Electron. Commun. Probab. **20** (2015), no. 27, 1–12.
DOI: 10.1214/ECP.v20-3992
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

Uniform estimates for averages of order statistics of matrices*

Richard Lechner† Markus Passenbrunner† Joscha Prochno†

Abstract

We prove uniform estimates for the expected value of averages of order statistics of matrices in terms of their largest entries. As an application, we obtain similar probabilistic estimates for ℓ_p norms via real interpolation.

Keywords: Order statistic; probabilistic inequality; combinatorial inequality; real interpolation.
AMS MSC 2010: 60E15; 05A20.
Submitted to ECP on December 16, 2014, final version accepted on March 17, 2015.

1 Introduction and Main Results

Combinatorial and probabilistic inequalities play an important role in a variety of areas of mathematics, especially in Banach space theory. In [8] and [9], S. Kwapień and C. Schütt studied combinatorial expressions involving matrices and obtained inequalities in terms of the average of the largest entries of the matrix. To be more precise, they showed that

$$\frac{1}{n!}\sum_{\pi\in\mathfrak{S}_n}\max_{1\le i\le n}|a_{i\pi(i)}| \simeq \frac{1}{n}\sum_{k=1}^{n} s(k), \tag{1.1}$$

where $s(k)$ is the k-th largest entry of the matrix a and $\mathfrak{S}_n$ the symmetric group. This estimate seems crucial if one wants to compute the projection constant of symmetric Banach spaces and related invariants. Among other things, the authors obtained estimates for the positive projection constant of finite dimensional Orlicz spaces and estimated the order of the projection constant of the Lorentz spaces $\ell_{2,1}^n$. Also, the symmetric sublattices of $\ell_1(c_0)$ as well as the finite dimensional symmetric subspaces of ℓ_1 were characterized. Further applications and extensions of (1.1) can be found in [9, 16, 17, 11, 13], just to mention a few.

The main result of this paper is a generalization of (1.1) in the sense that we study the expected value of averages of higher order statistics of a matrix in a more general setting described below. Our method of proof is purely probabilistic in nature, whereas the proof of (1.1) in [8] uses non-trivial combinatorial arguments.

In what follows, given a finite set G, we denote the normalized counting measure on G by $\mathbb{P}$, i.e.,

$$\mathbb{P}(E) = \frac{|E|}{|G|}, \qquad E\subseteq G,$$

*Support: Austrian Science Fund grants FWF P23987; FWF P22549; FWF P23987; FWFM 1628000.
†University of Linz, Austria. E-mail: richard.lechner,markus.passenbrunner,joscha.prochno@jku.at

where $|\cdot|$ denotes the cardinality. $\mathbb{E}$ will always denote the expectation with respect to the normalized counting measure. Moreover, for a vector $x \in \mathbb{R}^n$ with non-negative entries, we denote its k-th largest entry by

$$\operatorname*{k-max}_{1\le i\le n} x_i.$$

In particular, $\operatorname{1-max}_{1\le i\le n} x_i$ is the maximal value, $\operatorname{n-max}_{1\le i\le n} x_i$ the minimal value of x. Our main result is the following:

Theorem 1.1. *Let $n, N \in \mathbb{N}$ and $a \in \mathbb{R}^{n\times N}$. Let G be a collection of maps from $I = \{1,\dots,n\}$ to $J = \{1,\dots,N\}$ and $C_G > 0$ be a constant only depending on G. Assume that for all $i \in I$, $j \in J$ and all different pairs $(i_1, j_1), (i_2, j_2) \in I \times J$*

(i) $\mathbb{P}(\{g \in G : g(i) = j\}) = 1/N$,

(ii) $\mathbb{P}(\{g \in G : g(i_1) = j_1, g(i_2) = j_2\}) \le C_G/N^2$.

Then, for every $\ell \le n$,

$$\frac{c}{N}\sum_{j=1}^{\ell N} s(j) \le \int_G \sum_{k=1}^{\ell} \operatorname*{k-max}_{1\le i\le n} |a_{ig(i)}| \,\mathrm{d}\mathbb{P}(g) \le \frac{2}{N}\sum_{j=1}^{\ell N} s(j), \tag{1.2}$$

where $c = 2^{-5}(1+2C_G)^{-2}$.

Observe that estimate (1.1) [8, Theorem 1.1] is a special case of our result with the choice $\ell = 1$ and $G = \mathfrak{S}_n$, and that for $\ell = 1$ and $G = \{1,\dots,n\}^{\{1,\dots,n\}}$ we directly obtain [2, Lemma 7]. Note that in this general setting $\mathbb{E}\max_{1\le i\le n} |a_{ig(i)}|$ was already studied in [9]. In a slightly different setting, order statistics were considered also in [2, 3, 4, 5, 6].

We will now present two natural choices for the set G that appear frequently in the literature (cf. [8, 9, 16, 17, 15, 13, 2, 1, 7, 12]).

Example 1.2. If $N = n$ and $G = \mathfrak{S}_n$ is the group of permutations of the numbers $\{1,\dots,n\}$, then $\mathbb{P}(\pi(i) = j) = 1/n$ for all $1 \le i, j \le n$. Moreover, for $(i_1, j_1) \ne (i_2, j_2)$ we have $\mathbb{P}(\pi(i_1) = j_1, \pi(i_2) = j_2) \le 1/[n(n-1)] \le 2/n^2$. Hence, the assumptions of Theorem 1.1 are satisfied with $C_G \le 2$ and thus

$$\frac{1}{800}\frac{1}{n}\sum_{j=1}^{\ell n} s(j) \le \frac{1}{n!}\sum_{\pi\in\mathfrak{S}_n}\sum_{k=1}^{\ell} \operatorname*{k-max}_{1\le i\le n} |a_{i\pi(i)}| \le 2\frac{1}{n}\sum_{j=1}^{\ell n} s(j).$$

