Uniform estimates for order statistics and Orlicz functions *

Yehoram Gordon [†‡] Alexander E. Litvak [†] Carsten Schütt[†]
Elisabeth Werner[† §]

Abstract

We establish uniform estimates for order statistics: Given a sequence of independent identically distributed random variables $\xi_1, \dots, \xi_n$ and a vector of scalars $x = (x_1, \dots, x_n)$, and $1 \le k \le n$, we provide estimates for $\mathbb{E}\ k\text{-}\min_{1 \le i \le n} |x_i \xi_i|$ and $\mathbb{E}\, k\text{-}\max_{1 \le i \le n} |x_i \xi_i|$ in terms of the values k and the Orlicz norm $\|y_x\|_M$ of the vector $y_x = (1/x_1, \dots, 1/x_n)$. Here $M(t)$ is the appropriate Orlicz function associated with the distribution function of the random variable $|\xi_1|$, $G(t) = \mathbb{P}(\{|\xi_1| \le t\})$. For example, if ξ_1 is the standard $N(0,1)$ Gaussian random variable, then $G(t) = \sqrt{\frac{2}{\pi}} \int_0^t e^{-\frac{s^2}{2}} ds$ and $M(s) = \sqrt{\frac{2}{\pi}} \int_0^s e^{-\frac{1}{2t^2}} dt$. We would like to emphasize that our estimates do not depend on the length n of the sequence.

1 Introduction

In this paper we establish uniform estimates for order statistics. The k-th order statistic of a statistical sample of size n is equal to its k-th smallest value, or equivalently its $(n-k+1)$-th largest value. Order statistics are among the most fundamental tools in non-parametric statistics and inference and consequently there is extensive literature on order statistics. We only cite [1, 2, 7] and references therein.

Order statistics are more resilient to faulty sensor reading than max, min or average and thus they find applications when methods are needed to study configurations that take on a ranked order. To name only a few: wireless networks, signal processing, image processing, compressed sensing, data reconstruction, learning theory and data mining. A small sample of works done in this area are [3, 4, 5, 6, 9, 23, 25].

*Keywords: Order statistics, expectations, moments, normal distribution, exponential distribution, Orlicz norms. 2010 MSC: 60E15, 46B06, 46B09, 62G30.

†Partially supported by AIM, Palo Alto

‡Supported in part by "France-Israel Cooperation Grant #3-1350" and by the "Fund for the Promotion of Research at the Technion".

§Partially supported by a NSF Grant, by a FRG Grant, and by a BSF Grant

Order statistics on random sequences appear naturally in Banach space theory, in computations of various random parameters associated with the geometry of convex bodies in high dimensions, in random matrix theory (computing the distribution of eigenvalues, and in approximation theory (see e.g. [8, 10, 11, 12, 14, 17, 24, 28, 29, 30]). This list of course does not include the enormous quantity of published works which deal with evaluations and applications of *max* and *min* associated with various random parameters, e.g., smallest and largest eigenvalues of random matrices, as these are the extreme values in the scale of order statistics. For the important special cases of order statistics, the minimum and maximum value of a sample, very precise estimates were obtained in [13, 15, 16]. The new approach started there was to give estimates of the minimum and maximum value of the sample

$$\mathbb{E} \min_{1 \le i \le n} |x_i \xi_i| \quad \text{and} \quad \mathbb{E} \max_{1 \le i \le n} |x_i \xi_i|, \tag{1}$$

in terms of *Orlicz norms* (see the definition below). The expressions for the bounds on the expectations in (1) are relatively simple. For instance, it was shown in [13] that

$$c_1 \|x\|_M \le \mathbb{E} \max_{1 \le i \le n} |x_i \xi_i| \le c_2 \|x\|_M,$$

where c_1, c_2 are absolute positive constants and $\|\cdot\|_M$ is an Orlicz norm depending only on the distribution of ξ_1; and in [15, 16] that

$$c_3 \left(\sum_{i=1}^n \frac{1}{|x_i|} \right)^{-1} \le \mathbb{E} \min_{1 \le i \le n} |x_i \xi_i| \le c_4 \left(\sum_{i=1}^n \frac{1}{|x_i|} \right)^{-1},$$

where c_3, c_4 are absolute positive constants and the distribution of ξ_1 satisfies some natural conditions. In fact, in [13] much more general case was considered (see also [20] and [26]).

Here we study general order statistics for i.i.d. (independent identically distributed) random variables $\xi_1, \dots, \xi_n$ and scalars $x_1, \dots, x_n$

$$\mathbb{E}\, k\text{-}\min_{1 \le i \le n} |x_i \xi_i| \quad \text{and} \quad \mathbb{E}\, k\text{-}\max_{1 \le i \le n} |x_i \xi_i|, \tag{2}$$

where for a given sequence of real numbers $a_1, \dots, a_n$ we denote the k-th smallest one by $k\text{-}\min_{1 \le i \le n} a_i$. In particular, $1\text{-}\min_{1 \le i \le n} a_i = \min_{1 \le i \le n} a_i$ and $n\text{-}\min_{1 \le i \le n} a_i = \max_{1 \le i \le n} a_i$. In the same way we denote the k-th biggest number by $k\text{-}\max_{1 \le i \le n} a_i$. Thus, $k\text{-}\max_{1 \le i \le n} a_i = (n-k+1)\text{-}\min_{1 \le i \le n} a_i$. In fact, in the theory of order statistics the standard notation for $k\text{-}\min$ is $a_{k:n}$. In this paper such a notation could be misleading and we prefer to use $k\text{-}\min$.

Of course the expressions for the bounds on the expectations in (2) are much more involved than for expectations in (1). In view of possible applications we strive to keep them as simple as possible – at the expense of the constants involved. We show that for $1 \le k \le n/2$

$$c_1 \max_{1 \le j \le k} \|\, (1/x_i)_{i=j}^n \,\|^{-1}_{\frac{2e}{k-j+1} N} \le \mathbb{E}\ k\text{-}\min_{1 \le i \le n} |x_i \xi_i| \le c_2 \max_{1 \le j \le k} \|\, (1/x_i)_{i=j}^n \,\|^{-1}_{\frac{2e}{k-j+1} N}, \tag{3}$$

and for $1 \leq k \leq cn$

$$
\begin{aligned}
& c_3 \left(\max_{0\leq \ell\leq ck-1} \| (1/x_i)_{i=1}^{k+\ell} \|^{-1}_{\frac{2e}{\ell+1}N} + \|(x_{k+ck},\dots,x_n)\|_M \right) \\
(4) \qquad & \leq \mathbb{E} \operatorname*{k-max}_{1\leq i\leq n} |x_i\xi_i| \\
& \leq c_4 \left(\max_{0\leq \ell\leq ck-1} \| (1/x_i)_{i=1}^{k+\ell} \|^{-1}_{\frac{2e}{\ell+1}N} + \|(x_{k+ck},\dots,x_n)\|_M \right),
\end{aligned}
$$

where $\|\cdot\|_N$, $\|\cdot\|_M$ are Orlicz norms, corresponding to Orlicz functions N, M, which are computed in terms of the distribution function of the random variables under consideration and c, c_1, c_2, c_3, c_4 are positive constants that depend mildly on the distribution function and on k. But - and this is the main point here - the constants do not depend on n and on the scalars $x_1,\dots,x_n$. Moreover, the dependence on k is very mild: it is of order of $\ln k$ (or $1/\ln k$). The dependence on the distribution is essentially just normalization. The precise statements are given in Section 3.

We would like to emphasize that we do not think that the expressions in (3) and (4) can be simplified (at least in terms of Orlicz-type functions). In Section 3 we give an example (Example 3.4) that supports this.

In problems where only a small number of random variables is involved, numerical computations will give sufficient estimates for order statistics. However, in the case when a large number of random variables is involved, numerical computations may not be feasible. Our formulae allow easy computations also in that situation.

Finally let us mention that throughout this paper we use the following notation. For a random variable ξ on a probability space $(\Omega,\mathfrak{A},\mathbb{P})$ we denote its distribution function by G_ξ and $1-G_\xi$ by F_ξ

$$G_\xi(t) = \mathbb{P}(\{\xi \leq t\}) \qquad \text{and} \qquad F_\xi(t) = \mathbb{P}(\{\xi > t\}).$$

Acknowledgment. We would like to thank Hermann König, Kiel, for discussions.

2 Preliminaries. Orlicz functions and norms.

In this section we recall some facts about Orlicz functions and norms. For more details and other properties of Orlicz spaces we refer to [21, 22, 27].

