(I_k))^{1/d}]^* \leq n\}$.

Let us note some immediate consequences in this context of Theorem 3.3. Observe that in the above conjecture, the matrix variables \underline{A} are taken as approximants of \underline{a}_τ : 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 τ 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 \underline{a}_τ . 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 [HMS96] to handle noncommutative polynomials. All results were obtained on an Intel Xeon(R) E-2176M CPU (2.70GHz \times 6) with 32Gb of RAM. Our code is available online³.

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

$$\left[\limsup_{d \rightarrow \infty} \text{tr}_k(\kappa_{\tau, d}(\underline{A}, \underline{A}^*)(I_k))^{1/d} \right]^*.$$

For several values of the matrix size k , we sample N matrices from the following set

$$\tilde{\Omega}_{\tau, d}^{(k)} = \{\underline{A} : n - \varepsilon \leq \tilde{\Phi}_{\tau, d}^{(k)}(\underline{A}) \leq n + \varepsilon\},$$

with respect to a given matrix distribution. Overall, it boils down to sampling in an ε -perturbation of the n -level set of the approximate Siciak function $\tilde{\Phi}_{\tau, d}^{(k)}$. Then, we choose an $f \in \text{Sym } \mathbb{R}\langle \underline{X} \rangle$ and empirically check if (4.1) from Theorem 4.1 holds in average. This verification is performed by computing the expectation $E(\text{tr}_k(f(\underline{A})))$ over the N selected samples \underline{A} from $\tilde{\Omega}_{\tau, d}^{(k)}$ and comparing it with $\tau(f(\underline{A}))$. For all experiments, we select $N = 10^5$.

³<http://homepages.laas.fr/vmagron/files/NCCD.zip>

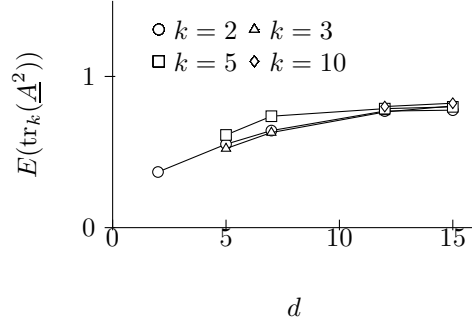


FIGURE 1. Averaging the normalized traces of \underline{A}^2 over $N = 10^5$ samples of $\tilde{\Omega}_{\tau,d}^{(k)}$ for a single free semicircular of variance 1.

4.1. Single free semicircular. Here, we consider a single standard free semicircular \underline{A} of variance 1. By [MS17, Definition 4], the odd moments of \underline{A} are 0 and the even moments are given by the Catalan numbers, i.e., $\tau(\underline{A}^{2d+1}) = 0$ and $\tau(\underline{A}^{2d}) = \frac{1}{d+1} \binom{2d}{d}$. The first four corresponding values of $\kappa_{\tau,d}$ are:

$$\begin{aligned}\kappa_{\tau,1}(\underline{A}, \underline{A})(1) &= 1 + \underline{A}^2, \\ \kappa_{\tau,2}(\underline{A}, \underline{A})(1) &= 2 - \underline{A}^2 + \underline{A}^4, \\ \kappa_{\tau,3}(\underline{A}, \underline{A})(1) &= 2 + 3\underline{A}^2 - 3\underline{A}^4 + \underline{A}^6, \\ \kappa_{\tau,4}(\underline{A}, \underline{A})(1) &= 3 - 3\underline{A}^2 + 8\underline{A}^4 - 5\underline{A}^6 + \underline{A}^8.\end{aligned}$$

