Canad. Math. Bull. Vol. 66 (3), 2023, pp. 808–826
http://dx.doi.org/10.4153/S0008439522000741

Norms on complex matrices induced by random vectors

Ángel Chávez, Stephan Ramon Garcia, and Jackson Hurley

Abstract. We introduce a family of norms on the $n \times n$ complex matrices. These norms arise from a probabilistic framework, and their construction and validation involve probability theory, partition combinatorics, and trace polynomials in noncommuting variables. As a consequence, we obtain a generalization of Hunter's positivity theorem for the complete homogeneous symmetric polynomials.

1 Introduction

This paper introduces norms on the space M_n of $n \times n$ complex matrices that are induced by random vectors in $\mathbb{R}^n$. Specifically, we construct a family of norms for each random vector $\boldsymbol{X}$ whose entries are independent and identically distributed (iid) random variables with sufficiently many moments. Initially, these norms are defined on complex Hermitian matrices as symmetric functions of their (necessarily real) eigenvalues. This contrasts with Schatten and Ky-Fan norms, which are defined in terms of singular values. To be more specific, our norms do not arise from the machinery of symmetric gauge functions [13, Section 7.4.7]. The random vector norms we construct are actually generalizations of the complete homogeneous symmetric (CHS) polynomial norms introduced in [1].

1.1 Preliminaries

Our main result (Theorem 1.1 on page 4) connects a wide range of topics, such as cumulants, Bell polynomials, partitions, and Schur convexity. We briefly cover the preliminary concepts and notation necessary to state our main results.

1.1.1 Numbers and matrices

In what follows, $\mathbb{N} = \{1, 2, \ldots\}$; the symbols $\mathbb{R}$ and $\mathbb{C}$ denote the real and complex number systems, respectively. Let M_n denote the set of $n \times n$ complex matrices and $\mathrm{H}_n \subset \mathrm{M}_n$ the subset of $n \times n$ Hermitian complex matrices. We reserve the letter A for Hermitian matrices (so $A = A^*$) and Z for arbitrary square complex matrices. The

Received by the editors November 18, 2022; revised December 2, 2022; accepted December 11, 2022.
Published online on Cambridge Core December 23, 2022.
S.R.G. was partially supported by the NSF (Grant No. DMS-2054002).
AMS subject classification: **47A30, 15A60, 16R30.**
Keywords: Norm, symmetric polynomial, partition, trace, positivity, convexity, expectation, complexification, trace polynomial, probability distribution.

eigenvalues of each $A \in \mathrm{H}_n$ are real and denoted $\lambda_1(A) \geq \lambda_2(A) \geq \cdots \geq \lambda_n(A)$. We may write $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_n)$ if A is understood.

1.1.2 Probability theory

A *probability space* is a measure space $(\Omega, \mathcal{F}, \mathbf{P})$, in which $\mathcal{F}$ is a σ-algebra on Ω, $\mathbf{P}$ is nonnegative, and $\mathbf{P}(\Omega) = 1$. A *random variable* is a measurable function $X : \Omega \to \mathbb{R}$. We assume that $\Omega \subseteq \mathbb{R}$ and X is nondegenerate, that is, nonconstant. The *expectation* of X is $\mathbf{E}[X] = \int_\Omega X \, d\mathbf{P}$, often written as $\mathbf{E}X$. For $p \geq 1$, let $L^p(\Omega, \mathcal{F}, \mathbf{P})$ denote the vector space of random variables such that $\|X\|_{L^p} = (\mathbf{E}|X|^p)^{1/p} < \infty$. The pushforward measure $X_*\mathbf{P}$ of X is the *probability distribution* of X. The *cumulative distribution* of X is $F_X(x) = \mathbf{P}(X \leq x)$, which is the pushforward measure of $(-\infty, x]$. If $X_*\mathbf{P}$ is absolutely continuous with respect to Lebesgue measure m, the Radon–Nikodym derivative $f_X = dX_*P/dm$ is the *probability density function* (PDF) of X [5, Chapter 1].

1.1.3 Random vectors

A *random vector* is a tuple $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$, in which $X_1, X_2, \ldots, X_n$ are real-valued random variables on a common probability space $(\Omega, \mathcal{F}, \mathbf{P})$; we assume $\Omega \subseteq \mathbb{R}$. A random vector $\boldsymbol{X}$ is *positive definite* if its *second-moment matrix* $\Sigma(\boldsymbol{X}) = [\mathbf{E}X_iX_j]_{i,j=1}^n$ exists and is positive definite. This occurs if the X_i are iid and belong to $L^2(\Omega, \mathcal{F}, \mathbf{P})$ (see Lemma 3.1).

1.1.4 Moments

For $k \in \mathbb{N}$, the kth *moment* of X is $\mu_k = \mathbf{E}[X^k]$, if it exists. If X has PDF f_X, then $\mu_k = \int_{-\infty}^{\infty} x^k f_X(x)\, dm(x)$. The *mean* of X is μ_1 and the *variance* of X is $\mu_2 - \mu_1^2$; Jensen's inequality ensures that the variance is positive since X is nondegenerate. The *moment generating function* (if it exists) of X is

$$M(t) = \mathbf{E}[e^{tX}] = \sum_{k=0}^{\infty} \mathbf{E}[X^k] \frac{t^k}{k!} = \sum_{k=0}^{\infty} \mu_k \frac{t^k}{k!}. \tag{1.1}$$

If $X_1, X_2, \ldots, X_n$ are independent, then $\mathbf{E}[X_1^{i_1} X_2^{i_2} \cdots X_n^{i_n}] = \prod_{k=1}^n \mathbf{E}[X_k^{i_k}]$, for all $i_1, i_2, \ldots, i_n \in \mathbb{N}$ whenever both sides exist.

1.1.5 Cumulants

If X admits a moment generating function $M(t)$, then the rth *cumulant* κ_r of X is defined by the *cumulant generating function*

$$K(t) = \log M(t) = \sum_{r=1}^{\infty} \kappa_r \frac{t^r}{r!}. \tag{1.2}$$

The first two cumulants are $\kappa_1 = \mu_1$ and $\kappa_2 = \mu_2 - \mu_1^2$. If X does not admit a moment generating function but $X \in L^d(\Omega, \mathcal{F}, \mathbf{P})$ for some $d \in \mathbb{N}$, we can define $\kappa_1, \kappa_2, \ldots, \kappa_d$ by the recursion $\mu_r = \sum_{\ell=0}^{r-1} \binom{r-1}{\ell} \mu_\ell \kappa_{r-\ell}$ for $1 \leq r \leq d$ (see [5, Section 9]).

1.1.6 Power-series coefficients

The coefficient c_k of t^k in $f(t) = \sum_{r=0}^{\infty} c_r t^r$ is denoted $[t^k]f(t)$, as is standard in combinatorics and the study of generating functions.

1.1.7 Complete Bell polynomials

The *complete Bell polynomials of degree* ℓ [4, Section II] are the polynomials $B_\ell(x_1, x_2, \ldots, x_\ell)$ defined by

$$\sum_{\ell=0}^{\infty} B_\ell(x_1, x_2, \ldots, x_\ell)\frac{t^\ell}{\ell!} = \exp\left(\sum_{j=1}^{\infty} x_j \frac{t^j}{j!}\right). \tag{1.3}$$

For example, $B_0 = 1$, $B_2(x_1, x_2) = x_1^2 + x_2$, and

$$B_4(x_1, x_2, x_3, x_4) = x_1^4 + 6x_1^2 x_2 + 4x_1 x_3 + 3x_2^2 + x_4. \tag{1.4}$$

1.1.8 Symmetric and positive functions

A function is *symmetric* if it is invariant under all permutations of its arguments. A continuous real-valued function on M_n or H_n is *positive definite* if it is everywhere positive, except perhaps at 0.

1.1.9 Partitions

A *partition* of $d \in \mathbb{N}$ is a tuple $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_r) \in \mathbb{N}^r$ such that $\pi_1 \geq \pi_2 \geq \cdots \geq \pi_r$ and $\pi_1 + \pi_2 + \cdots + \pi_r = d$ [20, Section 1.7]. We denote this $\boldsymbol{\pi} \vdash d$ and write $|\boldsymbol{\pi}| = r$ for the number of parts in the partition. Define

$$\kappa_{\boldsymbol{\pi}} = \kappa_{\pi_1}\kappa_{\pi_2}\cdots\kappa_{\pi_r} \quad \text{and} \quad y_{\boldsymbol{\pi}} = \prod_{i \geq 1} (i!)^{m_i} m_i!, \tag{1.5}$$

in which $m_i = m_i(\boldsymbol{\pi})$ is the multiplicity of i in $\boldsymbol{\pi}$. For example, $\boldsymbol{\pi} = (4,4,2,1,1,1)$ yields $\kappa_{\boldsymbol{\pi}} = \kappa_4^2 \kappa_2 \kappa_1^3$ and $y_{\boldsymbol{\pi}} = (1!^3 3!)(2!^1 1!)(4!^2 2!) = 13{,}824$. Note that $y_{\boldsymbol{\pi}}$ is not the quantity $z_{\boldsymbol{\pi}} = \prod_{i \geq 1} i^{m_i} m_i!$ from symmetric function theory [21, Proposition 7.7.6].

1.1.10 Power sums

For $\boldsymbol{\pi} \vdash d$, let $p_{\boldsymbol{\pi}}(x_1, x_2, \ldots, x_n) = p_{\pi_1} p_{\pi_1} \cdots p_{\pi_r}$, where $p_k(x_1, x_2, \ldots, x_n) = x_1^k + x_2^k + \cdots + x_n^k$ is a *power-sum symmetric polynomial*; we often simply write p_k. If $A \in \mathrm{H}_n$ has eigenvalues $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_n)$, we write

$$p_{\boldsymbol{\pi}}(\boldsymbol{\lambda}) = p_{\pi_1}(\boldsymbol{\lambda}) p_{\pi_2}(\boldsymbol{\lambda}) \cdots p_{\pi_r}(\boldsymbol{\lambda}) = (\operatorname{tr} A^{\pi_1})(\operatorname{tr} A^{\pi_2}) \cdots (\operatorname{tr} A^{\pi_r}). \tag{1.6}$$

1.1.11 Complete homogeneous symmetric polynomials

The *CHS polynomial* of degree d in $x_1, x_2, \ldots x_n$ is

$$h_d(x_1, x_2, \ldots, x_n) = \sum_{1 \leq i_1 \leq \cdots \leq i_d \leq n} x_{i_1} x_{i_2} \cdots x_{i_d}, \tag{1.7}$$

https://doi.org/10.4153/S0008439522000741 Published online by Cambridge University Press

the sum of all monomials of degree d in $x_1, x_2, \dots, x_n$ (see [21, Section 7.5]). For example, $h_0(x_1, x_2) = 1$, $h_2(x_1, x_2) = x_1^2 + x_1x_2 + x_2^2$, and $h_4(x_1, x_2) = x_1^4 + x_1^3x_2 + x_1^2x_2^2 + x_1x_2^3 + x_2^4$. Hunter proved that the even-degree CHS polynomials are positive definite [14]. This has been rediscovered many times [1, Theorem 1], [2, Lemma 3.1], [3], [6, Theorem 2], [9, Corollary 17], [19, Theorem 2.3], [22, Theorem 1].

