Canad. Math. Bull. 2023, pp. 1–11
http://dx.doi.org/10.4153/S0008439523000875

Norms on complex matrices induced by random vectors II: extension of weakly unitarily invariant norms

Ángel Chávez, Stephan Ramon Garcia, and Jackson Hurley

Abstract. We improve and expand in two directions the theory of norms on complex matrices induced by random vectors. We first provide a simple proof of the classification of weakly unitarily invariant norms on the Hermitian matrices. We use this to extend the main theorem in Chávez, Garcia, and Hurley (2023, *Canadian Mathematical Bulletin* 66, 808–826) from exponent $d \geq 2$ to $d \geq 1$. Our proofs are much simpler than the originals: they do not require Lewis' framework for group invariance in convex matrix analysis. This clarification puts the entire theory on simpler foundations while extending its range of applicability.

1 Introduction

A norm $\|\cdot\|$ on M_n, the space of $n \times n$ complex matrices, is *unitarily invariant* if $\|UAV\| = \|A\|$ for all $A \in \mathrm{M}_n$ and unitary $U, V \in \mathrm{M}_n$. A norm on $\mathbb{R}^n$ which is invariant under entrywise sign changes and permutations is a *symmetric gauge function*. A theorem of von Neumann asserts that any unitarily invariant norm on M_n is a symmetric gauge function applied to the singular values [10, Theorem 7.4.7.2]. For example, the Schatten norms are unitarily invariant and defined for $d \geq 1$ by

$$\|A\|_{S_d} = \left(|\sigma_1|^d + |\sigma_2|^d + \cdots + |\sigma_n|^d\right)^{1/d},$$

in which $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$ are the singular values of $A \in \mathrm{M}_n$.

A norm $\|\cdot\|$ on the $\mathbb{R}$-vector space H_n of $n \times n$ complex Hermitian matrices is *weakly unitarily invariant* if $\|U^*AU\| = \|A\|$ for all $A \in \mathrm{H}_n$ and unitary $U \in \mathrm{M}_n$. For example, the numerical radius

$$r(A) = \sup_{\mathbf{x} \in \mathbb{C}^n \setminus \{\mathbf{0}\}} \frac{\langle A\mathbf{x}, \mathbf{x}\rangle}{\langle \mathbf{x}, \mathbf{x}\rangle}$$

is a weakly unitarily invariant norm on H_n [12]. Lewis proved that any weakly unitarily invariant norm on H_n is a symmetric vector norm applied to the eigenvalues [11, Section 8].

Received by the editors October 12, 2023; revised October 25, 2023; accepted October 30, 2023.
Published online on Cambridge Core November 6, 2023.
S.R.G. was partially supported by the NSF (Grant No. DMS-2054002)
AMS subject classification: **47A30, 15A60, 16R30.**
Keywords: Norm, symmetric polynomial, trace, probability distribution, unitary invariance.

https://doi.org/10.4153/S0008439523000875 Published online by Cambridge University Press

Our first result is a short proof of Lewis' theorem that avoids his theory of group invariance in convex matrix analysis [11], the wonderful but complicated framework that underpins [1, 7]. Our new approach uses more standard techniques, such as Birkhoff's theorem on doubly stochastic matrices [6].

Theorem 1.1 *A norm $\|\cdot\|$ on H_n is weakly unitarily invariant if and only if there is a symmetric norm $f:\mathbb{R}^n\to\mathbb{R}$ such that $\|A\| = f(\lambda_1,\lambda_2,\ldots,\lambda_n)$ for all $A\in\mathrm{H}_n$. Here, $\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_n$ are the eigenvalues of A.*

The *random-vector norms* of the next theorem are weakly unitarily invariant norms on H_n that extend to weakly unitarily invariant norms on M_n (see Theorem 1.3). They appeared in [7], and they generalize the complete homogeneous symmetric polynomial norms of [1, Theorem 1]. The original proof of [7, Theorem 1.1(a)] requires $d\geq 2$ and relies heavily on Lewis' framework for group invariance in convex matrix analysis [11]. However, Theorem 1.2 now follows directly from Theorem 1.1. Moreover, Theorem 1.2 generalizes [7, Theorem 1.1(a)] to the case $d\geq 1$.

Theorem 1.2 *Let $d\geq 1$ be real and $\mathbf{X}$ be an independent and identically distributed (iid) random vector in $\mathbb{R}^n$, that is, the entries of $\mathbf{X}=(X_1,X_2,\ldots,X_n)$ are nondegenerate iid random variables. Then*

$$||A||_{\mathbf{X},d} = \left(\frac{\mathbb{E}|\langle\mathbf{X},\boldsymbol{\lambda}\rangle|^d}{\Gamma(d+1)}\right)^{1/d} \tag{1.1}$$

is a weakly unitarily invariant norm on H_n. Here, $\Gamma(\cdot)$ denotes the gamma function and $\boldsymbol{\lambda}=(\lambda_1,\lambda_2,\ldots,\lambda_n)$ denotes the vector of eigenvalues $\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_n$ of A. Moreover, if the entries of $\mathbf{X}$ each have at least m moments, then for all $A\in\mathrm{H}_n$ the function $f:[1,m]\to\mathbb{R}$ defined by $f(d)=\|A\|_{\mathbf{X},d}$ is continuous.