Example 1.3. If $N = n$ and G is the set of all mappings from $\{1,\dots,n\}$ into $\{1,\dots,n\}$, then $\mathbb{P}(g(i) = j) = 1/n$ for all $1 \le i, j \le n$. Moreover, for $(i_1, j_1) \ne (i_2, j_2)$ we have $\mathbb{P}(g(i_1) = j_1, g(i_2) = j_2) \le 1/n^2$. Hence, the assumptions of Theorem 1.1 are satisfied with $C_G = 1$ and thus

$$\frac{1}{288}\frac{1}{n}\sum_{j=1}^{\ell n} s(j) \le \frac{1}{n^n}\sum_{g\in G}\sum_{k=1}^{\ell} \operatorname*{k-max}_{1\le i\le n} |a_{ig(i)}| \le 2\frac{1}{n}\sum_{j=1}^{\ell n} s(j).$$

Another combinatorial inequality that was obtained in [8, Theorem 1.2] and which turned out to be crucial to study and characterize symmetric subspaces of L_1 (cf. [16, 17, 13]) states that for all $1 \le p \le \infty$

$$\frac{1}{n!}\sum_{\pi\in\mathfrak{S}_n}\Big(\sum_{i=1}^{n} |a_{i\pi(i)}|^p\Big)^{1/p} \simeq \frac{1}{n}\sum_{k=1}^{n} s(k) + \Big(\frac{1}{n}\sum_{k=n+1}^{n^2} s(k)^p\Big)^{1/p}. \tag{1.3}$$

In Section 5, we will use Theorem 1.1 to generalize this result and show that the lower bound in (1.3) can be naturally derived via real interpolation. The upper bound is quite easily obtained and we just follow [8]. Please note that averages of order statistics of matrices naturally appear, as they are strongly related to the K-functional of the interpolation couple (ℓ_1, ℓ_∞). Again, two typical choices for the set of maps G are $\mathfrak{S}_n$ and $\{1,\dots,n\}^{\{1,\dots,n\}}$. We will prove the following result:

Theorem 1.4. *Let $n, N \in \mathbb{N}$, $a \in \mathbb{R}^{n\times N}$, and $1 \le p < \infty$. Let G be a collection of maps from $I = \{1,\dots,n\}$ to $J = \{1,\dots,N\}$ and $C_G > 0$ be a constant only depending on G. Assume that for all $i \in I$, $j \in J$ and all different pairs $(i_1, j_1), (i_2, j_2) \in I \times J$*

(i) $\mathbb{P}(\{g \in G : g(i) = j\}) = 1/N$,

(ii) $\mathbb{P}(\{g \in G : g(i_1) = j_1, g(i_2) = j_2\}) \le C_G/N^2$.

Then

$$C\left[\frac{1}{N}\sum_{k=1}^{N} s(k) + \Big(\frac{1}{N}\sum_{k=N+1}^{nN} s(k)^p\Big)^{1/p}\right] \le \mathbb{E}\Big(\sum_{i=1}^{n} |a_{ig(i)}|^p\Big)^{1/p}$$
$$\le \frac{1}{N}\sum_{k=1}^{N} s(k) + \Big(\frac{1}{N}\sum_{k=N+1}^{nN} s(k)^p\Big)^{1/p},$$

where $C > 0$ is a constant only depending on C_G.

The organization of the paper is as follows. In Section 3, we will prove the lower estimate in (1.2). This is done by reducing the problem to the case of matrices only taking values in $\{0, 1\}$ and showing the estimate for this subclass of matrices. In Section 4, we establish the upper bound in (1.2) by passing from averages of order statistics to equivalent Orlicz norms and using an extreme point argument. Section 5 contains the proof of Theorem 1.4.

2 Notation and Preliminaries

Throughout this paper we will use $|E|$ to denote the cardinality of a finite set E. By $\mathfrak{S}_n$ we denote the symmetric group on the set $\{1,\dots,n\}$. We will denote by $\lfloor x\rfloor$ and $\lceil x\rceil$ the largest integer $m \le x$ and the smallest integer $m \ge x$, respectively.

For an arbitrary matrix $a = (a_{ij})_{i,j=1}^{n,N}$, we denote by $(s(k))_{k=1}^{nN}$ the decreasing rearrangement of $(|a_{ij}|)_{i,j=1}^{n,N}$. To avoid confusion, in certain cases we write $(s_a(k))_{k=1}^{nN}$ to emphasize the underlying matrix a.

Please recall that the Paley-Zygmund inequality states that for every non-negative random variable Z and all $0 < \theta < 1$

$$\mathbb{P}(Z \ge \theta \cdot \mathbb{E}Z) \ge (1-\theta)^2 \frac{(\mathbb{E}Z)^2}{\mathbb{E}Z^2}. \tag{2.1}$$

A convex function $M : [0,\infty) \to [0,\infty)$ is called an Orlicz function if $M(0) = 0$ and if M is not constant. Given an Orlicz function M, the Orlicz sequence space ℓ_M^n is $\mathbb{R}^n$ equipped with the Luxemburg norm

$$\|x\|_M = \inf\left\{\lambda > 0 : \sum_{i=1}^{n} M\left(\frac{|x_i|}{\lambda}\right) \le 1\right\}.$$

For example, the classical ℓ_p spaces are Orlicz spaces with $M(t) = p^{-1}t^p$. The closed unit ball of the space ℓ_M^n will be denoted by B_M^n. We write $\mathrm{ext}(B_M^n)$ for the set of extreme points of B_M^n and $\text{s-conv}(M)$ shall denote the set of points of strict convexity of M. We will make use of the following characterization of extreme points of B_M^n:

Lemma 2.1 ([18], Lemma 1). *Let M be an Orlicz function. Then $x \in \operatorname{ext}(B_M^n)$ if and only if*

(i) $\sum_{i=1}^n M(|x_i|) = 1$,

(ii) there exists at most one index $i_0 \in \mathbb{N}$, $1 \leq i_0 \leq n$ such that $x_{i_0} \notin \pm\,\text{s-conv}(M)$.