1.1.12 Schur convexity

Let $\widetilde{\boldsymbol{x}} = (\widetilde{x}_1, \widetilde{x}_2, \dots, \widetilde{x}_n)$ be the nondecreasing rearrangement of $\boldsymbol{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$. Then $\boldsymbol{y}$ *majorizes* $\boldsymbol{x}$, denoted $\boldsymbol{x} \prec \boldsymbol{y}$, if $\sum_{i=1}^n \widetilde{x}_i = \sum_{i=1}^n \widetilde{y}_i$ and $\sum_{i=1}^k \widetilde{x}_i \le \sum_{i=1}^k \widetilde{y}_i$ for $1 \le k \le n$. A function $f : \mathbb{R}^n \to \mathbb{R}$ is *Schur convex* if $f(\boldsymbol{x}) \le f(\boldsymbol{y})$ whenever $\boldsymbol{x} \prec \boldsymbol{y}$. This occurs if and only if $(x_i - x_j)(\frac{\partial}{\partial x_i} - \frac{\partial}{\partial x_j}) f(x_1, x_2, \dots, x_n) \ge 0$ for all $1 \le i < j \le n$, with equality if and only if $x_i = x_j$ [18, p. 259].

1.2 Statement of main results

With the preliminary concepts and notation covered, we can state our main theorem. In what follows, Γ is the gamma function and $\langle \cdot, \cdot \rangle$ is the Euclidean inner product on $\mathbb{R}^n$.

Theorem 1.1 *Let $d \ge 2$ and $\boldsymbol{X} = (X_1, X_2, \dots, X_n)$, in which $X_1, X_2, \dots, X_n \in L^d(\Omega, \mathcal{F}, \mathbf{P})$ are nondegenerate iid random variables.*

(a) $\|A\|_{\boldsymbol{X},d} = \left(\dfrac{\mathbf{E}|\langle \boldsymbol{X}, \boldsymbol{\lambda} \rangle|^d}{\Gamma(d+1)} \right)^{1/d}$ *is a norm on* H_n.

(b) *If the X_i admit a moment generating function $M(t)$ and $d \ge 2$ is even, then*

$$\|A\|_{\boldsymbol{X},d}^d = [t^d] M_\Lambda(t) \quad \textit{for all } A \in \mathrm{H}_n, \tag{1.8}$$

in which $M_\Lambda(t) = \prod_{i=1}^n M(\lambda_i t)$ is the moment generating function for the random variable $\Lambda = \langle \boldsymbol{X}, \boldsymbol{\lambda}(A) \rangle = \lambda_1 X_1 + \lambda_2 X_2 + \cdots + \lambda_n X_n$. In particular, $\|A\|_{\boldsymbol{X},d}$ is a positive definite, homogeneous, symmetric polynomial in the eigenvalues of A.

(c) *If the first d moments of X_i exist, then*

$$\|A\|_{\boldsymbol{X},d}^d = \frac{1}{d!} B_d(\kappa_1 \operatorname{tr} A, \kappa_2 \operatorname{tr} A^2, \dots, \kappa_d \operatorname{tr} A^d) \tag{1.9}$$

$$= \sum_{\boldsymbol{\pi} \vdash d} \frac{\kappa_{\boldsymbol{\pi}} p_{\boldsymbol{\pi}}(\boldsymbol{\lambda})}{y_{\boldsymbol{\pi}}} \quad \textit{for } A \in \mathrm{H}_n, \tag{1.10}$$

in which B_d is given by (1.3), and in which $\kappa_{\boldsymbol{\pi}}$ and $y_{\boldsymbol{\pi}}$ are defined in (1.5), $p_{\boldsymbol{\pi}}(\boldsymbol{\lambda})$ is defined in (1.6), and the second sum runs over all partitions $\boldsymbol{\pi}$ of d.

(d) *The function $\boldsymbol{\lambda}(A) \mapsto \|A\|_{\boldsymbol{X},d}$ is Schur convex.*

(e) *Let $\boldsymbol{\pi} = (\pi_1, \pi_2, \dots, \pi_r)$ be a partition of d. Define $\mathrm{T}_{\boldsymbol{\pi}} : \mathrm{M}_n \to \mathbb{R}$ by setting $\mathrm{T}_{\boldsymbol{\pi}}(Z)$ to be $1/\binom{d}{d/2}$ times the sum over the $\binom{d}{d/2}$ possible locations to place $d/2$ adjoints*

https://doi.org/10.4153/S0008439522000741 Published online by Cambridge University Press

* among the d copies of Z in $(\operatorname{tr}\underbrace{ZZ\cdots Z}_{\pi_1})(\operatorname{tr}\underbrace{ZZ\cdots Z}_{\pi_2})\cdots(\operatorname{tr}\underbrace{ZZ\cdots Z}_{\pi_r})$. Then

$$\|Z\|_{\boldsymbol{X},d} = \left(\sum_{\boldsymbol{\pi}\vdash d} \frac{\kappa_{\boldsymbol{\pi}} \mathrm{T}_{\boldsymbol{\pi}}(Z)}{y_{\boldsymbol{\pi}}}\right)^{1/d} \quad \text{for } Z \in \mathrm{M}_n, \tag{1.11}$$

in which $\kappa_{\boldsymbol{\pi}}$ and $y_{\boldsymbol{\pi}}$ are defined in (1.5) and the sum runs over all partitions $\boldsymbol{\pi}$ of d, is a norm on M_n that restricts to the norm on H_n above. In particular, $\|Z\|_{\boldsymbol{X},d}^d$ is a positive definite trace polynomial in Z and Z^*.

The independence of the X_i is not needed in (a) and (d) (see Remarks 3.4 and 3.5, respectively). A more precise definition of $\mathrm{T}_{\boldsymbol{\pi}}(Z)$ is in Section 3.5, although the examples in the next section better illustrate how to compute (1.11).

The positive definiteness of (1.8), (1.9), and (1.11) is guaranteed by Theorem 1.1; the triangle inequality is difficult to verify directly. Positivity is not obvious since we consider the eigenvalues of $A \in \mathrm{H}_n$ and not their absolute values in (a) and (b). Thus, these norms on H_n do not arise from singular values or symmetric gauge functions [13, Section 7.4.7]. Norms like ours can distinguish singularly cospectral graphs, unlike the operator, Frobenius, Schatten–von Neumann, and Ky Fan norms (see [1, Example 2]).

1.3 Organization

This paper is organized as follows: We first cover examples and applications in Section 2, including a generalization of Hunter's positivity theorem. The proof of Theorem 1.1, which is lengthy and involves a variety of ingredients, is contained in Section 3. We end this paper in Section 4 with a list of open questions that demand further exploration.

2 Examples and applications

We begin with computations for small d (Section 2.1). Gamma random variables (Section 2.2) lead to a generalization of Hunter's positivity theorem (Section 2.3). We examine norms arising from familiar distributions in Sections 2.4–2.10.

2.1 Generic computations

Let $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$, where the X_i are nondegenerate iid random variables such that the stated cumulants and moments exist. For $d = 2$ and 4, we obtain trace-polynomial representations of $\|Z\|_d$ in terms of cumulants or moments. This can also be done for $d = 6, 8, \ldots$, but we refrain from the exercise.

Example 2.1 The two partitions of $d = 2$ satisfy $\kappa_{(2)} = \kappa_2 = \mu_2 - \mu_1^2$, $\kappa_{(1,1)} = \kappa_1^2 = \mu_1^2$, and $y_{(2)} = y_{(1,1)} = 2$. There are $\binom{2}{1} = 2$ ways to place two adjoints * in a string of two Zs. Thus, $\mathrm{T}_{(2)}(Z) = \frac{1}{2}(\operatorname{tr}(Z^*Z) + \operatorname{tr}(ZZ^*)) = \operatorname{tr}(Z^*Z)$ and

https://doi.org/10.4153/S0008439522000741 Published online by Cambridge University Press

$\mathrm{T}_{(1,1)}(Z) = \frac{1}{2}((\operatorname{tr} Z^*)(\operatorname{tr} Z) + (\operatorname{tr} Z)(\operatorname{tr} Z^*)) = (\operatorname{tr} Z^*)(\operatorname{tr} Z)$, so

$$\|Z\|_{\mathbf{X},2}^2 = \sum_{\pi \vdash d} \frac{\kappa_\pi \mathrm{T}_\pi(A)}{y_\pi} = \frac{\mu_2 - \mu_1^2}{2} \operatorname{tr}(Z^* Z) + \frac{\mu_1^2}{2}(\operatorname{tr} Z^*)(\operatorname{tr} Z). \tag{2.1}$$

If $\mu_1 = 0$ (mean zero), then $\|\cdot\|_2$ is a nonzero multiple of the Frobenius norm since the variance $\mu_2 - \mu_1^2$ is positive by nondegeneracy. As predicted by Theorem 1.1, the norm (2.1) on M_n reduces to (1.9) on H_n since $B_2(x_1, x_2) = x_1^2 + x_2$ and

$$\begin{aligned}\|A\|_{\mathbf{X},2}^2 &= \frac{1}{2} B_2(\kappa_1 \operatorname{tr} A, \kappa_2 \operatorname{tr} A^2)\\ &= \frac{1}{2}\big[(\kappa_1 \operatorname{tr} A)^2 + \kappa_2 \operatorname{tr}(A^2)\big] = \frac{\mu_2 - \mu_1^2}{2} \operatorname{tr}(A^2) + \frac{\mu_1^2}{2}(\operatorname{tr} A)^2,\end{aligned}$$

which agrees with (2.1) if $Z = A = A^*$.