A left continuous convex function $M : [0,\infty) \to [0,\infty]$ is called Orlicz function, if $M(0) = 0$ and if M is neither the function that is constant 0 nor the function that takes the value 0 at 0 and is ∞ elsewhere. The corresponding Orlicz norm on $\mathbb{R}^n$ is defined by

$$\|x\|_M = \inf\left\{ \rho > 0 \,\middle|\, \sum_{i=1}^n M(|x_i|/\rho) \leq 1 \right\}. \tag{5}$$

Note that the expression for $\|\cdot\|_M$ makes also sense if the function M is merely positive and increasing. Although in that case the expression need not be a norm, we keep the same notation $\|\cdot\|_M$. We often use formula (5) in a slightly different form, namely

$$1/\|x\|_M = \sup\left\{\rho > 0 \,\middle|\, \sum_{i=1}^{n} M\left(\rho\,|x_i|\right) \le 1\right\}.$$

Clearly, $M \le \bar{M}$ implies $\|\cdot\|_M \le \|\cdot\|_{\bar{M}}$. Moreover, if M is an Orlicz function and $s \ge 1$, then

$$sM(t) \le M(st) \tag{6}$$

for every $t \ge 0$. In particular, this implies

$$\|\cdot\|_{sM} \le s\|\cdot\|_M. \tag{7}$$

The dual function M^* to an Orlicz function M is defined by

$$M^*(s) = \sup_{0 \le t < \infty} (t \cdot s - M(t)).$$

For instance, for $M(t) = \frac{1}{q}\, t^q$, $q \ge 1$, the dual function is $M^*(t) = \frac{1}{q^*}\, t^{q^*}$ with $\frac{1}{q} + \frac{1}{q^*} = 1$.

Let the function $p = p_M : [0,\infty) \to [0,\infty]$ be given by

$$p(t) = \begin{cases} 0 & t = 0 \\ M'(t) & M(t) < \infty \\ \infty & M(t) = \infty, \end{cases}$$

where M' is the left hand side derivative of M. Then p is increasing and the left hand side inverse q of the increasing function p is

$$q(s) = \inf\{t \in [0,\infty) \mid p(t) > s\}.$$

Then

$$M^*(s) = \int_0^s q(t)dt.$$

To a given random variable ξ we associate an Orlicz function $M = M_\xi$ in the following way:

$$M(s) = \int_0^s \int_{\frac{1}{t} \le |\xi|} |\xi| d\mathbb{P} dt = \int_{\frac{1}{s} \le |\xi|} \left(s|\xi| - 1\right) d\mathbb{P}. \tag{8}$$

The equality here follows by changing the order of integration and the convexity of M follows by the definition of convexity. We prefer to keep in mind both formulae for M. Note that equivalently one can write

$$M(s) = \mathbb{E}\left(s|\xi| - 1\right)_+,$$

where, as usual, $h_+(x)$ denotes $h(x)$ if $h(x) \ge 0$ and 0 otherwise.

We claim that the dual function $M^* = M^*_\xi$ is given on $[0, \int |\xi| d\mathbb{P}]$ by

$$M^*\left(\int_{t\le|\xi|} |\xi| d\mathbb{P}\right) = \mathbb{P}(|\xi| \ge t) \tag{9}$$

and $M^*(s) = \infty$ for $s > \int |\xi| d\mathbb{P}$.

Indeed, by definition

$$\begin{aligned} M^*(s) = \sup_{0\le w}(w\cdot s - M(w)) &= \sup_{0\le w}\left(w\cdot s - \int_0^w \int_{\frac{1}{u}\le|\xi|} |\xi| d\mathbb{P} du\right) \\ &= \sup_{0\le w} \int_0^w \left(s - \int_{\frac{1}{u}\le|\xi|} |\xi| d\mathbb{P}\right) du. \end{aligned}$$

If $s > \int |\xi| d\mathbb{P}$ then the supremum is equal to ∞. Now fix $t \ge 0$, set

$$s = \int_{t\le|\xi|} |\xi| d\mathbb{P}$$

and consider the function

$$\phi(w) := \int_0^w \left(s - \int_{\frac{1}{u}\le|\xi|} |\xi| d\mathbb{P}\right) du.$$

It is easy to see that ϕ is increasing on $[0, 1/t]$ and decreasing on $[1/t, \infty)$. Therefore,

$$M^*(s) = \sup_{0\le w} \phi(w) = \phi(1/t) = \int_0^{1/t} \int_{t\le|\xi|<1/u} |\xi| \; d\mathbb{P} \; du.$$

Changing the order of integration we obtain

$$M^*(s) = \int_{t\le|\xi|} \int_0^{1/|\xi|} |\xi| \; du \; d\mathbb{P} = \int_{t\le|\xi|} d\mathbb{P} = \mathbb{P}(|\xi| \ge t),$$

which proves (9). □

In the Gaussian case we have

$$F(t) = \mathbb{P}\left(\{|\xi| > t\}\right) = \sqrt{\tfrac{2}{\pi}} \int_t^\infty e^{-\frac{s^2}{2}} ds$$

and thus

$$M(s) = \sqrt{\tfrac{2}{\pi}} \int_0^s \int_{\frac{1}{t}}^\infty u e^{-\frac{u^2}{2}} du dt = \sqrt{\tfrac{2}{\pi}} \int_0^s e^{-\frac{1}{2t^2}} dt. \tag{10}$$

This implies that on the interval $[0, \sqrt{2/\pi}]$ M^* is given by

$$M^*(s) = \int_0^s \frac{1}{\sqrt{2\ln\left(\sqrt{\frac{2}{\pi}}\frac{1}{u}\right)}} du.$$

For $s > \sqrt{2/\pi}$, $M^*(s) = \infty$.

3 The main results

Now we consider certain functions associated with a random variable $\xi : \Omega \to \mathbb{R}$.

The function $F : [0,\infty) \to [0,\infty)$ is given by

$$F(t) = \mathbb{P}(|\xi| > t). \tag{11}$$

We assume that F is strictly decreasing on $[0,\infty)$ and $F(0) = 1$. In particular, F is invertible.

The function $N : [0,\infty) \to [0,\infty)$ is defined by

$$N(t) = \ln \frac{1}{F(t)} \tag{12}$$

and is assumed to be convex. In particular, N is an Orlicz function. For such a function N and $k \in \mathbb{N}$ we put

$$N_j = \frac{2e}{k-j+1} N, \qquad j = 1, \ldots, k. \tag{13}$$

Furthermore, let us observe that under assumptions above for all $t \geq 0$ and all $s \geq 1$ we have

$$F(st) \leq F(t)^s. \tag{14}$$

Indeed, by (6) we have $sN(t) \leq N(st)$, i.e. $-s \ln F(t) \leq -\ln F(st)$, which is equivalent to (14).

The following theorem generalizes results from [15, 16], where similar estimates were obtained for Gaussian distributions. Of course, the Gaussian case is simpler and the corresponding formulae are less involved. We discuss the details in Remark 3.2 following the theorem.

Theorem 3.1 *Let $1 \leq k \leq \frac{n}{2}$ and let $\xi_1, \ldots, \xi_n$ be i.i.d. copies of a random variable ξ. Let F, N and N_j, $j = 1, \ldots, k$, be as specified in (11), (12) and (13). Then for all $0 < x_1 \leq x_2 \leq \ldots \leq x_n$*

$$c_1 \max_{1\leq j\leq k} \| (1/x_i)_{i=j}^n \|_{N_j}^{-1} \leq \mathbb{E}\ k\text{-}\min_{1\leq i\leq n} |x_i \xi_i| \leq 16e^2\ C_N\ \ln(k+1) \max_{1\leq j\leq k} \| (1/x_i)_{i=j}^n \|_{N_j}^{-1},$$

where $c_1 = 1 - \frac{1}{\sqrt{2\pi}}$ and $C_N = \max\{N(1), 1/N(1)\}$.

Moreover, the lower estimate does not require the condition “N is an Orlicz function”.