For a detailed and thorough introduction to Orlicz spaces we refer the reader to [14] or [10].

3 The lower bound

In this section we will prove the lower bound in (1.2). We begin by recalling some notation and assumptions given in Theorem 1.1. Let $a \in \mathbb{R}^{n\times N}$, $I = \{1,\dots,n\}$, $J = \{1,\dots,N\}$, and G be a collection of maps from I to J. The matrix a will be fixed throughout the entire section. By $\mathbb{P}$ we denote the normalized counting measure on G, i.e., $\mathbb{P}(E) = |E|/|G|$ for $E \subset G$. We assume a uniform distribution of the random variable $g \mapsto g(i)$ for each $i \in I$, i.e.,

$$\mathbb{P}(g(i) = j) = \frac{1}{N}, \qquad i \in I, j \in J.$$

Moreover, we assume for all different pairs $(i_1, j_1), (i_2, j_2) \in I \times J$ that

$$\mathbb{P}(g(i_1) = j_1, g(i_2) = j_2) \leq \frac{C_G}{N^2},$$

with a constant $C_G \geq 1$ that depends on G, but not on n or N.

Without loss of generality, we will assume that a has only non-negative entries. It is enough to show the lower estimate in (1.2) for matrices a that consist of only the ℓN largest entries, while all others are equal to zero. This is because if we change any entry $a_{i_0 j_0} \leq s(\ell N + 1)$ by setting $a_{i_0 j_0} = 0$, the left hand side in (1.2) remains the same, while $\underset{1\leq i\leq n}{\text{k-max}} |a_{ig(i)}|$ does not increase for any $g \in G$.

3.1 The key ingredients

We will now introduce a bijective function h that determines the ordering of the values of a. The crucial point is that this function does not depend on the actual values of the matrix, but merely on their relative size. So let $h : \{1,\dots,nN\} \to I \times J$ be a bijective function satisfying

$$\begin{aligned} &a(h(j)) \geq a(h(j+1)), && 1 \leq j \leq \ell N, \\ &a(h(j)) = 0, && \ell N + 1 \leq j \leq nN. \end{aligned} \tag{3.1}$$

Observe that there is possibly more than one choice for h, since some of the entries of the matrix a might have the same value.

For all $j \in \mathbb{N}$, $1 \leq j \leq nN$, define the random variable

$$Y_j : G \to \{0,1\}, \qquad Y_j(g) = \begin{cases} 1, & \text{if } h(j) \in g, \\ 0, & \text{if } h(j) \notin g, \end{cases}$$

and given $m \in \mathbb{N}$, $1 \leq m \leq nN$, let

$$X_m : G \to \{0,1,\dots,n\}, \qquad X_m(g) := \sum_{j=1}^m Y_j(g) = |h(\{1,\dots,m\}) \cap g|,$$

where we identify g with its graph $\{(i, g(i)) : i \in I\}$. X_m counts the number of elements in the path $\{(i, g(i)) : i \in I\}$ that intersect with the positions of the m largest entries of a.

As we will see in Subsection 3.2, the random variables X_m are strongly related to order statistics.

In Lemma 3.1, Lemma 3.2, and Lemma 3.3, we investigate crucial properties of the distribution function of X_m.

Lemma 3.1. *For all $m \in \mathbb{N}$ with $1 \le m \le nN$, we have*

$$\mathbb{P}(X_m \ge 1) \ge \frac{m}{N}\Big(1 - C_G \frac{m-1}{2N}\Big). \tag{3.2}$$

In particular,

$$\mathbb{P}(X_{\lceil N/C_G \rceil} \ge 1) \ge \frac{1}{2C_G}.$$

Proof. By using the inclusion-exclusion principle, we obtain

$$\begin{aligned} \mathbb{P}(X_m \ge 1) &= \mathbb{P}\Big(\bigcup_{j=1}^m \{g \in G : Y_j(g) = 1\}\Big) \ge \sum_{j=1}^m \mathbb{P}(Y_j = 1) - \sum_{i<j} \mathbb{P}(Y_i = 1, Y_j = 1) \\ &\ge \frac{m}{N} - \frac{m(m-1)C_G}{2N^2} = \frac{m}{N}\Big(1 - C_G\frac{m-1}{2N}\Big), \end{aligned}$$

where the latter inequality is a direct consequence of conditions (i) and (ii) in Theorem 1.1. □

Lemma 3.2. *For all $m \in \mathbb{N}$ with $1 \le m \le nN$, and all $\theta \in (0,1)$, we have*

$$\mathbb{P}\Big(X_m \ge \theta \cdot \frac{m}{N}\Big) \ge (1-\theta)^2 \frac{m}{N + m \cdot C_G}. \tag{3.3}$$

Proof. The result follows as a consequence of Paley-Zygmund's inequality (cf. (2.1)). Therefore, we need to compute $\mathbb{E}X_m$ and $\mathbb{E}X_m^2$. Note that $\mathbb{E}Y_j = \mathbb{P}(Y_j = 1) = 1/N$ and thus $\mathbb{E}X_m = \sum_{j=1}^m Y_j = m/N$. Moreover, since $Y_j = Y_j^2$, we have

$$\mathbb{E}X_m^2 = \sum_{i,j=1}^m \mathbb{E}Y_iY_j = \sum_{j=1}^m \mathbb{E}Y_j + \sum_{i \ne j} \mathbb{E}Y_iY_j \le \frac{m}{N} + C_G\frac{m(m-1)}{N^2},$$

where the latter inequality is a direct consequence of conditions (i) and (ii) in Theorem 1.1. Inserting those estimates in (2.1), we obtain the result. □