Example 2.2 The five partitions of $d = 4$ satisfy

$$\begin{array}{lllll} \kappa_{(4)} = \kappa_4, & \kappa_{(3,1)} = \kappa_1\kappa_3, & \kappa_{(2,2)} = \kappa_2^2, & \kappa_{(2,1,1)} = \kappa_2\kappa_1^2, & \kappa_{(1,1,1,1)} = \kappa_1^4,\\ y_{(4)} = 24, & y_{(3,1)} = 6, & y_{(2,2)} = 8, & y_{(2,1,1)} = 4, & y_{(1,1,1,1)} = 24. \end{array}$$

There are $\binom{4}{2} = 6$ ways to place two adjoints * in a string of four Zs. For example,

$$\begin{aligned} 6\mathrm{T}_{(3,1)}(Z) &= (\operatorname{tr} Z^*Z^*Z)(\operatorname{tr} Z) + (\operatorname{tr} Z^*ZZ^*)(\operatorname{tr} Z) + (\operatorname{tr} Z^*ZZ)(\operatorname{tr} Z^*)\\ &\quad + (\operatorname{tr} ZZ^*Z^*)(\operatorname{tr} Z) + (\operatorname{tr} ZZ^*Z)(\operatorname{tr} Z^*) + (\operatorname{tr} ZZZ^*)(\operatorname{tr} Z^*)\\ &= 3\operatorname{tr}(Z^{*2}Z)(\operatorname{tr} Z) + 3(\operatorname{tr} Z^2Z^*)(\operatorname{tr} Z^*). \end{aligned}$$

Summing over all five partitions yields the following norm on M_n:

$$\begin{aligned} \|Z\|_{\mathbf{X},4}^4 = \tfrac{1}{72}\big(&3\kappa_1^4(\operatorname{tr} Z^*)^2(\operatorname{tr} Z)^2 + 3\kappa_2\kappa_1^2(\operatorname{tr} Z^*)^2\operatorname{tr}(Z^2) + 3\kappa_2\kappa_1^2\operatorname{tr}(Z^{*2})(\operatorname{tr} Z)^2\\ &+ 12\kappa_2\kappa_1^2(\operatorname{tr} Z^*)(\operatorname{tr} Z^*Z)(\operatorname{tr} Z) + 6\kappa_3\kappa_1\operatorname{tr}(Z^{*2}Z)(\operatorname{tr} Z)\\ &+ 6\kappa_3\kappa_1\operatorname{tr}(Z^*)\operatorname{tr}(Z^*Z^2) + 6\kappa_2^2(\operatorname{tr} Z^*Z)^2 + 3\kappa_2^2\operatorname{tr}(Z^2)\operatorname{tr}(Z^{*2})\\ &+ 2\kappa_4\operatorname{tr}(Z^2Z^{*2}) + \kappa_4\operatorname{tr}(ZZ^*ZZ^*)\big). \end{aligned} \tag{2.2}$$

If $Z = A \in \mathrm{H}_n$, Theorem 1.1.c and (1.4) ensure that the above reduces to

$$\tfrac{1}{24}\big(\kappa_1^4(\operatorname{tr} A)^4 + 6\kappa_1^2\kappa_2\operatorname{tr}(A^2)(\operatorname{tr} A)^2 + 4\kappa_1\kappa_3\operatorname{tr}(A^3)\operatorname{tr}(A) + 3\kappa_2^2\operatorname{tr}(A^2)^2 + \kappa_4\operatorname{tr}(A^4)\big).$$

2.2 Gamma random variables

Let $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$, in which the X_i are independent with probability density

$$f(t) = \begin{cases} \frac{1}{\beta^\alpha \Gamma(\alpha)} t^{\alpha-1} e^{-t/\beta}, & \text{if } t > 0,\\ 0, & \text{if } t \le 0. \end{cases} \tag{2.3}$$

Here, $\alpha, \beta > 0$ (note that $\alpha = k/2$ and $\beta = 2$ yield a chi-squared random variable with k degrees of freedom, and $\alpha = \beta = 1$ is the standard exponential distribution). Then $M(t) = (1 - \beta t)^{-\alpha}$ and $K(t) = -\alpha \log(1 - \beta t)$, so

$$\kappa_r = \alpha\beta^r(r-1)! \quad \text{for } r \in \mathbb{N}. \tag{2.4}$$

https://doi.org/10.4153/S0008439522000741 Published online by Cambridge University Press

For even $d \geq 2$,

$$\|A\|_{X,d}^d = [t^d] \prod_{i=1}^n \frac{1}{(1-\beta\lambda_i t)^\alpha} = [t^d] \left(\frac{1}{\beta^n t^n p_A(\beta^{-1} t^{-1})} \right)^\alpha \quad \text{for } A \in \mathrm{H}_n, \tag{2.5}$$

in which $p_A(t) = \det(tI - A)$ denotes the characteristic polynomial of A.

Example 2.3 Since $\kappa_1 = \alpha\beta$ and $\kappa_2 = \alpha\beta^2$, (2.1) becomes $\|Z\|_{X,2}^2 = \frac{1}{2}\alpha\beta^2 \operatorname{tr}(Z^* Z) + \frac{1}{2}\alpha^2\beta^2(\operatorname{tr} Z^*)(\operatorname{tr} Z)$ for $Z \in \mathrm{M}_n$. Similarly, (2.2) yields generalizations of [1, equations (8) and (9)] (which correspond to $\alpha = \beta = 1$):

$$\begin{aligned}
\|Z\|_{X,4}^4 = \tfrac{1}{24}\big(&\alpha^4\beta^4(\operatorname{tr} Z)^2(\operatorname{tr} Z^*)^2 + \alpha^3\beta^4(\operatorname{tr} Z^*)^2 \operatorname{tr}(Z^2) \\
&+ 4\alpha^3\beta^4(\operatorname{tr} Z)(\operatorname{tr} Z^*)(\operatorname{tr} Z^* Z) + 2\alpha^2\beta^4(\operatorname{tr} Z^* Z)^2 \\
&+ \alpha^3\beta^4(\operatorname{tr} Z)^2 \operatorname{tr}(Z^{*2}) + \alpha^2\beta^4 \operatorname{tr}(Z^2)\operatorname{tr}(Z^{*2}) \\
&+ 4\alpha^2\beta^4 \operatorname{tr}(Z^*)\operatorname{tr}(Z^* Z^2) + 4\alpha^2\beta^4 \operatorname{tr}(Z)\operatorname{tr}(Z^{*2} Z) \\
&+ 2\alpha\beta^4 \operatorname{tr}(Z^* Z Z^* Z) + 4\alpha\beta^4 \operatorname{tr}(Z^{*2} Z^2)\big).
\end{aligned}$$

2.3 A generalization of Hunter's positivity theorem

Examining the gamma distribution (Section 2.2) recovers Hunter's theorem [14] (Corollary 2.6) and establishes a powerful generalization (Theorem 2.5).

Example 2.4 Let $\alpha = \beta = 1$ in (2.3) and (2.5). Then

$$\|A\|_{X,d}^2 = [t^d] \prod_{i=1}^n \frac{1}{1-\lambda_i t} = [t^d] \frac{1}{t^n p_A(t^{-1})} \quad \text{for } A \in \mathrm{H}_n, \tag{2.6}$$

which is [1, Theorem 20]. Expand each factor $(1-\lambda_i t)^{-1}$ as a geometric series, multiply out the result, and deduce that for $d \geq 2$ even,

$$\|A\|_{X,d}^d = [t^d] \prod_{i=1}^n \frac{1}{1-\lambda_i t} = [t^d] \sum_{r=0}^\infty h_r(\lambda_1, \lambda_2, \ldots, \lambda_n) t^r. \tag{2.7}$$

From (2.4), we have $\kappa_i = (i-1)!$. Therefore,

$$\frac{\kappa_{\boldsymbol{\pi}}}{y_{\boldsymbol{\pi}}} = \frac{\prod_{i \geq 1} [(i-1)!]^{m_i}}{\prod_{i \geq 1} (i!)^{m_i} m_i!} = \frac{1}{\prod_{i \geq 1} i^{m_i} m_i!}$$

for any partition $\boldsymbol{\pi}$. Theorem 1.1 and (1.5) imply that for even $d \geq 2$ and $A \in \mathrm{H}_n$,

$$h_d(\lambda_1, \lambda_2, \ldots, \lambda_n) = \|A\|_{X,d}^d = \sum_{\boldsymbol{\pi} \vdash d} \frac{\kappa_{\boldsymbol{\pi}} p_{\boldsymbol{\pi}}}{y_{\boldsymbol{\pi}}} = \sum_{\boldsymbol{\pi} \vdash d} \frac{p_{\boldsymbol{\pi}}}{z_{\boldsymbol{\pi}}}, \tag{2.8}$$

in which $z_{\boldsymbol{\pi}} = \prod_{i \geq 1} i^{m_i} m_i!$ and $p_{\boldsymbol{\pi}}$ is given by (1.6). This recovers the combinatorial representation of even-degree CHS polynomials [21, Proposition 7.7.6] and establishes Hunter's positivity theorem since $\|\cdot\|_{X,d}^d$ is positive definite.

The next theorem generalizes Hunter's theorem [14], which is the case $\alpha = 1$.

Theorem 2.5 *For even $d \geq 2$ and $\alpha \in \mathbb{N}$,*

$$H_{d,\alpha}(x_1, x_2, \ldots, x_n) = \sum_{\substack{\pi \vdash d \\ |\pi| \leq \alpha}} c_\pi h_\pi(x_1, x_2, \ldots, x_n)$$

is positive definite on $\mathbb{R}^n$, in which the sum runs over all partitions $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_r)$ of d. Here, $h_{\boldsymbol{\pi}} = h_{\pi_1} h_{\pi_2} \cdots h_{\pi_r}$ is a product of CHS polynomials and

$$c_\pi = \frac{\alpha!}{(\alpha - |\boldsymbol{\pi}|)! \prod_{i=1}^r m_i!},$$

where $|\boldsymbol{\pi}|$ denotes the number of parts in $\boldsymbol{\pi}$ and m_i is the multiplicity of i in $\boldsymbol{\pi}$.