Remark 3.2 (The Gaussian case.) In [15, 16] it was shown that for $N(0,1)$ random variables g_i, $i = 1, \ldots, n$ and for all $0 < x_1 \leq x_2 \leq \ldots \leq x_n$

$$c_0 \max_{1\leq j\leq k} \frac{k+1-j}{\sum_{i=j}^n 1/x_i} \leq \mathbb{E}\ k\text{-}\min_{1\leq i\leq n} |x_i g_i| \leq 2\sqrt{2\pi}\ \ln(k+1) \max_{1\leq j\leq k} \frac{k+1-j}{\sum_{i=j}^n 1/x_i}. \tag{15}$$

where $c_0 = \left(1 - \frac{1}{\sqrt{2\pi}}\right) \frac{1}{2e} \sqrt{\frac{\pi}{2}}$. It is well known (and can be directly computed) that the Gaussian distribution satisfies the conditions of Theorem 3.1. Thus the estimate (15) can be obtained from Theorem 3.1 (with different absolute constants). In Proposition 3.7 at the end of this section we show that in the Gaussian case $N \sim H$, where

$$H(t) = \begin{cases} t & \text{for } \ 0 \le t < 1 \\ t^2 & \text{for } \ t \ge 1. \end{cases}$$

Remark 3.3 It does not seem possible to get simpler expressions (related to Orlicz functions) that would eliminate the dependence of the constants on k in the estimates of Theorem 3.1. In particular we tried the approach from [19], which seems to be related, but it does not work. An indication for that is the following crucial example that rules out many natural candidates for simpler expressions.

Example 3.4 Let $\xi_1, \dots, \xi_n$ be independent $N(0,1)$-random variables and $0 \le x_1 \le \dots \le x_n$. If $x_1, \dots, x_k$ are significantly smaller than $x_{k+1}, \dots, x_n$ then

$$\mathbb{E}\ k\text{-}\min_{1 \le i \le n} |x_i \xi_i| \sim \mathbb{E} \max_{1 \le i \le k} |x_i \xi_i| \sim \|(x_i)_{i=1}^k\|_M,$$

where M is given in Lemma 5.2.

Our second theorem provides bounds for expectations of k-max. As in Theorem 3.1 we assume that F is strictly decreasing, $F(0) = 1$, and that $N = -\ln F$ is a convex function, where F is given by (11). Note that such a function F satisfies

$$\int_{t \le |\xi_1|} |\xi_1| d\mathbb{P} \le \left(1 + \frac{1}{N(t)}\right) t \cdot F(t) \tag{16}$$

for all positive t. We verify this. Since $F = e^{-N}$ and N is convex

$$\int_{t \le |\xi_1|} |\xi_1| d\mathbb{P} = t \cdot F(t) + \int_t^\infty F(s) ds = t \cdot F(t) + \int_t^\infty e^{-N(s)} ds.$$

Using (6), we have $N(s) \ge \frac{s}{t} N(t)$ for $s \ge t$. Therefore

$$\begin{aligned} \int_{t \le |\xi_1|} |\xi_1| d\mathbb{P} &\le t \cdot F(t) + \int_t^\infty e^{-\frac{s}{t} N(t)} ds \\ &\le t \cdot F(t) + \frac{t}{N(t)} e^{-N(t)} = t \cdot F(t) + \frac{t}{N(t)} F(t), \end{aligned}$$

which implies (16).

Theorem 3.5 *Let $\xi_1, \dots, \xi_n$ be i.i.d. copies of a random variable ξ. Let F, M, and N be as specified in (11), (8), and (12). Let $1 < k \le n$ and $k_0 = \left[\frac{4(k-1)}{F(1)}\right]$. Assume that $k + k_0 \le n$. Then for all $x_1 \ge x_2 \ge \ldots \ge x_n > 0$*

$$\frac{1}{4}\left(\max_{0\le \ell \le k_0-1} \| (1/x_i)_{i=1}^{k+\ell} \|^{-1}_{\frac{2e}{\ell+1}N} + \left(1 + \frac{\ln(8(k-1))}{N(1)}\right)^{-1} \|(x_{k+k_0}, \dots, x_n)\|_M \right)$$

$$\le \mathbb{E} \operatorname*{k\text{-}max}_{1\le i\le n} |x_i\xi_i| \le c\left(C_N \ln(k+1) \max_{0\le \ell\le k_0-1} \| (1/x_i)_{i=1}^{k+\ell} \|^{-1}_{\frac{2e}{\ell+1}N} + \|(x_{k+k_0}, \dots, x_n)\|_M \right),$$

where $C_N = \max\{N(1), 1/N(1)\}$, and c is an absolute positive constant.

Remark 3.6 The case $k = 1$ was obtained in [13] (see also Lemma 5.2 below): Let

$$M(s) = \int_0^s \int_{\frac{1}{t}\le |\xi_1|} |\xi_1| d\mathbb{P} dt = \int_{\frac{1}{s}\le|\xi|} (s|\xi| - 1)\, d\mathbb{P}.$$

Then, for all $x \in \mathbb{R}^n$ one has

$$c_1\|x\|_M \le \int_\Omega \max_{1\le i\le n} |x_i \xi_i(\omega)| d\mathbb{P}(\omega) \le c_2 \|x\|_M,$$

In particular, formula (10) shows that in the Gaussian case one has

$$M(s) = \sqrt{\tfrac{2}{\pi}} \int_0^s e^{-\frac{1}{2t^2}} dt.$$

We conclude this section with the following proposition, which shows that in the Gaussian case the function N (and hence the functions N_j) is easily computable.

Proposition 3.7 *Let $N = -\ln F$, where*

$$F(t) = \sqrt{\tfrac{2}{\pi}} \int_t^\infty e^{-\frac{s^2}{2}} ds,$$

and

$$H(t) = \begin{cases} t & \text{for } 0 \le t < 1 \\ t^2 & \text{for } t \ge 1. \end{cases}$$

Then H is an Orlicz function and for every $t \ge 0$

$$(2\pi e)^{-1/2}\ H(t) \le N(t) \le 4.5\ H(t).$$

In particular, if $k \le n$ and N_j, $j \le k$, as in (13) then for every $t \ge 0$

$$\sqrt{\frac{2e}{\pi}}\ \frac{1}{k-j+1}\ H(t) \le N_j(t) = \frac{2e}{k-j+1}\ N(t) \le \frac{9e}{k-j+1}\ H(t).$$

Proof. Clearly, H is an Orlicz function.

For every $0 \le t \le \sqrt{\pi/8}$ we have

$$\frac{1}{2} \le 1 - \sqrt{\tfrac{2}{\pi}}\ t \le F(t) = 1 - \sqrt{\tfrac{2}{\pi}} \int_0^t e^{-\frac{s^2}{2}} ds \le 1 - \sqrt{\tfrac{2}{e\pi}}\ t.$$

Since $(x-1)/2 \le \ln x \le x - 1$ on $[1,2]$, we observe for $0 \le t \le \sqrt{\pi/8}$

$$N(t) = \ln \frac{1}{F(t)} \le \frac{1}{F(t)} - 1 \le \frac{1}{1 - \sqrt{\frac{2}{\pi}}\ t} - 1 \le \frac{\sqrt{\frac{2}{\pi}}\ t}{1 - \sqrt{\frac{2}{\pi}}\ t} \le \sqrt{\tfrac{8}{\pi}}\ t$$

and

$$N(t) = \ln \frac{1}{F(t)} \ge \frac{1}{2}\left(\frac{1}{F(t)} - 1\right) \ge \frac{1}{2}\left(\frac{1}{1 - \sqrt{\frac{2}{e\pi}}\ t} - 1\right) \ge \frac{t}{\sqrt{2e\pi}}.$$

This shows the desired result for $0 \le t \le \sqrt{\pi/8}$.

Consider now the function $f(t) = N(t) - t^2/2$ and observe that $f(0) = 0$. Since

$$F(t) \le \sqrt{\tfrac{2}{\pi}} \int_t^\infty \frac{s}{t}\ e^{-\frac{s^2}{2}} ds = \sqrt{\tfrac{2}{\pi}}\ \frac{1}{t}\ \exp\left(-\frac{t^2}{2}\right),$$

we obtain that $f'(t) \ge 0$ for $t \ge 0$. Thus, for every $t \ge 0$ one has $N(t) \ge t^2/2$.

Finally, note that for every $t > 0$ and every $A > 0$

$$F(t) \ge \sqrt{\tfrac{2}{\pi}} \int_t^{t+A} e^{-\frac{s^2}{2}} ds \ge \sqrt{\tfrac{2}{\pi}}\ A \exp\left(-\left(t + A\right)^2/2\right).$$

Applying this with $A = \sqrt{\pi/2}$ and $t \ge \sqrt{\pi/8}$ (then $t + A \le 3t$) we get that

$$F(t) \ge \exp\left(-9t^2/2\right).$$

This implies $t^2/2 \le N(t) \le 9t^2/2$ for $t \ge \sqrt{\pi/8}$. In particular, $H/\sqrt{2\pi e} \le N \le 9H/2$. □

4 k-min (Proof of Theorem 3.1)

In this section we prove Theorem 3.1. For reader convenience we split the proof into 2 subsections. Although for Theorem 3.1 we need only $p = 1$ in the some statements below, we prefer to formulate them in full generality for possible future applications.