The simplified proof of Theorem 1.1 and the extension of Theorem 1.2 from $d\geq 2$ to $d\geq 1$ permit the main results of [7], restated below as Theorem 1.3, to rest on simpler foundations while enjoying a wider range of applicability. The many perspectives offered in Theorem 1.3 explain the normalization in (1.1).

Theorem 1.3 *Let $\mathbf{X}=(X_1,X_2,\ldots,X_n)$, in which $X_1,X_2,\ldots,X_n\in L^d(\Omega,\mathcal{F},\mathbb{P})$ are nondegenerate iid random variables. Let $\boldsymbol{\lambda}=(\lambda_1,\lambda_2,\ldots,\lambda_n)$ denote the vector of eigenvalues $\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_n$ of $A\in\mathrm{H}_n$.*

(1) *For real $d\geq 1$, $\|A\|_{\mathbf{X},d} = \left(\frac{\mathbb{E}|\langle\mathbf{X},\boldsymbol{\lambda}\rangle|^d}{\Gamma(d+1)}\right)^{1/d}$ is a norm on H_n (now by Theorem 1.2).*

(2) *If the X_i admit a moment generating function $M(t)=\mathbb{E}[e^{tX}]=\sum_{k=0}^{\infty}\mathbb{E}[X^k]\frac{t^k}{k!}$ and $d\geq 2$ is an even integer, then $\|A\|_{\mathbf{X},d}^d$ is the coefficient of t^d in $M_\Lambda(t)$ for all $A\in\mathrm{H}_n$, in which $M_\Lambda(t)=\prod_{i=1}^n M(\lambda_i t)$ is the moment generating function for the random variable $\Lambda=\langle\mathbf{X},\boldsymbol{\lambda}(A)\rangle=\lambda_1X_1+\lambda_2X_2+\cdots+\lambda_nX_n$. In particular, $\|A\|_{\mathbf{X},d}$ is a positive definite, homogeneous, symmetric polynomial in the eigenvalues of A.*

https://doi.org/10.4153/S0008439523000875 Published online by Cambridge University Press

(3) *Let $d \geq 2$ be an even integer. If the first d moments of X_i exist, then*

$$\|A\|_{\mathbf{X},d}^d = \frac{1}{d!} B_d(\kappa_1 \operatorname{tr} A, \kappa_2 \operatorname{tr} A^2, \ldots, \kappa_d \operatorname{tr} A^d) = \sum_{\boldsymbol{\pi} \vdash d} \frac{\kappa_{\boldsymbol{\pi}} p_{\boldsymbol{\pi}}(\boldsymbol{\lambda})}{y_{\boldsymbol{\pi}}} \quad \text{for } A \in \mathrm{H}_n,$$

in which:

(a) $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_r) \in \mathbb{N}^r$ *is a partition of d; that is,* $\pi_1 \geq \pi_2 \geq \cdots \geq \pi_r$ *and* $\pi_1 + \pi_2 + \cdots + \pi_r = d$ *[13, Section 1.7]; we denote this* $\boldsymbol{\pi} \vdash d$;

(b) $p_{\boldsymbol{\pi}}(x_1, x_2, \ldots, x_n) = p_{\pi_1} p_{\pi_2} \cdots p_{\pi_r}$, *in which* $p_k(x_1, x_2, \ldots, x_n) = x_1^k + x_2^k + \cdots + x_n^k$ *is a power-sum symmetric polynomial;*

(c) B_d *is a complete Bell polynomial, defined by* $\sum_{\ell=0}^{\infty} B_\ell(x_1, x_2, \ldots, x_\ell) \frac{t^\ell}{\ell!} = \exp(\sum_{j=1}^{\infty} x_j \frac{t^j}{j!})$ *[2, Section II];*

(d) *The cumulants* $\kappa_1, \kappa_2, \ldots, \kappa_d$ *are defined by the recursion* $\mu_r = \sum_{\ell=0}^{r-1} \binom{r-1}{\ell} \mu_\ell \kappa_{r-\ell}$ *for* $1 \leq r \leq d$, *in which* $\mu_r = \mathbb{E}[X_1^r]$ *is the rth moment of* X_1 *[5, Section 9]; and*

(e) $\kappa_{\boldsymbol{\pi}} = \kappa_{\pi_1} \kappa_{\pi_2} \cdots \kappa_{\pi_r}$ *and* $y_{\boldsymbol{\pi}} = \prod_{i \geq 1} (i!)^{m_i} m_i!$, *in which* $m_i = m_i(\boldsymbol{\pi})$ *is the multiplicity of i in* $\boldsymbol{\pi}$.

