TENSOR NORMS ON ORDERED NORMED SPACES, POLARIZATION CONSTANTS, AND EXCHANGEABLE DISTRIBUTIONS

SVANTE JANSON

Abstract. We define new norms for symmetric tensors over ordered normed spaces; these norms are defined by considering linear combinations of tensor products or powers of positive elements only. Relations between the different norms are studied. The results are applied to the problem of representing a finitely exchangeable distribution as a mixture of powers, i.e, mixture of distributions of i.i.d. sequences, using a signed mixing measure.

1. Introduction

Let E be a normed space, and consider a tensor $\mathbf{x} \in E^{\otimes n}$. By definition $\mathbf{x}$ is a linear combination of elementary tensors $x_1 \otimes \cdots \otimes x_n$, and, roughly speaking, the projective tensor norm $\|\mathbf{x}\|_\pi$ measures how large such a linear combination has to be; see the definition (3.2) below for a formal statement.

If the tensor $\mathbf{x}$ is symmetric, it can also be written as a linear combination of tensor powers $x^{\otimes n}$, and the symmetric projective norm (3.12) introduced by [10] measures how large such a linear combination has to be.

In the case when the normed space E is an ordered space, it also make sense to ask about decompositions into tensor products or tensor powers of positive elements only. We define in (4.3) and (4.4) two norms on symmetric tensors that measure the size of such decompositions. This gives four different norms on the space $E^{\vee n}$ of symmetric tensors; they are all equivalent but, in general, different.

We study these norms and relations betweeen them in Sections 4–6. In particular, we study the norms of the identity operator between the four spaces obtained by equipping $E^{\vee n}$ with these norms, i.e., the best constants in the inequalities relating these norms to each other. (These constants depend on the space E and on the order n.) One of these constants is known as the *polarization constant* [9]; three other of them, defined in Section 4, are natural versions for ordered spaces, and we call them *positive polarization constants*. Among other results, we show that the space ℓ_1 is extreme for several of these polarization constants.

Date: 6 November, 2018; revised 14 January, 2020.

2010 *Mathematics Subject Classification.* 46B28, 60G09.

Partly supported by the Knut and Alice Wallenberg Foundation.

One motivation for the present paper is the problem of representing finitely exchangeable distributions of random vectors as mixtures of distributions of independent sequences. This problem is described more fully in Section 7. It is well known that, in contrast to de Finetti's theorem for infinite exchangeable sequences, such representations with a probability measure as the mixing measure are in general not possible for finitely exchangeable distributions; however, a substitute exists where the mixing measure is a signed measure [5; 15; 19; 16]. A natural question is how large the norm of this mixing measure has to be, and it is shown in Section 7 that this is essentially equivalent to studying one of the positive tensor norms defined in Section 4, in the special case when $E = \ell_1$ (or a finite-dimensional ℓ_1^m). We use this to derive several new results on the optimal norm of the mixing measure.

Section 8 gives some simple explicit examples in the case when E is a Euclidean space.

Sections 2 and 3 contain background material, surveying definitions and elementary properties of polarization, polarization constants and tensor products. These sections provide background and easy references to various facts for use in later sections. (There are no new results there.)

In the main part of the paper, starting with Section 4, we consider ordered normed spaces, and thus spaces over $\mathbb{R}$. However, in the introductory Sections 2 and 3, no ordering is considered, so E can be any normed space, with real or complex scalars.

1.1. **Some notation.** We consider linear spaces over $\mathbb{K}$, where $\mathbb{K}$ is either $\mathbb{R}$ or $\mathbb{C}$. In particular, E or F is always a normed space over $\mathbb{K}$. Furthermore, $n \geqslant 1$ is an integer, usually fixed but arbitrary. (We sometimes tacitly assume that the spaces have non-zero dimension.)

For $1 \leqslant m < \infty$, ℓ_p^m is $\mathbb{K}^m$ with the ℓ_p-norm. In particular, ℓ_2^m is the usual Euclidean space $\mathbb{R}^m$ or $\mathbb{C}^m$. We also write $\ell_p^\infty = \ell_p$, and let $\ell_p(S)$ denote the ℓ_p space with index set S, i.e., $\ell_p(S) := L^p(S, \mu)$ where μ is the counting measure on S. Thus $\ell_p = \ell_p(\mathbb{N})$ and $\ell_p^m = \ell_p([m])$, where $[m] := \{1, \dots, m\}$. The standard basis in ℓ_p or ℓ_p^m is denoted by (e_i).

For a normed space E, $B(E) := \{x \in E : \|x\| \leqslant 1\}$, the closed unit ball of E.

For a real number x, $\lfloor x \rfloor$ and $\lceil x \rceil$ are the integers obtained by rounding x downwards and upwards, respectively.

"Positive" should generally be interpreted as "non-negative".

2. Symmetric multilinear forms and polynomials

In this section, we review some basic theory of symmetric multilinear forms and operators, including the important polarization formula. See e.g. Dineen [9, Section 1.1] for further details. In this section, we allow both real and complex scalars; we therefore denote the scalar field by $\mathbb{K}$ ($= \mathbb{R}$ or $\mathbb{C}$).

$\mathcal{L}(^nE;F) = \mathcal{L}(E,\dots,E;F)$ denotes the space of all continuous n-linear operators $E^n \to F$. We will mainly consider the case $F = \mathbb{K}$: $\mathcal{L}(^nE;\mathbb{K}) = \mathcal{L}(E,\dots,E;\mathbb{K})$ is the space of all continuous n-linear forms $E^n \to \mathbb{K}$.

It is well-known that an n-linear operator $L : E^n \to F$ is continuous if and only if it is bounded, i.e., if the norm

$$\|L\| := \sup\bigl\{|L(x_1,\dots,x_n)| : \|x_1\|,\dots,\|x_n\| \leqslant 1\bigr\} \tag{2.1}$$

is finite. $\mathcal{L}(^nE;F)$ is a normed space with the norm $\| \, \|$ in (2.1). (It is a Banach space if F is complete, e.g. if $F = \mathbb{K}$.)

Definitions 2.1. Let $\mathfrak{S}_n$ be the symmetric group of the $n!$ permutations of $\{1,\dots,n\}$.

(i) If $L \in \mathcal{L}(^nE;F)$ and $\sigma \in \mathfrak{S}_n$, then $L_\sigma \in \mathcal{L}(^nE;F)$ is given by

$$L_\sigma(x_1,\dots,x_n) := L\bigl(x_{\sigma(1)},\dots,x_{\sigma(n)}\bigr). \tag{2.2}$$

(ii) $L \in \mathcal{L}(^nE;F)$ is *symmetric* if $L_\sigma = L$ for all $\sigma \in \mathfrak{S}_n$. Let

$$\mathcal{L}_{\mathsf{s}}(^nE;F) := \{L \in \mathcal{L}(^nE;F) : L \text{ is symmetric}\} \tag{2.3}$$

be the space of continuous (or, equivalently, bounded) symmetric n-linear operators.

(iii) If $L \in \mathcal{L}(^nE;F)$, then its *symmetrization* $\tilde{L} \in \mathcal{L}_{\mathsf{s}}(^nE;F)$ is given by

$$\tilde{L} := \frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n} L_\sigma. \tag{2.4}$$

Note that L is symmetric $\iff L = \tilde{L}$, and that the symmetrization map $L \mapsto \tilde{L}$ is a linear projection of of $\mathcal{L}(^nE;F)$ onto $\mathcal{L}_{\mathsf{s}}(^nE;F)$.

2.1. **Polynomials.** If $L : E^n \to F$ is an n-linear operator (or any function on E^n), we define $\hat{L} : E \to F$ by

$$\hat{L}(x) := L(x,\dots,x). \tag{2.5}$$

In other words, $\hat{L}$ is the restriction of L to the diagonal.

Definitions 2.2. (i) A function $q : E \to K$ is a *homogeneous polynomial of degree n* if $q = \hat{L}$ for some n-linear form $L : E^n \to \mathbb{K}$.

(ii) If p is a homogeneous polynomial on E, let

$$\|p\| := \sup\bigl\{|p(x)| : \|x\| \leqslant 1\bigr\}, \tag{2.6}$$

i.e., the usual sup-norm of the restriction of p to the unit ball of E. We say that p is *bounded* if $\|p\| < \infty$, i.e., if p is bounded on the unit ball of E.

(iii) Let $\mathcal{P}_n(E)$ be the space of bounded homogeneous polynomials of degree n. This is a normed space with the norm $\| \, \|$; furthermore, it is a Banach space (see Corollary 2.4 below).

2.2. **Polarization.** We have the following important *polarization identity*, see e.g. [9, Corollary 1.6].

Lemma 2.3. *If $L : E^n \to F$ is an n-linear operator (bounded or not), then*

$$\tilde{L}(x_1, \dots, x_n) = \frac{1}{2^n n!} \sum_{\varepsilon_1, \dots, \varepsilon_n = \pm 1} \varepsilon_1 \cdots \varepsilon_n \, \hat{L}\Bigl(\sum_{i=1}^n \varepsilon_i x_i\Bigr). \tag{2.7}$$

In particular, if L is symmetric, then

$$L(x_1, \dots, x_n) = \frac{1}{2^n n!} \sum_{\varepsilon_1, \dots, \varepsilon_n = \pm 1} \varepsilon_1 \cdots \varepsilon_n \, \hat{L}\Bigl(\sum_{i=1}^n \varepsilon_i x_i\Bigr). \tag{2.8}$$

□

Corollary 2.4. *The mapping $\pi : L \mapsto \hat{L}$ is a linear bijection of $\mathcal{L}_{\mathsf{s}}({}^nE; \mathbb{K})$ onto $\mathcal{P}_n(E)$. Hence, $\mathcal{P}_n(E)$ is isomorphic to $\mathcal{L}_{\mathsf{s}}({}^nE; \mathbb{K})$ as normed spaces, with equivalence of norms.* □

More precisely, (2.5) and (2.8) yield the following inequalities for $L \in \mathcal{L}_{\mathsf{s}}({}^nE; \mathbb{K})$.

$$\|\hat{L}\| \leqslant \|L\|, \tag{2.9}$$

$$\|L\| \leqslant \frac{n^n}{n!} \|\hat{L}\|. \tag{2.10}$$

Define, for a multilinear form L on E^n

$$\|L\|_\Delta := \|\hat{L}\| = \sup\bigl(|L(x, \dots, x)| : \|x\| \leqslant 1\bigr). \tag{2.11}$$

Then (2.9)–(2.10) can also be written

$$\|L\|_\Delta \leqslant \|L\|, \tag{2.12}$$

$$\|L\| \leqslant \frac{n^n}{n!} \|L\|_\Delta. \tag{2.13}$$

Hence, $\|\ \|$ and $\|\ \|_\Delta$ are two equivalent norms on $\mathcal{L}_{\mathsf{s}}({}^nE; \mathbb{K})$.

2.3. **Polarization constants.**

Definition 2.5. The polarization constant $c_{\mathsf{s}}(n, E)$ is defined by, see [9, Definition 1.40],

$$c_{\mathsf{s}}(n, E) := \sup_{L \in \mathcal{L}_{\mathsf{s}}({}^nE; \mathbb{K})} \frac{\|L\|}{\|\hat{L}\|} = \sup_{L \in \mathcal{L}_{\mathsf{s}}({}^nE; \mathbb{K})} \frac{\|L\|}{\|L\|_\Delta}, \tag{2.14}$$

where, as in similar suprema below, we define $\frac{0}{0} := 0$. Equivalently, $c_{\mathsf{s}}(n, E)$ is the norm of the linear operator $\pi^{-1} : \mathcal{P}_n(E) \to \mathcal{L}_{\mathsf{s}}({}^nE; \mathbb{K})$, see Corollary 2.4.

Since $\hat{L} = \hat{\tilde{L}}$, we also have

$$c_{\mathsf{s}}(n, E) = \sup_{L \in \mathcal{L}({}^nE; \mathbb{K})} \frac{\|\tilde{L}\|}{\|\hat{\tilde{L}}\|} = \sup_{L \in \mathcal{L}({}^nE; \mathbb{K})} \frac{\|\tilde{L}\|}{\|\hat{L}\|}. \tag{2.15}$$

By (2.14) and (2.9)–(2.10),

$$1 \leqslant c_{\mathsf{s}}(n,E) \leqslant \frac{n^n}{n!}. \tag{2.16}$$

Both inequalities in (2.16) can be attained. (The upper bound in (2.16) was conjectured by Mazur and Orlicz in "The Scottish Book", and proved in 1932 by Martin; see Harris [14] and the references there.)

Example 2.6. For any Hilbert space H (real or complex; of finite or infinite dimension) and any $n \geqslant 1$, $c_{\mathsf{s}}(n,H)=1$; see Banach [2]. See further [14].

Example 2.7. For any $n \geqslant 1$ and any $m \geqslant n$, $c_{\mathsf{s}}(n,\ell_1^m)=c_{\mathsf{s}}(n,\ell_1)=n^n/n!$.

To see this, let $n \leqslant m \leqslant \infty$ and define $L \in \mathcal{L}({}^n\ell_1^m,\mathbb{K})$ by

$$L(x_1,\dots,x_n)=\prod_{i=1}^n x_{ii}, \qquad \text{where } x_i=(x_{ij})_{j=1}^m. \tag{2.17}$$

L is not symmetric, so we consider its symmetrization $\tilde{L}$. We have, letting $e_1,e_2,\dots$ be the usual basis vectors in ℓ_p^m,

$$\|\tilde{L}\| \geqslant |\tilde{L}(e_1,\dots,e_n)| = \frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n} L\bigl(e_{\sigma(1)},\dots,e_{\sigma(n)}\bigr)=\frac{1}{n!} \tag{2.18}$$

and, by the arithmetic-geometric inequality, if $x=(x_i)_1^m$,

$$|\hat{L}(x)| = \prod_{i=1}^n |x_i| \leqslant \Bigl(\frac{1}{n}\sum_{i=1}^n |x_i|\Bigr)^n \leqslant n^{-n}\|x\|^n. \tag{2.19}$$

Hence, $\|\hat{L}\| \leqslant n^{-n}$, and by (2.15),

$$c_{\mathsf{s}}(n,\ell_1^m) \geqslant \frac{\|\tilde{L}\|}{\|\hat{L}\|} \geqslant \frac{n^n}{n!}. \tag{2.20}$$

The converse inequality follows by (2.16).

Consequently, recalling (2.16) again,

$$\sup_E c_{\mathsf{s}}(n,E)=c_{\mathsf{s}}(n,\ell_1^n)=c_{\mathsf{s}}(n,\ell_1)=\frac{n^n}{n!}. \tag{2.21}$$

Thus, ℓ_1 is extremal among all normed spaces, and so is ℓ_1^n when n is given.

See e.g. [9] and [24] for further examples.

Remark 2.8. Dineen [9, Definition 1.10] defines also

$$c_{\mathsf{s}}(E) := \limsup_{n\to\infty} c_{\mathsf{s}}(n,E)^{1/n}. \tag{2.22}$$

It is an obvious conjecture that the limit always exists, i.e., that lim sup can be replaced by lim in (2.22); however, this seems to be unproven so far. The same applies to the related quantities in Remark 5.4.

By (2.16) and Stirling's formula, for any normed space E,

$$1 \leqslant c_{\mathsf{s}}(E) \leqslant e, \tag{2.23}$$

with both bounds attained since $c_{\mathsf{s}}(H) = 1$ for a Hilbert space H and $c_{\mathsf{s}}(\ell_1) = e$ by Examples 2.6 and 2.7. As another example, [9, Proposition 1.43] implies that $c_{\mathsf{s}}(\ell_\infty) \leqslant e/2$.

Dimant, Galicer and Rodríguez [8] show that $c_{\mathsf{s}}(E) = 1$ for every finite-dimensional complex vector space, but not for every finite-dimensional real vector space.

Remark 2.9. Another type of polarization constants, called *linear polarization constants* has also been studied in the literature, see e.g. [3]. There are, as far as we know, no direct connections with the polarization constants studied here.

3. Tensor products

In this section we recall some basic properties of tensor products and symmetric tensor products. (These results are not new, but we present them in a form suitable for later use.) See e.g. Ryan [29, Chapters 1 and 2], Dineen [9, Chapter 1] and Floret [10] for basic definitions, further details and many other things not mentioned here. In particular, note that we only consider *tensor powers*, i.e., tensor products of a space with itself (one or several times). Again, we allow in this section both real and complex scalars.

3.1. **The projective tensor norm.** Let $E^{\otimes n} = E\otimes\cdots\otimes E$ be the algebraic n:th tensor power of E. Recall that an element $\mathbf{x} \in E^{\otimes n}$ can be written, non-uniquely, as a linear combination

$$\mathbf{x} = \sum_{k=1}^{N} a_k x_{1k} \otimes \cdots \otimes x_{nk} \tag{3.1}$$

of *elementary tensors* $x_{1k} \otimes \cdots \otimes x_{nk}$ for some $x_{ik} \in E$, $i = 1, \dots, n$, $k = 1, \dots, N$, and $a_k \in \mathbb{K}$. (Here and below, N is an arbitrary positive integer.)

The *projective tensor norm* $\| \, \|_\pi$ on $E^{\otimes n}$ is defined by

$$\|\mathbf{x}\|_\pi := \inf \left\{ \sum_{k=1}^{N} |a_k| \|x_{1k}\| \cdots \|x_{nk}\| : \mathbf{x} = \sum_{k=1}^{N} a_k x_{1k} \otimes \cdots \otimes x_{nk} \right\}. \tag{3.2}$$

This is a norm on $E^{\otimes n}$. We denote $E^{\otimes n}$ with this norm by $E_\pi^{\otimes n}$.

We use the notation $\|\mathbf{x}\|_{\pi;\,E}$ when we want to show the space E explicitly, but usually we omit E from the notation.

Remark 3.1. If E has infinite dimension, then $E_\pi^{\otimes n}$ is not complete even if E is. The *projective tensor power* $E_\pi^{\widehat{\otimes} n}$ of a Banach space E is defined as the completion of $E_\pi^{\otimes n}$. The norms defined below on $E^{\otimes n}$ or its subspace $E^{\vee n}$ (also defined below) are all equivalent to $\| \, \|_\pi$, and thus the completions with respect to these norms are the same, as vector spaces, as the completion $E_\pi^{\widehat{\otimes} n}$ or the corresponding completion of $E^{\vee n}$ (i.e., the closure of $E^{\vee n}$ in $E_\pi^{\widehat{\otimes} n}$). Hence, the results below on e.g. inequalities between the different norms extend trivially to the completed spaces.

While it often is natural to work with completed spaces, we have in the present paper not much need for them, and we will work with normed spaces such as $E_\pi^{\otimes n}$ without completing them. Hence, we leave extensions to completed tensor products to the reader.

Remark 3.2. It is not difficult to see that for an elementary tensor $\mathbf{x} = x_1 \otimes \cdots \otimes x_n$,

$$\|x_1 \otimes \cdots \otimes x_n\|_\pi = \|x_1\| \cdots \|x_n\|. \tag{3.3}$$

The projective norm is the largest norm on $E^{\otimes n}$ that satisfies (3.3).

Remark 3.3. Roughly speaking, the unit ball of $E_\pi^{\otimes n}$ is spanned by the elementary tensors $x_1 \otimes \cdots \otimes x_n$ with $x_1, \dots, x_n \in B(E)$. More precisely $B(E_\pi^{\otimes n})$ equals the closed convex hull of the set of these elementary tensors. If $\dim(E) < \infty$, we do not have to take the closure because the convex hull of a compact set is compact in a finite-dimensional space [28, Theorem 3.20(d)]; thus $B(E_\pi^{\otimes n})$ then equals the convex hull of the set of these elementary tensors. This means that the infimum in (3.2) is attained when $\dim(E) < \infty$.

Remark 3.4. It follows from (3.2) or from Remark 3.3 that for any linear operator $T : E_\pi^{\otimes n} \to F$, where F is a normed space,

$$\|T\| = \sup\bigl\{\|T(x_1 \otimes \cdots \otimes x_n)\| : x_1, \dots, x_n \in B(E)\bigr\}. \tag{3.4}$$

Conversely, this characterizes $\| \, \|_\pi$.

Example 3.5. In the finite-dimensional case $E = \mathbb{K}^m$ (with any norm), the space $E^{\otimes 2}$ is naturally identified with the m^2-dimensional space of $m \times m$ matrices. (We will use this without comment in some examples below.) We recall for later use two well-known examples of the projective tensor norm $\| \, \|_\pi$ in $E^{\otimes 2}$: If $E = \ell_1^m$, then the norm is the ℓ_1-norm, so $(\ell_1^m)_\pi^{\otimes 2} = \ell_1^{m\times m}$ [29, Exercise 2.6]. If $E = \ell_2^m$, then the norm in $(\ell_2^m)_\pi^{\otimes 2}$ of a matrix is its Trace class norm (also known as nuclear norm and Schatten S_1 norm, see e.g. [13, §3.8], [21, §30.2], [31, Chapter 48]); if A is a symmetric matrix (Hermitean in the complex case), then this norm equals the sum of the absolute values of the m eigenvalues.