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

4.2. Pair of free semicircular distributions. Here, we consider a pair $\underline{A} = (A_1, A_2)$ of standard free semicircular distributions of variance 1. The odd moments of \underline{A} are 0 and the even moments are given by the number of non-crossing pairings which respect the indices, e.g., $\tau(A_1 A_2 A_1) = 0$ and $\tau(A_1 A_1 A_2 A_2 A_1 A_1) = 2$. The first three corresponding values of $\kappa_{\tau,d}$ are:

$$\begin{aligned}\kappa_{\tau,1}(\underline{A}, \underline{A})(1) &= 1 + A_1^2 + A_2^2, \\ \kappa_{\tau,2}(\underline{A}, \underline{A})(1) &= 3 - A_1^2 - A_2^2 + A_1^4 + A_1 A_2^2 A_1 + A_2 A_1^2 A_2 + A_2^4, \\ \kappa_{\tau,3}(\underline{A}, \underline{A})(1) &= 3 + 5A_1^2 + 5A_2^2 - 3A_1^4 - 2A_1^2 A_2^2 - A_1 A_2^2 A_1 - A_2 A_2^1 A_2 - 2A_2^2 A_1^2 \\ &\quad - 3A_2^4 + A_1^6 + A_1^2 A_2^2 A_1^2 + A_1 A_2 A_1^2 A_2 A_1 + A_1 A_2^4 A_1 + A_2 A_1^4 A_2 \\ &\quad + A_2 A_1 A_2^2 A_1 A_2 + A_2^6.\end{aligned}$$

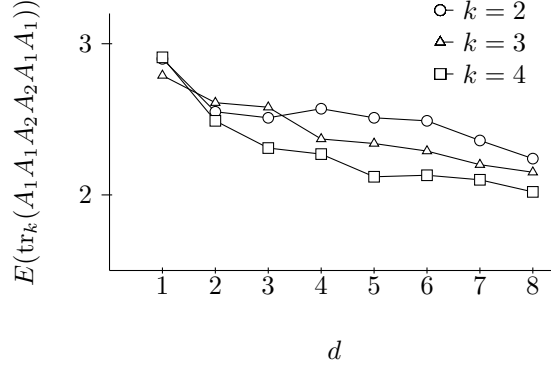


FIGURE 2. Averaging the normalized traces of $A_1 A_1 A_2 A_2 A_1 A_1$ over $N = 10^5$ samples of $\tilde{\Omega}_{\tau,d}^{(k)}$ for a pair of free semicirculars of variance 1.

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

4.3. Pair of free Poisson distributions. Finally, we consider a pair $\underline{A} = (A_1, A_2)$ 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:

$$\begin{aligned} \kappa_{\tau,1}(\underline{A}, \underline{A})(1) &= 1 + 2c - 2A_1 - 2A_2 + \frac{A_1^2 + A_2^2}{c}, \\ \kappa_{\tau,2}(\underline{A}, \underline{A})(1) &= 1 + 2c + 4c^2 - (4 + 8c)(A_1 + A_2) + \left(8 + \frac{5}{c} + \frac{1}{c^2}\right)(A_1^2 + A_2^2) \\ &\quad + 4(A_1 A_2 + A_2 A_1) - \left(\frac{2}{c^2} + \frac{4}{c}\right)(A_1^3 + A_2^3) \\ &\quad - \frac{A_1^2 A_2 + A_2 A_1^2 + A_1 A_2^2 + A_2^2 A_1}{c} - \frac{2(A_1 A_2 A_1 + A_2 A_1 A_2)}{c} \\ &\quad + \frac{A_1^4 + A_1 A_2^2 A_1 + A_2 A_1^2 A_2 + A_2^4}{c^2}. \end{aligned}$$