Lemma 3.3. *For all $m \in \mathbb{N}$ with $1 \le m \le nN$, we have*

$$\mathbb{P}(X_m \ge 1) \ge \min\Big\{\frac{m}{2N}, \frac{1}{2C_G}\Big\} \mathbb{P}(X_{\ell N} \ge 1),$$

and for all $k, m \in \mathbb{N}$ with $2kN \le m \le nN$

$$\mathbb{P}(X_m \ge k) \ge \frac{\mathbb{P}(X_{\ell N} \ge k)}{2 + 4C_G}. \tag{3.4}$$

Proof. Let $1 \le m \le nN$. If $m \le N/C_G$, Lemma 3.1 implies

$$\mathbb{P}(X_m \ge 1) \ge \frac{m}{2N} \ge \frac{m}{2N} \cdot \mathbb{P}(X_{\ell N} \ge 1) = \min\Big\{\frac{m}{2N}, \frac{1}{2C_G}\Big\} \mathbb{P}(X_{\ell N} \ge 1).$$

On the other hand, if $m \ge N/C_G$, Lemma 3.1 implies

$$\begin{aligned} \mathbb{P}(X_m \ge 1) &\ge \mathbb{P}(X_{\lceil N/C_G \rceil} \ge 1) \ge \frac{1}{2C_G} \\ &\ge \frac{1}{2C_G}\mathbb{P}(X_{\ell N} \ge 1) = \min\Big\{\frac{m}{2N}, \frac{1}{2C_G}\Big\} \mathbb{P}(X_{\ell N} \ge 1). \end{aligned}$$

Now we prove (3.4). Let $k \le n/2$ and m such that $2kN \le m \le nN$. Then Lemma 3.2 with $\theta = 1/2$ implies

$$\mathbb{P}(X_m \ge k) \ge \mathbb{P}(X_{2kN} \ge k) \ge \frac{1}{2+4C_G} \ge \frac{1}{2+4C_G}\mathbb{P}(X_{\ell N} \ge k). \qquad \square$$

3.2 Reduction to two valued matrices

We will now reduce the problem of estimating the expected value of averages of order statistics of general matrices to matrices only taking one value different from zero. To do so, we need some more definitions.

Let $\mathcal{A}_h$ be the collection of all non-negative real $n \times N$ matrices b that satisfy

$$\begin{aligned} b(h(j)) &\ge b(h(j+1)), & 1 \le j \le \ell N, \\ b(h(j)) &= 0, & \ell N + 1 \le j \le nN. \end{aligned} \tag{3.5}$$

For every $b \in \mathcal{A}_h$, we set

$$\widetilde{b}(h(j)) := \Big(\frac{1}{\ell N}\sum_{i=1}^{\ell N} b(h(i))\Big) \cdot \mathbb{1}_{\{1,\dots,\ell N\}}(j), \qquad 1 \le j \le nN,$$

as the matrix that contains the averaged entries of b. Note that $\widetilde{b} \in \mathcal{A}_h$. Moreover, we define

$$a_m(h(k)) := \mathbb{1}_{\{1,\dots,m\}}(k), \qquad 1 \le m \le nN.$$

Observe that $a_m \in \mathcal{A}_h$ for all $1 \le m \le nN$. For $b \in \mathcal{A}_h$ and $g \in G$ we put

$$S_k(b)(g) := \operatorname*{k-max}_{1 \le i \le n} b_{ig(i)} \qquad \text{and} \qquad S(b)(g) := \sum_{k=1}^{\ell} S_k(b)(g).$$

Lemma 3.4. *Let $m \in \mathbb{N}$, $1 \le m \le \ell N$. Then we have*

$$\mathbb{E}S(\widetilde{a_m}) \le (8 + 16C_G) \cdot \mathbb{E}S(a_m).$$

Proof. Observe that for every integer k with $1 \le k \le \ell$,

$$\begin{aligned} \mathbb{E}S_k(a_m) &= \mathbb{P}(S_k(a_m) = 1) = \mathbb{P}(X_m \ge k), \\ \mathbb{E}S_k(\widetilde{a_m}) &= \frac{m}{\ell N} \cdot \mathbb{P}\Big(S_k(\widetilde{a_m}) = \frac{m}{\ell N}\Big) = \frac{m}{\ell N} \cdot \mathbb{P}(X_{\ell N} \ge k). \end{aligned}$$

As a consequence,

$$\mathbb{E}S(\widetilde{a_m}) = \frac{m}{\ell N}\sum_{k=1}^{\ell} \mathbb{P}(X_{\ell N} \ge k) \qquad \text{and} \qquad \mathbb{E}S(a_m) = \sum_{k=1}^{\ell}\mathbb{P}(X_m \ge k).$$

Thus, in order to prove the lemma, it is enough to show that

$$\frac{m}{\ell N}\sum_{k=1}^{\ell}\mathbb{P}(X_{\ell N} \ge k) \le (8 + 16C_G) \cdot \sum_{k=1}^{\ell}\mathbb{P}(X_m \ge k)$$

for $1 \le m \le \ell N$. First, we assume $m \le 2N$. Then, Lemma 3.3 implies

$$\begin{aligned} \frac{m}{\ell N}\sum_{k=1}^{\ell}\mathbb{P}(X_{\ell N} \ge k) &\le \frac{m}{N} \cdot \mathbb{P}(X_{\ell N} \ge 1) \le \frac{2}{N}\max\{N, m \cdot C_G\}\mathbb{P}(X_m \ge 1) \\ &\le 4C_G \cdot \sum_{k=1}^{\ell}\mathbb{P}(X_m \ge k), \end{aligned}$$

i.e., the assertion of the lemma for $m \le 2N$.