(4) *For real $d \geq 1$, the function* $\boldsymbol{\lambda}(A) \mapsto \|A\|_{\mathbf{X},d}$ *is Schur convex; that is, it respects majorization* $\prec$ *(see (3.1)).*

(5) *Let $d \geq 2$ be an even integer. 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 * 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*

$$\|Z\|_{\mathbf{X},d} = \left(\sum_{\boldsymbol{\pi} \vdash d} \frac{\kappa_{\boldsymbol{\pi}} \mathrm{T}_{\boldsymbol{\pi}}(Z)}{y_{\boldsymbol{\pi}}} \right)^{1/d} \tag{1.2}$$

is a norm on M_n *that restricts to the norm on* H_n *above. In particular,* $\|Z\|_{\mathbf{X},d}^d$ *is a positive definite trace polynomial in Z and Z^*.*

The paper is structured as follows. Section 2 provides several examples afforded by the theorems above. The proofs of Theorems 1.1 and 1.2 appear in Sections 3 and 4, respectively. Section 5 concludes with some brief remarks.

2 Examples

The norm $\| \cdot \|_{\mathbf{X},d}$ defined in (1.1) is determined by its unit ball. This provides one way to visualize the properties of random vector norms. We consider a few examples here and refer the reader to [7, Section 2] for further examples and details.

2.1 Normal random variables

Suppose $d \geq 2$ is an even integer and $\mathbf{X}$ is a random vector whose entries are independent normal random variables with mean μ and variance σ^2. The example

https://doi.org/10.4153/S0008439523000875 Published online by Cambridge University Press

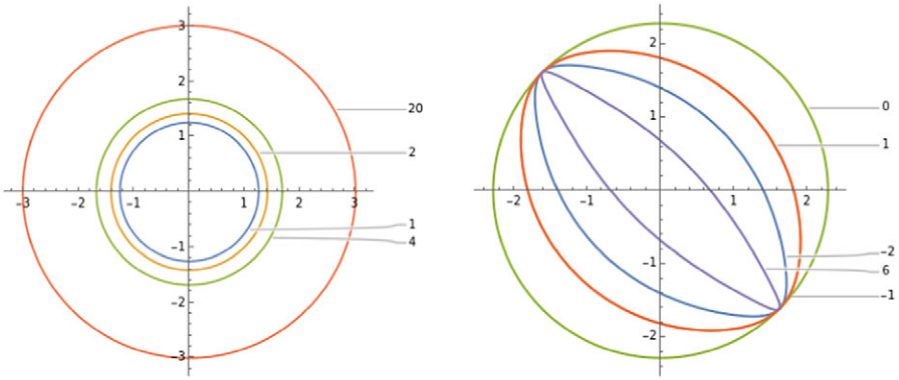

Figure 1: (Left) Unit circles for $\|\cdot\|_{\mathbf{X},d}$ with $d = 1, 2, 4, 20$, in which X_1 and X_2 are standard normal random variables. (Right) Unit circles for $\|\cdot\|_{\mathbf{X},10}$, in which X_1 and X_2 are normal random variables with means $\mu = -2, -1, 0, 1, 6$ and variance $\sigma^2 = 1$.

in [7, equation (2.12)] illustrates

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

in which $\|\cdot\|_{\mathrm{F}}$ is the Frobenius norm. For $d = 2$, the extension to M_n guaranteed by Theorem 1.3 is $\|Z\|_{\mathbf{X},2}^2 = \frac{1}{2}\sigma^2 \operatorname{tr}(Z^*Z) + \frac{1}{2}\mu^2(\operatorname{tr} Z^*)(\operatorname{tr} Z)$ [7, p. 816].

Now, let $n = 2$. If $\mu = 0$, the restrictions of $\|\cdot\|_{\mathbf{X},d}$ to $\mathbb{R}^2$ (whose elements are identified with diagonal matrices) reproduce multiples of the Euclidean norm. If $\mu \neq 0$, then the unit circles for $\|\cdot\|_{\mathbf{X},d}$ are approximately elliptical (see Figure 1).

2.2 Standard exponential random variables

If $d \geq 2$ is an even integer and $\mathbf{X}$ is a random vector whose entries are independent standard exponential random variables, then $\|A\|_{\mathbf{X},d}^d$ equals the *complete homogeneous symmetric polynomial* $h_d(\lambda_1, \lambda_2, \ldots, \lambda_n) = \sum_{1 \leq k_1 \leq \cdots \leq k_d \leq n} \lambda_{k_1}\lambda_{k_2}\cdots\lambda_{k_d}$ in the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ [1]. For $d = 4$, the extension to M_n guaranteed by Theorem 1.3 is [1, equation (9)]

$$\begin{aligned}\|Z\|_4^4 = \frac{1}{24}\big(&(\operatorname{tr} Z)^2 \operatorname{tr}(Z^*)^2 + \operatorname{tr}(Z^*)^2 \operatorname{tr}(Z^2) + 4\operatorname{tr}(Z)\operatorname{tr}(Z^*)\operatorname{tr}(Z^*Z)\\ &+ 2\operatorname{tr}(Z^*Z)^2 + (\operatorname{tr} Z)^2 \operatorname{tr}(Z^{*2}) + \operatorname{tr}(Z^2)\operatorname{tr}(Z^{*2}) + 4\operatorname{tr}(Z^*)\operatorname{tr}(Z^*Z^2)\\ &+ 4\operatorname{tr}(Z)\operatorname{tr}(Z^{*2}Z) + 2\operatorname{tr}(Z^*ZZ^*Z) + 4\operatorname{tr}(Z^{*2}Z^2)\big).\end{aligned}$$

The unit balls for these norms are illustrated in Figure 2 (left) .