The fundamental property of tensor products is that they linearize multilinear operators. More precisely, in our case, for any linear space F, there is a natural bijection between multilinear maps $L : E^n \to F$ and linear maps $\bar{L} : E^{\otimes n} \to F$ determined by

$$L(x_1, \dots, x_n) = \bar{L}(x_1 \otimes \cdots \otimes x_n). \tag{3.5}$$

In particular, taking $F = \mathbb{K}$, this gives a 1–1 correspondence between n-linear forms on E and linear forms on $E^{\otimes n}$. It follows from (3.5), the definition (2.1) and (3.4) that for an n-linear map $L \in \mathcal{L}({}^nE; \mathbb{K})$, the norm $\|\bar{L}\|_\pi^*$ of $\bar{L}$ as a linear functional on $E_\pi^{\otimes n}$ equals the norm $\|L\|$ of L.

In the sequel, we abuse notation by denoting also the map $E^{\otimes n} \to \mathbb{K}$ corresponding to $L : E^n \to \mathbb{K}$ as in (3.5) by the same symbol L (instead of

$\bar{L}$). We thus have

$$\|L\|_\pi^* = \|L\|. \tag{3.6}$$

The space $(E^{\otimes n})^*$ of bounded linear functionals on $E^{\otimes n}$ is thus identified (isometrically) with $\mathcal{L}({}^nE;\mathbb{K})$.

3.2. **Symmetric tensor products.** A permutation $\sigma \in \mathfrak{S}_n$ defines an automorphism ι_σ of $E^{\otimes n}$ that is defined on elementary tensors by $\iota_\sigma(x_1 \otimes \cdots \otimes x_n) := x_{\sigma(1)} \otimes \cdots \otimes x_{\sigma(n)}$ and extended by linearity. A tensor $\mathbf{x} \in E^{\otimes n}$ is *symmetric* if $\iota_\sigma(\mathbf{x}) = \mathbf{x}$ for every $\sigma \in \mathfrak{S}_n$. The *symmetric tensor product* $E^{\vee n}$ is the subspace of $E^{\otimes n}$ consisting of the symmetric tensors.

Define the *symmetrization operator* $\Lambda := \frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n} \iota_\sigma$. Then Λ is a linear projection of $E^{\otimes n}$ onto $E^{\vee n}$. We define the *elementary symmetric tensors*

$$x_1 \vee \cdots \vee x_n := \Lambda(x_1 \otimes \cdots \otimes x_n) = \frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n} x_{\sigma(1)} \otimes \cdots \otimes x_{\sigma(n)} \in E^{\vee n}. \tag{3.7}$$

Note that the *tensor powers* are elementary symmetric:

$$x^{\vee n} := x \vee \cdots \vee x = x \otimes \cdots \otimes x = x^{\otimes n}. \tag{3.8}$$

We will mainly use the notation $x^{\otimes n}$, also when discussing $E^{\vee n}$.

If $\mathbf{x} \in E^{\vee n}$ is a symmetric tensor with a representation (3.1), then also

$$\mathbf{x} = \Lambda(\mathbf{x}) = \sum_{k=1}^N a_k x_{1k} \vee \cdots \vee x_{nk}. \tag{3.9}$$

Hence, the linear space $E^{\vee n}$ is spanned by the tensors $x_1 \vee \cdots \vee x_n$.

Furthermore, $E^{\vee n}$ is also spanned by the (smaller) set of tensor powers $x^{\otimes n}$ in (3.8). This follows from the polarization identity (2.7) applied to the multilinear map $L : E^n \to E^{\otimes n}$ given by $L(x_1,\dots,x_n) := x_1 \otimes \cdots \otimes x_n$, which yields, using (3.7) and (2.4),

$$x_1 \vee \cdots \vee x_n = \tilde{L}(x_1,\dots,x_n) = \frac{1}{2^n n!}\sum_{\varepsilon_1,\dots,\varepsilon_n=\pm1} \varepsilon_1\cdots\varepsilon_n\Bigl(\sum_{i=1}^n \varepsilon_i x_i\Bigr)^{\otimes n}. \tag{3.10}$$

It follows easily, using symmetrization by Λ as in (3.9), that for a symmetric tensor $\mathbf{x} \in E^{\vee n}$, the projective norm in (3.2) is also given by

$$\|\mathbf{x}\|_\pi = \inf\Bigl\{\sum_{k=1}^N |a_k|\|x_{1k}\|\cdots\|x_{nk}\| : \mathbf{x} = \sum_{k=1}^N a_k x_{1k} \vee \cdots \vee x_{nk}\Bigr\}. \tag{3.11}$$

The *symmetric projective tensor norm* (or *projective s-tensor norm*) on $E^{\vee n}$, introduced by Floret [10], is defined by

$$\|\mathbf{x}\|_{\pi,\mathsf{s}} := \inf\Bigl\{\sum_{k=1}^N |a_k|\|x_k\|^n : \mathbf{x} = \sum_{k=1}^N a_k x_k^{\otimes n}\Bigr\}. \tag{3.12}$$

By (3.11), (3.12) and (3.10), $\|\mathbf{x}\|_\pi \leqslant \|\mathbf{x}\|_{\pi,\mathsf{s}} < \infty$, so $\|\ \|_{\pi,\mathsf{s}}$ is another norm on $E^{\vee n}$. We will see in (3.20) below that the norms are equivalent. We denote the normed spaces obtained by equipping $E^{\vee n}$ with the norms $\| \|_\pi$ and $\| \|_{\pi,\mathsf{s}}$ by $E_\pi^{\vee n}$ and $E_{\pi,\mathsf{s}}^{\vee n}$, respectively.

Remark 3.6. It follows from (3.12) and Remark 3.2 that for an elementary tensor power $\mathbf{x} = x^{\otimes n}$,

$$\|x^{\otimes n}\|_{\pi,\mathsf{s}} = \|x^{\otimes n}\|_\pi = \|x\|^n. \tag{3.13}$$

The projective s-tensor norm is the largest norm on $E^{\vee n}$ that satisfies (3.13).

Remark 3.7. In analogy with Remark 3.3, the unit balls $B(E_\pi^{\vee n})$ and $B(E_{\pi,\mathsf{s}}^{\vee n})$ equal the closed convex hull of the sets $\{x_1 \vee \cdots \vee x_n : x_1, \dots, x_n \in B(E)\}$ and $\{\pm x^{\vee n} : x \in B(E)\}$, respectively. Again, if $\dim(E) < \infty$, we do not have to take the closures, and thus the infima in (3.11) and (3.12) are attained.

Remark 3.8. Similarly, in analogy with (3.4), it follows from (3.11) and (3.12) that for any linear operator $T : E_\pi^{\vee n} \to F$, where F is a normed space,

$$\|T\|_{E_\pi^{\vee n},F} = \sup\bigl\{\|T(x_1 \vee \cdots \vee x_n)\| : x_1, \dots, x_n \in B(E)\bigr\}, \tag{3.14}$$

$$\|T\|_{E_{\pi,\mathsf{s}}^{\vee n},F} = \sup\bigl\{\|T(x^{\otimes n})\| : x \in B(E)\bigr\}. \tag{3.15}$$

Conversely, these properties characterize the norms $\| \|_\pi$ and $\| \|_{\pi,\mathsf{s}}$ on $E^{\vee n}$.

Similarly to the bijection in (3.5), there is a bijection between symmetric multilinear maps $E^n \to F$ and linear maps $E^{\vee n} \to F$ given by

$$L(x_1, \dots, x_n) = L(x_1 \vee \cdots \vee x_n), \tag{3.16}$$

where we again abuse notation by using the same symbol for both operators. In particular, taking $F = \mathbb{K}$, this yields a bijection between linear forms on $E^{\vee n}$ and symmetric multilinear forms $E^n \to K$.

Let L be a linear form on $E^{\vee n}$. The norm of L in the dual of $E_\pi^{\vee n}$ is by (3.11), (3.16) and (2.1),

$$\begin{aligned}\|L\|_\pi^* &= \sup\bigl\{|L(x_1 \vee \cdots \vee x_n)| : \|x_1\|, \dots, \|x_n\| \leqslant 1\bigr\}\\ &= \sup\bigl\{|L(x_1, \dots, x_n)| : \|x_1\|, \dots, \|x_n\| \leqslant 1\bigr\}\\ &= \|L\|.\end{aligned} \tag{3.17}$$

and the norm in the dual of $E_{\pi,\mathsf{s}}^{\vee n}$ is by (3.12), (3.16) and (2.11),

$$\begin{aligned}\|L\|_{\pi,\mathsf{s}}^* &= \sup\bigl\{|L(x^{\otimes n})| : \|x\| \leqslant 1\bigr\}\\ &= \sup\bigl\{|L(x, \dots, x)| : \|x\| \leqslant 1\bigr\}\\ &= \|L\|_\Delta.\end{aligned} \tag{3.18}$$

We obtain from (3.17)–(3.18) and the definition (2.14) immediately the following:

Lemma 3.9. *The polarization constant $c_{\mathsf{s}}(n, E)$ is given by*

$$c_{\mathsf{s}}(n, E) = \sup_{L \in (E^{\vee n})^*} \frac{\|L\|_\pi^*}{\|L\|_{\pi,\mathsf{s}}^*}. \tag{3.19}$$

In other words, $c_{\mathsf{s}}(n,E)$ *equals the norm of the identity map* $(E^{\vee n}_{\pi,\mathsf{s}})^* \to (E^{\vee n}_{\pi})^*$. ☐

Corollary 3.10. *The polarization constant* $c_{\mathsf{s}}(n,E)$ *equals the norm of the identity map* $E^{\vee n}_{\pi} \to E^{\vee n}_{\pi,\mathsf{s}}$. *In other words, for any* $\mathbf{x} \in E^{\vee n}$,

$$\|\mathbf{x}\|_{\pi} \leqslant \|\mathbf{x}\|_{\pi,\mathsf{s}} \leqslant c_{\mathsf{s}}(n,E)\|\mathbf{x}\|_{\pi} \tag{3.20}$$

and $c_{\mathsf{s}}(n,E)$ *is the smallest constant for which this holds for all* $\mathbf{x} \in E^{\vee n}$.

Proof. Lemma 3.9 and duality. ☐

By (3.14), Corollary 3.10 is also equivalent to

$$c_{\mathsf{s}}(n,E) = \sup\bigl\{\|x_1 \vee \cdots \vee x_n\|_{\pi,\mathsf{s}} : \|x_1\|,\dots,\|x_n\| \leqslant 1\bigr\}. \tag{3.21}$$

In other words, by (3.12), $c_{\mathsf{s}}(n,E)$ describes how efficiently a symmetric tensor $x_1 \vee \cdots \vee x_n$ with $x_1,\dots,x_n \in B(E)$ can be decomposed as a linear combination of tensor powers $y_j^{\otimes n}$.

Example 3.11. For a Hilbert space H, Banach [2] showed $c_{\mathsf{s}}(n,H) = 1$, as said in Example 2.6. Thus Corollary 3.10 yields $\|\mathbf{x}\|_{\pi,\mathsf{s}} = \|\mathbf{x}\|_{\pi}$ for any $\mathbf{x} \in H^{\vee n}$, $n \geqslant 1$; in other words, $H^{\vee n}_{\pi,\mathsf{s}} = H^{\vee n}_{\pi}$ isometrically. See [12, Section 5].

3.3. **Functorial properties.** If E and F are two normed spaces and $T : E \to F$ is a bounded linear operator, then T induces a linear operator $T^{\otimes n} : E^{\otimes n} \to F^{\otimes n}$ by $T^{\otimes n}(x_1 \otimes \cdots \otimes x_n) = Tx_1 \otimes \cdots \otimes Tx_n$; furthermore, $T^{\otimes n}$ restricts to $T^{\vee n} : E^{\vee n} \to F^{\vee n}$. We note the following well-known fact.

Theorem 3.12. *If* E *and* F *are normed spaces and* $T : E \to F$ *is a bounded linear operator, then* $T^{\otimes n} : E^{\otimes n}_{\pi} \to F^{\otimes n}_{\pi}$, $T^{\vee n} : E^{\vee n}_{\pi} \to F^{\vee n}_{\pi}$ *and* $T^{\vee n} : E^{\vee n}_{\pi,\mathsf{s}} \to F^{\vee n}_{\pi,\mathsf{s}}$ *all have norm* $\|T\|^n$.

Proof. An immediate consequence of (3.4), (3.14), (3.15) together with (3.3) and (3.13). ☐

There are some related simple results when we change the normed space. Recall that the *Banach–Mazur distance* between two isomorphic normed spaces (in particular, Banach spaces) is $\inf\{\|T\|\|T^{-1}\|\}$, taking the infimum over all isomorphisms $T : E \to F$.

Theorem 3.13. (i) *If* F *is a quotient space of* E, *then* $c_{\mathsf{s}}(n,F) \leqslant c_{\mathsf{s}}(n,E)$.

(ii) *If* F *is a* κ*-complemented subspace of* E, *i.e.,* F *is a subspace and there exists a projection* $P : E \to F$ *of norm* $\|P\| \leqslant \kappa$, *then* $c_{\mathsf{s}}(n,F) \leqslant \kappa^n c_{\mathsf{s}}(n,E)$. *In particular, if* F *is* 1*-complemented, then* $c_{\mathsf{s}}(n,F) \leqslant c_{\mathsf{s}}(n,E)$.

(iii) *If* E *and* F *are isomorphic normed spaces, then* $c_{\mathsf{s}}(n,F) \leqslant d(E,F)^n c_{\mathsf{s}}(n,E)$, *where* $d(E,F)$ *is the Banach–Mazur distance. In particular,* $c_{\mathsf{s}}(n,E) = c_{\mathsf{s}}(n,F)$ *when* E *and* F *are isometric.*

We omit the proof of Theorem 3.13, which is essentially identical to the proof of our ordered version Theorem 4.15 below.

Remark 3.14. It is not true in general that $c_{\mathsf{s}}(n, F) \leqslant c_{\mathsf{s}}(n, E)$ when F is a subspace of E. For example, ℓ_1 (as any separable Banach space) can be embedded isometrically as a subspace of ℓ_∞. However, by [9, Proposition 1.43] and (2.21), $c_{\mathsf{s}}(n, \ell_\infty) < c_{\mathsf{s}}(n, \ell_1)$ for any $n \geqslant 2$. Also, see [9, p. 52] and Remark 2.8, $c_{\mathsf{s}}(\ell_\infty) < c_{\mathsf{s}}(\ell_1)$.

4. Positive tensor products and polarization constants

In the remainder of the paper, we assume that $\mathbb{K} = \mathbb{R}$, and that E is an ordered normed space, i.e., a normed space that is also an ordered linear space. This means that there is given a closed cone E^+ of *positive* elements in E; the order is defined by $x \leqslant y \iff y - x \in E^+$, and, conversely, $E^+ := \{x : x \geqslant 0\}$.

We assume also that $E = E^+ - E^+$, i.e., that every $x \in E$ can be written as a difference $y - z$ of two positive elements. We define a new norm $\| \, \|_+$ on E by

$$\|x\|_+ := \inf\bigl\{\|y\| + \|z\| : x = y - z, \ y \geqslant 0, \ z \geqslant 0\bigr\} \tag{4.1}$$

and note that the triangle inequality implies $\|x\|_+ \geqslant \|x\|$. Finally, we assume that

$$c_+(E) := \sup\bigl\{\|x\|_+ : \|x\| \leqslant 1\bigr\} \tag{4.2}$$

is finite. Thus, $\| \, \|$ and $\| \, \|_+$ are equivalent norms on E. Let E_+ denote E equipped with the norm $\| \, \|_+$. Then, $c_+(E)$ is the norm of the identity operator $E \to E_+$.

Example 4.1. Some standard examples are ℓ_p^m and ℓ_p, for $1 \leqslant p \leqslant \infty$, and more generally $L^p(\mathcal{S}, \mathcal{F}, \mu)$ for any measure space $(\mathcal{S}, \mathcal{F}, \mu)$, with the standard definition of positive elements. It is easy to see that in these examples (for $m \geqslant 2$) $c_+(E) = 2^{1-1/p}$. In particular, $c_+(\ell_1) = c_+(L^1(\mathcal{S}, \mathcal{F}, \mu)) = 1$, so in these spaces the norms $\| \, \|_+$ and $\| \, \|$ coincide.

Example 4.2. The examples in Example 4.1 are examples of Banach lattices, which also include many other important Banach spaces, see e.g. [30] or [22] for definition and further examples. In a Banach lattice E, every $x \in E$ has a decomposition $x = x_+ - x_-$ with $x_\pm \in E^+$ and $\|x_\pm\| \leqslant \|x\|$; hence $1 \leqslant c_+(E) \leqslant 2$.

Example 4.3. If we go beyond Banach lattices, then $c_+(E)$ may be arbitrarily large. A simple example is provided by $E = \mathbb{R}^2$ with usual positive cone (the first quadrant) and the norm $\|(x, y)\| := |x - y| + C|x + y|$ for a large constant C; then $\|(1, -1)\| = 2$ and $\|(1, -1)\|_+ = 2C + 2$. Thus $c_+(E) \geqslant C + 1$. (In fact, equality holds.)

4.1. **Positive tensor products.**

We are interested in decompositions of tensors using tensor products of positive elements only. If E is an ordered normed space, define in analogy with (3.2) and (3.12) the tensor norms

$$\|\mathbf{x}\|_{\pi,+} := \inf \left\{ \sum_{k=1}^{N} |a_k| \|x_{1k}\| \cdots \|x_{nk}\| : \mathbf{x} = \sum_{k=1}^{N} a_k x_{1k} \otimes \cdots \otimes x_{nk},\ x_{ik} \geqslant 0 \right\}. \tag{4.3}$$

on $E^{\otimes n}$, and

$$\|\mathbf{x}\|_{\pi,\mathsf{s},+} := \inf \left\{ \sum_{k=1}^{N} |a_k| \|x_k\|^n : \mathbf{x} = \sum_{k=1}^{N} a_k x_k^{\otimes n},\ x_k \geqslant 0 \right\}. \tag{4.4}$$

on $E^{\vee n}$; these norms are thus defined using only positive elements in the decompositions. For a symmetric tensor $\mathbf{x} \in E^{\vee n}$, we have in analogy with (3.11) also

$$\|\mathbf{x}\|_{\pi,+} = \inf \left\{ \sum_{k=1}^{N} |a_k| \|x_{1k}\| \cdots \|x_{nk}\| : \mathbf{x} = \sum_{k=1}^{N} a_k x_{1k} \vee \cdots \vee x_{nk},\ x_k \geqslant 0 \right\}. \tag{4.5}$$

It is perhaps not obvious that $\|\mathbf{x}\|_{\pi,\mathsf{s},+}$ always is finite, i.e., that there always exists a decomposition as in (4.4); this is part of Lemma 4.6 below.

As in (3.2), we may add E as another subscript in these notations, and write e.g. $\|\mathbf{x}\|_{\pi,+;E}$, when we want to indicate the space E explicitly.

We first note that $\|\ \|_{\pi,+}$ is an ordinary projective tensor power norm $\|\ \|_{\pi}$, but for the (in general) differently normed space E_+.

Lemma 4.4. *The norm $\|\ \|_{\pi,+;E}$ equals the norm $\|\ \|_{\pi;\,E_+}$ in $(E_+)^{\otimes n}_{\pi}$.*

Proof. If $x \geqslant 0$, then $\|x\|_+ = \|x\|$. Hence (4.3) implies that $\|\mathbf{x}\|_{\pi;\,E_+} \leqslant \|\mathbf{x}\|_{\pi,+;E}$.

Conversely, it suffices to consider $\mathbf{x} = x_1 \otimes \cdots \otimes x_n$ with $x_1, \dots, x_n \in E$. Let $\varepsilon > 0$, and choose $x_{i0}, x_{i1} \in E$ such that $x_i = x_{i0} - x_{i1}$ and $\|x_{i0}\| + \|x_{i1}\| \leqslant \|x_i\|_+ + \varepsilon$, see (4.1). Then,

$$\mathbf{x} = x_1 \otimes \cdots \otimes x_n = \sum_{j_1=0}^{1} \cdots \sum_{j_n=0}^{1} (-1)^{\sum_i j_i} x_{1j_1} \otimes \cdots \otimes x_{nj_n} \tag{4.6}$$

and thus

$$\begin{aligned} \|\mathbf{x}\|_{\pi,+;E} &\leqslant \sum_{j_1=0}^{1} \cdots \sum_{j_n=0}^{1} \|x_{1j_1}\| \cdots \|x_{nj_n}\| = \prod_{i=1}^{n} \bigl(\|x_{i0}\| + \|x_{i1}\|\bigr) \\ &\leqslant \prod_{i=1}^{n} \bigl(\|x_i\|_+ + \varepsilon\bigr). \end{aligned} \tag{4.7}$$

Letting $\varepsilon \to 0$ yields $\|\mathbf{x}\|_{\pi,+;E} \leqslant \prod_{i=1}^{n} \|x_i\|_+ = \|\mathbf{x}\|_{\pi;\,E_+}$. ☐

Remark 4.5. Lemma 4.4 does not extend to the symmetric tensor products and norms. For an example, let $E = \ell^2_1$, so $E_+ = E$ by Example 4.1; however, $\|(1,-1)^{\otimes 2}\|_{\pi,\mathsf{s}} = \|(1,-1)\|^2 = 4$ by (3.13), while $\|(1,-1)^{\otimes 2}\|_{\pi,\mathsf{s},+} = 8$ by (6.2) and (6.19) below.