Now, let $m \ge 2N+1$ and choose the integer $t \ge 1$ such that $2tN+1 \le m \le 2(t+1)N$. The sequence $k \mapsto \mathbb{P}(X_{\ell N} \ge k)$ is decreasing, hence, noting that $t \le \ell$,

$$\frac{m}{\ell N}\sum_{k=1}^{\ell} \mathbb{P}(X_{\ell N} \ge k) \le \frac{m}{tN}\sum_{k=1}^{t} \mathbb{P}(X_{\ell N} \ge k).$$

Then, estimate (3.4) of Lemma 3.3 implies

$$\begin{aligned}\frac{m}{\ell N}\sum_{k=1}^{\ell} \mathbb{P}(X_{\ell N} \ge k) &\le \frac{m}{tN}\sum_{k=1}^{t}(2+4C_G)\cdot \mathbb{P}(X_m \ge k) \\ &\le \frac{(4+8C_G)(t+1)}{t}\sum_{k=1}^{t}\mathbb{P}(X_m \ge k) \\ &\le (8+16C_G)\cdot \sum_{k=1}^{\ell}\mathbb{P}(X_m \ge k). \end{aligned}$$ □

Lemma 3.5. *We have*

$$\mathbb{E}S(\widetilde{a}) \le (8+16C_G)\cdot \mathbb{E}S(a).$$

Proof. Recall that $X_j(g) = |h(\{1,\dots,j\}) \cap g|$. Hence, for all $b \in \mathcal{A}_h$,

$$\mathbb{E}S(b) = \sum_{k=1}^{\ell}\sum_{j=1}^{\ell N} b(h(j))\cdot \mathbb{P}(\{g : X_{j-1}(g) = k-1, h(j) \in g\}).$$

Defining

$$f(j) := \sum_{k=1}^{\ell}\mathbb{P}(\{g : X_{j-1}(g) = k-1, h(j)\in g\}), \qquad 1 \le j \le \ell N,$$

we can write

$$\mathbb{E}S(b) = \sum_{j=1}^{\ell N} f(j) b(h(j)).$$

Since $a, \widetilde{a} \in \mathcal{A}_h$, $a(h(j)) = s_a(j)$ and $\widetilde{a}(h(j)) = (\ell N)^{-1}\sum_{i=1}^{\ell N} s_a(i)$ for all $j \le \ell N$, we obtain

$$\mathbb{E}S(a) = \sum_{j=1}^{\ell N} f(j)s_a(j) \qquad \text{and} \qquad \mathbb{E}S(\widetilde{a}) = \sum_{j=1}^{\ell N}\widetilde{f}(j)s_a(j),$$

where for all $1 \le j \le \ell N$

$$\widetilde{f}(j) = \frac{1}{\ell N}\sum_{i=1}^{\ell N} f(i).$$

Note that the functions f and $\widetilde{f}$ only depend on h, i.e., only on the positions of the entries in the matrix and not on their values. Since $a_m(h(j)) = 1$ for $j \le m$ (zero otherwise) and $a_m, \widetilde{a_m} \in \mathcal{A}_h$, we have

$$\mathbb{E}S(a_m) = \sum_{j=1}^{m} f(j) \qquad \text{and} \qquad \mathbb{E}S(\widetilde{a_m}) = \sum_{j=1}^{m}\widetilde{f}(j).$$

Now we conclude with $C = 8 + 16C_G$ that

$$\begin{aligned} C\mathbb{E}S(a) - \mathbb{E}S(\widetilde{a}) &= s_a(1)[Cf(1) - \widetilde{f}(1)] + \sum_{j=2}^{\ell N}[Cf(j) - \widetilde{f}(j)]s_a(j) \\ &\geq s_a(2)\sum_{j=1}^{2}[Cf(j) - \widetilde{f}(j)] + \sum_{j=3}^{\ell N}[Cf(j) - \widetilde{f}(j)]s_a(j) \end{aligned}$$

where we used Lemma 3.4 for $m = 1$. Continuing in this fashion and using Lemma 3.4 for $m = 2, \ldots, \ell N$, we obtain

$$C\mathbb{E}S(a) - \mathbb{E}S(\widetilde{a}) \geq 0.$$

□

3.3 Conclusion

As we have seen, we can reduce the case of general a to multiples of matrices only taking values zero and one. Before we finally prove the lower bound in the main theorem, we will need another simple lemma.

Lemma 3.6. *Let $b \in \mathcal{A}_h$ be an $(n \times N)$-matrix consisting of ℓN ones and $(n - \ell)N$ zeros. Then, for all $1 \leq k \leq \ell/2$,*

$$\mathbb{E} \operatorname*{k-max}_{1\leq i\leq n} b_{ig(i)} \geq \frac{1}{2 + 4C_G}.$$

Proof. Let $k \leq \ell/2$. Using Lemma 3.2 with $\theta = 1/2$, we obtain

$$\begin{aligned} \mathbb{E} \operatorname*{k-max}_{1\leq i\leq n} b_{ig(i)} &\geq \int_{\{g: X_{2kN}(g)\geq k\}} \operatorname*{k-max}_{1\leq i\leq n} b_{ig(i)} \, \mathrm{d}\mathbb{P}(g) \\ &= \mathbb{P}(X_{2kN} \geq k) \geq \frac{k}{2(1 + 2kC_G)} \geq \frac{1}{2 + 4C_G}. \end{aligned}$$

□

Proof of the lower bound in Theorem 1.1. By Theorem 3.5 we obtain

$$\mathbb{E}\sum_{k=1}^{\ell} \operatorname*{k-max}_{1\leq i\leq n} a_{ig(i)} \geq \frac{1}{8(1 + 2C_G)}\mathbb{E}\sum_{k=1}^{\ell} \operatorname*{k-max}_{1\leq i\leq n} \widetilde{a}_{ig(i)}.$$