4.1 The lower bound in Theorem 3.1

The lower bound in Theorem 3.1 follows immediately from Proposition 4.4 below. To prove this Proposition we need the following two simple lemmas. Similar lemmas were used in [15, 16] (Lemmas 4 and 5 in [15] and Lemmas 8 and 9 in [16]). We omit the proofs.

Lemma 4.1 *Let $0 < x_1 \le x_2 \le ... \le x_n$. Let $\xi_1, \dots, \xi_n$ be identically distributed random variables. Let $F(t) = \mathbb{P}\{|\xi_1| > t\}$ and $G(t) = 1 - F(t)$. Then*

$$\mathbb{P}\left\{\min_{1\le i\le n} |x_i\xi_i| \le t\right\} \le \sum_{i=1}^{n} G\left(t/x_i\right).$$

Moreover, if the ξ_i's are independent then for every $t > 0$

$$\mathbb{P}\left\{\min_{1\le i\le n} |x_i\xi_i| > t\right\} = \prod_{i=1}^{n} F\left(t/x_i\right).$$

Lemma 4.2 *Let $1 \le k \le n$. Let $0 < x_1 \le x_2 \le ... \le x_n$ and $\xi_1, \dots, \xi_n$ be i.i.d. random variables. Let $G(t) = \mathbb{P}\{|\xi_1| \le t\}$ and*

$$a = a(t) = \frac{e}{k}\ \sum_{i=1}^{n} G\left(t/x_i\right).$$

Assume that t is such that $0 < a < 1$. Then

$$\mathbb{P}\left\{k\text{-}\min_{1\le i\le n} |x_i\xi_i| \le t\right\} \le \frac{1}{\sqrt{2\pi k}}\ \frac{a^k}{1-a}. \tag{17}$$

Remark 4.3 Note that if G is continuous and $G(s) = 0$ if and only if $s = 0$ then the condition on t in Lemma 4.2 above corresponds to the condition

$$0 < t < \|(1/x_i)_{i=1}^n\|_H^{-1},$$

where $H = \frac{e}{k}G$.

Proposition 4.4 *Let $p > 0$, $1 \le k \le n$, and $0 < x_1 \le x_2 \le ... \le x_n$. Let $\xi_1, \dots, \xi_n$ be i.i.d. random variables, $F(t) = \mathbb{P}\left(\{|\xi_1| > t\}\right)$, $N(t) = \ln\frac{1}{F(t)}$, and $N_j = \frac{2e}{k-j+1}N$, $j = 1, \dots, k$. Then*

$$\left(1 - \tfrac{1}{\sqrt{2\pi}}\right)\ \max_{1\le j\le k}\ \|\,(1/x_i)_{i=j}^n\,\|_{N_j}^{-p} \le \mathbb{E}\,k\text{-}\min_{1\le i\le n} |x_i\xi_i|^p.$$

Proof. Let $c = \left(1 - \frac{1}{\sqrt{2\pi}}\right)^{1/p}$. It is enough to show that for every $k \le n$

$$c \,\|\, (1/x_i)_{i=1}^n \,\|_{N_1}^{-1} \le \left(\mathbb{E}\, k\text{-} \min_{1 \le i \le n} |x_i \xi_i|^p \right)^{1/p}. \tag{18}$$

Indeed, assume that (18) is true. Fix $j \le k$. Since

$$\mathbb{E}\ k\text{-} \min_{1 \le i \le n} |x_i \xi_i|^p \ge \mathbb{E}\ (k - j + 1)\text{-} \min_{j \le i \le n} |x_i \xi_i|^p,$$

(18) implies

$$\left(\mathbb{E}\ k\text{-} \min_{1 \le i \le n} |x_i \xi_i|^p \right)^{1/p} \ge c \,\|\, (1/x_i)_{i=j}^n \,\|_{N_j}^{-1},$$

for all $1 \le j \le k$.

Now we show estimate (18). Fix $\varepsilon > 0$ small enough and put

$$A = \|\, (1/x_i)_{i=1}^n \,\|_{N_1}^{-1} - \varepsilon.$$

We use that $1 - t \le -\ln t$ for $t > 0$ and that $N_1 = \frac{2e}{k} N = \frac{2e}{k} \ln \frac{1}{F}$ and we obtain

$$\begin{aligned} a \ &:= \ \tfrac{e}{k} \sum_{i=1}^n G(A/x_i) = \tfrac{e}{k} \sum_{i=1}^n (1 - F(A/x_i)) \\ &\le \ \tfrac{e}{k} \sum_{i=1}^n \ln \frac{1}{F(A/x_i)} = \frac{1}{2} \sum_{i=1}^n N_1(A/x_i) \le 1/2. \end{aligned}$$

Applying Lemma 4.2, we get

$$\mathbb{P}\left\{ k\text{-} \min_{1 \le i \le n} |x_i \xi_i|^p \ge A^p \right\} \ge 1 - \frac{1}{\sqrt{2\pi k}} \ \frac{a^k}{1 - a} \ge 1 - \frac{1}{\sqrt{2\pi}},$$

as $a \le 1/2$. This implies

$$\mathbb{E}\ k\text{-} \min_{1 \le i \le n} |x_i \xi_i|^p \ge A^p \ \mathbb{P}\left\{ k\text{-} \min_{1 \le i \le n} |x_i \xi_i|^p \ge A^p \right\} \ge \left(1 - \tfrac{1}{\sqrt{2\pi}}\right) \ A^p.$$

Sending ε to 0 we obtain the desired result. □

4.2 The upper bound in Theorem 3.1

Here we prove the upper bound in Theorem 3.1, which will follow from Proposition 4.9 below. We start with a combinatorial lemma, partial case of which was proved in [16] (Lemma 4 there). Since Lemma 4.5 is of independent interest and is much more involved technically than Lemma 4 in [16], we provide the complete proof.

Lemma 4.5 *Let $H : \mathbb{R} \to \mathbb{R}$ be an Orlicz function. For every k with $1 \leq k \leq n$ and every $z_1 \geq z_2 \geq \ldots \geq z_n > 0$ there is a partition of nonempty sets $A_1, \ldots, A_k$ of the set $\{1, \ldots, n\}$ such that*

$$\min_{1\leq j\leq k} \left\| (z_i)_{i=j}^n \right\|_{\frac{H}{k-j+1}} \leq 4\max\{H(1), 1/H(1)\} \min_{1\leq j\leq k} \left\| (z_i)_{i\in A_j} \right\|_H . \tag{19}$$

We want to emphasize that it is important that the partition consists of exactly k sets. Our proof shows that the partition can be taken as intervals, that is $A_j = \{n_j + 1, \ldots, n_{j+1}\}$ for an increasing sequence $0 = n_0 < n_1 < \ldots < n_k = n$.

Proof. We may assume that $H(1) = 1$. Indeed, as H is convex and as $H(0) = 0$, $H(s) \leq \frac{s}{t}H(t)$ for all $0 < s < t$. Thus if $H(1) \leq 1$, then

$$H(1)\|y\|_{\frac{H}{H(1)}} \leq \|y\|_H \leq \|y\|_{\frac{H}{H(1)}}$$

for every $y \in \mathbb{R}^n$. Similarly, if $H(1) > 1$

$$\|y\|_{\frac{H}{H(1)}} \leq \|y\|_H \leq H(1)\|y\|_{\frac{H}{H(1)}}.$$

We consider three cases.

Case 1:

$$z_1 \leq \frac{1}{4} \left\|(z_i)_{i=1}^n\right\|_{\frac{H}{k}} . \tag{20}$$

Note that $H(1) = 1$ implies $t = \|(t, 0, \ldots, 0)\|_H$ for every $t > 0$, in particular $z_1 = \|(z_1, 0, \ldots, 0)\|_H$. We put $n_0 = 0$ and after having chosen $n_0, \ldots, n_\ell < n$ we define $n_{\ell+1} \leq n$ to be the largest integer such that

$$\left\|(z_i)_{i=n_\ell+1}^{n_{\ell+1}}\right\|_H \leq \frac{1}{2} \left\|(z_i)_{i=1}^n\right\|_{\frac{H}{k}} . \tag{21}$$

We define

$$B_\ell = \{n_{\ell-1} + 1, \ldots, n_\ell\}, \qquad \ell = 1, \ldots, L.$$

These sets are basically the partition we are looking for, except for a slight change that is necessary in order to get exactly k sets.