Lemma 4.6. (i) *For every* $\mathbf{x} \in E^{\otimes n}$,

$$\|\mathbf{x}\|_\pi \leqslant \|\mathbf{x}\|_{\pi,+} \leqslant c_+(E)^n \|\mathbf{x}\|_\pi. \tag{4.8}$$

(ii) *There exists a constant* $\gamma(n)$ *(not depending on* E*) such that for every* $\mathbf{x} \in E^{\vee n}$,

$$\|\mathbf{x}\|_{\pi,s} \leqslant \|\mathbf{x}\|_{\pi,s,+} \leqslant \gamma(n)\|\mathbf{x}\|_{\pi,s;\,E_+} \leqslant \gamma(n) c_+(E)^n \|\mathbf{x}\|_{\pi,s}. \tag{4.9}$$

Proof. (i): The first inequality in (4.8) is trivial. Since the identity map $I : E \to E_+$ has norm $c_+(E)$, the identity map $I^{\otimes n} : E_\pi^{\otimes n} \to (E_+)_\pi^{\otimes n}$ has norm $c_+(E)^n$, see Theorem 3.12, which yields the second inequality by Lemma 4.4.

(ii): Again, the first inequality is trivial. Furthermore, the argument just given for (i) shows also that $I^{\vee n} : E_{\pi,\mathsf{s}}^{\vee n} \to (E_+)_{\pi,\mathsf{s}}^{\vee n}$ has norm $c_+(E)^n$, which yields the third inequality in (4.9).

For the second inequality, by (3.12), it suffices to consider a tensor power $\mathbf{x} = x^{\otimes n}$. Decompose $x = y - z$ with $y, z \geqslant 0$. Define, for $t \in \mathbb{R}$, the tensor $w(t) \in E^{\vee n}$ by

$$w(t) := (y + tz)^{\otimes n} = \bigl(x + (1+t)z\bigr)^{\otimes n} = \sum_{i=0}^{n} \binom{n}{i} (t+1)^{n-i} x^{\vee i} \vee z^{\vee(n-i)}, \tag{4.10}$$

where we have used the binomial theorem in the commutative tensor algebra $\bigcup_{n \geqslant 0} E^{\vee n}$. Note that $x^{\otimes n} = w(-1)$, and that for $t \geqslant 0$ we have $y + tz \geqslant 0$ and thus

$$\|w(t)\|_{\pi,\mathsf{s},+} \leqslant \|y + tz\|^n \leqslant (\|y\| + t\|z\|)^n. \tag{4.11}$$

Now suppose that μ is a finite signed measure on $[0, \infty)$ such that

$$\int_0^\infty (t+1)^j \,\mathrm{d}\mu(t) = \begin{cases} 1, & j = 0, \\ 0, & j = 1, \dots, n. \end{cases} \tag{4.12}$$

Then (4.10) yields

$$\int_0^\infty w(t) \,\mathrm{d}\mu(t) = x^{\vee n} = x^{\otimes n}. \tag{4.13}$$

Suppose further that μ is supported at a finite number of points, i.e., μ is a linear combination of Dirac measures $\sum_k \lambda_k \delta_{t_k}$. Then the integral in (4.13) is a linear combination $\sum_k \lambda_k w(t_k) = \sum_k \lambda_k (y + t_k z)^{\otimes n}$ and thus (4.4) yields

$$\begin{aligned} \|x^{\otimes n}\|_{\pi,\mathsf{s},+} &\leqslant \sum_k |\lambda_k| \|y + t_k z\|^n \leqslant \sum_k |\lambda_k| (\|y\| + t_k\|z\|)^n \\ &\leqslant \sum_k |\lambda_k| \max(1, t_k)^n (\|y\| + \|z\|)^n. \end{aligned} \tag{4.14}$$

Taking the infimum over all decompositions $x = y - z$ we obtain

$$\|x^{\otimes n}\|_{\pi,\mathsf{s},+} \leqslant \sum_k |\lambda_k| \max(1, t_k)^n \|x\|_+^n \tag{4.15}$$

This implies the second inequality in (4.9) with

$$\gamma(n) = \sum_k |\lambda_k| \max(1, t_k)^n. \tag{4.16}$$

It remains to show that such a μ exists. For this we choose $t_1 < \cdots < t_{n+1}$ arbitrarily in $[0, \infty)$. The equations (4.12) become the system of linear equations

$$\sum_{k=1}^{n+1} \lambda_k (t_k + 1)^j = \begin{cases} 1, & j = 0, \\ 0, & j = 1, \dots, n. \end{cases} \tag{4.17}$$

The coefficient matrix is the Vandermonde matrix with entries $(t_k + 1)^j$, $k = 1, \dots, n+1$ and $j = 0, \dots, n$; this matrix is non-singular and thus (4.17) has a solution. $\square$

The decompositions used in the proof above are in general not optimal. Optimal decompositions may be much harder to find; two non-trivial examples are given in (6.32) and (6.34) with (6.38).

From now on, we let $\gamma(n)$ denote the smallest possible constant such that (4.9) holds for all E and all $\mathbf{x} \in E^{\vee n}$. We will show that $\gamma(n) = 2^{n-1}$ in Theorem 6.1, but until this is proved, we regard $\gamma(n)$ as an unknown constant.

Lemma 4.6 shows that for any normed space E, $\| \|_\pi$ and $\| \|_{\pi,+}$ are equivalent norms on $E^{\otimes n}$, and $\| \|_{\pi,\mathsf{s}}$ and $\| \|_{\pi,\mathsf{s},+}$ are equivalent norms on $E^{\vee n}$. We use $E^{\otimes n}_{\pi,+}$, $E^{\vee n}_{\pi,+}$ and $E^{\vee n}_{\pi,\mathsf{s},+}$ to denote $E^{\otimes n}$ with the norm $\| \|_{\pi,+}$ and $E^{\vee n}$ with the norms $\| \|_{\pi,+}$ and $\| \|_{\pi,\mathsf{s},+}$, respectively.

Remark 4.7. In analogy with (3.3) and (3.13), it follows that for a positive elementary tensor product $\mathbf{x} = x_1 \otimes \cdots \otimes x_n$ with $x_1, \dots, x_n \geqslant 0$,

$$\|x_1 \otimes \cdots \otimes x_n\|_{\pi,+} = \|x_1 \otimes \cdots \otimes x_n\|_\pi = \|x_1\| \cdots \|x_m\|, \tag{4.18}$$

and for a positive elementary tensor power $\mathbf{x} = x^{\otimes n}$ with $x \geqslant 0$,

$$\|x^{\otimes n}\|_{\pi,\mathsf{s},+} = \|x^{\otimes n}\|_{\pi,\mathsf{s}} = \|x^{\otimes n}\|_{\pi,+} = \|x^{\otimes n}\|_\pi = \|x\|^n. \tag{4.19}$$

The norms $\| \|_{\pi,+}$ and $\| \|_{\pi,\mathsf{s},+}$ are the largest norms on $E^{\otimes n}$ and $E^{\vee n}$, respectively, that satisfy (4.18) and (4.19).

However, note that (for $n \geqslant 2$), (4.18) and (4.19) in general are false for general $x \in E$; hence $\| \|_{\pi,\mathsf{s},+}$ are not tensor norms in the usual sense. In fact, by Lemma 4.4 and (3.3) applied to E_+,

$$\|x_1 \otimes \cdots \otimes x_n\|_{\pi,+} = \prod_{i=1}^n \|x_i\|_+, \qquad x_1, \dots, x_n \in E. \tag{4.20}$$

Another counterexample for (4.19) is given by the same example $E = \ell_1^2$ and $\|(1,-1)^{\otimes 2}\|_{\pi,\mathsf{s},+} = 8$ as in Remark 4.5, given by (6.2) and (6.19) below; see also (4.29).

Remark 4.8. Let $B^+(E) := B(E) \cap E^+$, the positive part of the unit ball. In analogy with Remark 3.3, the unit balls $B(E^{\vee n}_{\pi,+})$ and $B(E^{\vee n}_{\pi,\mathsf{s},+})$ equal the closed convex hull of the sets $\{\pm x_1 \vee \cdots \vee x_n : x_1, \dots, x_n \in B^+(E)\}$ and $\{\pm x^{\vee n} : x \in B^+(E)\}$, respectively. Again, if $\dim(E) < \infty$, these equal the convex hulls (which already are closed); hence, the infima in (4.3)–(4.5) are attained when $\dim(E) < \infty$.

Remark 4.9. Similarly, in analogy with (3.14)–(3.15), it follows from (4.5) and (4.4) that for any linear operator $T : E^{\vee n}_{\pi} \to F$, where F is a normed space,

$$\|T\|_{E^{\vee n}_{\pi,+},F} = \sup\bigl\{\|T(x_1 \vee \cdots \vee x_n)\| : x_1, \dots, x_n \in B^+(E)\bigr\}. \tag{4.21}$$

and

$$\|T\|_{E^{\vee n}_{\pi,\mathsf{s},+},F} = \sup\bigl\{\|T(x^{\otimes n})\| : x \in B^+(E)\bigr\}. \tag{4.22}$$

Conversely, these properties characterize the norms $\| \|_{\pi,+}$ and $\| \|_{\pi,\mathsf{s},+}$ on $E^{\vee n}$.

Remark 4.10. Even if E is a Banach lattice, $\| \|_{\pi,+}$ and $\| \|_{\pi,\mathsf{s},+}$ are in general not lattice norms, i.e., in general $|\mathbf{x}| \leqslant |\mathbf{y}|$ does not imply $\|\mathbf{x}\| \leqslant \|\mathbf{y}\|$. For example, consider (cf. Remarks 4.5 and 4.7) $E = \ell^2_1$ and let $\mathbf{x} = (1,-1)^{\otimes 2} = \bigl(\begin{smallmatrix} 1 & -1 \\ -1 & 1 \end{smallmatrix}\bigr) \in E^{\otimes 2}$ and $\mathbf{y} = |\mathbf{x}| = \bigl(\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\bigr) = (1,1)^{\otimes 2}$. Then, $|\mathbf{x}| = \mathbf{y}$ but, see (6.2), (6.19) and (4.19), $\|\mathbf{x}\|_{\pi,\mathsf{s},+} = 8$ and $\|\mathbf{y}\|_{\pi,\mathsf{s},+} = 4$.

For Banach lattices E and F, Fremlin [11] defined a positive projective tensor norm $\| \|_{|\pi|}$ on $E \otimes F$ such that the completion is a Banach lattice. In particular, for a Banach lattice E, $\| \|_{|\pi|}$ is defined on $E^{\otimes n}$, and there is also a symmetric version $\| \|_{s,|\pi|}$ on $E^{\vee n}$, inroduced by Bu and Buskes [4]. It is easily seen that

$$\|\mathbf{x}\|_{|\pi|} = \inf\bigl\{\|\mathbf{y}\|_{\pi,+} : \mathbf{y} \geqslant |\mathbf{x}|\bigr\}, \tag{4.23}$$

$$\|\mathbf{x}\|_{s,|\pi|} = \inf\bigl\{\|\mathbf{y}\|_{\pi,\mathsf{s},+} : \mathbf{y} \geqslant |\mathbf{x}|\bigr\}. \tag{4.24}$$

Remark 4.11. A related notion of *non-negative rank* of a non-negative tensor $\mathbf{x}$, meaning the smallest N in a decomposition (4.3) with $a_k \geqslant 0$, has been studied by several authors, see e.g. Qi, Comon and Lim [25, 26] and the references there. Note, however, that we consider arbitrary $\mathbf{x}$ above, and do not require $a_k \geqslant 0$.

4.2. **Positive polarization constants.** In analogy with Corollary 3.10, we define $c_{\mathsf{s},+}(n,E)$, $c_{\mathsf{s};\,\mathsf{s},+}(n,E)$, and $c_{+;\,\mathsf{s},+}(n,E)$ as the norms of the identity map $E^{\vee n}_{\pi} \to E^{\vee n}_{\pi,\mathsf{s},+}$, $E^{\vee n}_{\pi,\mathsf{s}} \to E^{\vee n}_{\pi,\mathsf{s},+}$, and $E^{\vee n}_{\pi,+} \to E^{\vee n}_{\pi,\mathsf{s},+}$, respectively, i.e.,

$$c_{\mathsf{s},+}(n,E) := \sup_{\mathbf{x} \in E^{\vee n}} \frac{\|\mathbf{x}\|_{\pi,\mathsf{s},+}}{\|\mathbf{x}\|_{\pi}}, \tag{4.25}$$

$$c_{\mathsf{s};\,\mathsf{s},+}(n,E) := \sup_{\mathbf{x} \in E^{\vee n}} \frac{\|\mathbf{x}\|_{\pi,\mathsf{s},+}}{\|\mathbf{x}\|_{\pi,\mathsf{s}}}, \tag{4.26}$$

$$c_{+;\,\mathsf{s},+}(n,E) := \sup_{\mathbf{x}\in E^{\vee n}} \frac{\|\mathbf{x}\|_{\pi,\mathsf{s},+}}{\|\mathbf{x}\|_{\pi,+}}. \tag{4.27}$$

By (3.14), (3.15) and (4.5), it suffices to consider elementary tensors $\mathbf{x} = x_1 \vee \cdots \vee x_n$ in (4.25) and (4.27) and $\mathbf{x} = x^{\otimes n}$ in (4.26), i.e.,

$$c_{\mathsf{s},+}(n,E) = \sup_{x_1,\dots,x_n\in E} \frac{\|x_1 \vee \cdots \vee x_n\|_{\pi,\mathsf{s},+}}{\|x_1\|\cdots\|x_n\|}, \tag{4.28}$$

$$c_{\mathsf{s};\,\mathsf{s},+}(n,E) = \sup_{x\in E} \frac{\|x^{\otimes n}\|_{\pi,\mathsf{s},+}}{\|x\|^n}. \tag{4.29}$$

$$c_{+;\,\mathsf{s},+}(n,E) = \sup_{x_1,\dots,x_n\in E^+} \frac{\|x_1 \vee \cdots \vee x_n\|_{\pi,\mathsf{s},+}}{\|x_1\|\cdots\|x_n\|} \tag{4.30}$$

Since $\|x\| = \|x\|_+$ when $x \geqslant 0$, it follows from (4.4) that

$$\|\mathbf{x}\|_{\pi,\mathsf{s},+;\,E} = \|\mathbf{x}\|_{\pi,\mathsf{s},+;\,E_+}, \tag{4.31}$$

and thus Lemma 4.4 implies

$$c_{+;\,\mathsf{s},+}(n,E) = c_{\mathsf{s},+}(n,E_+). \tag{4.32}$$

We will therefore usually ignore $c_{+;\,\mathsf{s},+}$, and leave it to the reader.

We may also consider the identity maps $E_\pi^{\otimes n} \to E_{\pi,+}^{\otimes n}$ and $E_\pi^{\vee n} \to E_{\pi,+}^{\vee n}$, but we then do not need any new notation since it was shown in the proof of Lemma 4.6 that both have norm $c_+(E)^n$, i.e.,

$$\sup_{\mathbf{x}\in E^{\otimes n}} \frac{\|\mathbf{x}\|_{\pi,+}}{\|\mathbf{x}\|_\pi} = \sup_{\mathbf{x}\in E^{\vee n}} \frac{\|\mathbf{x}\|_{\pi,+}}{\|\mathbf{x}\|_\pi} = c_+(E)^n. \tag{4.33}$$

Note also that the inverses of all identity maps considered here have norm 1. Thus, or directly from the definitions, $c_{\mathsf{s},+}(n,E) \geqslant 1$, $c_{\mathsf{s};\,\mathsf{s},+}(n,E) \geqslant 1$, $c_{+;\,\mathsf{s},+}(n,E) \geqslant 1$, and $c_+(E) \geqslant 1$.

Several inequalities between the different polarization constants follow directly from the definitions and Corollary 3.10, by considering compositions of the identity maps. For example,

$$\max\bigl(c_\mathsf{s}(n,E), c_{\mathsf{s};\,\mathsf{s},+}(n,E)\bigr) \leqslant c_{\mathsf{s},+}(n,E) \leqslant c_\mathsf{s}(n,E) c_{\mathsf{s};\,\mathsf{s},+}(n,E). \tag{4.34}$$

Similarly, by (4.26) and (4.9),

$$1 \leqslant c_{\mathsf{s};\,\mathsf{s},+}(n,E) \leqslant \gamma(n) c_+(E)^n. \tag{4.35}$$

Moreover, using (4.31), $\gamma(n)$ is the smallest constant such that $\|\mathbf{x}\|_{\pi,\mathsf{s},+;\,E_+} \leqslant \gamma(n)\|\mathbf{x}\|_{\pi,\mathsf{s};\,E_+}$ for all normed spaces E and all $\mathbf{x} \in E^{\vee n}$, i.e.,

$$\gamma(n) = \sup_E c_{\mathsf{s};\,\mathsf{s},+}(n,E_+). \tag{4.36}$$

Using (4.35), we thus also have

$$\gamma(n) = \sup_E \frac{c_{\mathsf{s};\,\mathsf{s},+}(n,E)}{c_+(E)^n}. \tag{4.37}$$

Example 4.12. For $n = 1$, $E^{\vee 1} = E^{\otimes 1} = E$. It is obvious that the norms $\|x\|_\pi = \|x\|_{\pi,\mathsf{s}} = \|x\|$ for any $x \in E$; furthermore, see (4.4) and (4.1), $\|x\|_{\pi,\mathsf{s},+} = \|x\|_+$. In particular, by (4.25)–(4.26) and (4.2),

$$c_{\mathsf{s},+}(1, E) = c_{\mathsf{s};\,\mathsf{s},+}(1, E) = c_+(E). \tag{4.38}$$

Thus $\gamma(1) = 1$.

We note also that the definitions (4.3)–(4.4) and (4.20) imply $\|x^{\otimes n}\|_{\pi,\mathsf{s},+} \geqslant \|x^{\otimes n}\|_{\pi,+} = \|x\|_+^n$. Thus, using (4.34), (4.29), and (4.2),

$$c_{\mathsf{s},+}(n, E) \geqslant c_{\mathsf{s};\,\mathsf{s},+}(n, E) \geqslant c_+(E)^n. \tag{4.39}$$

Example 4.13. If H is a Hilbert space H, then (4.34) and Example 3.11 yield

$$c_{\mathsf{s};\,\mathsf{s},+}(n, H) = c_{\mathsf{s},+}(n, H). \tag{4.40}$$

We will see in Examples 8.1 and 8.2 that the result by Banach [2] in Example 3.11 does not extend to the positive tensor norms, i.e., in general $\|\mathbf{x}\|_{\pi,+;\,H} \neq \|\mathbf{x}\|_{\pi,\mathsf{s},+;\,H}$, even when H is ℓ_2^2 with the usual ordering.

Furthermore, Example 8.2 also shows that for this example $c_{\mathsf{s},+}(2, \ell_2^2) = c_{\mathsf{s};\,\mathsf{s},+}(2, \ell_2^2) = 3$, and thus the second inequality in (4.39) is strict; recall that $c_+(\ell_2^2)^2 = 2$ by Example 4.1.

4.3. **Multilinear forms on ordered spaces.** If E is an ordered normed space, define for an n-linear form $L \in \mathcal{L}(^nE; \mathbb{R})$, in analogy with (2.1) and (2.11),

$$\|L\|_+ := \sup\bigl\{|L(x_1, \dots, x_n)| : \|x_1\| = \dots = \|x_n\| \leqslant 1,\ x_1, \dots, x_n \geqslant 0\bigr\}. \tag{4.41}$$

$$\|L\|_{\Delta,+} := \|\hat{L}\|_+ = \sup\bigl\{|L(x, \dots, x)| : \|x\| \leqslant 1,\ x \geqslant 0\bigr\}. \tag{4.42}$$

Then $\|L\|_+$ equals the norm $\|L\|_{\pi,+}^*$ in the dual of $E_{\pi,+}^{\otimes n}$. If L is symmetric, then also $\|L\|_+$ equals the norm $\|L\|_{\pi,+}^*$ in the dual of $E_{\pi,+}^{\vee n}$, and $\|L\|_{\Delta,+}$ equals the norm $\|L\|_{\pi,\mathsf{s},+}^*$ in the dual of $E_{\pi,\mathsf{s},+}^{\vee n}$.