2.3 Bernoulli random variables

A *Bernoulli* random variable is a discrete random variable X defined according to $\mathbb{P}(X = k) = q^k(1-q)^{1-k}$ for $k = 0, 1$ and $0 < q < 1$. Suppose d is an even integer and $\mathbf{X}$

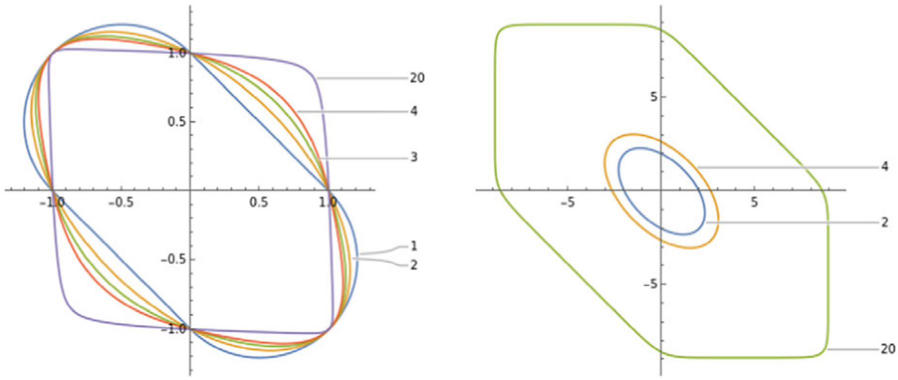

Figure 2: (Left) Unit circles for $\|\cdot\|_{\mathbf{X},d}$ with $d = 1, 2, 3, 4, 20$, in which X_1 and X_2 are standard exponentials. (Right) Unit circles for $\|\cdot\|_{\mathbf{X},d}$ with $d = 2, 4, 20$, in which X_1 and X_2 are Bernoulli with $q = 0.5$.

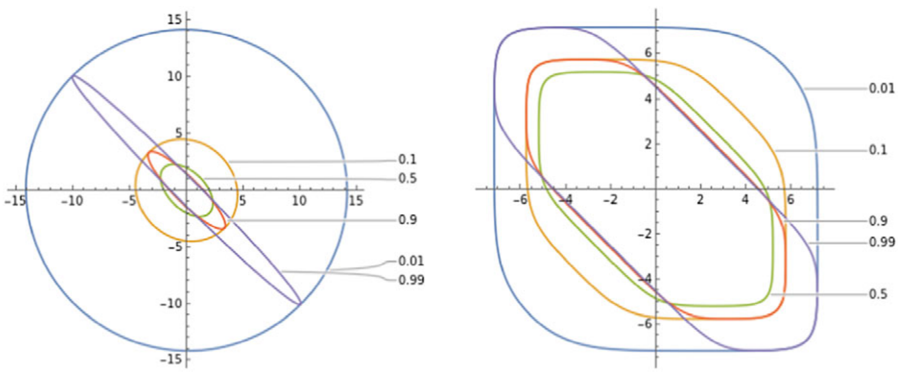

Figure 3: Unit circles for $\|\cdot\|_{\mathbf{X},d}$, in which X_1 and X_2 are Bernoulli with varying parameter q and with $d = 2$ (left) and $d = 10$ (right) .

is a random vector whose entries are independental Bernoulli random variables with parameter q.

Remark 2.1 An expression for $\|A\|_{\mathbf{X},d}^d$ appears in [7, Section 2.7]. However, there is a missing multinomial coefficient. The correct expression for $\|A\|_{\mathbf{X},d}^d$ is given by

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

in which $|I|$ is the number of nonzero i_k; that is, $I = \{k : i_k \neq 0\}$. We thank the anonymous referee for pointing out the typo in [7, Section 2.7]. Figures 2 (right) and 3 illustrate the unit balls for these norms in a variety of cases.