Recall that a Wishart matrix is of the form $A = \frac{1}{k} G G^*$, 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 . In our experiments, we sample $\tilde{\Omega}_{\tau,d}^{(k)}$ with

respect to the Wishart matrix distribution, for $c = k = 5$, with a perturbation of $\varepsilon = 10$. For $f(\underline{A}) = A_1 + A_2$, we obtain the successive averages of normalized traces: 10.42, 10.24, 10.16, 10.14 and $10.1 \simeq 10 = \tau(A_1 + A_2)$, for $d \in \{1, \dots, 5\}$, respectively. For $f(\underline{A}) = A_1 A_2$, we obtain 26.78, 25.78, 25.7, 25.6 and 25.4, while $\tau(A_1 A_2) = 5^2 = 25$. Eventually, for $f(\underline{A}) = A_1^3$, 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 and symmetry 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 $f \in \text{Sym } \mathbb{R}\langle \underline{X} \rangle$, a positive integer m and $G = \{g_1, \dots, g_m\} \subseteq \text{Sym } \mathbb{R}\langle \underline{X} \rangle$, the *semialgebraic* set D_G associated to G is defined as follows:

$$(4.3) \quad D_G := \bigcup_{k \in \mathbb{N}} \{ \underline{A} = (A_1, \dots, A_n) \in \mathbb{S}_k^n : g_j(\underline{A}) \succeq 0, \quad j = 1, \dots, m \}.$$

Now fix a possibly infinite dimensional complex Hilbert space \mathcal{H} endowed with a scalar product $\langle \cdot | \cdot \rangle$ and let $\mathcal{B}(\mathcal{H})$ be the algebra of all bounded linear operators on \mathcal{H} . As usual, positivity for an operator $C \in \mathcal{B}(\mathcal{H})$ 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 $C \succ 0$. For any set \mathcal{M} of bounded operators on \mathcal{H} , we define

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

and its noncommutative extension

$$(4.4) \quad D_{G, \text{nc}}^{\mathcal{M}} := \bigcup_{k \in \mathbb{N}} \{ \underline{A} = (A_1, \dots, A_n) \in \mathbb{M}_k^{\text{sa}}(\mathcal{M})^n : g_j(\underline{A}) \succeq 0 \text{ on } \mathcal{H}^k, \quad j = 1, \dots, m \}.$$

(We denote by $\mathbb{M}_k^{\text{sa}}(\mathcal{M})$ 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 tr (see [Tak79, Chapter V.2]). In particular, the set D_G defined in (4.3) corresponds, with the notation from (4.4), to $D_{G, \text{nc}}^{\mathbb{C}}$.

Given an arbitrary algebra \mathcal{A} and $g, h \in \mathcal{A}$, we denote by $[g, h] := gh - hg$ the *commutator* of g and h . In the particular case when $\mathcal{A} = \mathbb{C}\langle \underline{X} \rangle$, two nc polynomials g, h are called *cyclically equivalent* (denoted $g \stackrel{\text{cyc}}{\sim} h$) if $g - h$ is a sum of commutators. In particular, $\tau(g) = \tau(h)$.

Given the trace τ 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 $\langle \underline{X} \rangle_{d-d_j}$ 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 $\text{tr}_{\min}(f, G)$ as follows:

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

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

$$(4.6) \quad \mathrm{tr}_{\min}(f, G)^{\mathcal{M}} := \inf\{\mathrm{tr} f(\underline{A}) : \underline{A} \in D_{G, \mathrm{nc}}^{\mathcal{M}}\}.$$

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

$$(4.7) \quad \begin{aligned} \tau_d(f, K) &:= \inf_{\tau} \tau(f) \\ \text{s.t. } M_d(\tau)_{u,v} &= M_d(\tau)_{w,z}, \quad \text{for all } u^*v \stackrel{\mathrm{cyc}}{\approx} w^*z, \\ M_d(\tau)_{1,1} &= 1, \\ M_d(\tau) &\succeq 0, \quad M_d(\tau) \in \mathbb{S}_{\sigma(n,d)}. \\ M_d(g_j\tau) &\succeq 0, \quad M_d(g_j\tau) \in \mathbb{S}_{\sigma(n,d-d_j)}, \quad j = 1, \dots, m. \end{aligned}$$

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 τ from the inverse moment matrix (as stated in Proposition 3.5) 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 Skłodowska-Curie Actions, grant agreement 813211 (POEMA), as well as the FMJH Program PGMO (EPICS project).