By duality, $c_{\mathsf{s},+}(n, E)$ and $c_{\mathsf{s};\,\mathsf{s},+}(n, E)$ equal the norms of the identity operators $(E_{\pi,\mathsf{s},+}^{\vee n})^* \to (E_\pi^{\vee n})^*$ and $(E_{\pi,\mathsf{s},+}^{\vee n})^* \to (E_{\pi,\mathsf{s}}^{\vee n})^*$, respectively. Hence, using (3.17)–(3.18),

$$c_{\mathsf{s},+}(n, E) = \sup_{L \in \mathcal{L}_\mathsf{s}(^nE; \mathbb{K})} \frac{\|L\|}{\|L\|_{\Delta,+}}, \tag{4.43}$$

$$c_{\mathsf{s};\,\mathsf{s},+}(n, E) = \sup_{L \in \mathcal{L}_\mathsf{s}(^nE; \mathbb{K})} \frac{\|L\|_\Delta}{\|L\|_{\Delta,+}}. \tag{4.44}$$

4.4. **Functorial properties.** We have functorial properties similar to the ones in Theorem 3.12, but now only for positive operators.

Theorem 4.14. *If E and F are ordered normed spaces and $T : E \to F$ is a positive bounded linear operator, then $T^{\otimes n} : E_{\pi,+}^{\otimes n} \to F_{\pi,+}^{\otimes n}$, $T^{\vee n} : E_{\pi,+}^{\vee n} \to F_{\pi,+}^{\vee n}$ and $T^{\vee n} : E_{\pi,\mathsf{s},+}^{\vee n} \to F_{\pi,\mathsf{s},+}^{\vee n}$ all have norm $\|T\|^n$.*

Proof. An immediate consequence of the definitions (4.3) and (4.4) together with (4.19). □

It follows that there is a version of Theorem 3.13 for $c_{\mathsf{s},+}(n,E)$ and $c_{\mathsf{s};\,\mathsf{s},+}(n,E)$, but more restrictive.

Theorem 4.15. *Theorem 3.13 extends to both $c_{\mathsf{s},+}$ and $c_{\mathsf{s};\,\mathsf{s},+}$, with the following additional assumptions.*

(i) *The quotient mapping $E \to F$ is positive.*

(ii) *The projection $P : E \to F$ is positive.*

(iii) *A positive isomorphism $T : E \to F$ exists, and the Banach–Mazur distance $d(E,F)$ is replaced by the "positive Banach–Mazur distance" $d_+(E,F) := \inf\{\|T\|\|T^{-1}\|\}$, taking the infimum over all positive isomorphisms $T : E \to F$ only.*

The positive Banach–Mazur distance defined here is in general not symmetric. It can be replaced by the (generally larger) symmetric "ordered Banach–Mazur distance", defined by taking the infimum over order isomorphisms $T : E \to F$ only.

The three parts of Theorem 4.15 are proved by the same argument, which we state more generally as a lemma.

Lemma 4.16. *Suppose that E and F are normed spaces, that $T : E \to F$ is a positive bounded linear operator onto F, and that A is a constant such that for every $x \in F$ and $\varepsilon > 0$, there exists $y \in E$ with $Ty = x$ and $\|y\| \leqslant (A+\varepsilon)\|x\|$. Then $c_{\mathsf{s},+}(n,F) \leqslant (A\|T\|)^n c_{\mathsf{s},+}(n,E)$ and $c_{\mathsf{s};\,\mathsf{s},+}(n,F) \leqslant (A\|T\|)^n c_{\mathsf{s};\,\mathsf{s},+}(n,E)$.*

Proof. Consider first $c_{\mathsf{s},+}$. We use (4.28). Let $x_1,\dots,x_n \in B(F)$ and let $\varepsilon > 0$. By assumption, there exist $y_1,\dots,y_n \in E$ such that $\|y_i\| \leqslant A+\varepsilon$ and $Ty_i = x_i$. Then $T^{\vee n}(y_1 \vee \dots \vee y_n) = x_1 \vee \dots \vee x_n$ and thus, using Theorem 4.14 and (4.28),

$$\begin{aligned}\|x_1 \vee \dots \vee x_n\|_{\pi,\mathsf{s},+} &\leqslant \|T\|^n \|y_1 \vee \dots \vee y_n\|_{\pi,\mathsf{s},+} \leqslant \|T\|^n c_{\mathsf{s},+}(n,E)\|y_1\|\cdots\|y_n\| \\ &\leqslant \|T\|^n c_{\mathsf{s},+}(n,E)(A+\varepsilon)^n. \end{aligned} \tag{4.45}$$

Now let $\varepsilon \to 0$ and use (4.28) again.

The proof for $c_{\mathsf{s};\,\mathsf{s},+}$ is the same, now using (4.29) and taking $x_i = x$ and $y_i = y$ for all i. □

Proof of Theorem 4.15. We apply Lemma 4.16 as follows:

(i) Let T be the quotient mapping $E \to F$. Then $\|T\| = 1$, and, by definition of the quotient norm, the assumption of the lemma holds with $A = 1$.

(ii) Let $T = P$ and let $A = 1$; we can take $y = x$.

(iii) If $T : E \to F$ is a positive isomorphism, we take $y = T^{-1}x$ and the assumption of the lemma holds with $A = \|T^{-1}\|$. Thus $c_{\mathsf{s},+}(n,F) \leqslant (\|T\|\|T^{-1}\|)^n c_{\mathsf{s},+}(n,E)$ and similarly for $c_{\mathsf{s};\,\mathsf{s},+}$. Take the infimum over T. □

5. ℓ_1 IS EXTREME

We have seen in Example 2.7 that ℓ_1 and ℓ_1^n are extremal among all normed spaces for $c_{\mathsf{s}}(n,E)$. The next two theorems show that they are extremal also for $c_{\mathsf{s},+}$ and $c_{\mathsf{s};\,\mathsf{s},+}$, provided we compensate for $c_+(E)$; recall that Example 4.3 shows that $c_+(E)$ may be arbitrarily large, so (4.38) and (4.39) show that $\sup_E c_{\mathsf{s},+}(n,E) = \sup_E c_{\mathsf{s};\,\mathsf{s},+}(n,E) = \infty$ for any $n \geqslant 1$.

Theorem 5.1. *If $n \leqslant m \leqslant \infty$, then $c_{\mathsf{s},+}(n,\ell_1^m) = \kappa(n)$, where*

$$\kappa(n) := \|e_1 \vee \cdots \vee e_n\|_{\pi,\mathsf{s},+;\,\ell_1} = \|e_1 \vee \cdots \vee e_n\|_{\pi,\mathsf{s},+;\,\ell_1^n}. \tag{5.1}$$

Furthermore,

$$\sup_E \frac{c_{\mathsf{s},+}(n,E)}{c_+(E)^n} = \sup_{E\colon c_+(E)=1} c_{\mathsf{s},+}(n,E) = c_{\mathsf{s},+}(n,\ell_1) = c_{\mathsf{s},+}(n,\ell_1^n) = \kappa(n). \tag{5.2}$$

Proof. First, note that the natural injection $\ell_1^n \to \ell_1$ and projection $\ell_1 \to \ell_1^n$ have norm 1, and that this implies the equality of the two tensor norms in (5.1) by Theorem 4.14.

Let $x_1,\dots,x_n \in E$ with $x_i \geqslant 0$ and $\|x_i\| = 1$. Define a linear map $T : \ell_1^n \to E$ by $Te_i := x_i$. Then T is positive and $\|T\| = 1$, and thus, by Theorem 4.14, $T^{\vee n} : (\ell_1^n)^{\vee n}_{\pi,\mathsf{s},+} \to E^{\vee n}_{\pi,\mathsf{s},+}$ has norm 1. Hence,

$$\begin{aligned}\|x_1 \vee \cdots \vee x_n\|_{\pi,\mathsf{s},+;\,E} &= \|T^{\vee n}(e_1 \vee \cdots \vee e_n)\|_{\pi,\mathsf{s},+;\,E}\\ &\leqslant \|e_1 \vee \cdots \vee e_n\|_{\pi,\mathsf{s},+;\,\ell_1^n} = \kappa(n).\end{aligned} \tag{5.3}$$

It now follows from (4.5) that for any $\mathbf{x} \in E^{\vee n}$,

$$\|\mathbf{x}\|_{\pi,\mathsf{s},+} \leqslant \kappa(n)\|\mathbf{x}\|_{\pi,+}. \tag{5.4}$$

Combining (5.4) and (4.8) yields $\|\mathbf{x}\|_{\pi,\mathsf{s},+} \leqslant \kappa(n)\|\mathbf{x}\|_{\pi,+} \leqslant \kappa(n)c_+(E)^n\|\mathbf{x}\|_\pi$ and thus

$$c_{\mathsf{s},+}(n,E) \leqslant \kappa(n)c_+(E)^n. \tag{5.5}$$

It follows immediately from (5.5) and $c_+(\ell_1^m) = 1$ that $c_{\mathsf{s},+}(n,\ell_1^m)$ and all terms in (5.2) are at most $\kappa(n)$.

Conversely, if $n \leqslant m \leqslant \infty$, then, using the injection $\ell_1^m \to \ell_1$,

$$\begin{aligned}\kappa(n) &= \|e_1 \vee \cdots \vee e_n\|_{\pi,\mathsf{s},+;\,\ell_1} \leqslant \|e_1 \vee \cdots \vee e_n\|_{\pi,\mathsf{s},+;\,\ell_1^m}\\ &\leqslant c_{\mathsf{s},+}(\ell_1^m)\|e_1 \vee \cdots \vee e_n\|_{\pi;\,\ell_1^m} = c_{\mathsf{s},+}(\ell_1^m).\end{aligned} \tag{5.6}$$

Hence $c_{\mathsf{s},+}(\ell_1^m) = \kappa(n)$. Furthermore, (5.6) implies that each term in (5.2) is at least $\kappa(n)$, so equalities holds. ☐

Example 5.2. We show that $\kappa(2) = 3$. This can be shown using the general results (7.23) and (7.33) in Remark 7.8 and Section 7.3, but we give a direct proof.

For an upper bound, we use the decomposition

$$e_1 \vee e_2 = 2(\tfrac12 e_1 + \tfrac12 e_2)^{\otimes 2} - \tfrac12 e_1^{\otimes 2} - \tfrac12 e_2^{\otimes 2}. \tag{5.7}$$

For a lower bound, we consider the linear map $L : (\ell_1^2)^{\vee 2} \to \mathbb{R}$ given by $e_1^* \otimes e_1^* + e_2^* \otimes e_2^* - 6e_1^* \vee e_2^*$, i.e., $\left(\begin{smallmatrix} a & b \\ b & c \end{smallmatrix}\right) \mapsto a + c - 6b$. A positive unit vector in ℓ_1^2 is $(x, 1-x)$ for some $x \in [0,1]$, and

$$L\bigl((x,1-x)^{\otimes 2}\bigr) = x^2 + (1-x)^2 - 6x(1-x) = 1 - 8x(1-x). \tag{5.8}$$

Since $0 \leqslant x(1-x) \leqslant \frac{1}{4}$, $|L\bigl((x,1-x)^{\otimes 2}\bigr)| \leqslant 1$, and thus, by (4.22), $\|L\|^*_{\pi,\mathsf{s},+} \leqslant 1$. Furthermore, $L(e_1 \vee e_2) = -3$. Hence $\|e_1 \vee e_2\|_{\pi,\mathsf{s},+} \geqslant 3$.

Consequently,

$$\kappa(2) = \|e_1 \vee e_2\|_{\pi,\mathsf{s},+;\,\ell_1} = 3. \tag{5.9}$$

We study the constant $\kappa(n)$ further in Section 7, where it plays an important role.

Theorem 5.3. *If* $2 \leqslant m \leqslant \infty$, *then* $c_{\mathsf{s};\,\mathsf{s},+}(n, \ell_1^m) = \gamma(n)$. *Thus,*

$$\sup_E \frac{c_{\mathsf{s};\,\mathsf{s},+}(n,E)}{c_+(E)^n} = \sup_{E:\,c_+(E)=1} c_{\mathsf{s};\,\mathsf{s},+}(n,E) = c_{\mathsf{s};\,\mathsf{s},+}(n,\ell_1) = c_{\mathsf{s};\,\mathsf{s},+}(n,\ell_1^2) = \gamma(n). \tag{5.10}$$

We will find the explicit value 2^{n-1} in Theorem 6.1.

Proof. Since $c_+(\ell_1^m) = 1$, $c_{\mathsf{s};\,\mathsf{s},+}(n,\ell_1^m) \leqslant \gamma(n)$ by (4.35).

Conversely, suppose that $x = y - z$ with $y, z \in E^+$. Let $y_0 := y/\|y\|$ and $z_0 := z/\|z\|$ (with $0/0 := 0$). Further, assuming $m \geqslant 2$, let $u := (\|y\|, -\|z\|, 0, \dots) \in \ell_1^m$; then $\|u\| = \|y\| + \|z\|$.

Define the linear map $T : \ell_1^m \to E$ by $T(a_1, a_2, \dots) = a_1 y_0 + a_2 z_0$. Then $T(u) = y - z = x$. Furthermore, T has norm (at most) 1 and maps positive elements to positive, and therefore by Theorem 4.14, $T^{\otimes n}$ maps $(\ell_1^m)^{\vee n}_{\pi,\mathsf{s},+}$ into $E^{\vee n}_{\pi,\mathsf{s},+}$ with norm at most 1. Consequently, recalling (3.13),

$$\begin{aligned}\|x^{\otimes n}\|_{\pi,\mathsf{s},+;\,E} &= \|T^{\otimes n} u^{\otimes n}\|_{\pi,\mathsf{s},+;\,E} \leqslant \|u^{\otimes n}\|_{\pi,\mathsf{s},+;\,\ell_1^m} \leqslant c_{\mathsf{s};\,\mathsf{s},+}(n,\ell_1^m)\|u^{\otimes n}\|_{\pi,\mathsf{s};\,\ell_1^m} \\ &= c_{\mathsf{s};\,\mathsf{s},+}(n,\ell_1^m)\|u\|^n = c_{\mathsf{s};\,\mathsf{s},+}(n,\ell_1^m)(\|y\| + \|z\|)^n. \end{aligned} \tag{5.11}$$

Taking the infimum over all decompositions $x = y - z$ with $y, z \in E^+$ yields

$$\|x^{\otimes n}\|_{\pi,\mathsf{s},+} \leqslant c_{\mathsf{s};\,\mathsf{s},+}(n,\ell_1^m)\|x\|_+^n. \tag{5.12}$$

This holds for every $x \in E$, and hence, by (4.29) and (4.31),

$$c_{\mathsf{s};\,\mathsf{s},+}(n,E_+) \leqslant c_{\mathsf{s};\,\mathsf{s},+}(n,\ell_1^m). \tag{5.13}$$

This holds for every normed space E, and thus (4.36) shows $\gamma(n) \leqslant c_{\mathsf{s};\,\mathsf{s},+}(n,\ell_1^m)$. Hence, each term in (5.10) is at least $\gamma(n)$. On the other hand, $c_{\mathsf{s};\,\mathsf{s},+}(n,\ell_1^m)$ and all terms in (5.10) are at most $\gamma(n)$ by (4.35). Hence, equalities hold. $\square$

Remark 5.4. In analogy with (2.22), we can define

$$c_{\mathsf{s},+}(E) := \limsup_{n\to\infty} c_{\mathsf{s},+}(n,E)^{1/n}, \tag{5.14}$$

$$c_{\mathsf{s};\,\mathsf{s},+}(E) := \limsup_{n\to\infty} c_{\mathsf{s};\,\mathsf{s},+}(n,E)^{1/n}. \tag{5.15}$$

By (4.34), (2.23), (4.35) and Theorem 6.1 below,

$$1 \leqslant c_{\mathsf{s};\,\mathsf{s},+}(E) \leqslant c_{\mathsf{s},+}(E) \leqslant c_{\mathsf{s}}(E)c_{\mathsf{s};\,\mathsf{s},+}(E) \leqslant ec_{\mathsf{s};\,\mathsf{s},+}(E), \tag{5.16}$$

$$c_{\mathsf{s};\,\mathsf{s},+}(E) \leqslant 2c_+(E). \tag{5.17}$$

For example, by Theorem 6.1, $c_{\mathsf{s};\,\mathsf{s},+}(\ell_1^m) = 2$ for $2 \leqslant m \leqslant \infty$. By (5.16), $2 \leqslant c_{\mathsf{s},+}(\ell_1^m) \leqslant 2e$; we do not know the exact value.

6. The value of $\gamma(n)$

The proof of Lemma 4.6 yields an upper bound for $\gamma(n)$ in (4.16)–(4.17). However, it seems difficult to evaluate this exactly in general, and we do not know whether this method yields an upper bound is optimal. We thus find $\gamma(n)$ by a different method, using Theorem 5.3. (This gives another proof of Lemma 4.6(ii).)

Theorem 6.1. (i) *For $n \geqslant 1$, $\gamma(n) = 2^{n-1}$.*
(ii) *If $2 \leqslant m \leqslant \infty$ and $n \geqslant 1$, then $c_{\mathsf{s};\,\mathsf{s},+}(n, \ell_1^m) = 2^{n-1}$.*

Proof. By Theorem 5.3, $\gamma(n) = c_{\mathsf{s};\,\mathsf{s},+}(n, \ell_1^m)$, for any $m \geqslant 2$. Hence, the two parts are equivalent, and it suffices to prove (ii) with $m = 2$. Thus, let $E = \ell_1^2$ and use (4.29), which yields

$$\gamma(n) = c_{\mathsf{s};\,\mathsf{s},+}(n, \ell_1^2) = \sup_{a,b\in\mathbb{R}} \frac{\|(a,b)^{\otimes n}\|_{\pi,\mathsf{s},+}}{(|a|+|b|)^n}. \tag{6.1}$$

Fix $n \geqslant 1$ and write, for convenience,

$$\psi(a,b) := \|(a,b)^{\otimes n}\|_{\pi,\mathsf{s},+;\,\ell_1^2}. \tag{6.2}$$

Since $-(a,b) = (-a,-b)$, it suffices to consider $a \geqslant 0$. Obviously, if $a, b \geqslant 0$, then $(a,b) \in (\ell_1^2)_+$ and thus, by (4.19),

$$\psi(a,b) := \|(a,b)\|^n = (a+b)^n, \qquad a, b \geqslant 0. \tag{6.3}$$

Hence, the interesting case is $a > 0 > b$. However, we continue to consider general $a, b \in \mathbb{R}$.

The unit vectors in $(\ell_1^2)_+$ are $(x, 1-x)$, $x \in [0,1]$. Consequently, the definition (4.4) can be written as

$$\psi(a,b) := \|(a,b)^{\otimes n}\|_{\pi,\mathsf{s},+} := \inf \|\mu\| = \inf \int_0^1 |\mathrm{d}\mu|(x), \tag{6.4}$$

taking the infimum over all signed measures of the type $\mu = \sum_{k=1}^N a_k\delta_{x_k}$ on $[0,1]$ such that

$$\int_0^1 (x, 1-x)^{\otimes n}\,\mathrm{d}\mu(x) = (a,b)^{\otimes n}. \tag{6.5}$$

In other words, we take the infimum over all signed measures with finite support in $[0,1]$ that satisfy (6.5). On the other hand, for any signed measure on $[0,1]$,

$$\Bigl\|\int_0^1 (x,1-x)^{\otimes n}\,\mathrm{d}\mu(x)\Bigr\|_{\pi,\mathsf{s},+} \leqslant \int_0^1 \bigl\|(x,1-x)^{\otimes n}\bigr\|_{\pi,\mathsf{s},+}|\,\mathrm{d}\mu|(x) = \int_0^1 |\,\mathrm{d}\mu|(x) \tag{6.6}$$

since the integral exists as a Bochner integral in $(\ell_1^2)^{\vee n}_{\pi,\mathsf{s},+}$. (Recall that the spaces are finite-dimensional, so there is no problem with convergence.) Consequently, we can just as well take the infima in (6.4) over all signed measures μ on $[0,1]$ satisfying (6.5).