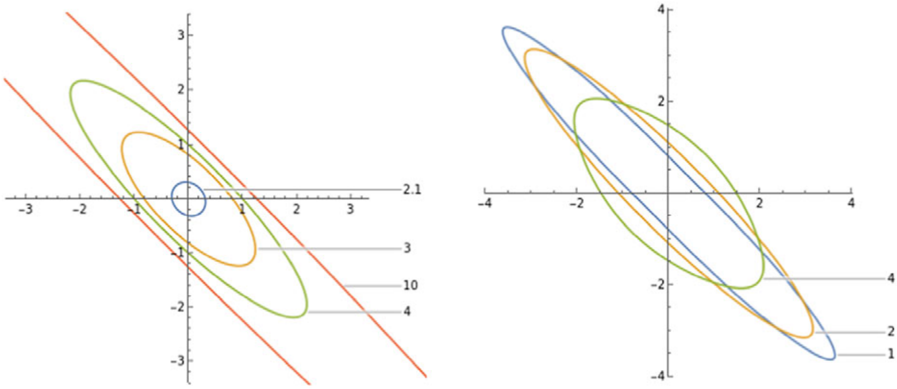

Figure 4: (Left) Unit circles for $\|\cdot\|_{\mathbf{X},2}$, in which X_1 and X_2 are independent Pareto random variables with $\alpha = 2.1, 3, 4, 10$ and $x_m = 1$. (Right) Unit circles for $\|\cdot\|_{\mathbf{X},d}$, in which X_1 and X_2 are independent Pareto random variables with $\alpha = 5$ and $p = 1, 2, 4$.

2.4 Pareto random variables

Suppose $\alpha, x_m > 0$. A random variable X distributed according to the probability density function

$$f_X(t) = \begin{cases} \dfrac{\alpha x_m^\alpha}{t^{\alpha+1}}, & \text{if } t \geq x_m, \\ 0, & \text{if } t < x_m, \end{cases}$$

is a *Pareto* random variable with parameters α and x_m. Suppose $\mathbf{X}$ is a random vector whose entries are Pareto random variables. Then $\|A\|_{\mathbf{X},d}$ exists whenever $\alpha > d$ [7, Section 2.10].

Suppose $d = 2$ and $\mathbf{X}$ is a random vector whose entries are independent Pareto random variables with $\alpha > 2$ and $x_m = 1$. If $n = 2$, then

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

Figure 4 (left) illustrates the unit circles for $\|\cdot\|_{\mathbf{X},2}$ with varying α. As $\alpha \to \infty$, the unit circles approach the parallel lines at $\lambda_2 = \pm\sqrt{2} - \lambda_1$; that is, $|\operatorname{tr} A|^2 = 2$. Figure 4 (right) depicts the unit circles for $\|\cdot\|_{\mathbf{X},d}$ with fixed α and varying d.

3 Proof of Theorem 1.1

The proof of Theorem 1.1 follows from Propositions 3.1 and 3.5.

Proposition 3.1 *If $\|\cdot\|$ is a weakly unitarily invariant norm on H_n, then there is a symmetric norm f on $\mathbb{R}^n$ such that $\|A\| = f(\boldsymbol{\lambda}(A))$ for all $A \in \mathrm{H}_n$.*

https://doi.org/10.4153/S0008439523000875 Published online by Cambridge University Press

Proof Hermitian matrices are unitarily diagonalizable. Since $\|\cdot\|$ is weakly unitarily invariant, $\|A\| = \|D\|$, in which D is a diagonalization of A. Consequently, $\|A\|$ must be a function in the eigenvalues of A. Moreover, any permutation of the entries in D is obtained by conjugating D by a permutation matrix, which is unitary. Therefore, $\|A\|$ is a symmetric function in the eigenvalues of A. In particular, $\|A\| = f(\boldsymbol{\lambda}(A))$ for some symmetric function f. Given $\mathbf{a} = (a_1, a_2, \ldots, a_n) \in \mathbb{R}^n$, define the Hermitian matrix

$$\operatorname{diag}\mathbf{a} = \begin{bmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{bmatrix}.$$

Then $\boldsymbol{\lambda}(\operatorname{diag}\mathbf{a}) = P\mathbf{a}$ for some permutation matrix P. Symmetry of f implies

$$f(\mathbf{a}) = f(P\mathbf{a}) = f\big(\boldsymbol{\lambda}(\operatorname{diag}\mathbf{a})\big) = \|\operatorname{diag}\mathbf{a}\|.$$

Consequently, f inherits the defining properties of a norm on $\mathbb{R}^n$. ∎

Let $\widetilde{\mathbf{x}} = (\widetilde{x}_1, \widetilde{x}_2, \ldots, \widetilde{x}_n)$ denote the nondecreasing rearrangement of $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$. Then $\mathbf{y}$ *majorizes* $\mathbf{x}$, denoted $\mathbf{x} \prec \mathbf{y}$, if

$$\sum_{i=1}^{n} \widetilde{x}_i = \sum_{i=1}^{n} \widetilde{y}_i \quad \text{and} \quad \sum_{i=1}^{k} \widetilde{x}_i \leq \sum_{i=1}^{k} \widetilde{y}_i \quad \text{for } 1 \leq k \leq n-1. \tag{3.1}$$

Recall that a matrix with nonnegative entries is *doubly stochastic* if each row and column sums to 1. The next result is due to Hardy, Littlewood, and Pólya [9].