We verify first that such a partition exists. For this we have to show that each B_ℓ contains at least one element, i.e. $B_\ell \neq \emptyset$. In other words, we show that $0 = n_0 < n_1 < \ldots < n_L = n$. Indeed, if $n_{l-1} < n$, then $n_{\ell-1} + 1 \in B_\ell$ because

$$\frac{1}{4} \left\|(z_i)_{i=1}^n\right\|_{\frac{H}{k}} \geq z_1 \geq z_{n_{\ell-1}+1} = \left\|\left(0, \ldots, 0, z_{n_{\ell-1}+1}, 0, \ldots, 0\right)\right\|_H .$$

In the last equality we used again that $H(1) = 1$. Thus $B_\ell \neq \emptyset$ and $n_L = n$ which means that the partition is well defined.

We show now that $L > k$. By (21) for every $\varepsilon \in (0,1)$ and for $\ell = 0, \dots, L-1$ we have

$$\sum_{i=n_\ell+1}^{n_{\ell+1}} H\left((2-\varepsilon)\,\|(z_i)_{i=1}^n\|_{\frac{H}{k}}^{-1} z_i\right) \le 1,$$

which implies

$$\sum_{i=1}^{n} H\left((2-\varepsilon)\,\|(z_i)_{i=1}^n\|_{\frac{H}{k}}^{-1} z_i\right) \le L.$$

Therefore

$$\|(z_i)_{i=1}^n\|_{\frac{H}{L}} \le \frac{1}{2}\,\|(z_i)_{i=1}^n\|_{\frac{H}{k}}\,.$$

This implies $L > k$ and below we use that the inequality is strict.

We claim that for all $\ell = 1, \dots, k$ one has

$$\left\|(z_i)_{i\in B_\ell}\right\|_H \ge \frac{1}{4}\,\|(z_i)_{i=1}^n\|_{\frac{H}{k}}\,. \tag{22}$$

Suppose that there is ℓ with $1 \le \ell \le k$ such that

$$\left\|(z_i)_{i\in B_\ell}\right\|_H < \frac{1}{4}\,\|(z_i)_{i=1}^n\|_{\frac{H}{k}}\,. \tag{23}$$

Since $L > k \ge \ell$ we have $n_\ell + 1 \le n$. As $\|\cdot\|_H$ is a norm and since $H(1) = 1$,

$$\left\|(z_i)_{i=n_{\ell-1}+1}^{n_\ell+1}\right\|_H \le \left\|(z_i)_{i\in B_\ell}\right\|_H + \|(0,\dots,0,z_{n_\ell+1})\|_H = \left\|(z_i)_{i\in B_\ell}\right\|_H + z_{n_\ell+1}.$$

By (20) and (23)

$$\left\|(z_i)_{i=n_{\ell-1}+1}^{n_\ell+1}\right\|_H < \frac{1}{2}\,\|(z_i)_{i=1}^n\|_{\frac{H}{k}}\,.$$

This contradicts the definition of n_ℓ.

Now we define the partition $A_1, \dots, A_k$. We put $A_\ell = B_\ell$ for $1 \le \ell \le k-1$ and

$$A_k = \bigcup_{\ell=k}^{L} B_\ell.$$

Then, by (22),

$$\begin{aligned} \min_{1\le j\le k} \left\|(z_i)_{i=j}^n\right\|_{\frac{H}{k-j+1}} &\le \|(z_i)_{i=1}^n\|_{\frac{H}{k}} \\ &\le 4 \min_{1\le \ell\le k} \left\|(z_i)_{i\in B_\ell}\right\|_H \le 4 \min_{1\le \ell\le k} \left\|(z_i)_{i\in A_\ell}\right\|_H\,, \end{aligned} \tag{24}$$

which proves (19).

Case 2:

$$z_1 > \frac{1}{4}\,\|(z_i)_{i=1}^n\|_{\frac{H}{k}} \qquad \text{and for all } \ j \le k \ \text{ one has } \ z_j > \frac{1}{4}\left\|(z_i)_{i=j}^n\right\|_{\frac{H}{k+1-j}}\,.$$

We choose $A_j = \{j\}$ for $j = 1, \dots, k-1$ and $A_k = \{k, \dots, n\}$. Then for every $j \le k$

$$\left\| (z_i)_{i \in A_j} \right\|_H \ge z_j > \frac{1}{4} \left\| (z_i)_{i=j}^n \right\|_{\frac{H}{k+1-j}} ,$$

which proves (19).

Case 3:

$$z_1 > \frac{1}{4} \|(z_i)_{i=1}^n\|_{\frac{H}{k}} \qquad \text{and there exists } \ j \le k \ \text{ such that } \ z_j \le \frac{1}{4} \left\| (z_i)_{i=j}^n \right\|_{\frac{H}{k+1-j}} .$$

Let m be the smallest integer such that $m > 1$ and

$$z_m \le \frac{1}{4} \|(z_i)_{i=m}^n\|_{\frac{H}{k+1-m}} . \tag{25}$$

For $1 \le \ell < m$ we choose $A_\ell = \{\ell\}$. Then

$$\left\| (z_i)_{i \in A_\ell} \right\|_H = z_\ell > \frac{1}{4} \|(z_i)_{i=\ell}^n\|_{\frac{H}{k+1-\ell}}$$

and therefore

$$\min_{1 \le j < m} \left\| (z_i)_{i=j}^n \right\|_{\frac{H}{k-j+1}} \le 4 \min_{1 \le j < m} \left\| (z_i)_{i \in A_j} \right\|_H .$$

Now we consider the sequence $z_m \ge z_{m+1} \ge \dots \ge z_n > 0$ and proceed as in *Case 1.* The assumption of *Case 1* is fulfilled by (25). The procedure of *Case 1* gives a partition $A_m, \dots, A_k$ of $\{m, \dots, n\}$ satisfying (24)

$$4 \min_{m \le \ell \le k} \left\| (z_i)_{i \in A_\ell} \right\|_H \ge \|(z_i)_{i=m}^n\|_{\frac{H}{k+1-m}} .$$

This completes the proof. $\Box$

We will also use the following lemma.

Lemma 4.6 *Let $p > 0$ and $0 < x_1 \le x_2 \le ... \le x_n$ be real numbers. Let $\xi_1, \dots, \xi_n$ be i.i.d. random variables. Let $F(t) = \mathbb{P}(|\xi_1| > t)$ be strictly decreasing and $N = -\ln F$ be an Orlicz function. Then*

$$\left(1 - \tfrac{1}{\sqrt{2\pi}}\right) \left\| \left(\frac{1}{x_i}\right)_{i=1}^n \right\|_{2eN}^{-p} \le \mathbb{E} \min_{1 \le i \le n} |x_i \xi_i|^p \le (1 + \Gamma(1+p)) \left\| \left(\frac{1}{x_i}\right)_{i=1}^n \right\|_N^{-p} .$$

Remark 4.7 If N is an Orlicz function then by (7)

$$(2e)^{-p} \| \cdot \|_N^{-p} \le \| \cdot \|_{2eN}^{-p} .$$

Remark 4.8 The left hand side inequality does not require the condition "N is an Orlicz function."

Proof of Lemma 4.6. The left hand inequality follows from Proposition 4.4.

To prove the right hand side inequality we choose

$$t_0 = \left\| \left(\frac{1}{x_i} \right)_{i=1}^{n} \right\|_N^{-p}.$$

Then for all $t \geq t_0$

$$\sum_{i=1}^{n} \ln \left(1/F\left(t^{1/p}/x_i \right) \right) \geq 1.$$

By (14) for all $t \geq t_0$ and all x_i

$$(t/t_0)^{\frac{1}{p}} \ln \frac{1}{F(t_0^{\frac{1}{p}}/x_i)} \leq \ln \frac{1}{F(t^{\frac{1}{p}}/x_i)}.$$

By Lemma 4.1,

$$\mathbb{P}\left\{ \min_{1\leq i\leq n} |x_i\xi_i|^p > t \right\} = \prod_{i=1}^{n} F\left(t^{\frac{1}{p}}/x_i \right) = \exp\left(-\sum_{i=1}^{n} \ln \frac{1}{F\left(t^{\frac{1}{p}}/x_i \right)} \right),$$

and thus for all $t \geq t_0$

$$\mathbb{P}\left\{ \min_{1\leq i\leq n} |x_i\xi_i|^p > t \right\} \leq \exp\left(-(t/t_0)^{\frac{1}{p}} \sum_{i=1}^{n} \ln \frac{1}{F\left(t_0^{\frac{1}{p}}/x_i \right)} \right) \leq \exp\left(-\left(\frac{t}{t_0} \right)^{\frac{1}{p}} \right).$$

Therefore

$$\begin{aligned}
\mathbb{E} \min_{1\leq i\leq n} |x_i\xi_i|^p &= \int_0^\infty \mathbb{P}\left\{ \min_{1\leq i\leq n} |x_i\xi_i| > t^{\frac{1}{p}} \right\} dt \\
&= \int_0^{t_0} \mathbb{P}\left\{ \min_{1\leq i\leq n} |x_i\xi_i| > t^{\frac{1}{p}} \right\} dt + \int_{t_0}^\infty \mathbb{P}\left\{ \min_{1\leq i\leq n} |x_i\xi_i| > t^{\frac{1}{p}} \right\} dt \\
&\leq t_0 + \int_{t_0}^\infty \exp\left(-\left(\frac{t}{t_0} \right)^{\frac{1}{p}} \right) dt.
\end{aligned}$$

We substitute $t = t_0 s^p$, then

$$\mathbb{E} \min_{1\leq i\leq n} |x_i\xi_i|^p \leq t_0 + t_0 p \int_1^\infty s^{p-1} e^{-s} ds \leq t_0 \left(1 + p\Gamma(p) \right),$$

which completes the proof. □

Since by (7)

$$\| \cdot \|_{N_j} \leq 2e \| \cdot \|_{\frac{N}{k-j+1}},$$

the following Proposition implies the upper estimate in Theorem 3.1.