Expanding the tensor products in (6.5) in $(\ell_1^2)^{\otimes n}$, we see that (6.5) is equivalent to the system of equations

$$\int_0^1 x^{n-k}(1-x)^k\,\mathrm{d}\mu(x) = a^{n-k}b^k, \qquad k=0,\dots,n. \tag{6.7}$$

The coefficients of the $n+1$ polynomials $q_k(x) := x^{n-k}(1-x)^k$, $k=0,\dots,n$, form a triangular matrix which is non-singular; consequently these polynomials form a basis in the $(n+1)$-dimensional space $P_{\leqslant n}$ of polynomials (of a real variable) of degree at most n. Hence, there exists a unique linear functional $\chi_{a,b}$ on $P_{\leqslant n}$ such that

$$\chi_{a,b}(q_k) = a^{n-k}b^k, \qquad k=0,\dots,n, \tag{6.8}$$

and (6.7) is equivalent to $\int_0^1 q_k(x)\,\mathrm{d}\mu(x) = \chi_{a,b}(q_k)$, $k=0,\dots,n$, and thus to

$$\int_0^1 p(x)\,\mathrm{d}\mu(x) = \chi_{a,b}(p), \qquad p\in P_{\leqslant n}. \tag{6.9}$$

For a compact interval $[c,d]\subset\mathbb{R}$, let $C[c,d]$ be the standard space of (real) continuous functions on $[c,d]$ with the norm

$$\|f\| := \sup_{x\in[c,d]} |f(x)|, \tag{6.10}$$

and let $P_{\leqslant n}[c,d]$ denote $P_{\leqslant n}$ regarded as a subspace of $C[c,d]$, i.e., equipped with the norm (6.10). The dual space of $C[c,d]$ is the space of signed measures on $[c,d]$, with the total variation norm as in (6.4). Hence (6.4) and (6.9) yield

$$\psi(a,b) = \inf\bigl\{\|\mu\|_{C[0,1]^*} : \mu(p) = \chi_{a,b}(p) \text{ for } p\in P_{\leqslant n}[0,1]\bigr\}, \tag{6.11}$$

which by the Hahn–Banach theorem yields

$$\psi(a,b) = \|\chi_{a,b}\|_{P_{\leqslant n}[0,1]^*}. \tag{6.12}$$

We next identify $\chi_{a,b}$. The definition (6.8) and the binomial theorem yield, for $k=0,\dots,n$,

$$\chi_{a,b}\bigl(x^{n-k}\bigr) = \chi_{a,b}\bigl(x^{n-k}(x+1-x)^k\bigr) = \sum_{j=0}^{k}\binom{k}{j}\chi_{a,b}\bigl(x^{n-k+k-j}(1-x)^j\bigr)$$

$$
\begin{aligned}
&= \sum_{j=0}^{k} \binom{k}{j} a^{n-k+k-j} b^j = a^{n-k}(a+b)^k \\
&= (a+b)^n \Bigl(\frac{a}{a+b}\Bigr)^{n-k}, \qquad (6.13)
\end{aligned}
$$

where the last equality assumes that $a+b \neq 0$. Consequently, if $a+b \neq 0$, then

$$\chi_{a,b}\bigl(p\bigr) = (a+b)^n p\Bigl(\frac{a}{a+b}\Bigr) \qquad (6.14)$$

for the monomials $p(x) = x^{n-k}$, and thus for all $p \in P_{\leqslant n}$. It can also be seen immediately that (6.14) defines a linear functional on $P_{\leqslant n}$ that satisfies (6.8). Hence, in this case $\chi_{a,b}$ is essentially a point evaluation at $a/(a+b)$, and $\psi(a,b)$ is by (6.12) given by the optimization problem

$$\psi(a,b) = |a+b|^n \sup\Bigl\{\Bigl|p\Bigl(\frac{a}{a+b}\Bigr)\Bigr| : \max_{x\in[0,1]} |p(x)| = 1\Bigr\}, \qquad a+b\neq 0. \quad (6.15)$$

Note that if $a,b \geqslant 0$ (with $a+b>0$), then $a/(a+b) \in [0,1]$, so the supremum in (6.11) is trivially 1, and thus $\psi(a,b) = (a+b)^n$, as seen directly in (6.3). In contrast, in the case $a > 0 > b$, $a/(a+b) \notin [0,1]$, so (6.15) becomes an extrapolation problem.

In the case $a+b=0$, (6.13) yields instead $\chi_{a,b}(x^{n-k}) = 0$ for $k \geqslant 1$ and $\chi_{a,b}(x^n) = a^n$. Hence, letting $[x^k]p(x)$ denote the coefficient of x^k in the polynomial $p(x)$,

$$\chi_{a,-a}\bigl(p(x)\bigr) = a^n [x^n] p(x). \qquad (6.16)$$

In other words, apart from a constant factor, $\chi_{a,-a}$ extracts the coefficient of x^n. (This can also be seen as a limiting case of (6.14), with $a/(a+b) \to \infty$.)

We consider the two cases separately, beginning with the case $b=-a$. By homogeneity, it suffices to consider $a=1$. By (6.12) and (6.16),

$$\psi(1,-1) = \bigl\|p \mapsto [x^n]p(x)\bigr\|_{P_{\leqslant n}[0,1]^*}. \qquad (6.17)$$

The mapping $p(x) \mapsto p(2x-1)$ is an isometric bijection of $P_{\leqslant n}[-1,1]$ onto $P_{\leqslant n}[0,1]$. Since $[x^n]p(2x-1) = 2^n[x^n]p(x)$, it follows that we have

$$\psi(1,-1) = 2^n \bigl\|p \mapsto [x^n]p(x)\bigr\|_{P_{\leqslant n}[-1,1]^*}. \qquad (6.18)$$

We thus want to find the largest possible coefficient of x^n for a polynomial of degree n that is bounded by 1 on $[-1,1]$; equivalently, we want to find the polynomial $p(x)$ with leading coefficient x^n such that $\|p\|_{C[-1,1]} = \sup_{-1,1} |p(x)|$ is minimal. This is a classical problem in approximation theory, which is solved by a multiple of the Chebyshev polynomial $T_n(x) := \cos(n \arccos x)$, see e.g. [23, 18.38(i)] or Rivlin [27, Theorem 2.1]. Since T_n has norm 1 in $P_{\leqslant n}[-1,1]$ and its leading coefficient is 2^{n-1}, it follows that $p \mapsto [x^n]p(x)$ has norm 2^{n-1} on $P_{\leqslant n}[-1,1]$, and thus (6.18) yields

$$\psi(1,-1) = 2^{2n-1}. \qquad (6.19)$$

Consequently, (6.1) yields

$$\gamma(n) \geqslant \frac{\psi(1,-1)}{2^n} = 2^{n-1}. \tag{6.20}$$

In order to see that equality holds in (6.20), we now consider the case $a+b \neq 0$, where we have shown (6.15). It suffices to consider the case $|a| > |b|$ and $a > 0 > b$; then $\frac{a}{a+b} > 1$. We transfer again to $P_{\leqslant n}[-1,1]$ by the mapping $p(x) \mapsto p(2x-1)$ and see that $\chi_{a,b}$ in (6.14) then corresponds to

$$p \mapsto (a+b)^n p\Bigl(2\frac{a}{a+b} - 1\Bigr) = (a+b)^n p\Bigl(\frac{a-b}{a+b}\Bigr). \tag{6.21}$$

Let $\xi := \frac{a-b}{a+b} > 1$. The problem is now to maximize $p(\xi)$ for $p \in P_{\leqslant n}$ with $\sup_{-1\leqslant x\leqslant 1} |p(x)| \leqslant 1$. Again, the (unique) extremal polynomial is the Chebyshev polynomial $T_n(x)$, see [27, 2.7.1]; hence (6.12) and (6.21) yield

$$\psi(a,b) = (a+b)^n T_n\Bigl(\frac{a-b}{a+b}\Bigr), \qquad a > 0 > b \text{ and } a+b > 0. \tag{6.22}$$

Finally, we note that if $x > 1$ and $y := \operatorname{arccosh} x$, then $T_n(x) = T_n(\cosh y) = \cosh(ny)$, and thus

$$T_n(x) = \frac{1}{2}\bigl(e^{ny} + e^{-ny}\bigr) \leqslant \frac{1}{2}\bigl(e^{y} + e^{-y}\bigr)^n = 2^{n-1}x^n. \tag{6.23}$$

Consequently, (6.22) implies, for $a > 0 > b$ and $a+b > 0$,

$$\psi(a,b) \leqslant (a+b)^n 2^{n-1}\Bigl(\frac{a-b}{a+b}\Bigr)^n = 2^{n-1}(a-b)^n = 2^{n-1}(|a|+|b|)^n. \tag{6.24}$$

It follows from (6.24) and (6.19) (which is a limiting case that also follows from (6.22) by continuity), together with the trivial case $a, b \geqslant 0$ treated earlier, that $\psi(a,b) \leqslant 2^{n-1}(|a|+|b|)^n$ for all real a and b. Consequently, (6.1) yields

$$\gamma(n) = \sup_{a,b\in\mathbb{R}} \frac{\psi(a,b)}{(|a|+|b|)^n} \leqslant 2^{n-1}. \tag{6.25}$$

By (6.19) and (6.20), we have also the opposite inequality, and Theorem 6.1 is proved. □

Remark 6.2. Since $T_n(x) = \frac{1}{2}\bigl((x+\sqrt{x^2-1})^n + (x-\sqrt{x^2-1})^n\bigr)$, the formula (6.22) in the proof can be written (changing the sign of b)

$$\begin{aligned}\|(a,-b)^{\otimes n}\|_{\pi,\mathrm{s},+} = \psi(a,-b) &= (a-b)^n T_n\Bigl(\frac{a+b}{a-b}\Bigr)\\ &= \frac{\bigl(a+b+2\sqrt{ab}\bigr)^n + \bigl(a+b-2\sqrt{ab}\bigr)^n}{2}\\ &= \frac{\bigl(\sqrt{a}+\sqrt{b}\bigr)^{2n} + \bigl(\sqrt{a}-\sqrt{b}\bigr)^{2n}}{2},\end{aligned} \tag{6.26}$$

valid for any $a, b \geqslant 0$ by symmetry, with the case $a = b$ following by continuity or by (6.19).

Example 6.3. We used in the proof of Theorem 6.1 the classical fact that $T_n(x)$ is extremal for (6.18). This can be seen as follows, which also yields an explicit decomposition of the tensor product $(a,b)^{\otimes n}$. (See Rivlin [27] for further details and related results.)

We substitute $x=\cos\theta$; this yields an isometry $p\mapsto p(\cos\theta)$ of $P_{\leqslant n}[-1,1]$ onto the space of trigonometric polynomials

$$\mathcal{T}_n := \Bigl\{\sum_{k=0}^n a_k\cos^k\theta : a_0,\dots,a_n\in\mathbb{R}\Bigr\} = \Bigl\{\sum_{k=-n}^n b_{|k|}e^{\mathrm{i}k\theta} : b_0,\dots,b_n\in\mathbb{R}\Bigr\} \tag{6.27}$$

with the norm $\|q\|_{\mathcal{T}_n}=\sup_\theta|q(\theta)|$. The linear functional $p\mapsto[x^n]p(x)$ on $P_{\leqslant n}[-1,1]$ corresponds to the linear functional χ mapping a trigonometric polynomial $q(\theta)=\sum_{k=0}^n a_k\cos^k\theta=\sum_{k=-n}^n b_{|k|}e^{\mathrm{i}k\theta}$ to $a_n=2^nb_n$. A simple calculation (a Fourier inversion in $\mathbb{Z}_{2n}$) yields

$$\frac{1}{2n}\sum_{j=0}^{2n-1}(-1)^jq\Bigl(\frac{j\pi}{n}\Bigr)=\frac{1}{2n}\sum_{j=0}^{2n-1}\sum_{k=-n}^{n}b_{|k|}e^{\mathrm{i}j(k+n)\pi/n}=2b_n \tag{6.28}$$

and thus

$$|b_n|\leqslant\frac12\|q\|, \tag{6.29}$$

with equality for $q(\theta)=\cos(n\theta)$. Consequently, the linear functional $q\mapsto b_n$ has norm $\frac12$ on $\mathcal{T}_n$, so the linear functional $q\mapsto a_n=2^nb_n$ has norm 2^{n-1}. As said above, this corresponds by an isometry to the linear functional $[x^n]p(x)$ on $P_{\leqslant n}[-1,1]$, so this functional too has norm 2^{n-1} and (6.19) follows.

We see also from (6.28) that for any $p\in P_{\leqslant n}$, with $q(\theta)=p(\cos\theta)$,

$$[x^n]p(x)=2^nb=\frac{2^{n-1}}{2n}\sum_{j=0}^{2n-1}(-1)^jq\Bigl(\frac{j\pi}{n}\Bigr)=\frac{2^{n-1}}{2n}\sum_{j=0}^{2n-1}(-1)^jp\Bigl(\cos\frac{j\pi}{n}\Bigr). \tag{6.30}$$

Transforming back to [0,1], this yields

$$\begin{aligned}\chi_{1,-1}(p)=[x^n]p(x)&=\frac{2^{2n-1}}{2n}\sum_{j=0}^{2n-1}(-1)^jp\Bigl(\frac{1+\cos\frac{j\pi}{n}}{2}\Bigr)\\ &=\frac{2^{2n-1}}{2n}\sum_{j=0}^{2n-1}(-1)^jp\Bigl(\cos^2\frac{j\pi}{2n}\Bigr).\end{aligned} \tag{6.31}$$

This yields an optimal representation of $\chi_{1,-1}$ as a signed measure μ on [0,1], which by the argument above corresponds to an optimal decomposition of $(1,-1)^{\otimes n}$ into positive tensor powers:

$$(1,-1)^{\otimes n}=\frac{2^{2n-1}}{2n}\sum_{j=0}^{2n-1}(-1)^j\Bigl(\cos^2\frac{j\pi}{2n},\sin^2\frac{j\pi}{2n}\Bigr)^{\otimes n}. \tag{6.32}$$

(Note that there are only $n+1$ different tensor powers on the right-hand side, since the terms for j and $2n-j$ are equal in (6.32), as well as in (6.30) and (6.31).) Moreover, it follows also from this argument that this optimal decomposition is unique.

Example 6.4. We can similarly find an optimal decomposition of $(a,-b)^{\otimes n}$ for arbitrary $a,b>0$. Assume $a-b\neq 0$; then (6.21) (with $-b$ instead of b) and the arguments above show that we want to represent the linear functional $p\mapsto p(\xi)$ on $P_{\leqslant n}[-1,1]$ for a given $\xi=\frac{a+b}{a-b}$ with $|\xi|>1$. Again we seek a representation as a linear combination of $p\bigl(\cos\frac{j\pi}{n}\bigr)$, $j=0,\dots,n$, since these are the points where $|T_n(x)|$ attains its maximum on $[-1,1]$, Thus, again extending the summation to $j=0,\dots,2n-1$ for convenience, we want to find $c_j(\xi)$, with $c_{2n-j}(\xi)=c_j(\xi)$, such that

$$p(\xi)=\sum_{j=0}^{2n-1}c_j(\xi)p\Bigl(\cos\frac{j\pi}{n}\Bigr),\qquad p\in P_{\leqslant n}. \tag{6.33}$$

In fact, if (6.33) holds, then it extends to vector-valued polynomials (by considering each component separately); taking p to be the vector-valued polynomial $\bigl(\frac{1+x}{2},\frac{1-x}{2}\bigr)^{\otimes n}$ then yields

$$(a,-b)^{\otimes n}=\sum_{j=0}^{2n-1}(a-b)^n c_j\Bigl(\frac{a+b}{a-b}\Bigr)\Bigl(\cos^2\frac{j\pi}{2n},\sin^2\frac{j\pi}{2n}\Bigr)^{\otimes n}. \tag{6.34}$$

Since $P_{\leqslant n}$ has dimension $n+1$, there exists a unique such representation (6.33). Moreover, the general theory, see [27, Chapter 2] for details, or alternatively the calculations at the end of this example, shows that the representation (6.33) is optimal in the sense that $\sum_j|c_j(\xi)|$ equals the norm of $p\mapsto p(\xi)$ on $P_{\leqslant n}[-1,1]$; furthermore, this is the unique optimal representation. Consequently, (6.34) yields the unique optimal decomposition of $(a,-b)^{\otimes n}$.

In order to find $c_j(\xi)$, we take $p(x)=T_k(x)=\cos(k\arccos x)$ in (6.33) and find

$$T_k(\xi)=\sum_{j=0}^{2n-1}c_j(\xi)\cos\frac{jk\pi}{n},\qquad k=0,\dots n. \tag{6.35}$$

Furthermore, by our choice $c_{2n-j}(\xi)=c_j(\xi)$, $\sum_{j=0}^{2n-1}c_j(\xi)\sin\frac{jk\pi}{n}=0$ for any k; hence (6.35) yields

$$\sum_{j=0}^{2n-1}c_j(\xi)e^{-\mathrm{i}jk\pi/n}=T_{|k|}(\xi),\qquad k=-n,\dots n. \tag{6.36}$$

A Fourier inversion (on $\mathbb{Z}_{2n}$) now yields

$$c_j(\xi)=\frac{1}{2n}\sum_{k=-n}^{n-1}e^{\mathrm{i}jk\pi/n}T_{|k|}(\xi). \tag{6.37}$$

Substituting this in (6.34) yields the optimal decomposition of $(a,-b)^{\otimes n}$ for any $a,b>0$, with the case $a=b$ in (6.32) interpreted as a limit.

We can calculate the coefficients $c_j(\xi)$ in (6.37) more explicitly. Suppose that $a>b>0$, so $\xi=\frac{a+b}{a-b}>1$, and let $y:=\operatorname{arccosh}\frac{a+b}{a-b}>0$. Then $T_k(\xi)=\cosh(ky)$, and thus (6.37) yields

$$\begin{aligned}
c_j\Bigl(\frac{a+b}{a-b}\Bigr)&=\frac{1}{2n}\sum_{k=-n}^{n-1}e^{\mathrm{i}jk\pi/n}\cosh(ky)=\frac{1}{4n}\sum_{k=-n}^{n-1}\bigl(e^{\mathrm{i}jk\pi/n+ky}+e^{\mathrm{i}jk\pi/n-ky}\bigr)\\
&=\frac{1}{4n}\frac{e^{\mathrm{i}j\pi}\bigl(e^{ny}-e^{-ny}\bigr)}{e^{\mathrm{i}j\pi/n+y}-1}+\frac{1}{4n}\frac{e^{\mathrm{i}j\pi}\bigl(e^{-ny}-e^{ny}\bigr)}{e^{\mathrm{i}j\pi/n-y}-1}\\
&=(-1)^j\frac{\sinh(ny)}{2n}\Bigl(\frac{1}{e^{\mathrm{i}j\pi/n+y}-1}-\frac{1}{e^{\mathrm{i}j\pi/n-y}-1}\Bigr)\\
&=(-1)^j\frac{\sinh(ny)}{2n}\frac{e^{\mathrm{i}j\pi/n-y}-e^{\mathrm{i}j\pi/n+y}}{e^{2\mathrm{i}j\pi/n}+1-e^{\mathrm{i}j\pi/n}\bigl(e^{y}+e^{-y}\bigr)}\\
&=(-1)^j\frac{\sinh(ny)}{2n}\frac{\sinh y}{\cosh y-\cos(j\pi/n)}.
\end{aligned}\tag{6.38}$$

In particular, note that $\operatorname{sign}\bigl(c_j(n)\bigr)=(-1)^j$, so $c_j(\xi)$ alternates in sign. This shows by (6.33) and the fact that $T_n(\cos(j\pi/n))=(-1)^j$,

$$\bigl\|p\mapsto p(\xi)\bigr\|_{P_{\leqslant n}[-1,1]^*}=\sum_{j=0}^{2n-1}\bigl|c_j(n)\bigr|\tag{6.39}$$

and thus, by (6.12) and (6.21) (still with b replaced by $-b$)

$$\psi(a,-b)=(a-b)^n\sum_{j=0}^{2n-1}\bigl|c_j(n)\bigr|.\tag{6.40}$$

This verifies directly that the decomposition (6.34) is optimal, without the general theory referred to above.

Remark 6.5. Another expression for $c_j(\xi)$ can be obtained using the Lagrange interpolation polynomials $\ell_k(x)$ for the points $x_j=\cos\frac{j\pi}{n}$, $j=0,\dots,n$, see [23, §3.3]; these are given by $\ell_k(x)=\prod_{j\neq k}\frac{x-x_j}{x_k-x_j}$ and are characterized as the polynomials in $P_{\leqslant n}$ satisfying $\ell_k(x_j)=\delta_{jk}$, and thus, for any polynomial $p\in P_{\leqslant n}$ and any real (or complex) ξ,

$$p(\xi)=\sum_{j=0}^{n}\ell_j(\xi)p(x_j).\tag{6.41}$$

Consequently, $c_j(\xi)=\ell_j(\xi)$, now summing for $j=0,\dots,n$ only.

7. Exchangeable random variables

7.1. **More notation.** Let $S=(S,\mathcal{S})$ be an arbitrary measurable space. $\mathcal{M}(S)$ denotes the Banach space of (finite) signed measures on S, with $\|\mu\|$

defined to be the total variation of μ. Furthermore, $\mathcal{P}(S)$ is the subset of probability measures on S, i.e., the positive measures with norm 1. We regard $\mathcal{M}(S)$ and $\mathcal{P}(S)$ as measurable spaces with the σ-fields generated by the evaluations $\mu \mapsto \mu(A)$ for measurable $A \subseteq S$ (i.e., $A \in \mathcal{S}$). Recall that if X is a random element of S, then its distribution is a measure in $\mathcal{P}(S)$.

If $x \in S$, then δ_x denotes the Dirac measure, i.e., unit point mass, at x. (This is the distribution of the non-random $X := x$.)

For a finite (or countable) set S, we identify the space $\mathcal{M}(S)$ of signed measures on S with $\ell_1(S)$. In particular, δ_x is identified with the vector $(\mathbf{1}_{\{y=x\}})_{y\in S} \in \ell_1(S)$, and thus $\delta_i = e_i$ when $S = \mathbb{N}$. Note that if X is a random variable with distribution $\mu \in \mathcal{P}(\mathcal{S}) \subset \mathcal{M}(S)$, then δ_X is a random vector in $\ell_1(S)$, and for every $s \in S$, $\mathbb{E}\,\delta_X(s) = \mathbb{P}(X = s) = \mu\{s\}$; hence, with the identification $\mathcal{M}(S) = \ell_1(S)$ just made,

$$\mathbb{E}\,\delta_X = \mu \in \ell_1(S). \tag{7.1}$$

Recall that $[n] := \{1, \dots, n\}$.