Proof Let $\alpha \in \mathbb{N}$ and define polynomials $P_\ell^{(\alpha)}(x_1, x_2, \ldots, x_\ell)$ by

(2.9) $$P_0^{(\alpha)} = x_0 = 1 \quad \text{and} \quad \Big(1 + \sum_{r=1}^\infty x_r t^r\Big)^\alpha = \sum_{\ell=0}^\infty P_\ell^{(\alpha)}(x_1, x_2, \ldots, x_\ell) t^\ell.$$

Then

(2.10) $$P_\ell^{(\alpha)}(x_1, x_2, \ldots, x_\ell) = \sum_{\substack{i_1, i_2, \ldots, i_\alpha \leq \ell \\ i_1 + i_2 + \cdots + i_\alpha = \ell}} x_{i_1} x_{i_2} \cdots x_{i_\alpha} = \sum_{\substack{\pi \vdash \ell \\ |\pi| \leq \alpha}} c_\pi x_\pi.$$

Let $\boldsymbol{X}$ be a random vector whose n components are iid and distributed according to (2.3) with $\beta = 1$. Let $A \in \mathrm{H}_n$ have eigenvalues $x_1, x_2, \ldots, x_n$. For even $d \geq 2$,

$$\begin{aligned}
\|A\|_{\boldsymbol{X},d}^d &\overset{(2.5)}{=} [t^d] \Big(\prod_{i=1}^k \frac{1}{1 - x_i t} \Big)^\alpha \overset{(2.7)}{=} [t^d] \Big(1 + \sum_{r=1}^\infty h_r(x_1, x_2, \ldots, x_n) t^r \Big)^\alpha \\
&\overset{(2.9)}{=} [t^d] \sum_{\ell=0}^\infty P_\ell^{(\alpha)}(h_1, h_2, \ldots, h_\ell) t^\ell \\
&\overset{(2.10)}{=} [t^d] \sum_{\ell=0}^\infty \Big(\sum_{\substack{\pi \vdash \ell \\ |\pi| \leq \alpha}} c_\pi h_\pi(x_1, x_2, \ldots, x_n) \Big) t^\ell.
\end{aligned}$$

Consequently, $\sum_{\substack{\pi \vdash d \\ |\pi| \leq \alpha}} c_\pi h_\pi(x_1, x_2, \ldots, x_n) = \|A\|_{\boldsymbol{X},d}^d$, which is positive definite. ■

Corollary 2.6 (Hunter [14]) *For even $d \geq 2$, the complete symmetric homogeneous polynomial $h_d(x_1, x_2, \ldots, x_n)$ is positive definite.*

Example 2.7 If $\alpha = 2$, then we obtain the positive definite symmetric polynomial $H_{d,2}(x_1, x_2, \ldots, x_n) = \sum_{i=0}^d h_i(x_1, x_2, \ldots, x_n) h_{d-i}(x_1, x_2, \ldots, x_n)$.

Example 2.8 The relation $\sum_{\ell=0}^\infty H_{\ell,\alpha} t^\ell = (\sum_{\ell=0}^\infty h_\ell t^\ell)(\sum_{\ell=0}^\infty H_{\ell,\alpha-1} t^\ell)$ implies that the sequence $\{H_{d,\alpha}\}_{\alpha \geq 1}$ satisfies the recursion

(2.11) $$H_{d,\alpha} = \sum_{i=0}^d h_i H_{d-i,\alpha-1}.$$

https://doi.org/10.4153/S0008439522000741 Published online by Cambridge University Press

For example, let $j = 4$ and $\alpha = 3$. There are four partitions $\boldsymbol{\pi}$ of j with $|\boldsymbol{\pi}| \le 3$. These are $(1,1,2)$, $(1,3)$, $(2,2)$, and (4). Therefore,

$$\begin{aligned} H_{4,3}(x_1,x_2,x_3,x_4) &= c(1,1,2)h_1^2h_2 + c(1,3)h_1h_3 + c(2,2)h_2^2 + c(4)h_4 \\ &= \frac{3!}{0!2!1!}h_1^2h_2 + \frac{3!}{1!1!1!}h_1h_3 + \frac{3!}{1!2!}h_2^2 + \frac{3!}{2!1!}h_4 \\ &= 3h_1^2h_2 + 6h_1h_3 + 3h_2^2 + 3h_4 \end{aligned}$$

is a positive definite symmetric polynomial. In light of (2.11), we can also write $H_{4,3}(x_1,x_2,x_3,x_4) = \sum_{i=0}^4 h_i H_{4-i,2} = H_{4,2} + h_1H_{3,2} + h_2H_{2,2} + h_3H_{1,2} + h_4$.

2.4 Normal random variables

Let $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$, in which the X_i are independent normal random variables with mean μ and variance $\sigma^2 > 0$. Then $M(t) = \exp(t\mu + \frac{\sigma^2t^2}{2})$ and $K(t) = \frac{\sigma^2t^2}{2} + \mu t$; in particular, $\kappa_1 = \mu$ and $\kappa_2 = \sigma^2$ and all higher cumulants are zero. Then

$$M_{\boldsymbol{X},\boldsymbol{\lambda}}(t) = \prod_{i=1}^n \exp\Big(\lambda_i t\mu + \frac{\sigma^2\lambda_i^2t^2}{2}\Big) = \exp\Big(t\mu \operatorname{tr} A + \frac{\sigma^2 \operatorname{tr}(A^2)t^2}{2}\Big).$$

Theorem 1.1 and the above tell us that

$$\|A\|_{\boldsymbol{X},d}^d = \sum_{k=0}^{\frac{d}{2}} \frac{\mu^{2k}(\operatorname{tr} A)^{2k}}{(2k)!} \cdot \frac{\sigma^{d-2k}\|A\|_{\mathrm{F}}^{d-2k}}{2^{\frac{d}{2}-k}(\frac{d}{2}-k)!} \quad \text{for } A \in \mathrm{H}_n, \tag{2.12}$$

in which $\|A\|_{\mathrm{F}}$ is the Frobenius norm of A. For $d \ge 2$ even, Theorem 1.1 yields

$$\begin{aligned} \|Z\|_{\boldsymbol{X},2}^2 &= \tfrac{1}{2}\sigma^2 \operatorname{tr}(Z^*Z) + \tfrac{1}{2}\mu^2(\operatorname{tr} Z^*)(\operatorname{tr} Z), \\ \|Z\|_{\boldsymbol{X},4}^4 &= \tfrac{1}{24}\big(\mu^4(\operatorname{tr} Z)^2(\operatorname{tr} Z^*)^2 + \mu^2\sigma^2 \operatorname{tr}(Z^*)^2 \operatorname{tr}(Z^2) \\ &\quad + 4\mu^2\sigma^2(\operatorname{tr} Z)(\operatorname{tr} Z^*)(\operatorname{tr} Z^*Z) + 2\sigma^4(\operatorname{tr} Z^*Z)^2 \\ &\quad + \mu^2\sigma^2(\operatorname{tr} Z)^2 \operatorname{tr}(Z^{*2}) + \sigma^4 \operatorname{tr}(Z^2)\operatorname{tr}(Z^{*2})\big). \end{aligned}$$

Since $\kappa_r = 0$ for $r \ge 3$, we see that $\|Z\|_{\boldsymbol{X},4}^4$ does not contain summands like $\operatorname{tr}(Z^*)\operatorname{tr}(Z^*Z^2)$ and $\operatorname{tr}(Z^{*2}Z^2)$, in contrast to the formula in Example 2.3.

2.5 Uniform random variables

Let $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$, where the X_i are independent and uniformly distributed on $[a,b]$. Each X_i has probability density $f(x) = (b-a)^{-1}\mathbb{1}_{[a,b]}$, where $\mathbb{1}_{[a,b]}$ is the indicator function of $[a,b]$. Then

$$\mu_k = \mathbf{E}[X_i^k] = \int_{-\infty}^{\infty} x^k f(x)\,dx = \frac{h_k(a,b)}{k+1}, \tag{2.13}$$

in which $h_k(a,b)$ is the CHS polynomial of degree k in the variables a,b. The moment and cumulant generating functions of each X_i are $M(t) = \frac{e^{bt}-e^{at}}{t(b-a)}$ and $K(t) = \log(\frac{e^{t(b-a)}-1}{t(b-a)}) + at$. The cumulants are

https://doi.org/10.4153/S0008439522000741 Published online by Cambridge University Press

$$\kappa_r = \begin{cases} \frac{a+b}{2}, & \text{if } r = 1, \\ \frac{B_r}{r}(b-a)^r, & \text{if } r \text{ is even}, \\ 0, & \text{otherwise}, \end{cases}$$

in which B_r is the rth Bernoulli number [10]. Theorem 1.1 ensures that

$$\|A\|_{\mathbf{X},d}^d = [t^d] \prod_{i=1}^n \frac{e^{b\lambda_i t} - e^{a\lambda_i t}}{\lambda_i t(b-a)} \quad \text{for } A \in \mathrm{H}_n. \tag{2.14}$$

Example 2.9 If $[a,b] = [-1,1]$, then

$$\|Z\|_{\mathbf{X},4}^4 = \tfrac{1}{1{,}080}\big(10(\operatorname{tr} Z^*Z)^2 + 5\operatorname{tr}(Z^2)\operatorname{tr}(Z^{*2}) - 4(\operatorname{tr} Z^2 Z^{*2}) - 2\operatorname{tr}(ZZ^*ZZ^*)\big)$$

for $Z \in \mathrm{M}_n$, which is not obviously positive, let alone a norm. Indeed, $\operatorname{tr} Z^2 Z^{*2}$ and $\operatorname{tr}(ZZ^*ZZ^*)$ appear with negative scalars in front of them! Similarly,

$$\|A\|_{\mathbf{X},6}^6 = \tfrac{1}{45{,}360}\big(35(\operatorname{tr} A^2)^3 - 42\operatorname{tr}(A^4)\operatorname{tr}(A^2) + 16\operatorname{tr}(A^6)\big) \quad \text{for } A \in \mathrm{H}_6$$

has a nonpositive summand. Since $M_{\mathbf{X},\boldsymbol{\lambda}}(t) = \prod_{i=1}^n \frac{\sinh(\lambda_i t)}{\lambda_i t}$ is an even function of each λ_i, the corresponding norms are polynomials in even powers of the eigenvalues (so positive definiteness is no surprise, although the triangle inequality is nontrivial).

Example 2.10 If $[a,b] = [0,1]$, then $M_{\mathbf{X},\boldsymbol{\lambda}}(t) = \prod_{i=1}^n \frac{e^{\lambda_i t}-1}{\lambda_i t}$, and hence for $A \in \mathrm{H}_n$,

$$\|A\|_{\mathbf{X},2}^2 = \tfrac{1}{12}(2\lambda_1^2 + 3\lambda_1\lambda_2 + 2\lambda_2^2),$$
$$\|A\|_{\mathbf{X},4}^4 = \tfrac{1}{720}(6\lambda_1^4 + 15\lambda_1^3\lambda_2 + 20\lambda_1^2\lambda_2^2 + 15\lambda_1\lambda_2^3 + 6\lambda_2^4).$$

Unlike the previous example, these symmetric polynomials are not obviously positive definite since $\lambda_1^3\lambda_2$ and $\lambda_1\lambda_2^3$ need not be nonnegative.