REFERENCES

- [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, 6

- [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, 6, 30
- [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, 41
- [BKP16] Sabine Burgdorf, Igor Klep, and Janez Povh. *Optimization of polynomials in non-commuting variables*. SpringerBriefs in Mathematics. Springer, [Cham], 2016. 2, 42
- [Blo97] Thomas Bloom. Orthogonal polynomials in \mathbb{C}^n . *Indiana Univ. Math. J.*, 46(2):427–452, 1997. 21
- [BLPW15] Thomas Bloom, Norman Levenberg, Federico Piazzon, and Franck Wielonsky. Bernstein-Markov: a survey. *arXiv:1512.00739v1 [math.CV]*, 2015. 21, 26
- [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, 4, 14, 25
- [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/0205333v1 [math.FA]*, 2002. 2
- [DX14] Charles F Dunkl and Yuan Xu. *Orthogonal polynomials of several variables*. Number 155. Cambridge University Press, 2014. 1, 5, 10
- [GZ17] Vincent Guedj and Ahmed Zeriahi. *Degenerate Complex Monge-Ampère Equations*. Number 26 in *Tracts in Mathematics*. European Mathematical Society, 2017. 3, 13, 19, 34, 36, 37
- [Hel02] J. William Helton. “Positive” noncommutative polynomials are sums of squares. *Ann. of Math. (2)*, 156(2):675–694, 2002. 2
- [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
- [HLP+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. 41
- [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. 38
- [Kli91] Maciej Klimek. *Pluripotential theory*. Number 6 in *London Mathematical Society Monographs*. Oxford: Clarendon Press, 1991. 13, 19, 29, 30
- [KMP21] Igor Klep, Victor Magron, and Janez Povh. Sparse noncommutative polynomial optimization. *Mathematical Programming*, pages 1–41, 2021. 2
- [KMV20] I. Klep, V. Magron, and J. Volcic. Optimization over trace polynomials. *arXiv preprint arXiv:2006.12510*, 2020. Submitted. 2
- [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, 10, 12
- [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 Nevai. Bernstein’s inequality in L_p for and $(c, 1)$ bounds for orthogonal polynomials. *Annals of Mathematics*, pages 145–154, 1980. 1
- [MNT91] Attila Máté, Paul Nevai, 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. 39, 40
- [Nev86] Paul Nevai. Géza Freud, orthogonal polynomials and Christoffel functions. A case study. *Journal of Approximation Theory*, 48(1):3–167, 1986. 10
- [NPA08] Miguel Navascués, Stefano Pironio, and Antonio Acín. A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations. *New J. Phys.*, 10(7):073013, 2008. 2
- [Pau02] Vern Paulsen. *Completely bounded maps and operator algebras.*, volume 78 of *Cambridge studies in advanced mathematics*. Cambridge University Press, 2002. 2, 7, 10
- [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 1663–1672, 2018. 1
- [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. 2
- [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. 1
- [Put93] Mihai Putinar. Positive polynomials on compact semi-algebraic sets. *Indiana Univ. Math. J.*, 42(3):969–984, 1993. 2
- [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. 1
- [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. 13, 17, 18, 27, 34, 37
- [Tak79] Masamichi Takesaki. *Theory of Operator Algebras I*. Springer-Verlag, New York Inc., 1979. 41
- [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. 1
- [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. 6, 27
- [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. 35
- [Voi00] Dan Voiculescu. The Coalgebra of the Free Difference Quotient and Free Probability. *International Mathematics Research Notices*, 2000(2):79–106, 2000. 35
- [WM20] Jie Wang and Victor Magron. Exploiting term sparsity in noncommutative polynomial optimization. *arXiv preprint arXiv:2010.06956*, 2020. 2