7.2. **Finitely exchangeable distributions.** Let $S = (S, \mathcal{S})$ be a measurable space. A random vector $\mathbf{X} = (X_1, \dots, X_n)$ with values in S^n is *(finitely) exchangeable* if its distribution is symmetric under permutations. See e.g. Aldous [1] for a survey of both finite and infinite exchangeability.

For an infinite exchangeable sequence $\mathbf{X} = (X_i)_1^\infty$, the well-known de Finetti's theorem says that under weak technical conditions on S (for example that S is a Borel space), the distribution is a mixture of product (power) measures, see e.g. [1, §2] or [17, Theorem 1.1]. In formulas, this says that if $\mathcal{P}(S)$ is the space of probability measures on S, and $\mu_{\mathbf{X}} \in \mathcal{P}(S^\infty)$ is the distribution of $\mathbf{X}$, then there exists a probability measure λ on $\mathcal{P}(S)$ such that

$$\mu_{\mathbf{X}} = \int_{\mathcal{P}(S)} \nu^\infty \,\mathrm{d}\lambda(\nu). \tag{7.2}$$

It is also well-known that this, in general, fails for finitely exchangeable sequences, see e.g. [6; 7]. A substitute in the finite case is that there always exists such a representation with a *signed* measure λ. To be precise, see [5, V.52], [15], [19], [16], if $\mathbf{X} = (X_1, \dots, X_n)$ is exchangeable, with values in an arbitrary measurable space S, then there exists a signed measure λ on $\mathcal{P}(S)$, i.e., $\lambda \in \mathcal{M}(\mathcal{P}(S))$, such that

$$\mu_{\mathbf{X}} = \int_{\mathcal{P}(S)} \nu^n \,\mathrm{d}\lambda(\nu). \tag{7.3}$$

A natural question (posed in [16]) is how large the total variation $\|\lambda\|$ of λ has to be. Since $\mu_{\mathbf{X}}$ is a probability measure, we always have $\int \mathrm{d}\lambda(\nu) = 1$, and thus $\|\lambda\| \geqslant 1$, with equality if and only if λ is a probability measure (as in de Finetti's theorem (7.2)). Hence, $\|\lambda\|$ is a measure of how far the representation is from the ideal representation as a mixture of powers. Note that λ is not unique, so we are interested in the optimal λ, or more generally $\inf \|\lambda\|$ over all possible representing λ in (7.3).

Partial answers to this question are given by the following theorems, which connect this problem to the tensor norms studied above.

Theorem 7.1. (i) *If $\mathbf{X} = (X_1, \dots, X_n)$ is exchangeable, with values in an arbitrary measurable space S, then its distribution $\mu_{\mathbf{X}} \in \mathcal{P}(S^n)$ has a representation* (7.3) *with a signed measure λ on $\mathcal{P}(S)$ such that*

$$\|\lambda\|_{\mathcal{M}(\mathcal{P}(S))} \leqslant \kappa(n), \tag{7.4}$$

where, as in (5.1),

$$\kappa(n) := c_{\mathsf{s},+}(n, \ell_1) = \|e_1 \vee \cdots \vee e_n\|_{\pi,\mathsf{s},+;\ell_1^n}. \tag{7.5}$$

The constant $\kappa(n)$ given in (7.5) *is, in general, the best possible. We have*

$$\frac{n^n}{n!} \leqslant \kappa(n) \leqslant 2^{n-1}\frac{n^n}{n!}. \tag{7.6}$$

(ii) *If furthermore S is finite, with $|S| = m$, then* (7.4) *can be replaced by*

$$\|\lambda\|_{\mathcal{M}(\mathcal{P}(S))} \leqslant c_{\mathsf{s},+}(n, \ell_1^m). \tag{7.7}$$

Moreover, this constant is the best possible for the given S. If $m \geqslant n$, then this constant equals $\kappa(n)$.

By Example 5.2, $\kappa(2) = 3$; hence neither of the bounds in (7.6) is sharp.

Problem 7.2. What is the exact value of $\kappa(n)$?

It follows from (7.6) and Stirling's formula that, recalling (5.14),

$$e \leqslant \limsup_{n\to\infty} \kappa(n)^{1/n} = c_{\mathsf{s},+}(\ell_1) \leqslant 2e. \tag{7.8}$$

Problem 7.3. What is $\limsup_{n\to\infty} \kappa(n)^{1/n}$? Does $\lim_{n\to\infty} \kappa(n)^{1/n}$ exist?

Before proving Theorem 7.1, consider first for simplicity the case when S is finite. As said above, we then identify signed measures on S^n with vectors in $\ell_1(S^n)$; thus a distribution (i.e., probability measure) μ on S^n is the same as a positive element of norm 1 in $\ell_1(S^n)$. Since S is finite, $\ell_1(S^n) = \ell_1(S)^{\otimes n}_\pi$ isometrically, cf. Example 3.5. Thus, a distribution μ on S^n is the same as a positive element of norm 1 in $\ell_1(S)^{\otimes n}_\pi$. Furthermore, by definition, μ is exchangeable if it is invariant under permutations of the coordinates, which is the same as saying that μ, regarded as a tensor in $\ell_1(S)^{\otimes n}$, is a symmetric tensor. Hence, an exchangeable distribution μ is a positive element of $\ell_1(S)^{\vee n}$ with $\|\mu\|_\pi = 1$.

Consider now representations as in (7.3) of an exchangeable distribution $\mu_{\mathbf{X}}$. If λ has finite support, then (7.3) becomes a representation as in (4.4), and thus $\|\mu_{\mathbf{X}}\|_{\pi,\mathsf{s},+;\ell_1(s)} \leqslant \|\lambda\|$. Furthermore, this extends to arbitrary measures λ since (7.3) implies

$$\|\mu_{\mathbf{X}}\|_{\pi,\mathsf{s},+} \leqslant \int \|\nu^n\|_{\pi,\mathsf{s},+} \,\mathrm{d}|\lambda|(\nu) = \|\lambda\|. \tag{7.9}$$

(The spaces are finite-dimensional and there are no problems with measurablilty or convergence.) Conversely, a representation as in (4.4) yields a representation (7.3) with $\lambda = \sum_k a_k \|x_k\|^n \delta_{x_k/\|x_k\|}$ and thus $\|\lambda\| \leqslant \sum_k |a_k| \|x_k\|^n$. Recall also that by Remark 4.8, the infimum in (4.4) is attained. Consequently, when S is finite, we have the following result.

Theorem 7.4. *Let $\mathbf{X} = (X_1, \dots, X_n)$ be exchangeable, with values in a finite set S, and regard its distribution $\mu_{\mathbf{X}} \in \mathcal{P}(S^n)$ as a vector in $\ell_1(S^n)$. Then*

$$\inf\{\|\lambda\| : (7.3) \text{ holds}\} = \|\mu_{\mathbf{X}}\|_{\pi,\mathsf{s},+;\ell_1(S)}. \tag{7.10}$$

Moreover, the infimum in (7.10) *is attained by some signed measure λ on $\mathcal{P}(S)$; in fact, by some λ with finite support.* □

Theorem 7.4 shows that if S is finite, then (7.3) holds with $\|\lambda\| = \|\mu_{\mathbf{X}}\|_{\pi,\mathsf{s},+}$. A special case is to take $S = [n] := \{1, \dots, n\}$ and let the random vector $(X_1, \dots, X_n)$ be a uniformly random permutation of $\{1, \dots, n\}$, which means that $\mu_{\mathbf{X}} := e_1 \vee \dots \vee e_n$. This case is easily seen to be extreme. In fact, if S is any finite set and $\mathbf{x} = (x_1, \dots, x_n) \in S^n$, then

$$\varphi_{\mathbf{x}}(e_i) := \delta_{x_i} \tag{7.11}$$

defines a positive linear operator $\varphi_{\mathbf{x}} : \ell_1^n \to \mathcal{M}(S) = \ell_1(S)$ with $\|\varphi_{\mathbf{x}}\| = 1$, and thus, by Theorem 4.14, $\|\delta_{x_1} \vee \dots \vee \delta_{x_n}\|_{\pi,\mathsf{s},+} \leqslant \|e_1 \vee \dots \vee e_n\|_{\pi,\mathsf{s},+}$. Furthermore, every exchangeable distribution $\mu_{\mathbf{X}}$ on S^n is a convex combination of tensors of the type $\delta_{x_1} \vee \dots \vee \delta_{x_n}$. Consequently,

$$\|\mu_{\mathbf{X}}\|_{\pi,\mathsf{s},+} \leqslant \|e_1 \vee \dots \vee e_n\|_{\pi,\mathsf{s},+}. \tag{7.12}$$

This proves, together with (7.10), the main assertion in Theorem 7.1 when S is finite.

The general proof uses the same idea; we only have to add some technicalities, which we borrow from [16], where further details may be found if necessary; see also [19].

Proof of Theorem 7.1. (i): Fix a representation

$$e_1 \vee \dots \vee e_n = \sum_{k=1}^{N} a_k \eta_k^{\otimes n}, \tag{7.13}$$

where $a_k \in \mathbb{R}$ and $\eta_k \geqslant 0$ are unit vectors in $\ell_1^n = \mathcal{M}([n])$; thus $\eta_k \in \mathcal{P}([n])$.

For any $\mathbf{x} = (x_1, \dots, x_n) \in S^n$, define again the linear map $\varphi_{\mathbf{x}} : \mathcal{M}([n]) \to \mathcal{M}(S)$ by (7.11) and linearity, and note that $\varphi_{\mathbf{x}}$ maps $\mathcal{P}([n])$ into $\mathcal{P}(S)$. ($\varphi_{\mathbf{x}}$ is the natural push-forward of measures induced by the mapping $[n] \to S$ given by $i \mapsto x_i$.) Furthermore, $\varphi_{\mathbf{x}}^{\otimes n} : \mathcal{M}([n])^{\otimes n} \to \mathcal{M}(S)^{\otimes n}$ and we may regard $\mathcal{M}(S)^{\otimes n}$ as a subspace of $\mathcal{M}(S^n)$ also when S is infinite.

Define further, using the decomposition (7.13),

$$\psi_{\mathbf{x}} := \sum_{k=1}^{N} a_k \delta_{\varphi_{\mathbf{x}}(\eta_k)} \in \mathcal{M}(\mathcal{P}(S)). \tag{7.14}$$

Then, for any $\mathbf{x} \in S^n$, using (7.13) and (7.11),

$$\begin{aligned}\int_{\mathcal{P}(S)} \nu^n \,\mathrm{d}\psi_{\mathbf{x}}(\nu) &= \sum_{k=1}^N a_k \varphi_{\mathbf{x}}(\eta_k)^{\otimes n} = \sum_{k=1}^N a_k \varphi_{\mathbf{x}}^{\otimes n}\bigl(\eta_k^{\otimes n}\bigr) \\ &= \varphi_{\mathbf{x}}^{\otimes n}\bigl(e_1 \vee \cdots \vee e_n\bigr) = \varphi_{\mathbf{x}}(e_1) \vee \cdots \vee \varphi_{\mathbf{x}}(e_n) \\ &= \delta_{x_1} \vee \cdots \vee \delta_{x_n}. \end{aligned} \tag{7.15}$$

Furthermore, for each fixed $\eta \in \mathcal{P}(S)$, the map $\mathbf{x} \mapsto \varphi_{\mathbf{x}}(\eta)$ is measurable $S^n \to \mathcal{P}(S)$, and thus the map $\mathbf{x} \mapsto \psi_{\mathbf{x}}$ is measurable $S^n \to \mathcal{M}(\mathcal{P}(S))$. Hence, $\psi_{\mathbf{X}}$ is a random measure in $\mathcal{M}(\mathcal{P}(S))$. Moreover, by (7.14),

$$\|\psi_{\mathbf{X}}\|_{\mathcal{M}(\mathcal{P}(S))} \leqslant K := \sum_{k=1}^N |a_k|. \tag{7.16}$$

Hence, we can define the expectation $\lambda := \mathbb{E}\,\psi_{\mathbf{X}} \in \mathcal{M}(\mathcal{P}(S))$, cf. [18, Lemma 2.4]. Furthermore, (7.16) implies $\|\lambda\|_{\mathcal{M}(\mathcal{P}(S))} \leqslant K$, and (7.15) implies

$$\int_{\mathcal{P}(S)} \nu^n \,\mathrm{d}\lambda(\nu) = \mathbb{E}\int_{\mathcal{P}(S)} \nu^n \,\mathrm{d}\psi_{\mathbf{X}}(\nu) = \mathbb{E}\bigl(\delta_{X_1} \vee \cdots \vee \delta_{X_n}\bigr) = \mu_{\mathbf{X}}. \tag{7.17}$$

This shows the existence of a representation (7.3) with $\|\lambda\| \leqslant K$, given by (7.16).

We may, by Remark 4.8, choose the decomposition (7.13) such that $K = \|e_1 \vee \cdots \vee e_n\|_{\pi,\mathsf{s},+} = \kappa(n)$, and thus (7.4) holds.

To see that $\kappa(n)$ is best possible, it suffices to take $S = [n]$ and $\mu_{\mathbf{X}} = e_1 \vee \cdots \vee e_n$, as in the discussion before the proof. Then (7.10) shows that every representating measure λ satisfies $\|\lambda\| \geqslant \kappa(n)$.

Finally, (7.6) follows from (4.34), Theorem 6.1 and (2.21).

(ii): By (7.10) and the comment after it, we can find λ with

$$\|\lambda\| \leqslant \|\mu_{\mathbf{X}}\|_{\pi,\mathsf{s},+;\,\ell_1(S)} \leqslant c_{\mathsf{s},+}\bigl(n, \ell_1(S)\bigr) = c_{\mathsf{s},+}\bigl(n, \ell_1^m\bigr). \tag{7.18}$$

On the other hand, if M is a constant such that there always exists a λ with $\|\lambda\| \leqslant M$, then (7.10) shows that $\|\mu\|_{\pi,\mathsf{s},+} \leqslant M$ for every positive $\mu \in \ell_1(S)^{\vee n}$ with $\|\mu\|_\pi = 1$. This extends to all $\mu \in \ell_1(S)^{\vee n}$ with $\|\mu\|_\pi = 1$, by decomposing them in their positive and negative parts, and thus $c_{\mathsf{s},+}(n, \ell_1^m) = c_{\mathsf{s},+}(n, \ell_1(S)) \leqslant M$.

Finally, if $m \geqslant n$ then $c_{\mathsf{s},+}(n, \ell_1^m) = \kappa(n)$ by Theorem 5.1. □

Remark 7.5. The proof in [16] of the representation (7.3) used the argument above, with a decomposition (7.13) where η_k ranged over the $\binom{2n-1}{n-1}$ probability measures ν in $\mathcal{P}([n])$ such that $n\mu$ is integer-valued; it was shown in [16] by an algebraic argument that there always exists a unique such decomposition. No attempt was made in [16] to evaluate the best constant; in fact, a numerical calculation (using Maple) of the constant $K = K_n$ in (7.16) for the decomposition in [16] yields e.g. $K_2 = 3$, $K_3 = 20$, $K_4 = 210$, $K_5 = 3024$. These values are thus upper bounds for $\kappa(n)$; we see that for $n = 2$, we obtain the sharp constant $\kappa(2) = 3$ (see Example 5.2), but already for $n = 3$, this K_n is larger than the upper bound in (7.6) ($\kappa(3) \leqslant 18$). In other words (not surprisingly), the decomposition used in [16] is not optimal.

Remark 7.6. Note that the proof uses the σ-field on $\mathcal{M}(\mathcal{P}(S))$ defined in Section 7.1, and not the (in general larger) Borel σ-field on the Banach space $\mathcal{M}(\mathcal{P}(S))$; in general, the mapping $\mathbf{x} \to \psi_{\mathbf{x}}$ is not measurable if $\mathcal{M}(\mathcal{P}(S))$ is given the latter σ-field.

Remark 7.7. We have considered representations (7.3) where λ is a signed measure but ν ranges over probability measures. An alternative is to allow also ν to be a signed measure, i.e., to consider representations

$$\mu_{\mathbf{X}} = \int_{B(\mathcal{M}(S))} \nu^n \lambda(\mathrm{d}\nu) \tag{7.19}$$

where $B(\mathcal{M}(S))$ denotes the unit ball in the Banach space $M(S)$ of signed neasures on S. The arguments above are easily modified to this case and show that there always exists such a representation with

$$\|\lambda\|_{\mathcal{M}(B(\mathcal{M}(S)))} \leqslant \|e_1 \vee \dots \vee e_n\|_{\pi,\mathsf{s};\,\ell_1} = c_{\mathsf{s}}(n, \ell_1) = \frac{n^n}{n!}, \tag{7.20}$$

where we used Example 2.7 for the explicit value; moreover, this constant is the best possible. In particular, this shows that if $n \geqslant 2$, then we cannot in general find a representation (7.19) where λ is a probability measure on $B(\mathcal{M}(S))$.

Remark 7.8. The upper bound in (7.6) can be improved a little as follows.

By (3.10),

$$\kappa(n) = \|e_1 \vee \dots \vee e_n\|_{\pi,\mathsf{s},+} \leqslant \frac{1}{2^n n!} \sum_{\varepsilon_1,\dots,\varepsilon_n = \pm 1} \Bigl\| \Bigl(\sum_{i=1}^n \varepsilon_i e_i \Bigr)^{\otimes n} \Bigr\|_{\pi,\mathsf{s},+}. \tag{7.21}$$

Consider one of the terms in the sum, and suppose that $\varepsilon_i = 1$ for k indices i. The argument in the beginning of the proof of Theorem 5.3, up to the first inequality in (5.11), with $u := (k, -(n-k)) \in \ell_1^2$, show that, using (6.2),

$$\Bigl\| \Bigl(\sum_{i=1}^n \varepsilon_i e_i \Bigr)^{\otimes n} \Bigr\|_{\pi,\mathsf{s},+;\,\ell_1^n} \leqslant \|u^{\otimes n}\|_{\pi,\mathsf{s},+;\,\ell_1^2} = \psi\bigl(k, -(n-k)\bigr). \tag{7.22}$$

This is evaluated in (6.26), and thus (7.21) yields, by counting terms,

$$\begin{aligned} \kappa(n) &\leqslant \frac{1}{2^n n!} \sum_{k=0}^n \binom{n}{k} \psi\bigl(k, -(n-k)\bigr) \\ &= \frac{1}{2^{n+1} n!} \sum_{k=0}^n \binom{n}{k} \Bigl(\bigl(\sqrt{k} + \sqrt{n-k} \bigr)^{2n} + \bigl(\sqrt{k} - \sqrt{n-k} \bigr)^{2n} \Bigr). \end{aligned} \tag{7.23}$$

For $n = 2$, (7.23) yields the correct value 3. We have no reason to believe that the bound is sharp for larger n.

The improvement from the upper bound in (7.6) lies in that we here use the exact value (6.26) for each term, while the proof of (7.6) estimates each

$\psi(k,-(n-k))$ by the worst case $k=n/2$. However, the improvement is slight, since most terms in (7.21) have k close to $n/2$. In fact, simple asymptotic estimates (which we omit) show that asymptotically, (7.23) improves the upper bound only by a factor $\sqrt{2/3} \doteq 0.816$. Numerically, the improvement factor is close to this value also for small n, with a factor 0.75 for $n=2$ and 3.

7.3. **Binary variables.** If S is finite with $|S|=m<n$, we may hope that the bound $c_{\mathsf{s},+}(n,\ell_1^m)$ in (7.7) is better than $\kappa(n)$. We consider here only the simplest case $|S|=2$, for example $S=\{0,1\}$.

Let, for $0\leqslant j\leqslant n$, μ_j be the distribution of a random vector $\mathbf{X}\in S^n$ consisting of j 0's and $n-j$ 1's in random order; thus,

$$\mu_j=\Lambda\bigl(\delta_0^{\otimes j}\otimes\delta_1^{\otimes n-j}\bigr). \tag{7.24}$$

Evidently, μ_j is exchangeable. Moreover, every exchangeable distribution on S^n is a mixture of these measures μ_j, and it follows that

$$c_{\mathsf{s},+}(n,\ell_1^2)=c_{\mathsf{s},+}(n,\ell_1(S))=\sup_{0\leqslant j\leqslant n}\|\mu_j\|_{\pi,\mathsf{s},+;\,\ell_1(S)}. \tag{7.25}$$

We thus want to find $\|\mu_j\|_{\pi,\mathsf{s},+;\,\ell_1(S)}$.