Now take $b \in \mathcal{A}_h$ consisting of ℓN ones and $(n - \ell)N$ zeros such that

$$\left(\frac{1}{\ell N}\sum_{i=1}^{\ell N} s_a(i)\right) \cdot b = \widetilde{a}.$$

Then, by Lemma 3.6,

$$\begin{aligned} \mathbb{E}\sum_{k=1}^{\ell} \operatorname*{k-max}_{1\leq i\leq n} \widetilde{a}_{ig(i)} &= \left[\mathbb{E}\sum_{k=1}^{\ell} \operatorname*{k-max}_{1\leq i\leq n} b_{ig(i)}\right] \frac{1}{\ell N}\sum_{i=1}^{\ell N} s_a(i) \\ &\geq \left[\mathbb{E}\sum_{k=1}^{\ell/2} \operatorname*{k-max}_{1\leq i\leq n} b_{ig(i)}\right] \frac{1}{\ell N}\sum_{i=1}^{\ell N} s_a(i) \\ &\geq \frac{1}{4 + 8C_G}\frac{1}{N}\sum_{j=1}^{\ell N} s(j). \end{aligned}$$

Combining the above estimates, we obtain

$$\mathbb{E}\sum_{k=1}^{\ell} \operatorname*{k-max}_{1\leq i\leq n} a_{ig(i)} \geq \frac{1}{32(1 + 2C_G)^2}\frac{1}{N}\sum_{j=1}^{\ell N} s(j). \tag{3.6}$$

□

4 The upper bound

We will now prove the upper bound of Theorem 1.1 via an extreme point argument. To do so, we first use the fact that the average of the $j \le nN$ largest entries of a matrix $a \in \mathbb{R}^{n\times N}$ is equivalent to an Orlicz norm $\|a\|_{M_j}$ (cf. Lemma 4.1). Then, since the expected value of the average of order statistics defines a norm on $\mathbb{R}^{n\times N}$ as well, it is enough to prove the upper bound in Theorem 1.1 for the extreme points of $B^n_{M_j}$.

Recall that, for a vector $(x_i)_{i=1}^n \in \mathbb{R}^n$, we denote the decreasing rearrangement of $(|x_i|)_{i=1}^n$ by $(x_i^*)_{i=1}^n$. We start with the approximation of sums of decreasing rearrangements of vectors $x \in \mathbb{R}^n$ by equivalent Orlicz norms.

The following result is due to C. Schütt (private communication) and follows by direct computation. With his permission we include it here.

Lemma 4.1. *Let $j \in \mathbb{N}$, $1 \le j \le n$. Then, for all $x \in \mathbb{R}^n$, we have*

$$\frac{1}{2}\sum_{i=1}^{j} x_i^* \le \|x\|_{M_j} \le \sum_{i=1}^{j} x_i^*,$$

where

$$M_j(t) := \begin{cases} 0, & 0 \le t \le 1/j, \\ t - 1/j, & 1/j < t. \end{cases} \tag{4.1}$$

We are now able to prove the upper bound of Theorem 1.1.

Proposition 4.2. *Let $a \in \mathbb{R}^{n\times N}$. Then, for all $\ell \le n$,*

$$\mathbb{E}\sum_{k=1}^{\ell} \operatorname*{k-max}_{1\le i\le n} a_{ig(i)} \le \frac{2}{N}\|a\|_{M_{\ell N}}. \tag{4.2}$$

Proof. It is sufficient to show (4.2) for all $a \in \operatorname{ext}(B^{nN}_{M_\ell})$. Therefore, by Lemma 2.1 (2), we only need to consider matrices $a \in \mathbb{R}^{n\times N}$ that are of the form

$$a_{ij} := \begin{cases} \frac{1}{\ell N}, & (i,j) \ne (i_0, j_0), \\ 1 + \frac{1}{\ell N}, & (i,j) = (i_0, j_0) \end{cases} \tag{4.3}$$

for some index pair $(i_0, j_0) \in I \times J$ or satisfy $a_{ij} = \frac{1}{\ell N}$ for all $i = 1, \dots, n$, $j = 1, \dots, N$. However, the latter choice of a does not satisfy condition (1) in Lemma 2.1, since in that case $\sum_{i=1}^{nN} M_{\ell N}(s_a(i)) = 0$. So the extreme points of $B^{nN}_{M_{\ell N}}$ with positive entries are given by (4.3). Now, let a be such a point in $B^{nN}_{M_{\ell N}}$. Then

$$\begin{aligned}
&\mathbb{E}\sum_{k=1}^{\ell} \operatorname*{k-max}_{1\le i\le n} a_{ig(i)} \\
&= \int_{\{g:g(i_0)=j_0\}} \sum_{k=1}^{\ell} \operatorname*{k-max}_{1\le i\le n} a_{ig(i)}\, \mathrm{d}\mathbb{P}(g) + \int_{\{g:g(i_0)\ne j_0\}} \sum_{k=1}^{\ell} \operatorname*{k-max}_{1\le i\le n} a_{ig(i)}\, \mathrm{d}\mathbb{P}(g) = \frac{2}{N}.
\end{aligned}$$

On the other hand, we also have $\sum_{j=1}^{\ell N} s(j) = 2$. Therefore,

$$\mathbb{E}\sum_{k=1}^{\ell} \operatorname*{k-max}_{1\le i\le n} a_{ig(i)} = \frac{1}{N}\sum_{j=1}^{\ell N} s(j),$$

The result follows, since by Lemma 4.1

$$\sum_{j=1}^{\ell N} s(j) \le 2\|a\|_{M_{\ell N}}. \qquad \square$$

Proof of Theorem 1.1. Combining Lemma 4.1 and Proposition 4.2, we get

$$\mathbb{E}\sum_{k=1}^{\ell} \operatorname*{k-max}_{1\le i\le n} a_{ig(i)} \le \frac{2}{N}\sum_{i=1}^{\ell N} s(i), \tag{4.4}$$

which is the upper estimate in Theorem 1.1. Inequalities (3.6) and (4.4) together complete the proof. □

5 An application of Theorem 1.1

We now present an application and use Theorem 1.1 to prove Theorem 1.4. The proof uses real interpolation and is, what we find, a natural approach to combinatorial inequalities such as (1.3) that were obtained in [8]. Please notice that [8, Theorem 1.2] is a special case of Theorem 1.4 when $G = \mathfrak{S}_n$.