Proposition 4.9 *Let $1 \le k \le n$ and $0 < x_1 \le x_2 \le ... \le x_n$. Let $\xi_1, \dots, \xi_n$ be i.i.d. random variables. Let $F(t) = \mathbb{P}(|\xi_1| > t)$ be strictly decreasing, and let $N = -\ln F$ be an Orlicz function. Then*

$$\mathbb{E}\ k\text{-}\min_{1\le i\le n} |x_i\xi_i| \le 8e\ln(k+1)\ C_N \quad \max_{1\le j\le k} \left\| \left(\frac{1}{x_i}\right)_{i=j}^{n} \right\|_{\frac{N}{k-j+1}}^{-1}$$

where $C_N = \max\{N(1), 1/N(1)\}$.

Proof. The case $k = 1$ follows form Lemma 4.6. We assume $k \ge 2$.

Let $A_1, \dots A_k$ be the partition of $\{1 \dots n\}$ given by Lemma 4.5. Then for all $q \ge 1$

$$\begin{aligned}
\mathbb{E}\ k\text{-}\min_{1\le i\le n} |x_i\xi_i| &\le \mathbb{E}\ \max_{1\le j\le k} \min_{i\in A_j} |x_i\xi_i| \le \mathbb{E}\left(\sum_{j=1}^{k} \left|\min_{i\in A_j} |x_i\xi_i|\right|^q\right)^{\frac{1}{q}} \\
&\le \left(\mathbb{E}\sum_{j=1}^{k} \left|\min_{i\in A_j} |x_i\xi_i|\right|^q\right)^{\frac{1}{q}} = \left(\sum_{j=1}^{k} \mathbb{E}\left|\min_{i\in A_j} |x_i\xi_i|\right|^q\right)^{\frac{1}{q}}.
\end{aligned}$$

By Lemma 4.6 the latter expression is less than

$$(1+\Gamma(1+q))^{\frac{1}{q}} \left(\sum_{j=1}^{k} \left\| \left(\frac{1}{x_i}\right)_{i\in A_j} \right\|_N^{-q}\right)^{\frac{1}{q}} \le 2qk^{1/q} \max_{1\le j\le k} \left\| \left(\frac{1}{x_i}\right)_{i\in A_j} \right\|_N^{-1}.$$

The choice $q = \ln(k+1)$ gives

$$\mathbb{E}\ k\text{-}\min_{1\le i\le n} |x_i\xi_i| \le 2e\ln(k+1) \max_{1\le j\le k} \left\| \left(\frac{1}{x_i}\right)_{i\in A_j} \right\|_N^{-1}.$$

By Lemma 4.5 applied with $z_i = x_i$, $i \le n$, this expression is smaller than

$$8e\ C_N \ \max_{1\le j\le k} \left\| \left(\frac{1}{x_i}\right)_{i=j}^{n} \right\|_{N_j}^{-1}.$$

□

5 k-max (Proof of Theorem 3.5)

In this section we prove Theorem 3.5. As in the previous section we separate proofs of upper and lower bounds. We start with the upper bound.

5.1 The upper bound in Theorem 3.5

Here we prove the upper bound in Theorem 3.1. It will follow from Lemma 5.4 (see Remark 5.5 following the lemma).

We require a result from [13]. Let f be a random variable with continuous distribution and such that $\mathbf{E}|f| < \infty$. Let $t_n = t_n(f) = 0$, $t_0 = t_0(f) = \infty$, and for $j = 1, \ldots, n-1$

$$t_j = t_j(f) = \sup\left\{t \mid \mathbb{P}\{\omega \mid |f(\omega)| > t\} \geq \tfrac{j}{n}\right\}. \tag{26}$$

Since f has the continuous distribution, we have for every $j \geq 1$

$$\mathbb{P}\{\omega \mid |f(\omega)| \geq t_j\} = \tfrac{j}{n}.$$

For $j = 1, \ldots, n$ define the sets

$$\Omega_j = \Omega_j(f) = \{\omega \mid t_j \leq |f(\omega)| < t_{j-1}\}. \tag{27}$$

For all $j = 1, \ldots, n$ we have

$$\Omega_j = \{\omega \mid t_j \leq |f(\omega)| < t_{j-1}\} = \{\omega \mid t_j \leq |f(\omega)|\} \setminus \{\omega \mid t_{j-1} \leq |f(\omega)|\}.$$

Therefore

$$\mathbb{P}(\Omega_j) = \tfrac{j}{n} - \tfrac{j-1}{n} = \tfrac{1}{n}.$$

For $j = 1, \ldots, n$ let

$$y_j = y_j(f) := \int_{\Omega_j} |f(\omega)| d\mathbb{P}(\omega). \tag{28}$$

Then

$$\sum_{j=1}^{n} y_j = \mathbf{E}|f| \qquad \text{and} \qquad t_j \leq n y_j < t_{j-1} \qquad \text{for all } \ j = 1, \ldots, n.$$

In [13], Corollary 2 we proved the following statement.

Lemma 5.1 *Let $f_1, \ldots, f_n$ be i.i.d. random variables such that $\int |f_i(\omega)| d\mathbb{P}(\omega) = 1$. Let M be an Orlicz function such that for all $k = 1, \ldots, n$*

$$M^*\left(\sum_{j=1}^{k} y_j\right) = \tfrac{k}{n}.$$

Then, for all $x \in \mathbb{R}^n$

$$c_1 \|x\|_M \leq \int_{\Omega} \max_{1 \leq i \leq n} |x_i f_i(\omega)| d\mathbb{P}(\omega) \leq c_2 \|x\|_M,$$

where c_1 and c_2 are absolute positive constants.

This can be reformulated in the following way.

Lemma 5.2 *Let $\xi_1, \ldots, \xi_n$ be i.i.d. random variables such that $\int |\xi_i(\omega)| d\mathbb{P}(\omega) = 1$. Let M be the Orlicz function such that for all $s \geq 0$*

$$M(s) = \int_0^s \int_{\frac{1}{t} \leq |\xi_1|} |\xi_1| d\mathbb{P} dt = \int_{\frac{1}{s} \leq |\xi|} (s|\xi| - 1)\, d\mathbb{P}.$$

Then, for all $x \in \mathbb{R}^n$

$$c_1 \|x\|_M \leq \int_\Omega \max_{1 \leq i \leq n} |x_i \xi_i(\omega)| d\mathbb{P}(\omega) \leq c_2 \|x\|_M,$$

where c_1 and c_2 are absolute positive constants.

Proof. By definition

$$\sum_{i=1}^k y_i = \int_{t_k \leq \xi_1} |\xi_1(\omega)|\ d\mathbb{P}(\omega)$$

and

$$\mathbb{P}\left(\{\omega \ \mid \ t_k \leq \xi_1(\omega)\}\right) = \frac{k}{n}.$$

Therefore the Orlicz function M^* defined by

$$M^* \left(\int_{t \leq |\xi_1|} |\xi_1(\omega)| d\mathbb{P}(\omega) \right) = \mathbb{P}\{\omega | t \leq |\xi_1(\omega)|\}$$

satisfies the condition of Lemma 5.1. It is left to observe that the dual function to M^* is

$$M(s) = \int_0^s \int_{\frac{1}{t} \leq |\xi_1|} |\xi_1| d\mathbb{P} dt.$$

This has been verified in Section 2 (see formulae (8) and (9)). □

For the next lemma we need the following simple claim.