2.6 Laplace random variables

Let $\mathbf{X} = (X_1, X_2, \ldots, X_n)$, where the X_i are independent random variables distributed according to the probability density $f(x) = \frac{1}{2\beta} e^{-\frac{|x-\mu|}{\beta}}$, in which $\mu \in \mathbb{R}$ and $\beta > 0$. The moment and cumulant generating functions of the X_i are $M(t) = \frac{e^{\mu t}}{1-\beta^2 t^2}$ and $K(t) = \mu t - \log(1-\beta^2 t^2)$, respectively. The cumulants are

$$\kappa_r = \begin{cases} \mu, & \text{if } r = 1, \\ 2\beta^r (r-1)!, & \text{if } r \text{ is even}, \\ 0, & \text{otherwise}. \end{cases}$$

For even $d \geq 2$, it follows that $\|A\|_{\mathbf{X},d}^d$ is the dth term in the Taylor expansion of

$$\|A\|_{\mathbf{X},d}^d = [t^d] \prod_{i=1}^n \frac{e^{\mu t}}{1-\beta^2\lambda_i^2 t^2}. = e^{\mu \operatorname{tr} At}[t^d] \prod_{i=1}^n \frac{1}{1-\beta^2\lambda_i^2 t^2}. \tag{2.15}$$

https://doi.org/10.4153/S0008439522000741 Published online by Cambridge University Press

Example 2.11 Let $\mu = \beta = 1$. Expanding the terms in (2.15) gives

$$M_{\mathbf{X},\boldsymbol{\lambda}}(t) = e^{\operatorname{tr} At} \prod_{i=1}^{n} \frac{1}{1-\lambda_i^2 t^2} = \Big(\sum_{k=0}^{\infty} (\operatorname{tr} A)^k \frac{t^k}{k!} \Big) \Big(\sum_{k=0}^{\infty} h_k(\lambda_1^2, \lambda_2^2, \ldots, \lambda_n^2) t^{2k} \Big),$$

which implies $\|A\|_{\mathbf{X},d}^d = \sum_{k=0}^{d/2} \frac{(\operatorname{tr} A)^{2k}}{(2k)!} h_{\frac{d}{2}-k}(\lambda_1^2, \lambda_2^2, \ldots, \lambda_n^2)$.

2.7 Bernoulli random variables

Let $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$, in which the X_i are independent Bernoulli random variables. Each X_i takes values in $\{0,1\}$ with $\mathbf{P}(X_i = 1) = q$ and $\mathbf{P}(X_i = 0) = 1 - q$ for some fixed $0 < q < 1$. Each X_i satisfies $\mathbf{E}[X_i^k] = \sum_{j\in\{0,1\}} j^k \mathbf{P}(X_i = j) = q$ for $k \in \mathbb{N}$. We have $M(t) = 1 - q + qe^t$ and $K(t) = \log(1 - q + qe^t)$. The first few cumulants are

$$q, \qquad q - q^2, \qquad 2q^3 - 3q^2 + q, \qquad -6q^4 + 12q^3 - 7q^2 + q, \ldots.$$

For even $d \geq 2$, the multinomial theorem and independence imply that

$$\|A\|_{\mathbf{X},d}^d = \frac{1}{d!} \sum_{i_1+i_2+\cdots+i_n=d} q^{|I|} \lambda_1^{i_1} \lambda_2^{i_2} \cdots \lambda_n^{i_n},$$

in which $|I|$ denotes the cardinality of $I = \{i_1, i_2, \ldots, i_n\}$. We can write this as

$$\|A\|_{\mathbf{X},d}^d = \sum_{\boldsymbol{\pi} \vdash d} \frac{|\boldsymbol{\pi}|!}{d!} q^{|\boldsymbol{\pi}|} m_{\boldsymbol{\pi}}(\boldsymbol{\lambda}),$$

in which $m_{\boldsymbol{\pi}}$ denotes the *monomial symmetric polynomial* corresponding to the partition $\boldsymbol{\pi}$ of d [21, Section 7.3]. To be more specific,

$$m_{\boldsymbol{\pi}}(\boldsymbol{x}) = \sum_{\boldsymbol{\alpha}} x^{\boldsymbol{\alpha}},$$

in which the sum is taken over all distinct permutations $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_r)$ of the entries of $\boldsymbol{\pi} = (i_1, i_2, \ldots, i_r)$ and $x^{\boldsymbol{\alpha}} = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_r^{\alpha_r}$. For example, $m_{(1)} = \sum_i x_i$, $m_{(2)} = \sum_i x_i^2$, and $m_{(1,1)} = \sum_{i<j} x_i x_j$.

2.8 Finite discrete random variables

Let X be supported on $\{a_1, a_2, \ldots, a_\ell\} \subset \mathbb{R}$, with $\mathbf{P}(X = a_j) = q_j > 0$ for $1 \leq j \leq \ell$ and $q_1 + q_2 + \cdots + q_\ell = 1$. Then $\mathbf{E}[X^k] = \sum_{i=1}^{\ell} a_i^k q_i$, and hence

$$M(t) = \sum_{j=1}^{\ell} q_j \left(\sum_{k=0}^{\infty} a_j^k \frac{t^k}{k!} \right) = \sum_{j=1}^{\ell} q_j e^{a_j t}. \tag{2.16}$$

Let $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$, in which $X_1, X_2, \ldots, X_n \sim X$ are iid random variables.

Example 2.12 Let $\ell = 2$ and $a_1 = -a_2 = 1$ with $q_1 = q_2 = \frac{1}{2}$. The X_i are *Rademacher* random variables. Identity (2.16) yields $M(t) = \cosh t$, so $M_{\mathbf{X},\boldsymbol{\lambda}}(t) = \prod_{i=1}^n \cosh(\lambda_i t)$. For $n = 2$, we have $\|A\|_{\mathbf{X},2}^2 = \frac{1}{2}(\lambda_1^2 + \lambda_2^2)$, $\|A\|_{\mathbf{X},4}^4 = \frac{1}{24}(\lambda_1^4 + 6\lambda_2^2\lambda_1^2 + \lambda_2^4)$, and

$$\|A\|_{\mathbf{X},6}^6 = \tfrac{1}{720}(\lambda_1^6 + 15\lambda_2^2\lambda_1^4 + 15\lambda_2^4\lambda_1^2 + \lambda_2^6).$$

https://doi.org/10.4153/S0008439522000741 Published online by Cambridge University Press

Let $\gamma_p = \sqrt{2}(\sqrt{\pi})^{-1/p}\Gamma(\frac{p+1}{2})^{1/p}$ denote the pth moment of a standard normal random variable. Let $X_1, X_2, \ldots, X_n$ be independent Rademacher random variables (see Example 2.12). The classic Khintchine inequality asserts that

$$\left(\mathbf{E}\Big|\sum_{i=1}^{n}\lambda_i X_i\Big|^2\right)^{1/2} \leq \left(\mathbf{E}\Big|\sum_{i=1}^{n}\lambda_i X_i\Big|^p\right)^{1/p} \leq a_p\left(\mathbf{E}\Big|\sum_{i=1}^{n}\lambda_i X_i\Big|^2\right)^{1/2}, \tag{2.17}$$

for all $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}$ and $p \geq 2$, with $a_2 = 1$ and $a_p = \gamma_p$ for $p > 2$. Moreover, these constants are optimal [11]. Immediately, we obtain the equivalence of norms

$$\|A\|_{\mathrm{F}} \leq \Gamma(p+1)^{1/p}\|A\|_{\boldsymbol{X},p} \leq a_p\|A\|_{\mathrm{F}}, \tag{2.18}$$

for all $A \in \mathrm{H}_n(\mathbb{C})$ and $p \geq 2$. The proof of Theorem 1.e implies that $\|Z\|_{\mathrm{F}} \leq \Gamma(p+1)^{1/p}\|Z\|_{\boldsymbol{X},p} \leq a_p\|Z\|_{\mathrm{F}}$ for all $Z \in \mathrm{M}_n$ and $p \geq 2$.

In general, suppose that $X_1, X_2, \ldots, X_n$ are iid random variables. A comparison of the form (2.17) is a *Khintchine-type inequality*. Establishing a Khintchine-type inequality here is equivalent to establishing an equivalence of norms as in (2.18). This is always possible since $\mathrm{H}_n(\mathbb{C})$ is finite dimensional. However, establishing Khintchine-type inequalities is, in general, a nontrivial task (see [7, 8, 12, 15]).

2.9 Poisson random variables

Let $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$, in which the X_i are independent random variables on $\{0, 1, 2, \ldots\}$ distributed according to $f(t) = \frac{e^{-\alpha}\alpha^t}{t!}$, in which $\alpha > 0$. The moment and cumulant generating functions of the X_i are $M(t) = e^{\alpha(e^t-1)}$ and $K(t) = \alpha(e^t - 1)$, respectively. Therefore, $\kappa_i = \alpha$ for all $i \in \mathbb{N}$ and hence

$$\|A\|_{\boldsymbol{X},d}^d = \sum_{\pi \vdash d}\frac{\alpha^{|\pi|}p_\pi}{y_\pi}.$$

For example, if $A \in \mathrm{H}_n$ we have

$$4!\|A\|_{\boldsymbol{X},4}^4 = \alpha^4(\operatorname{tr} A)^4 + 6\alpha^3(\operatorname{tr} A)^2\operatorname{tr} A^2 + 4\alpha^2\operatorname{tr} A\operatorname{tr} A^3 + 3\alpha^2(\operatorname{tr} A^2)^2 + \alpha\operatorname{tr} A^4.$$

2.10 Pareto random variables

Let $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$, in which the X_i are independent random variables distributed according to the probability density

$$f(x) = \begin{cases} \frac{\alpha}{x^{\alpha+1}}, & x \geq 1, \\ 0, & x < 1. \end{cases}$$