Let us first recall some basic notions from interpolation theory. A pair (X_0, X_1) of Banach spaces is called a compatible couple if there is some Hausdorff topological space $\mathcal{H}$, in which each of X_0 and X_1 is continuously embedded. For example, (L_1, L_∞) is a compatible couple, since L_1 and L_∞ are continuously embedded into the space of measurable functions that are finite almost everywhere. Of course, any pair (X, Y) for which one of the spaces is continuously embedded in the other is a compatible couple.

For a compatible couple (X_0, X_1) (with corresponding Hausdorff space $\mathcal{H}$), we equip $X_0 + X_1$ with the norm

$$\|x\|_{X_0+X_1} := \inf_{x=x_0+x_1} \big(\|x_0\|_{X_0} + \|x_1\|_{X_1}\big), \tag{5.1}$$

under which this space becomes a Banach space. This definition is independent of the particular space $\mathcal{H}$.

The K-functional is constructed from the expression (5.1) by introducing a positive weighting factor $t > 0$, as follows: let (X_0, X_1) be a compatible couple. The K-functional is defined for each $f \in X_0 + X_1$ and $t > 0$ by

$$K(f,t) = K(f,t;X_0,X_1) := \inf_{f=f_0+f_1} (\|f_0\|_{X_0} + t\|f_1\|_{X_1}),$$

where the infimum extends over all representations $f = f_0 + f_1$ of f with $f_0 \in X_0$ and $f_1 \in X_1$.

Now, let (X_0, X_1) be a compatible couple and suppose $0 < \theta < 1$, $1 \le q < \infty$ or $0 \le \theta \le 1$ and $q = \infty$. The space $(X_0, X_1)_{\theta,q}$ consists of all $f \in X_0 + X_1$ for which the functional

$$\|f\|_{\theta,q} := \begin{cases} \left(\int_0^\infty \left[t^{-\theta} K(f,t;X_0,X_1)\right]^q \frac{\mathrm{d}t}{t}\right)^{1/q}, & 0<\theta<1,\ 1\le q<\infty, \\ \sup_{t>0} t^{-\theta} K(f,t;X_0,X_1), & 0\le\theta\le 1,\ q=\infty, \end{cases}$$

is finite.

Proof of Theorem 1.4

The proof of the upper bound is the same as in [8, Theorem 1.2] so we just present a proof of the lower bound. Let $1 \le p < \infty$, $\theta = 1 - 1/p$ and recall that

$$\|a\|_{\theta,p} = \left(\int_0^\infty \left[t^{-\theta} K\big(a,t;L_1^{|G|}(\ell_1^n), L_1^{|G|}(\ell_\infty^n)\big)\right]^p \frac{\mathrm{d}t}{t}\right)^{1/p}.$$

Observe that

$$K\big(a,t;L_1^{|G|}(\ell_1^n), L_1^{|G|}(\ell_\infty^n)\big) = \int_G K\big(a(g),t;\ell_1^n,\ell_\infty^n\big)\,\mathrm{d}\mathbb{P}(g). \tag{5.2}$$

This in combination with the triangle inequality for integrals yields

$$\begin{aligned}\|a\|_{\theta,p} &= \Big(\int_0^\infty [\int_G t^{-\theta} K\big(a(g),t;\ell_1^n,\ell_\infty^n)\big)\,\mathrm{d}\mathbb{P}(g)]^p \frac{\mathrm{d}t}{t}\Big)^{1/p} \\ &\le \int_G \Big(\int_0^\infty \Big[t^{-\theta} K\big(a(g),t;\ell_1^n,\ell_\infty^n)\big)\Big]^p \frac{\mathrm{d}t}{t}\Big)^{1/p} \mathrm{d}\mathbb{P}(g) \\ &= \int_G \|a(g)\|_{\ell_p^n}\,\mathrm{d}\mathbb{P}(g).\end{aligned}$$

Therefore, we have

$$\|a\|_{L_1^{|G|}(\ell_p^n)} \ge \|a\|_{\theta,p},$$

for all $a : G \to \mathbb{R}^n$, $a(g)(i) = a_{ig(i)}$. Now we compute the K-functional of the interpolation couple $(L_1^{|G|}(\ell_1^n), L_1^{|G|}(\ell_\infty^n))$ using (5.2). Observe that

$$\begin{aligned}K(a(g),t;\ell_1^n,\ell_\infty^n) &= \int_0^{\lfloor t\rfloor} (a(g))^*(s)\,\mathrm{d}s + \int_{\lfloor t\rfloor}^t (a(g))^*(s)\,\mathrm{d}s \\ &= \sum_{k=1}^{\lfloor t\rfloor} \text{k-max}\,|a(g)| + (t-\lfloor t\rfloor)\cdot \lceil \text{t}\rceil\text{-max}\,|a(g)|.\end{aligned}$$