Claim 5.3 *Let $(x_i)_{i=1}^n$ be a sequence. Then for every $j \leq n - k$ one has*

$$k\text{-}\max_{1 \leq i \leq n} |x_i| \leq j\text{-}\min_{1 \leq i \leq k+j-1} |x_i| + \max_{k+j \leq i \leq n} |x_i|.$$

Proof. If the numbers $|x_1|, \ldots, |x_{k+j-1}|$ contain the k biggest of the numbers $|x_1|, \ldots, |x_n|$, then

$$j\text{-}\min_{1 \leq i \leq k+j-1} |x_i| = k\text{-}\max_{1 \leq i \leq k+j-1} |x_i| = k\text{-}\max_{1 \leq i \leq n} |x_i|.$$

On the other hand, if the numbers $|x_1|, \ldots, |x_{k+j-1}|$ do not contain the k biggest of the numbers $|x_1|, \ldots, |x_n|$, then at least one of those is contained in the numbers $|x_{k+j}|, \ldots, |x_n|$ and therefore

$$\max_{k+j \leq i \leq n} |x_i| \geq k\text{-}\max_{1 \leq i \leq n} |x_i|.$$

□

Lemma 5.4 *Let $x_1 \geq x_2 \geq \ldots \geq x_n > 0$. Let $\xi_1, \ldots, \xi_n$ be i.i.d. random variables and $F(t) = \mathbb{P}(\{|\xi_1| > t\})$. Suppose that F is strictly decreasing and $N = -\ln F$ is an Orlicz function. Assume that M is the Orlicz function such that for all $s \geq 0$*

$$M(s) = \int_0^s \int_{\frac{1}{t} \leq |\xi|} |\xi| d\mathbb{P} dt.$$

Then we have

$$\mathbb{E}\, k\text{-}\max_{1 \leq i \leq n} |x_i \xi_i| \leq c \min_{1 \leq j \leq n-k} \left\{ C_N \ln(k+1) \max_{0 \leq \ell \leq j-1} \| (1/x_i)_{i=1}^{k+\ell} \|_{\frac{2e}{\ell+1} N}^{-1} + \|(x_{k+j}, \ldots, x_n)\|_M \right\},$$

where c is an absolute constant and $C_N = \max\{N(1), 1/N(1)\}$.

Remark 5.5 Note that Lemma 5.4 applied with $j = k_0 = \left[\frac{4(k-1)}{F(1)}\right]$ implies the upper bound in Theorem 3.5.

Proof of Lemma 5.4. Claim 5.3 implies

$$\mathbb{E}\, k\text{-}\max_{1 \leq i \leq n} |x_i \xi_i| \leq \min_{1 \leq j \leq n-k} \left(\mathbb{E}\, j\text{-}\min_{1 \leq i \leq k+j-1} |x_i \xi_i| + \mathbb{E} \max_{k+j \leq i \leq n} |x_i \xi_i| \right).$$

Applying Theorem 3.1 to the sequence $x_{k+j-1} \leq x_{k+j-2} \leq \ldots \leq x_1$, we observe that

$$\mathbb{E}\, j\text{-}\min_{1 \leq i \leq k+j-1} |x_i \xi_i| \leq 16 e^2 C_N \ln(k+1) \max_{1 \leq \ell \leq j} \| (1/x_i)_{i=1}^{k+j-\ell} \|_{\frac{2e}{j-\ell+1} N}^{-1}.$$

This is the same as

$$\mathbb{E}\, j\text{-}\min_{1 \leq i \leq k+j-1} |x_i \xi_i| \leq 16 e^2 C_N \ln(k+1) \max_{0 \leq \ell \leq j-1} \| (1/x_i)_{i=1}^{k+\ell} \|_{\frac{2e}{\ell+1} N}^{-1}.$$

On the other hand, Lemma 5.2 implies

$$\mathbb{E} \max_{k+j \leq i \leq n} |x_i \xi_i| \leq c\, \|(x_{k+j}, \ldots, x_n)\|_M,$$

where c is an absolute constant. This completes the proof. □

5.2 The lower bound in Theorem 3.5

The proof of the lower bound in Theorem 3.5 consists of two lemmas.

Lemma 5.6 *Let $x_1 \geq x_2 \geq \ldots \geq x_n > 0$. Let $\xi_1, \ldots, \xi_n$ be i.i.d. random variables and $F(t) = \mathbb{P}(|\xi_1| > t)$. For $k > 1$ let*

$$N_{F,k}(t) = \frac{F(1/t)}{4(k-1)}.$$

Then

$$\mathbb{E}\, k\text{-}\max_{1\le i\le n} |x_i\xi_i| \ge \max\left\{\frac{1}{2}\ \|(x_k,\dots,x_n)\|_{N_{F,k}},\ \max_{1\le \ell\le n-k} \mathbb{E}\,\ell\text{-}\min_{1\le i\le k+\ell-1} |x_i\xi_i|\right\}.$$

In particular, if $N(t)=\ln\frac{1}{F(t)}$*, then*

$$\mathbb{E}\, k\text{-}\max_{1\le i\le n} |x_i\xi_i| \ge \max\left\{\frac{1}{2}\ \|(x_k,\dots,x_n)\|_{N_{F,k}},\ (1-\tfrac{1}{\sqrt{2\pi}})\max_{1\le \ell\le n-k}\ \max_{1\le j\le \ell}\ \|\,(1/x_i)_{i=1}^{k+\ell-j}\,\|^{-1}_{\frac{2e}{\ell-j+1}N}\right\}.$$

Proof. First we show

$$\mathbb{E}\, k\text{-}\max_{1\le i\le n} |x_i\xi_i| \ge \frac{1}{2}\ \|(x_k,\dots,x_n)\|_{N_{F,k}}.$$

We have

$$\mathbb{P}\left\{\omega\ \middle|\ k\text{-}\max_{1\le i\le n} |x_i\xi_i(\omega)|\le t\right\}\le \sum_{j=0}^{k-1}\ \sum_{\substack{A\subset\{1,\dots n\}\\ |A|=j}}\ \prod_{i\in A} F\left(\frac{t}{x_i}\right)\prod_{i\notin A}\left(1-F\left(\frac{t}{x_i}\right)\right).$$

Since $x_1\ge x_2\ge\ldots\ge x_n>0$ and $|A^c|\ge n-k+1$, we observe

$$\mathbb{P}\left\{\omega\ \middle|\ k\text{-}\max_{1\le i\le n} |x_i\xi_i(\omega)|\le t\right\}\le \sum_{j=0}^{k-1}\ \sum_{|A|=j}\ \prod_{i\in A} F\left(\frac{t}{x_i}\right)\prod_{i=k}^{n}\left(1-F\left(\frac{t}{x_i}\right)\right).$$

Now we apply the Hardy-Littlewood-Polya inequality ([18]), which states that for non-negative numbers $a_1,\dots,a_m$ one has

$$\sum_{\substack{A\subset\{1,\dots m\}\\ |A|=j}}\ \prod_{i\in A} a_i\le \binom{m}{j}\left(\frac{1}{m}\ \sum_{i=1}^{m} a_i\right)^j\le \frac{1}{j!}\left(\sum_{i=1}^{m} a_i\right)^j.$$

This implies

$$\mathbb{P}\left\{\omega\ \middle|\ k\text{-}\max_{1\le i\le n} |x_i\xi_i(\omega)|\le t\right\}\le \sum_{j=0}^{k-1}\frac{1}{j!}\left(\sum_{i=1}^{n} F\left(\frac{t}{x_i}\right)\right)^j\prod_{i=k}^{n}\left(1-F\left(\frac{t}{x_i}\right)\right).$$

Since $F\left(\frac{t}{x_i}\right)\le 1$ and $1-x\le e^{-x}$ for $x\ge 0$, one has

$$\mathbb{P}\left\{\omega\ \middle|\ k\text{-}\max_{1\le i\le n} |x_i\xi_i(\omega)|\le t\right\}\le \sum_{j=0}^{k-1}\frac{1}{j!}\left(k-1+\sum_{i=k}^{n} F\left(\frac{t}{x_i}\right)\right)^j\exp\left(-\sum_{i=k}^{n} F\left(\frac{t}{x_i}\right)\right).$$