The moments that exist are $\mu_k = \frac{\alpha}{\alpha-k}$ for $k < \alpha$. For even $d \geq 2$ with $d < \alpha$, the multinomial theorem and independence yield

$$\begin{aligned} d!\|A\|_{\boldsymbol{X},d}^d &= \mathbf{E}[\langle \boldsymbol{X}, \boldsymbol{\lambda}\rangle^d] = \mathbf{E}\big[(\lambda_1 X_1 + \lambda_2 X_2 + \cdots + \lambda_n X_n)^d\big] \\ &= \mathbf{E}\left[\sum_{\substack{k_1+k_2+\cdots+k_n=d \\ k_1,k_2,\ldots,k_n \geq 0}}\binom{d}{k_1, k_2, \ldots, k_n}\prod_{i=1}^{n}(\lambda_i X_i)^{k_i}\right] \end{aligned}$$

https://doi.org/10.4153/S0008439522000741 Published online by Cambridge University Press

$$= \sum_{\substack{k_1+k_2+\cdots+k_n=d \\ k_1,k_2,\ldots,k_n\geq 0}} \binom{d}{k_1, k_2, \ldots, k_n} \prod_{i=1}^{n} \lambda_i^{k_i} \mathbf{E}[X_i^{k_i}]$$

$$= \sum_{\substack{k_1+k_2+\cdots+k_n=d \\ k_1,k_2,\ldots,k_n\geq 0}} \binom{d}{k_1, k_2, \ldots, k_n} \prod_{i=1}^{n} \frac{\alpha\lambda_i^{k_i}}{\alpha - k_i}.$$

In particular, $\lim_{\alpha\to\infty} d!\|A\|_{\mathbf{X}_\alpha,d}^d = (\operatorname{tr} A)^d$ and

$$\lim_{\alpha\to d^+} (\alpha - d)d!\|A\|_{\mathbf{X}_\alpha,d}^d = \lim_{\alpha\to d^+} (\alpha - d) \sum_{\substack{k_1+k_2+\cdots+k_n=d \\ k_1,k_2,\ldots,k_n\geq 0}} \binom{d}{k_1, k_2, \ldots, k_n} \prod_{i=1}^{n} \frac{\alpha\lambda_i^{k_i}}{\alpha - k_i}$$

$$= \lim_{\alpha\to d^+} (\alpha - d) \sum_{i=1}^{n} \binom{d}{d} \frac{d\lambda_i^d}{\alpha - d} = d \sum_{i=1}^{n} \lambda_i^d = d\|A\|_d^d,$$

in which $\|A\|_d$ is the Schatten d-norm on H_n.

Example 2.13 For $n = 2$,

$$\|A\|_{\mathbf{X},2}^2 = \frac{1}{2}\alpha\left(\frac{\lambda_1^2}{\alpha-2} + \frac{2\alpha\lambda_1\lambda_2}{(\alpha-1)^2} + \frac{\lambda_2^2}{\alpha-2}\right) \quad \text{and}$$

$$\|A\|_{\mathbf{X},4}^4 = \frac{1}{24}\alpha\left(\frac{\lambda_1^4}{\alpha-4} + \frac{4\alpha\lambda_1^3\lambda_2}{\alpha^2-4\alpha+3} + \frac{6\alpha\lambda_2^2\lambda_1^2}{(\alpha-2)^2} + \frac{4\alpha\lambda_1\lambda_2^3}{\alpha^2-4\alpha+3} + \frac{\lambda_2^4}{\alpha-4}\right).$$

3 Proof of Theorem 1.1

Let $d \geq 2$ be arbitrary, and let $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ be a random vector in $\mathbb{R}^n$, in which $X_1, X_2, \ldots, X_n \in L^d(\Omega, \mathcal{F}, \mathbf{P})$ are iid random variables. Independence is not needed for (a) (see Remark 3.4). We let $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_n)$ denote the vector of eigenvalues of $A \in \mathrm{H}_n$. As before, A denotes a typical Hermitian matrix and $Z \in \mathrm{M}_n$ an arbitrary square matrix.

The proofs of (a)–(e) of Theorem 1.1 are placed in separate subsections below. Before we proceed, we require an important lemma.

Lemma 3.1 *$\mathbf{X}$ is positive definite.*

Proof Hölder's inequality shows that each $X_i \in L^2(\Omega, \mathcal{F}, \mathbf{P})$, so μ_1 and μ_2 are finite. Jensen's inequality yields $\mu_1^2 \leq \mu_2$; nondegeneracy of the X_i ensures the inequality is strict. Independence implies that $\mathbf{E}[X_iX_j] = \mathbf{E}[X_i]\mathbf{E}[X_j]$ for $i \neq j$, so

$$\Sigma(\mathbf{X}) = [\mathbf{E}X_iX_j] = \begin{bmatrix} \mu_2 & \mu_1^2 & \cdots & \mu_1^2 \\ \mu_1^2 & \mu_2 & \cdots & \mu_1^2 \\ \vdots & \vdots & \ddots & \vdots \\ \mu_1^2 & \mu_1^2 & \cdots & \mu_2 \end{bmatrix} = (\mu_2 - \mu_1^2)I + \mu_1^2 J,$$

in which $\mu_2 - \mu_1^2 > 0$ and J is the all-ones matrix. Thus, $\Sigma(\mathbf{X})$ is the sum of a positive definite and a positive semidefinite matrix, so it is positive definite. ∎

https://doi.org/10.4153/S0008439522000741 Published online by Cambridge University Press

3.1 Proof of Theorem 1.1.a

Since $X_1, X_2, \ldots, X_n \in L^d(\Omega, \mathcal{F}, \mathbf{P})$ for some $d \geq 2$, Hölder's inequality implies the random variable $\Lambda = \langle \boldsymbol{X}, \boldsymbol{\lambda} \rangle$ satisfies

$$\langle \boldsymbol{\lambda}, \Sigma(\boldsymbol{X})\boldsymbol{\lambda} \rangle = \mathbf{E}[|\Lambda|^2] \leq (\mathbf{E}|\Lambda|^d)^{2/d}. \tag{3.1}$$

For $A \in \mathrm{H}_n$, consider the nonnegative function

$$\mathfrak{N}(A) = \left(\frac{\mathbf{E}|\langle \boldsymbol{X}, \boldsymbol{\lambda} \rangle|^d}{\Gamma(d+1)} \right)^{1/d}. \tag{3.2}$$

It is clearly homogeneous: $\mathfrak{N}(\alpha A) = |\alpha| \mathfrak{N}(A)$ for all $\alpha \in \mathbb{R}$. Lemma 3.1 ensures that $\Sigma(\boldsymbol{X})$ is positive definite, so (3.1) implies $\mathfrak{N}(A) = 0$ if and only if $A = 0$.

We must show that $\mathfrak{N}$ satisfies the triangle inequality. Our approach parallels that of [1, Theorem 1]. We first show that $\mathfrak{N}$ satisfies the triangle inequality on $\mathrm{D}_n(\mathbb{R})$, the space of real diagonal matrices. Then, we use Lewis' framework for convex matrix analysis [17] to establish the triangle inequality on H_n.

Let $\mathcal{V}$ be a finite-dimensional real vector space with inner product $\langle \cdot, \cdot \rangle_{\mathcal{V}}$. The adjoint ϕ^* of a linear map $\phi : \mathcal{V} \to \mathcal{V}$ satisfies $\langle \phi^*(A), B \rangle = \langle A, \phi(B) \rangle$ for all $A, B \in \mathcal{V}$. We say that ϕ is *orthogonal* if $\phi^* \circ \phi$ is the identity. Let $\mathrm{O}(\mathcal{V})$ denote the set of orthogonal linear maps on $\mathcal{V}$. If $\mathcal{G} \subset \mathrm{O}(\mathcal{V})$ is a subgroup, then $f : \mathcal{V} \to \mathbb{R}$ is $\mathcal{G}$-*invariant* if $f(\phi(A)) = f(A)$ for all $\phi \in \mathcal{G}$ and $A \in V$.

Definition 3.1 (Definition 2.1 of [17]) $\delta : \mathcal{V} \to \mathcal{V}$ is a *G-invariant normal form* if

(a) δ is $\mathcal{G}$-invariant.
(b) For each $A \in \mathcal{V}$, there is an $\phi \in \mathrm{O}(\mathcal{V})$ such that $A = \phi(\delta(A))$.
(c) $\langle A, B \rangle_{\mathcal{V}} \leq \langle \delta(A), \delta(B) \rangle_{\mathcal{V}}$ for all $A, B \in \mathcal{V}$.

Such a triple $(\mathcal{V}, G, \delta)$ is a *normal decomposition system* (NDS). Let $(\mathcal{V}, \mathcal{G}, \delta)$ be an NDS and $\mathcal{W} \subseteq \mathcal{V}$ a subspace. The *stabilizer* of $\mathcal{W}$ in $\mathcal{G}$ is $\mathcal{G}_{\mathcal{W}} = \{\phi \in \mathcal{G} : \phi(\mathcal{W}) = \mathcal{W}\}$. We restrict the domain of $\phi \in \mathcal{G}_{\mathcal{W}}$ and consider $\mathcal{G}_{\mathcal{W}}$ as a subset of $\mathrm{O}(\mathcal{W})$.

Lemma 3.2 (Theorem 4.3 of [17]) *Let* $(\mathcal{V}, \mathcal{G}, \delta)$ *and* $(\mathcal{W}, \mathcal{G}_{\mathcal{W}}, \delta|_{\mathcal{W}})$ *be NDSs with* $\operatorname{ran} \delta \subset \mathcal{W}$. *Then a* $\mathcal{G}$-*invariant function* $f : \mathcal{V} \to \mathbb{R}$ *is convex if and only if its restriction to* $\mathcal{W}$ *is convex.*

Let $\mathcal{V} = \mathrm{H}_n$ be the $\mathbb{R}$-vector space of complex Hermitian $(A = A^*)$ matrices equipped with the Frobenius inner product $(A, B) \mapsto \operatorname{tr} AB$. Let U_n denote the group of $n \times n$ complex unitary matrices. For $U \in \mathrm{U}_n$, define $\phi_U : \mathcal{V} \to \mathcal{V}$ by $\phi_U(A) = UAU^*$. Then $\mathcal{G} = \{\phi_U : U \in \mathrm{U}_n\}$ is a group under composition. We may regard it is a subgroup of $\mathrm{O}(\mathcal{V})$ since $\phi_U^* = \phi_{U^*}$.

Let $\mathcal{W} = \mathrm{D}_n(\mathbb{R}) \subset \mathcal{V}$ denote the set of real diagonal matrices. Then $\mathcal{G}_{\mathcal{W}} = \{\phi_P : P \in \mathcal{P}_n\}$, in which $\mathcal{P}_n$ is the group of $n \times n$ permutation matrices. Define $\delta : \mathcal{V} \to \mathcal{V}$ by $\delta(A) = \operatorname{diag}(\lambda_1(A), \lambda_2(A), \ldots, \lambda_n(A))$, the $n \times n$ diagonal matrix with $\lambda_1(A), \lambda_2(A), \ldots, \lambda_n(A)$ on its diagonal. Observe that $\operatorname{ran} \delta \subset \mathcal{W}$ since the eigenvalues of a Hermitian matrix are real. We maintain this notation below.

https://doi.org/10.4153/S0008439522000741 Published online by Cambridge University Press

Lemma 3.3 *$(\mathcal{V}, \mathcal{G}, \delta)$ and $(\mathcal{W}, \mathcal{G}_{\mathcal{W}}, \delta|_{\mathcal{W}})$ are NDSs.*

Proof We claim that $(\mathcal{V}, \mathcal{G}, \delta)$ is an NDS. (a) Eigenvalues are similarity invariant, so δ is $\mathcal{G}$-invariant. (b) For $A \in \mathcal{V}$, the spectral theorem gives a $U \in \mathrm{U}_n$ such that $A = U\delta(A)U^* = \phi_U(\delta(A))$. (c) For $A, B \in \mathcal{V}$, note that $\operatorname{tr} AB \le \operatorname{tr} \delta(A)\delta(B)$ [16, Theorem 2.2] (see [1, Remark 10] for further references).