Then, using (5.2), we obtain

$$\begin{aligned}\|a\|_{1-\frac{1}{p},p}^p &= \int_0^\infty t^{-p}\Big(\int_G K(a(g),t;\ell_1^n,\ell_\infty^n)\,\mathrm{d}\mathbb{P}(g)\Big)^p \mathrm{d}t \\ &= \int_0^\infty t^{-p}\Big(\int_G \sum_{k=1}^{\lfloor t\rfloor} \text{k-max}\,|a(g)| + (t-\lfloor t\rfloor)\cdot\lceil \text{t}\rceil\text{-max}\,|a(g)|)\,\mathrm{d}\mathbb{P}(g)\Big)^p \mathrm{d}t \\ &\ge \int_0^1 \big(\mathbb{E}a(g)^*(1)\big)^p\,\mathrm{d}t + \int_1^{n+1} t^{-p}\Big(\mathbb{E}\sum_{k=1}^{\lfloor t\rfloor} a(g)^*(k)\Big)^p \mathrm{d}t \\ &\ge c_1\Big[\big(\mathbb{E}a(g)^*(1)\big)^p + \sum_{\ell=1}^n \Big(\mathbb{E}\frac{1}{\ell}\sum_{k=1}^{\ell} a(g)^*(k)\Big)^p\Big],\end{aligned}$$

where c_1 is a positive absolute constant. By Theorem 1.1, we get

$$\begin{aligned}\Big[\mathbb{E}a(g)^*(1)\Big]^p + \sum_{\ell=1}^n\Big(\mathbb{E}\frac{1}{\ell}\sum_{k=1}^{\ell} a(g)^*(k)\Big)^p &\ge c_2\Big[\Big(\frac{1}{N}\sum_{j=1}^N s(j)\Big)^p + \sum_{\ell=1}^n\Big(\frac{1}{\ell N}\sum_{j=1}^{\ell N} s(j)\Big)^p\Big] \\ &\ge c_2\Big[\Big(\frac{1}{N}\sum_{j=1}^N s(j)\Big)^p + \sum_{\ell=1}^n \frac{1}{N}\sum_{j=\ell N+1}^{(\ell+1)N} s(j)^p\Big] \\ &= c_2\Big[\Big(\frac{1}{N}\sum_{j=1}^N s(j)\Big)^p + \frac{1}{N}\sum_{\ell=N+1}^{nN} s(\ell)^p\Big],\end{aligned}$$

where c_2 is a positive constant only depending on C_G. Taking the p-th root concludes the proof.

Acknowledgments. We would like to thank our colleague Erhard Aichinger for helpful discussions.

References

[1] E. D. Gluskin. Estimates of the norms of certain p-absolutely summing operators. *Funktsional. Anal. i Prilozhen.*, 12(2):24–31, 95, 1978.

[2] Y. Gordon, A. Litvak, C. Schütt, and E. Werner. Orlicz norms of sequences of random variables. *Ann. Probab.*, 30(4):1833–1853, 2002.

[3] Y. Gordon, A. Litvak, C. Schütt, and E. Werner. Geometry of spaces between polytopes and related zonotopes. *Bull. Sci. Math.*, 126(9):733–762, 2002.

[4] Y. Gordon, A. Litvak, C. Schütt, and E. Werner. Minima of sequences of Gaussian random variables. *C. R. Math. Acad. Sci. Paris*, 340(6):445–448, 2005.

[5] Y. Gordon, A. E. Litvak, C. Schütt, and E. Werner. On the minimum of several random variables. *Proc. Amer. Math. Soc.*, 134(12):3665–3675 (electronic), 2006.

[6] Y. Gordon, A. E. Litvak, C. Schütt, and E. Werner. Uniform estimates for order statistics and Orlicz functions. *Positivity*, 16(1):1–28, 2012.

[7] W. B. Johnson, B. Maurey, G. Schechtman, and L. Tzafriri. Symmetric structures in Banach spaces. *Mem. Amer. Math. Soc.*, 19(217):v+298, 1979.

[8] S. Kwapień and C. Schütt. Some combinatorial and probabilistic inequalities and their application to Banach space theory. *Studia Math.*, 82(1):91–106, 1985.

[9] S. Kwapień and C. Schütt. Some combinatorial and probabilistic inequalities and their application to Banach space theory. II. *Studia Math.*, 95(2):141–154, 1989.

[10] J. Lindenstrauss and L. Tzafriri. *Classical Banach spaces. I.* Springer-Verlag, Berlin-New York, 1977. Sequence spaces, Ergebnisse der Mathematik und ihrer Grenzgebiete, Vol. 92.

[11] S. Montgomery-Smith and E. Semenov. Random rearrangements and operators. In *Voronezh Winter Mathematical Schools*, volume 184 of *Amer. Math. Soc. Transl. Ser. 2*, pages 157–183. Amer. Math. Soc., Providence, RI, 1998.

[12] J. Prochno. A combinatorial approach to Musielak-Orlicz spaces. *Banach J. Math. Anal.*, 7(1):132–141, 2013.

[13] J. Prochno and C. Schütt. Combinatorial inequalities and subspaces of L_1. *Studia Math.*, 211(1):21–39, 2012.

[14] M. M. Rao and Z. D. Ren. *Theory of Orlicz spaces*, volume 146 of *Monographs and Textbooks in Pure and Applied Mathematics*. Marcel Dekker, Inc., New York, 1991.

[15] C. Schütt. On the positive projection constant. *Studia Math.*, 78(2):185–198, 1984.

[16] C. Schütt. Lorentz spaces that are isomorphic to subspaces of L^1. *Trans. Amer. Math. Soc.*, 314(2):583–595, 1989.

[17] C. Schütt. On the embedding of 2-concave Orlicz spaces into L^1. *Studia Math.*, 113(1):73–80, 1995.

[18] M. Wisła. Extreme points and stable unit balls in Orlicz sequence spaces. *Arch. Math. (Basel)*, 56(5):482–490, 1991.