Lemma 3.2 *If $\mathbf{x} \prec \mathbf{y}$, then there exists a doubly stochastic matrix D such that $\mathbf{y} = D\mathbf{x}$.*

The next lemma is Birkhoff's [6]; $n^2 - n + 1$ works in place of n^2 [10, Theorem 8.7.2].

Lemma 3.3 *If $D \in \mathrm{M}_n$ is doubly stochastic, then there exist permutation matrices $P_1, P_2, \ldots, P_{n^2} \in \mathrm{M}_n$ and nonnegative numbers $c_1, c_2, \ldots, c_{n^2}$ satisfying $\sum_{i=1}^{n^2} c_i = 1$ such that $D = \sum_{i=1}^{n^2} c_i P_i$.*

For each $A \in \mathrm{H}_n$, recall that $\boldsymbol{\lambda}(A) = (\lambda_1(A), \lambda_2(A), \ldots, \lambda_n(A))$ denotes the vector of eigenvalues $\lambda_1(A) \geq \lambda_2(A) \geq \cdots \geq \lambda_n(A)$. We regard $\boldsymbol{\lambda}(A)$ as a column vector for purposes of matrix multiplication.

Lemma 3.4 *If $A, B \in \mathrm{H}_n$, then there exist permutation matrices $P_1, P_2, \ldots, P_{n^2} \in \mathrm{M}_n$ and $c_1, c_2, \ldots, c_{n^2} \geq 0$ such that*

$$\boldsymbol{\lambda}(A+B) = \sum_{i=1}^{n^2} c_i P_i (\boldsymbol{\lambda}(A) + \boldsymbol{\lambda}(B)) \quad \textit{and} \quad \sum_{i=1}^{n^2} c_i = 1.$$

Proof The Ky Fan eigenvalue inequality [8] asserts that

$$\sum_{i=1}^{k} \lambda_i(A+B) \leq \sum_{i=1}^{k} \big(\lambda_i(A) + \lambda_i(B)\big) \quad \text{for all } 1 \leq k \leq n. \tag{3.2}$$

https://doi.org/10.4153/S0008439523000875 Published online by Cambridge University Press

The sum of the eigenvalues of a matrix is its trace. Consequently,

$$\sum_{i=1}^{n} \lambda_i(A+B) = \operatorname{tr}(A+B) = \operatorname{tr} A + \operatorname{tr} B = \sum_{i=1}^{n} \big(\lambda_i(A) + \lambda_i(B)\big),$$

so equality holds in (3.2) for $k = n$. Thus, $\boldsymbol{\lambda}(A+B) \prec \boldsymbol{\lambda}(A) + \boldsymbol{\lambda}(B)$. Lemma 3.2 provides a doubly stochastic matrix D such that $\boldsymbol{\lambda}(A+B) = D(\boldsymbol{\lambda}(A) + \boldsymbol{\lambda}(B))$. Lemma 3.3 provides the desired permutation matrices and nonnegative scalars. ∎

The following proposition completes the proof of Theorem 1.1.

Proposition 3.5 *If f is a symmetric norm on $\mathbb{R}^n$, then $\|A\| = f(\boldsymbol{\lambda}(A))$ defines a weakly unitarily invariant norm on H_n.*

Proof The function $\|A\| = f(\boldsymbol{\lambda}(A))$ is symmetric in the eigenvalues of A, so it is weakly unitarily invariant. It remains to show that $\|\cdot\|$ defines a norm on H_n.

Positive definiteness. A Hermitian matrix $A = 0$ if and only if $\boldsymbol{\lambda}(A) = 0$. Thus, the positive definiteness of f implies the positive definiteness of $\|\cdot\|$.

Homogeneity. If $c \geq 0$, then $\boldsymbol{\lambda}(cA) = c\boldsymbol{\lambda}(A)$. If $c < 0$, then

$$\boldsymbol{\lambda}(cA) = c \begin{bmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{bmatrix} \boldsymbol{\lambda}(A).$$

Then the homogeneity and symmetry of f imply that

$$\|cA\| = f\big(\boldsymbol{\lambda}(cA)\big) = f\big(c\boldsymbol{\lambda}(A)\big) = |c| f\big(\boldsymbol{\lambda}(A)\big) = |c|\,\|A\|.$$

Triangle inequality. Suppose that $A, B \in \mathrm{H}_n$. Lemma 3.4 ensures that there exist permutation matrices $P_1, P_2, \ldots, P_{n^2} \in \mathrm{M}_n$ and nonnegative numbers $c_1, c_2, \ldots, c_{n^2}$ satisfying $\sum_{i=1}^{n^2} c_i = 1$ such that $D = \sum_{i=1}^{n^2} c_i P_i$. Thus,

$$\|A+B\| = f\big(\boldsymbol{\lambda}(A+B)\big) = f\left(\sum_{i=1}^{n^2} c_i P_i\big(\boldsymbol{\lambda}(A) + \boldsymbol{\lambda}(B)\big)\right).$$