We claim that $(\mathcal{W}, \mathcal{G}_{\mathcal{W}}, \delta|_{\mathcal{W}})$ is an NDS. (a) $\delta|_{\mathcal{W}}$ is $\mathcal{G}_{\mathcal{W}}$-invariant since $\delta(\phi_P(A)) = \delta(PAP^*) = \delta(A)$ for all $A \in \mathcal{W}$ and $P \in \mathcal{P}_n$. (b) If $A \in \mathcal{W}$, then there is a $P \in \mathcal{P}_n$ such that $A = P\delta(A)P^* = \phi_P(\delta(A))$. (c) The diagonal elements of a diagonal matrix are its eigenvalues. Thus, this property is inherited from $\mathcal{V}$. ■

The function $\mathfrak{N} : \mathcal{V} \to \mathbb{R}$ is $\mathcal{G}$-invariant since it is a symmetric function of $\lambda_1(A), \lambda_2(A), \ldots, \lambda_n(A)$ (see Remark 3.4). If $A, B \in \mathcal{W}$, define random variables $X = \langle \boldsymbol{X}, \boldsymbol{\lambda}(A)\rangle$ and $Y = \langle \boldsymbol{X}, \boldsymbol{\lambda}(B)\rangle$. Since A and B are diagonal, $\boldsymbol{\lambda}(A+B) = \boldsymbol{\lambda}(A) + \boldsymbol{\lambda}(B)$ and hence Minkowski's inequality for $L^d(\Omega, \mathcal{F}, \mathbf{P})$ yields

$$\left(\mathbf{E}\left|\langle \boldsymbol{X}, \boldsymbol{\lambda}(A+B)\rangle\right|^d\right)^{1/d} = \left(\mathbf{E}|X+Y|^d\right)^{1/d} \le \left(\mathbf{E}|X|^d\right)^{1/d} + \left(\mathbf{E}|Y|^d\right)^{1/d}.$$

Thus, $\mathfrak{N}(A+B) \le \mathfrak{N}(A) + \mathfrak{N}(B)$ for all $A, B \in \mathcal{W}$, and hence $\mathfrak{N}$ is convex on $\mathcal{W}$. Lemma 3.2 implies that $\mathfrak{N}$ is convex on $\mathcal{V}$. Therefore, $\frac{1}{2}\mathfrak{N}(A+B) = \mathfrak{N}(\frac{1}{2}A + \frac{1}{2}B) \le \frac{1}{2}\mathfrak{N}(A) + \frac{1}{2}\mathfrak{N}(B)$ for all $A, B \in \mathcal{V}$, so (3.2) defines a norm on $\mathcal{V} = \mathrm{H}_n$.

Remark 3.4 Independence is not used in the proof of (a). Our proof only requires that the function $\|A\|_{\boldsymbol{X},d}$ be invariant with respect to unitary conjugation. If the X_i are assumed to be iid, but not necessarily independent, then $\|A\|_{\boldsymbol{X},d}$ is a homogeneous symmetric function of the eigenvalues of A. Any such function is invariant with respect to unitary conjugation.

3.2 Proof of Theorem 1.1.b

Let $d \ge 2$ be even, and let $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ be a random vector, in which $X_1, X_2, \ldots, X_n$ are iid random variables which admit a moment generating function $M(t)$. Let $A \in \mathrm{H}_n$ have eigenvalues $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n$. If $\Lambda = \langle \boldsymbol{X}, \boldsymbol{\lambda}\rangle = \lambda_1 X_1 + \lambda_2 X_2 + \cdots + \lambda_n X_n$, then independence ensures that $M_\Lambda(t) = \prod_{i=1}^n M(\lambda_i t)$. Thus, $\|A\|_{\boldsymbol{X},d}^d = \mathbf{E}[\Lambda^d]/d! = [t^d]M_\Lambda(t)$.

3.3 Proof of Theorem 1.1.c

Maintain the same notation as in the proof of (b). However, we only assume existence of the first d moments of the X_i. In this case, $M_\Lambda(t)$ is a formal series with $\kappa_1, \kappa_2, \ldots, \kappa_d$ determined and the remaining cumulants treated as formal variables. Then

$$\begin{aligned} M_\Lambda(t) &= \prod_{i=1}^n M(\lambda_i t) \overset{(1.2)}{=} \exp\left(\sum_{i=1}^n K(\lambda_i t)\right) \overset{(1.2)}{=} \exp\left(\sum_{j=1}^\infty \kappa_j(\lambda_1^j + \lambda_2^j + \cdots + \lambda_n^j)\frac{t^j}{j!}\right) \\ &= \exp\left(\sum_{j=1}^\infty \kappa_j \operatorname{tr}(A^j)\frac{t^j}{j!}\right) \overset{(1.3)}{=} \sum_{\ell=0}^\infty B_\ell(\kappa_1 \operatorname{tr} A, \kappa_2 \operatorname{tr} A^2, \ldots, \kappa_\ell \operatorname{tr} A^\ell)\frac{t^\ell}{\ell!}. \end{aligned}$$

Expanding the right side of (1.3) yields

$$B_\ell(x_1, x_2, \ldots, x_\ell) = \ell! \sum_{\substack{j_1, j_2, \ldots, j_\ell \geq 0 \\ j_1 + 2j_2 + \cdots + \ell j_\ell = \ell}} \prod_{r=1}^{\ell} \frac{x_r^{j_r}}{(r!)^{j_r} j_r!} = \ell! \sum_{\pi \vdash \ell} \frac{x_\pi}{y_\pi}, \tag{3.3}$$

in which $x_\pi = x_{i_1} x_{i_2} \cdots x_{i_j}$ for a each partition $\boldsymbol{\pi} = (i_1, i_2, \ldots, i_j)$ of ℓ. Substitute $x_i = \kappa_i \operatorname{tr}(A^i)$ above and obtain

$$d! \|A\|_{\mathbf{X},d}^d = [t^d] M_\Lambda(t) = B_d(\kappa_1 \operatorname{tr} A, \kappa_2 \operatorname{tr} A^2, \ldots, \kappa_d \operatorname{tr} A^d).$$

Finally, (3.3) and the above ensure that $\|A\|_{\mathbf{X},d}^d = \sum_{\pi \vdash d} \frac{\kappa_\pi p_\pi}{y_\pi}$ for $A \in \mathrm{H}_n$.

3.4 Proof of Theorem 1.1.d

Recall that a convex function $f : \mathbb{R}^n \to \mathbb{R}$ is Schur convex if and only if it is symmetric [18, p. 258]. Suppose that $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$. Let $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ be a random vector, in which $X_1, X_2, \ldots, X_n \in L^d(\Omega, \mathcal{F}, \mathbf{P})$ are iid. Define random variables $X = \langle \boldsymbol{X}, \boldsymbol{x} \rangle$ and $Y = \langle \boldsymbol{X}, \boldsymbol{y} \rangle$.

Define $\mathfrak{N} : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ by $\mathfrak{N}(\boldsymbol{x}) = \left(\frac{\mathbf{E}|\langle \boldsymbol{X}, \boldsymbol{x} \rangle|^d}{\Gamma(d+1)} \right)^{1/d}$. This function satisfies

$$\mathfrak{N}(\boldsymbol{x} + \boldsymbol{y}) = \left(\frac{\mathbf{E}|\langle \boldsymbol{X}, \boldsymbol{x} + \boldsymbol{y} \rangle|^d}{\Gamma(d+1)} \right)^{1/d} = \left(\frac{\mathbf{E}|X + Y|^d}{\Gamma(d+1)} \right)^{1/d} \leq \mathfrak{N}(\boldsymbol{x}) + \mathfrak{N}(\boldsymbol{y})$$

as seen in the proof of Theorem 1.1.a. Homogeneity implies that $\mathfrak{N}$ is convex on $\mathbb{R}^n$. Finally, $\mathfrak{N}$ is symmetric since $X_1, X_2, \ldots, X_n$ are iid. It follows that $\mathfrak{N}$ is Schur convex. Thus, $\boldsymbol{\lambda}(A) \mapsto \mathfrak{N}(\lambda_1, \lambda_2, \ldots, \lambda_n) = \|A\|_{X,d}$ is Schur convex.

Remark 3.5 Note that independence is not required in the previous argument.

3.5 Proof of Theorem 1.1.e

The initial details parallel those of [1, Theorem 3]. Let $\mathcal{V}$ be a $\mathbb{C}$-vector space with conjugate-linear involution $*$ and suppose that the real-linear subspace $\mathcal{V}_\mathbb{R} = \{v \in \mathcal{V} : v = v^*\}$ of $*$-fixed points has the norm $\|\cdot\|$. Then $e^{it}v + e^{-it}v^* \in \mathcal{V}_\mathbb{R}$ for each $v \in \mathcal{V}$ and $t \in \mathbb{R}$, and $t \mapsto \|e^{it}v + e^{-it}v^*\|$ is continuous for each $v \in \mathcal{V}$. The following is [1, Proposition 15].