We argue as in the proof of Theorem 6.1. This yields, cf. (6.4)–(6.5), that

$$\|\mu_j\|_{\pi,\mathsf{s},+;\,\ell_1(S)}=\inf\|\mu\|=\inf\int_0^1|\mathrm{d}\mu|(x), \tag{7.26}$$

taking the infimum over all signed measures μ on $[0,1]$ such that

$$\int_0^1(x,1-x)^{\otimes n}\,\mathrm{d}\mu(x)=\mu_j, \tag{7.27}$$

which is equivalent to, by expanding into coordinates in $(\mathbb{R}^2)^{\otimes n}$,

$$\int_0^1 x^k(1-x)^{n-k}\,\mathrm{d}\mu(x)=\frac{1}{\binom{n}{j}}\delta_{kj},\qquad k=0,\dots,n. \tag{7.28}$$

Let again T_n be the Chebyshev polynomial. Let $x\in[0,1]$, write $y:=1-x$, $t:=\arccos(2x-1)$ and $s:=t/2$. Then, $\cos^2 s=(1+\cos t)/2=x$, $\sin^2 s=1-\cos^2 s=y$, and thus $e^{\mathrm{i}s}=x^{1/2}+\mathrm{i}y^{1/2}$. Consequently,

$$\begin{aligned}T_n(2x-1)&=\cos(nt)=\cos(2ns)=\operatorname{Re}e^{\mathrm{i}2ns}=\operatorname{Re}\bigl(x^{1/2}+\mathrm{i}y^{1/2}\bigr)^{2n}\\&=\sum_{k=0}^{n/2}\binom{2n}{2k}x^k(-y)^{n-k}\\&=\sum_{k=0}^{n/2}\binom{2n}{2k}(-1)^{n-k}x^k(1-x)^{n-k}.\end{aligned} \tag{7.29}$$

Hence, if μ satisfies (7.28), then

$$\int_0^1 T_n(2x-1)\,\mathrm{d}\mu(x)=\frac{(-1)^{n-j}}{\binom{n}{j}}\binom{2n}{2j} \tag{7.30}$$

which implies, since $|T_n(2x-1)| \leqslant 1$ for $x \in [0,1]$,

$$\int_0^1 |\mathrm{d}\mu|(x) \geqslant \frac{\binom{2n}{2j}}{\binom{n}{j}}. \tag{7.31}$$

Recalling (7.26), we have shown that

$$\|\mu_j\|_{\pi,\mathsf{s},+;\,\ell_1(S)} \geqslant \frac{\binom{2n}{2j}}{\binom{n}{j}}. \tag{7.32}$$

Thus, by (7.25), using an elementary calculation to optimize j,

$$c_{\mathsf{s},+}\bigl(n,\ell_1^2\bigr) \geqslant \max_{0\leqslant j\leqslant n} \frac{\binom{2n}{2j}}{\binom{n}{j}} = \frac{\binom{2n}{2\lfloor n/2\rfloor}}{\binom{n}{\lfloor n/2\rfloor}}. \tag{7.33}$$

We conjecture that $T_n(2x-1)$ is extremal here too, so that equality holds in (7.32) and (7.33), but we leave that as an open problem.

In any case, (7.33) is a lower bound. Stirling's formula yields the asymptotic estimate

$$c_{\mathsf{s},+}\bigl(n,\ell_1^2\bigr) \geqslant 2^{n-1/2+o(1)}. \tag{7.34}$$

Hence, the constants $c_{\mathsf{s},+}\bigl(n,\ell_1^2\bigr)$ also grow exponentially, but possibly (presumably) at a slower rate than $\kappa(n) = c_{\mathsf{s},+}\bigl(n,\ell_1\bigr)$, see (7.8).

However, a numerical calculation reveals that for $2 \leqslant n \leqslant 4$, the lower bound $n^n/n!$ in (7.6) is smaller than the bound in (7.33). We thus have, using also Example 5.2 or Remark 7.8 for $n=2$, the improved bounds

$$\kappa(2) = c_{\mathsf{s},+}\bigl(2,\ell_1^2\bigr) = 3, \tag{7.35}$$

$$\kappa(3) \geqslant c_{\mathsf{s},+}\bigl(3,\ell_1^2\bigr) \geqslant 5, \tag{7.36}$$

$$\kappa(4) \geqslant c_{\mathsf{s},+}\bigl(4,\ell_1^2\bigr) \geqslant \frac{35}{3}. \tag{7.37}$$

Problem 7.9. Find a non-trivial upper bound for $c_{\mathsf{s},+}(n,\ell_1^2)$. Is, as conjectured above, (7.33) an equality?

Problem 7.10. Extend this to $c_{\mathsf{s},+}(n,\ell_1^m)$ for other fixed values of m.

7.4. **Extendible finitely exchangeable variables.** Let n and N be positive integers with $N \geqslant n$. An exchangeable random vector $\mathbf{X}_n = (X_1,\dots,X_n)$ in S^n is N*-extendible* if it can be extended to an exchangeable random vector $\mathbf{X}_N = (X_1,\dots,X_N)$. We similarly say that an exchangeable distribution on S^n is N-extendible if it is the distribution of an N-extendible vector. Note that by de Finetti's theorem (7.2), at least if S is a Borel space, a distribution is ∞-extendible if and only if it has a representation (7.3) with a probability measure λ. However, we will here consider the case of finite N. See e.g. [6; 7; 20] for various aspects of extendibility.

Let $\mathcal{E}_n = \mathcal{E}_n(S)$ be the set of exchangeable distributions on S^n, and let $\mathcal{E}_{n,N} = \mathcal{E}_{n,N}(S)$ be the subset of N-extendible distributions. Let $\Pi_{N,n}$:

$\mathcal{P}(S^N) \to \mathcal{P}(S^n)$ be the map induced by projecting a random vector $(X_1, \dots, X_N)$ onto its first n coordinates. Thus $\mathcal{E}_{n,N} = \Pi_{N,n}(\mathcal{E}_N) \subseteq \mathcal{E}_n$.

Consider again first the case when S is finite. Then, as discussed above, $\mathcal{E}_n$ is the set of positive unit elements in $\ell_1(S)^{\vee n}_{\pi}$.

Consider the special case $S = [N]$, and define

$$\chi_{n,N} := \Pi_{N,n}\big(e_1 \vee \dots \vee e_N\big) \in \mathcal{E}_{n,N}([N]). \tag{7.38}$$

This is thus the distribution of $(X_1, \dots, X_n)$ when $(X_1, \dots, X_N)$ is a uniformly random permutation of $[N]$; in other words, $\chi_{n,N}$ is the distribution of the random vector obtain by drawing n elements of $S = [N]$ without replacement. We will see that this is, not surprisingly, an extreme case, cf. e.g. [7]. Let

$$\kappa(n, N) := \|\chi_{n,N}\|_{\pi,\mathsf{s},+;\,\ell_1^N}. \tag{7.39}$$

For an arbitrary S and $\mathbf{x} = (x_1, \dots, x_N) \in S^N$, define $\varphi_{\mathbf{x}} : \mathcal{M}([N]) \to \mathcal{M}(S)$ by (7.11) and linearity. Then, $\varphi_{\mathbf{x}}^{\otimes n}(\chi_{n,N}) \in \mathcal{M}(S)^{\otimes n} \subseteq \mathcal{M}(S^n)$ is the distribution of the random vector obtained by drawing n elements of $x_1, \dots, x_N$ without replacement, see (7.43) below.

Theorem 7.11. (i) *Let $1 \leqslant n \leqslant N$. If $\mathbf{X} = (X_1, \dots, X_n)$ is exchangeable and N-extendible, with values in an arbitrary measurable space S, then its distribution $\mu_{\mathbf{X}} \in \mathcal{P}(S^n)$ has a representation* (7.3) *with a signed measure λ on $\mathcal{P}(S)$ such that*

$$\|\lambda\|_{\mathcal{M}(\mathcal{P}(S))} \leqslant \kappa(n, N). \tag{7.40}$$

The constant $\kappa(n, N)$ given in (7.39) *is, in general, the best possible.*

(ii) *If furthermore S is finite with $|S| = m$, then* (7.40) *can be replaced by*

$$\|\lambda\|_{\mathcal{M}(\mathcal{P}(S))} \leqslant \kappa(n, N; m) := \max_{\mathbf{x}\in S^N} \|\varphi_{\mathbf{x}}^{\otimes n}(\chi_{n,N})\|_{\pi,\mathsf{s},+;\,\ell_1(S)} \tag{7.41}$$

Moreover, this constant is the best possible for the given S. If $m \geqslant N$, then $\kappa(n, N; m) = \kappa(n, N)$.

Proof. The proof of Theorem 7.1 extends with minor changes as follows; we omit some details.

(i): Fix a representation

$$\chi_{n,N} = \sum_{k=1}^{M} a_k \eta_k^{\otimes n}, \tag{7.42}$$

where $a_k \in \mathbb{R}$ and $\eta_k \geqslant 0$ are unit vectors in $\ell_1^N = \mathcal{M}([N])$ and $K := \sum_k |a_k| = \kappa(n, N)$ (see Remark 4.8). Thus, $\eta_k \in \mathcal{P}([N])$. Define again $\psi_{\mathbf{x}}$ by (7.14). Then, similarly to (7.15),

$$\int_{\mathcal{P}(S)} \nu^n \,\mathrm{d}\psi_{\mathbf{x}}(\nu) = \varphi_{\mathbf{x}}^{\vee n}(\chi_{n,N}) = \Pi_{N,n}\big(\delta_{x_1} \vee \dots \vee \delta_{x_N}\big)$$

$$= \frac{1}{N!} \sum_{\sigma \in \mathfrak{S}_N} \delta_{x_{\sigma(1)}} \otimes \dots \otimes \delta_{x_{\sigma(n)}}. \tag{7.43}$$

Again, $\psi_{\mathbf{X}}$ is a bounded random measure in $\mathcal{M}(\mathcal{P}(S))$, and we define $\lambda := \mathbb{E}\,\psi_{\mathbf{X}} \in \mathcal{M}(\mathcal{P}(S))$. Then $\|\lambda\|_{\mathcal{M}(\mathcal{P}(S))} \leqslant K = \kappa(n,N)$, so (7.40) holds, and similarly to (7.17), using (7.43) and exchangeability,

$$\begin{aligned} \int_{\mathcal{P}(S)} \nu^n \,\mathrm{d}\lambda(\nu) &= \mathbb{E}\Bigl(\frac{1}{N!} \sum_{\sigma \in \mathfrak{S}_N} \delta_{X_{\sigma(1)}} \otimes \dots \otimes \delta_{X_{\sigma(n)}}\Bigr) \\ &= \mathbb{E}\bigl(\delta_{X_1} \otimes \dots \otimes \delta_{X_n}\bigr) = \mu_{\mathbf{X}}. \end{aligned} \tag{7.44}$$

The case $\mu_{\mathbf{X}} = \chi_{n,N}$ shows that the constant $\kappa(n,N)$ is best possible, using (7.9) as earlier.

(ii): When S is finite, every distribution in $\mathcal{E}_{n,N}$ is a convex combination of the distributions $\varphi_{\mathbf{x}}^{\otimes n}(\chi_{n,N})$, and thus (7.41) follows from (7.10).

Conversely, each $\varphi_{\mathbf{x}}^{\otimes n}(\chi_{n,N}) \in \mathcal{E}_{n,N}$, and thus (7.10) shows that (7.41) is best possible. □

Remark 7.12. Since $(N+1)$-extendible implies N-extendible, it follows from Theorem 7.11 that

$$\kappa(n) = \kappa(n,n) \geqslant \kappa(n,n+1) \geqslant \ldots \geqslant 1, \tag{7.45}$$

$$\kappa(n,N;m) \geqslant \kappa(n,n+1;m) \geqslant \ldots \geqslant 1. \tag{7.46}$$

One can also see (7.45) directly from (7.39), since $\chi_{n,N+1}$ is the average of $\varphi_{\mathbf{x}}^{\otimes n}(\chi_{n,N})$ over all sequences $\mathbf{x}$ of N distinct elements of $[N+1]$.

Diaconis and Freedman [7] showed (with precise estimates) that if N is large, then a distribution $\mu_{\mathbf{X}} \in \mathcal{E}_{n,N}$ is close to a distribution as in (7.2), in the sense of total variation. This implies similar results in terms of the constants in Theorem 7.11. In particular, for fixed n, the following theorem shows that $\kappa(n,N) \to 1$ as $N \to \infty$; more precisely, $\kappa(n,N) = 1 + O(1/N)$ for fixed n, and this rate is exact. However, there is a wide gap between the "constants" (depending on n) in the upper and lower bounds given by the theorem.

Theorem 7.13. *The following inequalities hold.*

(i) *If* $N > n(n-1)/2$, *then*

$$\kappa(n,N) \leqslant 1 + \frac{n(n-1)}{2N - n(n-1)}\bigl(\kappa(n)+1\bigr). \tag{7.47}$$

(ii) *If* $N \geqslant n$, *then*

$$\kappa(n,N) \geqslant e^{\frac{n-1}{2\lceil N/n \rceil}} \geqslant 1 + \frac{n(n-1)}{2(N+n)}. \tag{7.48}$$

(iii) *If* $N \geqslant n \geqslant m$, *then*

$$\kappa(n,N;m) \leqslant 1 + c_{\mathsf{s},+}(n,\ell_1^m)\frac{2mn}{N} \leqslant 1 + \frac{2mn\kappa(n)}{N}. \tag{7.49}$$

(iv) *If* $N \geqslant n \geqslant m$, *then*

$$\kappa(n,N;m) \geqslant e^{\frac{m-1}{2\lceil N/n\rceil}} \geqslant 1+\frac{(m-1)n}{2(N+n)}. \tag{7.50}$$

Proof. (i): Let ν_N be the uniform distribution on $[N]$. Then ν_N^n is the distribution of a random vector $(X_1,\dots,X_n)$ obtained by drawing randomly from $[N]$ with replacement. Conditioned on the event $\mathcal{D}$ that $X_1,\dots,X_n$ are distinct, this yields the distribution $\chi_{n,N}$, see (7.38) and the comments after it. Hence, if $q := \mathbb{P}(\mathcal{D})$, then, cf. [7],

$$\nu_N^n = q\chi_{n,N} + (1-q)\mu', \tag{7.51}$$

for some probability measure $\mu' \in \mathcal{P}([N]^n)$. Clearly, μ' is symmetric, i.e., exchangeable. Consequently,

$$\begin{aligned} q\|\chi_{n,N}\|_{\pi,\mathsf{s},+} &\leqslant q\|\nu_N^n\|_{\pi,\mathsf{s},+} + (1-q)\|\mu'\|_{\pi,\mathsf{s},+} \\ &\leqslant 1+(1-q)c_{\mathsf{s},+}(n,\ell_1^N)\|\mu'\|_\pi = 1+(1-q)\kappa(n) \end{aligned} \tag{7.52}$$

and thus

$$\kappa(n,N) = \|\chi_{n,N}\|_{\pi,\mathsf{s},+} \leqslant 1+\big(\kappa(n)+1\big)\frac{1-q}{q}. \tag{7.53}$$

Furthermore,

$$q=\mathbb{P}(\mathcal{D}) = \prod_{i=1}^{n-1}\Big(1-\frac{i}{N}\Big) \geqslant 1-\sum_{i=1}^{n-1}\frac{i}{N} = 1-\frac{n(n-1)}{2N}. \tag{7.54}$$

Hence, (7.47) follows

(ii): We modify Example 2.7. Partition $[N]$ into n sets $S_1,\dots,S_n$ and let $N_i := |S_i|$. Define a multilinear operator $L : (\ell_1^N)^n \to \mathbb{R}$ by, writing $x_i = (x_{ij})_{j=1}^N$,

$$L(x_1,\dots,x_n) = \prod_{i=1}^n \sum_{j\in S_i} x_{ij}. \tag{7.55}$$

Let $(X_1,\dots,X_n)$ be a random vector with distribution $\chi_{n,N}$; then by (7.1) (for $S=[N]^n$),

$$\chi_{n,N} = \mathbb{E}\,\delta_{(X_1,\dots,X_n)} = \mathbb{E}\big(\delta_{X_1}\otimes\cdots\otimes\delta_{X_n}\big) \tag{7.56}$$

and thus, regarding L also as a linear operator $L : (\ell_1^N)^{\otimes n} \to \mathbb{R}$,

$$\begin{aligned} L(\chi_{n,N}) &= \mathbb{E}\,L\big(\delta_{X_1}\otimes\ldots\otimes\delta_{X_n}\big) = \mathbb{E}\,L(\delta_{X_1},\dots,\delta_{X_n}) \\ &= \mathbb{E}\prod_{i=1}^n \mathbf{1}\{X_i\in S_i\} = \frac{N_1}{N}\cdot\frac{N_2}{N-1}\cdots\frac{N_n}{N-n+1}. \end{aligned} \tag{7.57}$$

Furthermore, for any $x=(x_j)_1^N \in \ell_1^N$ with $\|x\|\leqslant 1$, if $s_i := \sum_{j\in S_i}|x_j|$, then by the arithmetic-geometric inequality,

$$\big|L\big(x^{\otimes n}\big)\big| \leqslant s_1\cdots s_n \leqslant \Big(\frac{\sum_i s_i}{n}\Big)^n \leqslant n^{-n}. \tag{7.58}$$

Consequently, by (4.22), $\|L\|^*_{\pi,\mathsf{s},+} \leqslant n^{-n}$, and thus, recalling (7.54),

$$\kappa(n,N) = \|\chi_{n,N}\|_{\pi,\mathsf{s},+} \geqslant n^n L(\chi_{n,N}) = \prod_{i=1}^n \frac{nN_i}{N-i+1} = \frac{1}{q}\prod_{i=1}^n \frac{nN_i}{N}. \quad (7.59)$$

Suppose first that N is a multiple of n; $N = \ell n$ for an integer ℓ. Then we may choose $N_i = N/n = \ell$ for each i, and thus (7.59) yields,

$$\log \kappa(n,N) \geqslant -\log q = -\sum_{i=1}^{n-1} \log\Bigl(1 - \frac{i}{N}\Bigr) \geqslant \sum_{i=1}^{n-1} \frac{i}{N} = \frac{n(n-1)}{2N}. \quad (7.60)$$

For a general $N \geqslant n$ we let $\ell := \lceil N/n \rceil$ and $N_1 := \ell n$. Then $N \leqslant N_1 < N + n$, and (7.60) yields, using (7.45),

$$\log \kappa(n,N) \geqslant \log \kappa(n,N_1) \geqslant \frac{n(n-1)}{2N_1} = \frac{n-1}{2\lceil N/n \rceil} \geqslant \frac{n(n-1)}{2(N+n)} \quad (7.61)$$

and (7.48) follows.

(iii): Suppose that $\mu_{\mathbf{X}} \in \mathcal{E}_{n,N}$. Then, by Diaconis and Freedman [7, Theorem (3)], there exists a probability measure λ such that if $\mu_0 := \int_{\mathcal{P}(S)} \nu^n \,\mathrm{d}\lambda(\nu)$, then $\|\mu_{\mathbf{X}} - \mu_0\| \leqslant 2mn/N$. Consequently,

$$\begin{aligned}\|\mu_{\mathbf{X}}\|_{\pi,\mathsf{s},+;\ell_1(S)} &\leqslant \|\mu_0\|_{\pi,\mathsf{s},+} + \|\mu_{\mathbf{X}} - \mu_0\|_{\pi,\mathsf{s},+} \leqslant 1 + c_{\mathsf{s},+}(n,\ell_1(S))\|\mu_{\mathbf{X}} - \mu_0\|_{\pi} \\ &\leqslant 1 + c_{\mathsf{s},+}(n,\ell_1^m)\frac{2mn}{N}. \end{aligned} \quad (7.62)$$

The result follows by (7.10) and Theorem 5.1.

(iv): We modify Example 2.7 again. We may assume $S = [m]$. Let $n_1,\dots,n_m$ and $N_1,\dots,N_m$ be positive integers with $\sum_1^m n_k = n$ and $\sum_1^m N_k = N$. Partition $[N]$ and $[n]$ into sets S_k and T_k, respectively, with $|S_k| = N_k$ and $|T_k| = n_k$. Define a multilinear operator $L : (\ell_1^m)^n \to \mathbb{R}$ by, writing $x_i = (x_{ij})_{j=1}^m$,

$$L(x_1,\dots,x_n) = \prod_{k=1}^m \prod_{i\in T_k} x_{ik}. \quad (7.63)$$

If $x = (x_j)_1^m \in \ell_1^m$ with $\|x\| \leqslant 1$, then by the arithmetic-geometric inequality,

$$\bigl|L\bigl(x^{\otimes n}\bigr)\bigr| = \prod_{k=1}^m |x_k|^{n_k} = \prod_{k=1}^m n_k^{n_k} \prod_{k=1}^m \Bigl(\frac{|x_k|}{n_k}\Bigr)^{n_k} \leqslant \prod_{k=1}^m n_k^{n_k} \Bigl(\frac{\|x\|}{n}\Bigr)^n. \quad (7.64)$$

Consequently, by (4.22),

$$\|L\|^*_{\pi,\mathsf{s},+} \leqslant n^{-n} \prod_{k=1}^m n_k^{n_k}. \quad (7.65)$$

Let $\mathbf{x} = (x_1,\dots,x_N) \in [m]^N$ with $x_i = k$ when $i \in S_k$, and let $(X_1,\dots,X_n)$ be a random vector obtained by drawing without replacement from $x_1,\dots,x_N$.