The triangle inequality and homogeneity of f yield

$$\|A+B\| \leq \sum_{i=1}^{n^2} c_i f\big(P_i(\boldsymbol{\lambda}(A) + \boldsymbol{\lambda}(B))\big). \tag{3.3}$$

Since f is permutation invariant and $\sum_{i=1}^{n^2} c_i = 1$,

$$\sum_{i=1}^{n^2} c_i f\big(P_i(\boldsymbol{\lambda}(A) + \boldsymbol{\lambda}(B))\big) = \sum_{i=1}^{n^2} c_i f\big(\boldsymbol{\lambda}(A) + \boldsymbol{\lambda}(B)\big) = f\big(\boldsymbol{\lambda}(A) + \boldsymbol{\lambda}(B)\big).$$

Thus, the triangle inequality for f and (3.3) yield

$$\|A+B\| \leq f\big(\boldsymbol{\lambda}(A) + \boldsymbol{\lambda}(B)\big) \leq f\big(\boldsymbol{\lambda}(A)\big) + f\big(\boldsymbol{\lambda}(B)\big) = \|A\| + \|B\|.$$ ∎

4 Proof of Theorem 1.2

Let $\mathbf{X}$ be an iid random vector and define $f_{\mathbf{X},d} : \mathbb{R}^n \to \mathbb{R}$ by

$$f_{\mathbf{X},d}(\boldsymbol{\lambda}) = \left(\frac{\mathbb{E}|\langle \mathbf{X}, \boldsymbol{\lambda} \rangle|^d}{\Gamma(d+1)} \right)^{1/d} \quad \text{for } d \geq 1. \tag{4.1}$$

Since the entries of $\mathbf{X}$ are iid, $f_{\mathbf{X},d}$ is symmetric. In light of Theorem 1.1, it suffices to show that $f_{\mathbf{X},d}$ is a norm on $\mathbb{R}^n$; the continuity remark at the end of Theorem 1.2 is Proposition 4.2.

Proposition 4.1 *The function $f_{\mathbf{X},d}$ in (4.1) defines a norm on $\mathbb{R}^n$ for all $d \geq 1$.*

Proof The proofs for homogeneity and the triangle inequality in [7, Section 3.1] are valid for $d \geq 1$. However, the proof for positive definiteness in [7, Lemma 3.1] requires $d \geq 2$. The proof below holds for $d \geq 1$ and is simpler than the original.

Positive definiteness. If $f_{\mathbf{X},d}(\boldsymbol{\lambda}) = 0$, then $\mathbb{E}|\langle \mathbf{X}, \boldsymbol{\lambda} \rangle|^d = 0$. The nonnegativity of $|\langle \mathbf{X}, \boldsymbol{\lambda} \rangle|^d$ ensures that

$$\lambda_1 X_1 + \lambda_2 X_2 + \cdots + \lambda_n X_n = 0 \tag{4.2}$$

almost surely. Assume (4.2) has a nontrivial solution $\boldsymbol{\lambda}$ with nonzero entries $\lambda_{i_1}, \lambda_{i_2}, \ldots, \lambda_{i_k}$. If $k = 1$, then $X_{i_k} = 0$ almost surely, which contradicts the nondegeneracy of our random variables. If $k > 1$, then (4.2) implies that

$$X_{i_1} = a_{i_2} X_{i_2} + a_{i_3} X_{i_3} + \cdots + a_{i_k} X_{i_k} \tag{4.3}$$

almost surely, in which $a_{i_j} = -\lambda_{i_j}/\lambda_{i_1}$. The independence of $X_{i_1}, X_{i_2}, \ldots, X_{i_k}$ contradicts (4.3). Relation (4.2) therefore has no nontrivial solutions.

Homogeneity. This follows from the bilinearity of the inner product and linearity of expectation:

$$f_{\mathbf{X},d}(c\boldsymbol{\lambda}) = \left(\frac{\mathbb{E}|c\langle \mathbf{X}, \boldsymbol{\lambda} \rangle|^d}{\Gamma(d+1)} \right)^{1/d} = \left(\frac{|c|^d \mathbb{E}|\langle \mathbf{X}, \boldsymbol{\lambda} \rangle|^d}{\Gamma(d+1)} \right)^{1/d} = |c| f_{\mathbf{X},d}(\boldsymbol{\lambda}).$$

Triangle inequality. For $\boldsymbol{\lambda}, \boldsymbol{\mu} \in \mathbb{R}^n$, define random variables $X = \langle \mathbf{X}, \boldsymbol{\lambda} \rangle$ and $Y = \langle \mathbf{X}, \boldsymbol{\mu} \rangle$. Minkowski's inequality implies

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

The triangle inequality for $f_{\mathbf{X},d}$ follows. ∎

Proposition 4.2 *Suppose $\mathbf{X}$ is an iid random vector whose entries have at least m moments. The function $f : [1, m] \to \mathbb{R}$ defined by $f(d) = \|A\|_{\mathbf{X},d}$ is continuous for all $A \in \mathrm{H}_n$.*