Lemma 3.6 *For even $d \geq 2$, the following is a norm on $\mathcal{V}$ that extends $\|\cdot\|$:*

$$\mathfrak{N}_d(v) = \left(\frac{1}{2\pi \binom{d}{d/2}} \int_0^{2\pi} \|e^{it}v + e^{-it}v^*\|^d \, dt \right)^{1/d}. \tag{3.4}$$

Let $\langle x, x^* \rangle$ be the free monoid generated by x and x^*. Let $|w|$ denote the length of a word $w \in \langle x, x^* \rangle$, and let $|w|_x$ count the occurrences of x in w. For $Z \in \mathrm{M}_n$, let $w(Z) \in \mathrm{M}_n$ be the natural evaluation of w at Z. For example, if $w = xx^*x^2$, then $|w| = 4$, $|w|_x = 3$, and $w(Z) = ZZ^*Z^2$. The next lemma is [1, Lemma 16].

https://doi.org/10.4153/S0008439522000741 Published online by Cambridge University Press

Lemma 3.7 *Let $d \geq 2$ be even and $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_r)$ be a partition of d. For $Z \in \mathrm{M}_n$,*

$$\begin{aligned} &\frac{1}{2\pi}\int_0^{2\pi} \operatorname{tr}(e^{it}Z + e^{-it}Z^*)^{\pi_1} \cdots \operatorname{tr}(e^{it}Z + e^{-it}Z^*)^{\pi_r}\, dt \\ &\qquad = \sum_{\substack{w_1,\ldots,w_r \in \langle x, x^* \rangle: \\ |w_j| = \pi_j \ \forall j \\ |w_1 \cdots w_r|_x = \frac{d}{2}}} \operatorname{tr} w_1(Z) \cdots \operatorname{tr} w_r(Z). \end{aligned} \tag{3.5}$$

Given a partition $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_r)$ of d and $Z \in \mathrm{M}_n$, let

$$\mathrm{T}_{\boldsymbol{\pi}}(Z) = \frac{1}{\binom{d}{d/2}} \sum_{\substack{w_1,\ldots,w_r \in \langle x, x^* \rangle: \\ |w_j| = \pi_j \ \forall j \\ |w_1 \cdots w_r|_x = \frac{d}{2}}} \operatorname{tr} w_1(Z) \cdots \operatorname{tr} w_r(Z), \tag{3.6}$$

that is, $\mathrm{T}_{\boldsymbol{\pi}}(Z)$ is $1/\binom{d}{d/2}$ times the sum over the $\binom{d}{d/2}$ possible locations to place $d/2$ adjoints * among the d copies of Z in $(\operatorname{tr} \underbrace{ZZ\cdots Z}_{\pi_1})(\operatorname{tr} \underbrace{ZZ\cdots Z}_{\pi_2})\cdots(\operatorname{tr} \underbrace{ZZ\cdots Z}_{\pi_r})$.

Consider the conjugate transpose $*$ on $\mathcal{V} = \mathrm{M}_n$. The corresponding real subspace of $*$-fixed points is $\mathcal{V}_{\mathbb{R}} = \mathrm{H}_n$. Apply Proposition 3.6 to the norm $\|\cdot\|_d$ on H_n and obtain the extension $\mathfrak{N}_d(\cdot)$ to M_n defined by (3.4).

If $Z \in \mathrm{M}_n$ and $\mathfrak{N}_d(A) = \|A\|_d$ is the norm for $A \in \mathrm{H}_n$, then Proposition 3.6 ensures that the following is a norm on M_n:

$$\begin{aligned} \mathfrak{N}_d(Z) &\overset{(3.4)}{=} \left(\frac{1}{2\pi\binom{d}{d/2}} \int_0^{2\pi} \|e^{it}Z + e^{-it}Z\|_{\mathbf{X},d}^d \, dt \right)^{1/d} \\ &\overset{(1.10)}{=} \left(\frac{1}{2\pi\binom{d}{d/2}} \int_0^{2\pi} \sum_{\boldsymbol{\pi} \vdash d} \frac{\kappa_{\boldsymbol{\pi}} p_{\boldsymbol{\pi}}(\boldsymbol{\lambda}(e^{it}Z + e^{-it}Z^*))}{y_{\boldsymbol{\pi}}} \, dt \right)^{1/d} \\ &\overset{(1.6)}{=} \left(\frac{1}{\binom{d}{d/2}} \sum_{\boldsymbol{\pi} \vdash d} \frac{\kappa_{\boldsymbol{\pi}}}{y_{\boldsymbol{\pi}}} \cdot \frac{1}{2\pi} \int_0^{2\pi} \operatorname{tr}(e^{it}Z + e^{-it}Z^*)^{\pi_1} \cdots \operatorname{tr}(e^{it}Z + e^{-it}Z^*)^{\pi_r} \, dt \right)^{1/d} \\ &\overset{(3.5)}{=} \left(\frac{1}{\binom{d}{d/2}} \sum_{\boldsymbol{\pi} \vdash d} \frac{\kappa_{\boldsymbol{\pi}}}{y_{\boldsymbol{\pi}}} \sum_{\substack{w_1,\ldots,w_r \in \langle x, x^* \rangle: \\ |w_j| = \pi_j \ \forall j \\ |w_1 \cdots w_r|_x = \frac{d}{2}}} \operatorname{tr} w_1(Z) \cdots \operatorname{tr} w_r(Z) \right)^{1/d} \\ &\overset{(3.6)}{=} \left(\sum_{\boldsymbol{\pi} \vdash d} \frac{\kappa_{\boldsymbol{\pi}} \mathrm{T}_{\boldsymbol{\pi}}(Z)}{y_{\boldsymbol{\pi}}} \right)^{1/d}. \end{aligned}$$

4 Open questions

If $\|\cdot\|$ is a norm on M_n, then there is a scalar multiple of it (which may depend upon n) that is submultiplicative. One wonders which of the norms $\|\cdot\|_{\mathbf{X},d}$ are submultiplicative, or perhaps are when multiplied by a constant independent of n.

https://doi.org/10.4153/S0008439522000741 Published online by Cambridge University Press

For example, (2.1) ensures that for $d = 2$, a mean-zero distribution leads to a multiple of the Frobenius norm. If $\mu_2 = 2$, then the norm is submultiplicative.

Problem 1 Characterize those $\boldsymbol{X}$ that give rise to submultiplicative norms.

For the standard exponential distribution, [1, Theorem 31] provides an answer to the next question. An answer to the question in the general setting eludes us.

Problem 2 Characterize the norms $\|\cdot\|_{X,d}$ that arise from an inner product.

Several other unsolved questions come to mind.

Problem 3 Identify the extreme points with respect to random vector norms.

Problem 4 Characterize norms on M_n or H_n that arise from random vectors.

Acknowledgment We thank Bruce Sagan for remarks about symmetric functions.

References

[1] K. Aguilar, Á. Chávez, S. R. Garcia, and J. Volčič, *Norms on complex matrices induced by complete homogeneous symmetric polynomials*. Bull. Lond. Math. Soc. **54**(2022), 2078–2100. https://doi.org/10.1112/blms.12679
[2] A. I. Barvinok, *Low rank approximations of symmetric polynomials and asymptotic counting of contingency tables*. Preprint, 2005. arXiv:0503170
[3] V. J. Baston, *Two inequalities for the complete symmetric functions*. Math. Proc. Cambridge Philos. Soc. **84**(1978), no. 1, 1–3.
[4] E. T. Bell, *Exponential polynomials*. Ann. of Math. (2) **35**(1934), no. 2, 258–277.
[5] P. Billingsley, *Probability and measure*, Wiley Series in Probability and Statistics, John Wiley & Sons, Inc., Hoboken, NJ, 2012, Anniversary edition [of MR1324786], with a foreword by Steve Lalley and a brief biography of Billingsley by Steve Koppes.
[6] A. Böttcher, S. R. Garcia, M. Omar, and C. O'Neill, *Weighted means of B-splines, positivity of divided differences, and complete homogeneous symmetric polynomials*. Linear Algebra Appl. **608**(2021), 68–83.
[7] A. Eskenazis, P. Nayar, and T. Tkocz, *Gaussian mixtures: Entropy and geometric inequalities*. Ann. Probab. **46**(2018), no. 5, 2908–2945.
[8] A. Eskenazis, P. Nayar, and T. Tkocz, *Sharp comparison of moments and the log-concave moment problem*. Adv. Math. **334**(2018), 389–416.
[9] S. R. Garcia, M. Omar, C. O'Neill, and S. Yih, *Factorization length distribution for affine semigroups II: asymptotic behavior for numerical semigroups with arbitrarily many generators*. J. Combin. Theory Ser. A **178**(2021), Article no. 105358, 34 pp.
[10] H. W. Gould, *Explicit formulas for Bernoulli numbers*. Amer. Math. Monthly **79**(1972), 44–51.
[11] U. Haagerup, *The best constants in the Khintchine inequality*. Stud. Math. **70**(1981), no. 3, 231–283 (1982).
[12] A. Havrilla and T. Tkocz, *Sharp Khinchin-type inequalities for symmetric discrete uniform random variables*. Israel J. Math. **246**(2021), no. 1, 281–297.
[13] R. A. Horn and C. R. Johnson, *Matrix analysis*, 2nd ed., Cambridge University Press, Cambridge, 2013.
[14] D. B. Hunter, *The positive-definiteness of the complete symmetric functions of even order*. Math. Proc. Cambridge Philos. Soc. **82**(1977), no. 2, 255–258.
[15] R. Latała and K. Oleszkiewicz, *A note on sums of independent uniformly distributed random variables*. Colloq. Math. **68**(1995), no. 2, 197–206.
[16] A. S. Lewis, *Convex analysis on the Hermitian matrices*. SIAM J. Optim. **6**(1996), no. 1, 164–177.

[17] A. S. Lewis, *Group invariance and convex matrix analysis*. SIAM J. Matrix Anal. Appl. **17**(1996), no. 4, 927–949.
[18] A. W. Roberts and D. E. Varberg, *Convex functions*, Pure and Applied Mathematics, 57, Academic Press [Harcourt Brace Jovanovich], New York–London, 1973.
[19] I. Rovenţa and L. E. Temereancă, *A note on the positivity of the even degree complete homogeneous symmetric polynomials*. Mediterr. J. Math. **16**(2019), no. 1, Article no. 1, 16 pp.
[20] R. P. Stanley, *Enumerative combinatorics*. Vol. 1, Cambridge Studies in Advanced Mathematics, 49, Cambridge University Press, Cambridge, 1997, with a foreword by Gian-Carlo Rota, Corrected reprint of the 1986 original.
[21] R. P. Stanley, *Enumerative combinatorics.* Vol. 2, Cambridge Studies in Advanced Mathematics, 62, Cambridge University Press, Cambridge, 1999, with a foreword by Gian-Carlo Rota and Appendix 1 by Sergey Fomin.
[22] T. Tao, *Schur convexity and positive definiteness of the even degree complete homogeneous symmetric polynomials*, https://terrytao.wordpress.com/2017/08/06/

Department of Mathematics and Statistics, Pomona College, 610 North College Avenue, Claremont, CA 91711, USA
e-mail: angel.chavez@pomona.edu stephan.garcia@pomona.edu jacksonwhurley@gmail.com

https://doi.org/10.4153/S0008439522000741 Published online by Cambridge University Press