Then $(X_1,\dots,X_n)$ has distribution $\varphi_{\mathbf{x}}^{\otimes n}\chi_{n,N}$, and thus, with the notation $(N)_n := N(N-1)\cdots(N-n+1)$, similarly to (7.56)–(7.57),

$$\begin{aligned}L\bigl(\varphi_{\mathbf{x}}^{\otimes n}\chi_{n,N}\bigr) &= \mathbb{E}\,L\bigl(\delta_{X_1}\otimes\ldots\otimes\delta_{X_n}\bigr) = \mathbb{E}\,L(\delta_{X_1},\dots,\delta_{X_n})\\ &= \mathbb{E}\prod_{k=1}^{m}\prod_{i\in T_k}\mathbf{1}\{X_i=k\} = \frac{(N_1)_{n_1}\cdots(N_m)_{n_m}}{(N)_n}. \end{aligned}\tag{7.66}$$

Consequently, by (7.41), (7.65) and (7.66),

$$\begin{aligned}\kappa(n,N;m) &\geqslant \|\varphi_{\mathbf{x}}^{\otimes n}\chi_{n,N}\|_{\pi,\mathbf{s},+} \geqslant \frac{L(\varphi_{\mathbf{x}}^{\otimes n}\chi_{n,N})}{\|L\|^*_{\pi,\mathbf{s},+}} \geqslant \frac{n^n\prod_{k=1}^m(N_k)_{n_k}}{(N)_n\prod_{k=1}^m n_k^{n_k}}\\ &= \frac{\prod_{k=1}^m\prod_{j=1}^{n_k-1}(1-j/N_k)}{\prod_{j=1}^{n-1}(1-j/N)}\prod_{k=1}^m\Bigl(\frac{nN_k}{Nn_k}\Bigr)^{n_k} \end{aligned}\tag{7.67}$$

Suppose first again that $N=\ell n$ is a multiple of n. Then, given any $n_1,\dots,n_k$ with sum n, we may choose $N_k=\ell n_k$ for each k. Then the final product in (7.67) is 1, and (7.67) yields, using Lemma 7.14 below with $t=1/\ell$,

$$\log\kappa(n,N;m) \geqslant \frac{m-1}{2\ell} = \frac{(m-1)n}{2N}. \tag{7.68}$$

For a general $N\geqslant n$ we let $\ell := \lceil N/n\rceil$ and $N_1 := \ell n$. Then $N\leqslant N_1 < N+n$, and (7.68) yields

$$\log\kappa(n,N;m) \geqslant \log\kappa(n,N_1;m) \geqslant \frac{m-1}{2\lceil N/n\rceil} \geqslant \frac{(m-1)n}{2(N+n)}. \tag{7.69}$$

and (7.50) follows. □

Lemma 7.14. *Let $n\geqslant m\geqslant 1$ and let $n_1,\dots,n_m$ be positive integers with $\sum_1^m n_k = n$. Then, for every $t\in[0,1]$,*

$$\sum_{k=1}^m\sum_{i=0}^{n_k-1}\log\Bigl(1-t\frac{i}{n_k}\Bigr) - \sum_{i=0}^{n-1}\log\Bigl(1-t\frac{i}{n}\Bigr) \geqslant \frac{m-1}{2}t. \tag{7.70}$$

Proof. Define two positive measures on $[0,1)$ by

$$\nu_1 := \sum_{k=1}^m\sum_{i=0}^{n_k-1}\delta_{i/n_k}, \qquad\qquad \nu_2 := \sum_{i=0}^{n-1}\delta_{i/n}. \tag{7.71}$$

Both ν_1 and ν_2 are integer-valued and have total mass n. Furthermore, for any $x\in[0,1]$, the number of integers $i\geqslant 0$ such that $i/n_k < x$ equals $\lceil n_k x\rceil$. Hence,

$$\nu_1[0,x) = \sum_{k=1}^m\lceil n_k x\rceil \geqslant \sum_{k=1}^m n_k x = nx, \tag{7.72}$$

$$\nu_2[0,x) = \lceil nx\rceil. \tag{7.73}$$

Since ν_1 is integer-valued, it follows that

$$\nu_1[0,x) \geqslant \lceil nx \rceil = \nu_2[0,x), \qquad x \in [0,1]. \tag{7.74}$$

This implies, by a standard argument using integration by parts, that if $f(x)$ is any decreasing function on $[0,1)$, then

$$\int_0^1 f(x)\,\mathrm{d}\nu_1(x) \geqslant \int_0^1 f(x)\,\mathrm{d}\nu_2(x). \tag{7.75}$$

Choose $f(x) := \log(1-tx)+tx$. Then (7.75) implies

$$\begin{aligned}
&\sum_{k=1}^{m}\sum_{i=0}^{n_k-1} \log\Bigl(1-t\frac{i}{n_k}\Bigr) - \sum_{i=0}^{n-1}\log\Bigl(1-t\frac{i}{n}\Bigr) \\
&\qquad = \int_0^1 \log(1-tx)\bigl(\mathrm{d}\nu_1(x)-\mathrm{d}\nu_2(x)\bigr) \\
&\qquad \geqslant -\int_0^1 tx\bigl(\mathrm{d}\nu_1(x)-\mathrm{d}\nu_2(x)\bigr) = -\sum_{k=1}^{m}\sum_{i=0}^{n_k-1} t\frac{i}{n_k} + \sum_{i=0}^{n-1} t\frac{i}{n} \\
&\qquad = -t\sum_{k=1}^{m}\frac{n_k-1}{2} + t\frac{n-1}{2} = t\frac{m-1}{2}.
\end{aligned} \tag{7.76}$$

□

Remark 7.15. The proofs of the lower bounds in (7.48) and (7.50) really yields lower bounds for $\|\ \|_{\pi,\mathsf{s}}$ and not just the larger $\|\ \|_{\pi,\mathsf{s},+}$ in (7.59) and (7.67). Hence, the lower bounds cannot be expected to be close to the true values.

8. Further examples

Example 8.1. Let $E = \ell_2^2$, i.e., $\mathbb{R}^2$ with the usual Euclidean norm. If $A \in E^{\vee 2}$, so A is a symmetric 2×2 matrix, then, by Examples 3.5 and 3.11,

$$\|A\|_{\pi} = \|A\|_{\pi,\mathsf{s}} = |\lambda_1| + |\lambda_2|, \tag{8.1}$$

where λ_1, λ_2 are the eigenvalues of A.

In particular, taking $A := \left(\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right)$,

$$\bigl\|\left(\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right)\bigr\|_{\pi} = \bigl\|\left(\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right)\bigr\|_{\pi,\mathsf{s}} = 2. \tag{8.2}$$

Furthermore, $A = e_1\otimes e_2 + e_2\otimes e_1 = 2e_1 \vee e_2$, and thus (4.3) (or (4.5)) yields, together with (4.8),

$$\bigl\|\left(\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right)\bigr\|_{\pi,+} = 2. \tag{8.3}$$

A positive unit vector in ℓ_2^2 is $(\cos t, \sin t)$ for some $t \in [0,\frac{\pi}{2}]$. Hence, a representation of A as in (4.4) can be written

$$A = \int_0^{\pi/2} (\cos t, \sin t)^{\otimes 2}\,\mathrm{d}\mu(t) \tag{8.4}$$

for a signed measure μ on $[0,\frac{\pi}{2}]$ with finite support. Thus, $\|A\|_{\pi,\mathsf{s},+}$ is the infimum of $\|\mu\|$ over all such μ satisfying (8.4).

With $A = \left(\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right)$ as above, (8.4) says

$$\int_0^{\pi/2} \cos^2 t \,\mathrm{d}\mu(t) = \int_0^{\pi/2} \sin^2 t \,\mathrm{d}\mu(t) = 0, \qquad \int_0^{\pi/2} \cos t \sin t \,\mathrm{d}\mu(t) = 1, \tag{8.5}$$

and thus

$$\int_0^{\pi/2} \bigl(1-2\sin 2t\bigr)\,\mathrm{d}\mu(t) = \int_0^{\pi/2} \bigl(\cos^2 t + \sin^2 t - 4\sin t\cos t\bigr)\,\mathrm{d}\mu(t) = -4. \tag{8.6}$$

Since $\bigl|1-2\sin 2t\bigr| \leqslant 1$ on $[0,\frac{\pi}{2}]$, (8.6) implies $\|\mu\| \geqslant 4$, which is attained by $\mu = 2\delta_{\pi/4} - \delta_0 - \delta_{\pi/2}$. Hence,

$$\bigl\|\left(\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right)\bigr\|_{\pi,\mathsf{s},+} = 4 > \bigl\|\left(\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right)\bigr\|_{\pi,+} = 2. \tag{8.7}$$

In particular, (8.7) shows that the result by Banach [2], see again Example 3.11, does not extend to the positive tensor norms.

Example 8.2. Consider as in the previous example $E=\ell_2^2$. The different norms in $E^{\vee 2}$ can be described geometrically as follows.

We give a matrix $A \in E^{\vee 2}$ the coordinates (u,v,w) defined by

$$A = \frac{1}{2}\begin{pmatrix} u+w & v \\ v & u-w \end{pmatrix}. \tag{8.8}$$

In these coordinates, we have $(\cos t,\sin t)^{\otimes 2} = (1,\sin 2t,\cos 2t)$.

The unit ball of $E^{\vee 2}_{\pi} = E^{\vee 2}_{\pi,\mathsf{s}}$ (see Example 3.11 again) is by Remark 3.7 thus the convex hull of

$$\bigl\{\pm(1,\sin 2t,\cos 2t) : t\in[0,2\pi]\bigr\} = \bigl\{\pm(1,\sin s,\cos s) : s\in[0,2\pi]\bigr\}. \tag{8.9}$$

This is the convex hull of the union of two symmetric circles, and thus the unit ball is the cylinder $\{|u|\leqslant 1,\ v^2+w^2\leqslant 1\}$. In other words, $\|(u,v,w)\|_{\pi,\mathsf{s}} = \max\bigl\{|u|,\sqrt{v^2+w^2}\bigr\}$, which also easily is seen from (8.1).

For $E^{\vee 2}_{\pi,\mathsf{s},+}$ we are by (4.4) only allowed to use positive vectors $(\cos t,\sin t)$, i.e., $t\in[0,\frac{\pi}{2}]$. Consequently, the unit ball of $E^{\vee 2}_{\pi,\mathsf{s},+}$ is the convex hull of the union of two symmetric half-circles:

$$\bigl\{\pm(1,\sin s,\cos s) : s\in[0,\pi]\bigr\}. \tag{8.10}$$

Finally, in our coordinates, $(\cos s,\sin s)\vee(\cos t,\sin t) = \bigl(\cos(s-t),\sin(s+t),\cos(s+t)\bigr)$. When $s,t\in[0,\frac{\pi}{2}]$, we have $s+t\in[0,\pi]$ and $|s-t|\leqslant \min\{s+t,\pi-s-t\}$. It follows from (4.3) that the unit ball of $E^{\vee 2}_{\pi,+}$ is the convex hull of the union of two half-circles (the same as in (8.10)) and four elliptic arcs given by

$$\bigl\{\pm(1,\sin s,\cos s) : s\in[0,\pi]\bigr\} \cup \bigl\{\pm(|\cos s|,\sin s,\cos s) : s\in[0,\pi]\bigr\}. \tag{8.11}$$

Note that the three sets in (8.9), (8.10) and (8.11) are the sets of extreme points of the unit balls.

To help visualizing these three unit balls, we consider their orthogonal projections onto the plane $Q := \{w = 0\}$, which are the same as their intersections with Q since they all are symmetric with respect to reflection in this plane. It follows easily from (8.9), (8.10) and (8.11) that these projections all are polygons, with corners (extreme points)

$$B(E_\pi^{\vee 2}) : \{(\pm 1, \pm 1, 0)\}, \tag{8.12}$$

$$B(E_{\pi,\mathsf{s},+}^{\vee 2}) : \{\pm(1,1,0),\ \pm(1,0,0)\}, \tag{8.13}$$

$$B(E_{\pi,+}^{\vee 2}) : \{\pm(1,1,0),\ \pm(1,0,0),\ \pm(0,1,0)\}. \tag{8.14}$$

Equivalently, recalling (8.8) and taking $u = 2a$, $v = 2b$, for any $a, b \in \mathbb{R}$,

$$\left\|\left(\begin{smallmatrix} a & b \\ b & a \end{smallmatrix}\right)\right\|_\pi = \left\|\left(\begin{smallmatrix} a & b \\ b & a \end{smallmatrix}\right)\right\|_{\pi,\mathsf{s}} = \|(u,v,0)\|_\pi = \max\{|u|,|v|\} = 2\max\{|a|,|b|\}, \tag{8.15}$$

$$\left\|\left(\begin{smallmatrix} a & b \\ b & a \end{smallmatrix}\right)\right\|_{\pi,\mathsf{s},+} = \|(u,v,0)\|_{\pi,\mathsf{s},+} = \max\{|u|,|u-2v|\} = 2\max\{|a|,|a-2b|\}, \tag{8.16}$$

$$\left\|\left(\begin{smallmatrix} a & b \\ b & a \end{smallmatrix}\right)\right\|_{\pi,+} = \|(u,v,0)\|_{\pi,+} = \max\{|u|,|v|,|u-v|\} = 2\max\{|a|,|b|,|a-b|\}. \tag{8.17}$$

In particular, we find again (8.2), (8.3) and (8.7).

Conversely, (8.15)–(8.17) can be found by the analytic method in Example 8.1.

We find also, as another specific example,

$$\left\|\left(\begin{smallmatrix} 1 & -1 \\ -1 & 1 \end{smallmatrix}\right)\right\|_\pi = \left\|\left(\begin{smallmatrix} 1 & -1 \\ -1 & 1 \end{smallmatrix}\right)\right\|_{\pi,\mathsf{s}} = 2, \tag{8.18}$$

$$\left\|\left(\begin{smallmatrix} 1 & -1 \\ -1 & 1 \end{smallmatrix}\right)\right\|_{\pi,\mathsf{s},+} = 6, \tag{8.19}$$

$$\left\|\left(\begin{smallmatrix} 1 & -1 \\ -1 & 1 \end{smallmatrix}\right)\right\|_{\pi,+} = 4. \tag{8.20}$$

We claim that

$$c_{\mathsf{s},+}(2,\ell_2^2) = c_{\mathsf{s};\,\mathsf{s},+}(2,\ell_2^2) = 3. \tag{8.21}$$

In fact, the two polarization constants are equal by (4.40). They are at least 3 by (8.18) and (8.19). Finally, to show that they are at most 3, it suffices by (8.9) and (8.10) to consider $\mathbf{x} = (1, \sin s \cos s)$ with $s \in (\pi, 2\pi)$. Then, with $s' := s - \pi$,

$$\mathbf{x} = -(1, \sin s', \cos s') + (1,0,1) + (1,0,-1), \tag{8.22}$$

which by (8.10) shows that $\|\mathbf{x}\|_{\pi,\mathsf{s},+} \leqslant 3$; it then follows from (4.25) that $c_{\mathsf{s},+}(2,\ell_2^2) \leqslant 3$.

By a similar argument, using (8.7) for the lower bound and (8.11) and (8.10) for the upper, we obtain (omitting the details), recalling (4.27),

$$c_{+;\,\mathsf{s},+}(2,\ell_2^2) = 2. \tag{8.23}$$

We can also see that, as shown in (4.33), the norm of the identity $E_\pi^{\vee 2} \to E_{\pi,+}^{\vee 2}$ is $c_+(\ell_2^2)^2 = 2$, cf. Example 4.1.

Acknowledgement. I thank Takis Konstantopoulos for interesting discussions, and two anonymous refrees for helpful comments on the presentation.

References

[1] David J. Aldous: *Exchangeability and related topics.* École d'été de probabilités de Saint-Flour, XIII–1983, 1–198, Lecture Notes in Math., 1117, Springer, Berlin, 1985.

[2] S. Banach: Über homogene Polynome in (L^2). *Studia Math.* **7** (1938), 36–44.

[3] Carlos Benítez, Yannis Sarantopoulos & Andrew Tonge: Lower bounds for norms of products of polynomials. *Math. Proc. Cambridge Philos. Soc.* **124** (1998), no. 3, 395–408.

[4] Qingying Bu & Gerard Buskes: Polynomials on Banach lattices and positive tensor products. *J. Math. Anal. Appl.* **388** (2012), no. 2, 845–862.

[5] Claude Dellacherie & Paul-André Meyer: *Probabilities and Potential B.* (Translated from French.) North-Holland, Amsterdam, 1982.

[6] Persi Diaconis: Finite forms of de Finetti's theorem on exchangeability. *Synthese* **36** (1977), no. 2, 271–281.

[7] Persi Diaconis & David Freedman: Finite exchangeable sequences. *Ann. Probab.* **8** (1980), no. 4, 745–764.

[8] Verónica Dimant, Daniel Galicer & Jorge Tomás Rodríguez: The polarization constant of finite dimensional complex spaces is one. Preprint, 2019. `arXiv:1908.08107`

[9] Seán Dineen: *Complex Analysis on Infinite Dimensional Spaces.* Springer-Verlag London, 1999.

[10] Klaus Floret: Natural norms on symmetric tensor products of normed spaces. *Note Mat.* **17** (1997), 153–188.

[11] D. H. Fremlin: Tensor products of Banach lattices. *Math. Ann.* **211** (1974), 87–106.

[12] Shmuel Friedland & Lek-Heng Lim: Nuclear norm of higher-order tensors. *Math. Comp.* **87** (2018), no. 311, 1255–1281.

[13] I. C. Gohberg & M. G. Kreĭn: *Introduction to the Theory of Linear Nonselfadjoint Operators.* (Translated from Russian.) Amer. Math. Soc., Providence, R.I., 1969.

[14] Lawrence A. Harris: Commentary on Problem 73, *The Scottish Book: Mathematics from the Scottish Café*, ed. R. Daniel Maudlin, Birkhäuser, Boston, 1981, 143–146.

[15] Edwin T. Jaynes: Some applications and extensions of the de Finetti representation theorem. *Bayesian Inference and Decision Techniques*, 31–42, North-Holland, Amsterdam, 1986.

[16] Svante Janson, Takis Konstantopoulos & Linglong Yuan: On a representation theorem for finitely exchangeable random vectors. *J. Math. Anal. Appl.* **442** (2016), 703–714.

[17] Olav Kallenberg: *Probabilistic Symmetries and Invariance Principles.* Springer, New York, 2005.

[18] Olav Kallenberg: *Random Measures, Theory and Applications.* Springer, Cham, Switzerland, 2017.

[19] G. Jay Kerns & Gábor J. Székely: De Finetti's theorem for abstract finite exchangeable sequences, *J. Theoret. Probab.* **19** (2006), no. 3, 589–608.

[20] Takis Konstantopoulos & Linglong Yuan: On the extendibility of finitely exchangeable probability measures. *Trans. Amer. Math. Soc.*, to appear.

[21] Peter D. Lax: *Functional Analysis.* Wiley, New York, 2002.

[22] Peter Meyer-Nieberg: *Banach Lattices.* Springer-Verlag, Berlin, 1991.

[23] *NIST Handbook of Mathematical Functions.* Edited by Frank W. J. Olver, Daniel W. Lozier, Ronald F. Boisvert & Charles W. Clark. Cambridge Univ. Press, 2010.
Also available as *NIST Digital Library of Mathematical Functions*, `http://dlmf.nist.gov/`

[24] Alexandros Pappas, Andreas Kavadjiklis & Michael Karamolengos: Polarization constants of polynomials on Banach spaces. *Nonlinear Funct. Anal. Appl.* **14** (2009), no. 4, 551–562.

[25] Yang Qi, Pierre Comon & Lek-Heng Lim: Uniqueness of nonnegative tensor approximations. *IEEE Trans. Inform. Theory* **62** (2016), no. 4, 2170–2183.

[26] Yang Qi, Pierre Comon & Lek-Heng Lim: Semialgebraic geometry of nonnegative tensor rank. *SIAM J. Matrix Anal. Appl.* **37** (2016), no. 4, 1556–1580.

[27] Theodore J. Rivlin: *The Chebyshev Polynomials.* Wiley, New York, 1974.

[28] Walter Rudin: *Functional Analysis.* 2nd ed., McGraw-Hill, New York, 1991.

[29] Raymond A. Ryan: *Introduction to Tensor Products on Banach Spaces.* Springer-Verlag, London, 2002.

[30] Helmut H. Schaefer: *Banach Lattices and Positive Operators.* Springer-Verlag, New York-Heidelberg, 1974.

[31] François Trèves: *Topological Vector Spaces, Distributions and Kernels.* Academic Press, New York–London 1967.

Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden

Email address: `svante.janson@math.uu.se`

URL: `http://www2.math.uu.se/~svante/`