Proof Define the random variable $Y = \langle \mathbf{X}, \boldsymbol{\lambda} \rangle$, in which $\boldsymbol{\lambda}$ denotes the vector of eigenvalues of A. The random variable Y is a measurable function defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The pushforward measure of Y is the probability measure

https://doi.org/10.4153/S0008439523000875 Published online by Cambridge University Press

μ_Y on $\mathbb{R}$ defined by $\mu_Y(E) = \mathbb{P}(Y^{-1}(E))$ for all Borel sets E. Consequently,

$$\Gamma(d+1)\big(f(d)\big)^d = \mathbb{E}|Y|^d = \int |x|^d \, d\mu_Y.$$

The bound $|x|^d \le |x| + |x|^m$ holds for all $x \in \mathbb{R}$ and $1 \le d \le m$. Therefore,

$$\int |x|^d \, d\mu_Y \le \int |x| \, d\mu_Y + \int |x|^m \, d\mu_Y.$$

If $d_i \to d$, then $\int |x|^{d_i} d\mu_Y \to \int |x|^d d\mu_Y$ by the dominated convergence theorem. Consequently, $\Gamma(d_i+1)(f(d_i))^{d_i} \to \Gamma(d+1)(f(d))^d$ whenever $d_i \to d$. The function $\Gamma(d+1)(f(d))^d$ is therefore continuous in d. The continuity of the gamma function establishes continuity for f^d and f. ■

5 Remarks

Remark 5.1 A norm $\|\cdot\|$ on M_n is *weakly unitarily invariant* if $\|A\| = \|U^*AU\|$ for all $A \in \mathrm{M}_n$ and unitary $U \in \mathrm{M}_n$. A norm Φ on the space $C(S)$ of continuous functions on the unit sphere $S \subset \mathbb{C}^n$ is a *unitarily invariant function norm* if $\Phi(f \circ U) = \Phi(f)$ for all $f \in C(S)$ and unitary $U \in \mathrm{M}_n$. Every weakly unitarily invariant norm $\|\cdot\|$ on M_n is of the form $\|A\| = \Phi(f_A)$, in which $f_A \in C(S)$ is defined by $f_A(\mathbf{x}) = \langle A\mathbf{x}, \mathbf{x}\rangle$ and Φ is a unitarily invariant function norm [4], [3, Theorem 2.1].

Remark 5.2 Remark 3.4 of [7] is somewhat misleading. We state there that the entries of $\mathbf{X}$ are required to be identically distributed but not independent. To clarify, the entries of $\mathbf{X}$ being identically distributed guarantee that $\|\cdot\|_{\mathbf{X},d}$ satisfies the triangle inequality on H_n. The additional assumption of independence guarantees that $\|\cdot\|_{\mathbf{X},d}$ is also positive definite.

Acknowledgment We thank the referee for many helpful comments.

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), no. 6, 2078–2100.

[2] E. T. Bell, *Exponential polynomials*. Ann. Math. **35**(1934), no. 2, 258–277.

[3] R. Bhatia, *Matrix analysis*, Graduate Texts in Mathematics, 169, Springer, New York, 1997.

[4] R. Bhatia and J. A. R. Holbrook, *Unitary invariance and spectral variation*. Linear Algebra Appl. **95**(1987), 43–68.

[5] P. Billingsley, *Probability and measure*. 3rd ed., Wiley Series in Probability and Mathematical Statistics, Wiley, New York, 1995.

[6] G. Birkhoff, *Three observations on linear algebra*. Univ. Nac. Tucumán. Revista A 5(1946), 147–151.

[7] Á. Chávez, S. R. Garcia, and J. Hurley, *Norms on complex matrices induced by random vectors*. Canad. Math. Bull. **66**(2023), no. 3, 808–826.

[8] K. Fan, *On a theorem of Weyl concerning eigenvalues of linear transformations. I*. Proc. Natl. Acad. Sci. USA **35**(1949), 652–655.

[9] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Some simple inequalities satisfied by convex functions*. Messenger Math. **58**(1929), 145–152.

[10] R. A. Horn and C. R. Johnson, *Matrix analysis*. 2nd ed., Cambridge University Press, Cambridge, 2013.

https://doi.org/10.4153/S0008439523000875 Published online by Cambridge University Press

[11] A. S. Lewis, *Group invariance and convex matrix analysis*. SIAM J. Matrix Anal. Appl. **17**(1996), no. 4, 927–949.

[12] C.-K. Li, *Inequalities relating norms invariant under unitary similarities*. Linear Multilinear Algebra **29**(1991), nos. 3–4, 155–167.

[13] 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.

Mathematics Department, Regis University, 3333 Regis Boulevard, Denver, CO 80221 D-16, United States
e-mail: chave360@regis.edu

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

https://doi.org/10.4153/S0008439523000875 Published online by Cambridge University Press