Let

$$\alpha=\alpha(t)=\frac{1}{k-1}\ \sum_{i=k}^{n} F\left(\frac{t}{x_i}\right).$$

Then

$$\begin{aligned}
\mathbb{P}\left\{\omega \,\middle|\, k\text{-}\max_{1\le i\le n}|x_i\xi_i(\omega)|\le t\right\} &\le e^{-\alpha(k-1)}\sum_{j=0}^{k-1}\frac{1}{j!}\left((\alpha+1)(k-1)\right)^j \\
&\le e^{-\alpha(k-1)}(1+\alpha)^{k-1}\sum_{j=0}^{k-1}\frac{1}{j!}(k-1)^j \\
&\le e^{-\alpha(k-1)}(1+\alpha)^{k-1}e^{k-1} \\
&= \exp\left((k-1)\left(-\alpha+1+\ln(1+\alpha)\right)\right).
\end{aligned}$$

Now put

$$t_0 := \|(x_k,\ldots,x_n)\|_{N_{F,k}} \ge 0.$$

If $t_0=0$ we are done. If $t_0>0$ then for every $0<\varepsilon<t_0$

$$\alpha(t_0-\varepsilon)=\frac{1}{k-1}\sum_{i=k}^{n}F\left(\frac{t_0-\varepsilon}{x_i}\right)>4.$$

Since $k>1$, this implies

$$\mathbb{P}\left\{\omega \,\middle|\, k\text{-}\max_{1\le i\le n}|x_i\xi_i(\omega)|\le t_0-\varepsilon\right\}\le \exp((k-1)(-3+\ln 5))\le 1/2.$$

Thus

$$\mathbb{E}\,k\text{-}\max_{1\le i\le n}|x_i\xi_i|\ge (t_0-\varepsilon)\ \mathbb{P}\left\{\omega \,\middle|\, k\text{-}\max_{1\le i\le n}|x_i\xi_i(\omega)|\ge t_0-\varepsilon\right\}\ge\frac{t_0-\varepsilon}{2}.$$

Letting ε tend to 0 we obtain the first part of the desired estimate.

Now we show the second part of the estimate. We observe that for all $l\le n-k+1$

$$\ell\text{-}\min_{1\le i\le k+\ell-1}|x_i\xi_i(\omega)| = k\text{-}\max_{1\le i\le k+\ell-1}|x_i\xi_i(\omega)|.$$

This implies

$$\mathbb{E}\,k\text{-}\max_{1\le i\le n}|x_i\xi_i|\ge\max_{1\le\ell\le n-k+1}\mathbb{E}\,\ell\text{-}\min_{1\le i\le k+\ell-1}|x_i\xi_i|.$$

Finally, the "In particular" part of the theorem follows from Proposition 4.4. Note that in Proposition 4.4 the sequence (x_i) is in increasing order while in the lemma we are proving now it is in decreasing order. □

In the next lemma we provide a lower estimate on $\|\cdot\|_{N_{F,k}}$, appearing in Lemma 5.6.

Lemma 5.7 *Let $1<k\le n$. Let $\xi_1,\ldots,\xi_n$ be i.i.d. random variables with $\mathbb{E}|\xi_i|=1$. Let $F(t)=\mathbb{P}(|\xi_1|>t)$ be a strictly decreasing function such that $N=-\ln F$ is an Orlicz function. Let M be the Orlicz function defined by*

$$M(s)=\int_0^s\int_{\frac{1}{t}\le|\xi_1|}|\xi_1|d\mathbb{P}dt.$$

Let

$$N_{F,k}(t) = \frac{F\left(1/t\right)}{4(k-1)}.$$

Let $k_0 = \left[\frac{4(k-1)}{F(1)}\right]$ *and assume that* $k + k_0 \le n$. *Then for all* $x_1 \ge \ldots \ge x_n > 0$

$$\|(x_{k+k_0}, \ldots, x_n)\|_M \le \left(1 + \frac{\ln(8(k-1))}{N(1)}\right) \|(x_k, \ldots, x_n)\|_{N_{F,k}}.$$

Proof. Since both functions $\|\cdot\|_M$ and $\|\cdot\|_{N_{F,k}}$ are homogeneous, we may assume that $\|(x_{k+k_0}, \ldots, x_n)\|_M = 1$. Thus, without loss of generality, we can assume that $\sum_{i=k+k_0}^n M(x_i) = 1$ (otherwise we pass from the sequence $\{x_i\}_i$ to $\{x_i/(1+\varepsilon)\}_i$ for an suitably small $\varepsilon > 0$).

We put

$$A := F^{-1}(\alpha) = 1 + \frac{\ln(8(k-1))}{N(1)}.$$

Note that by (6), $N(A) \ge AN(1) \ge \ln 8 > 2$.

Case 1: $x_{k+k_0} \ge 1/A$. Then $x_k \ge x_{k+1} \ge \ldots \ge x_{k+k_0} \ge 1/A$.

Since F is a decreasing function, $1/F^{-1}$ is increasing and

$$\sum_{i=k}^{n} F\left((x_i A)^{-1}\right) \ge \sum_{i=k}^{k+k_0} F\left((x_i A)^{-1}\right) \ge (k_0+1)F(1) > 4(k-1).$$

This means that

$$\|(x_k, \ldots, x_n)\|_{N_{F,k}} \ge 1/A.$$

Case 2: $x_{k+k_0} < 1/A$. Then $1/A > x_{k+k_0} \ge \ldots \ge x_n$.

Since $\int_{\frac{1}{t} \le |\xi_1|} |\xi_1| d\mathbb{P}$ is an increasing function of t, we observe

$$M(s) = \int_0^s \int_{\frac{1}{t} \le |\xi_1|} |\xi_1| d\mathbb{P} dt \le s \int_{\frac{1}{s} \le |\xi_1|} |\xi_1| d\mathbb{P}.$$

By (16), applied with $t = 1/s$, we obtain that for all positive s

$$\int_{1/s \le |\xi_1|} |\xi_1| d\mathbb{P} \le \left(1 + \frac{1}{N(1/s)}\right) \frac{1}{s} F(1/s).$$

Recall that N is increasing and $N(A) > 2$. Thus for all $s \le 1/A$ we have

$$M(s) \le s \int_{\frac{1}{s} \le |\xi_1|} |\xi_1| d\mathbb{P} \le 2 F(1/s).$$

By the condition of Case 2, $x_i \le 1/A$ for $i \ge k + k_0$. This implies

$$1 = \sum_{i=k+k_0}^{n} M(x_i) \le 2 \sum_{i=k+k_0}^{n} F\left(\frac{1}{x_i}\right). \tag{29}$$

Now, by (6), we have $N(y) \ge \beta N(y/\beta)$ for every $y \ge 0$ and $\beta \ge 1$. Since $N = -\ln F$, we observe

$$F(y) \le F(y/\beta)^{\beta}$$

for every $y \ge 0$ and $\beta \ge 1$. Since F is decreasing, it implies

$$F(y) \le F(y/\beta)\ F(1)^{\beta-1}$$

for every $y \ge \beta \ge 1$. Applying the last inequality with $y = 1/x_i$ and $\beta = A$, we obtain for every $i \ge k + k_0$

$$F(1/x_i) \le F(1/(Ax_i))\ F(1)^{A-1}.$$

By (29),

$$\sum_{i=k}^{n} F\left(1/(Ax_i)\right) \ge \frac{1}{F(1)^{A-1}} \sum_{i=k+k_0}^{n} F\left(1/x_i\right) \ge \frac{1}{2F(1)^{A-1}}.$$

Now, by the choice of A,

$$A - 1 = \frac{\ln(8(k-1))}{\ln(1/F(1))},$$

and hence

$$2F(1)^{A-1} = \frac{1}{4(k-1)}.$$

Thus,

$$\sum_{i=k}^{n} F\left(1/(Ax_i)\right) \ge \frac{1}{4(k-1)},$$

which implies

$$\|(x_k, \ldots, x_n)\|_{N_{F,k}} \ge 1/A.$$

This completes the proof. $\square$

Finally we complete the proof of Theorem 3.5.

Proof of the lower bound in Theorem 3.5. Let $k_0 = \left[\frac{4(k-1)}{F(1)}\right]$. Applying Lemma 5.6 with $l = k_0$ we obtain

$$\mathbb{E}\, k\text{-}\max_{1 \le i \le n} |x_i \xi_i| \ge \max\left\{\frac{1}{2}\ \|(x_k, \ldots, x_n)\|_{N_{F,k}}, \left(1 - \tfrac{1}{\sqrt{2\pi}}\right) \max_{1 \le j \le k_0}\ \|\ (1/x_i)_{i=1}^{k+k_0-j}\ \|^{-1}_{\frac{2e}{k_0-j+1}N}\right\}.$$

By Lemma 5.7 we have for all x with $x_1 \ge \ldots \ge x_n > 0$

$$\|(x_{k+k_0}, \ldots, x_n)\|_M \le A\ \|(x_k, \ldots, x_n)\|_{N_{F,k}},$$

where $A = 1 + \frac{\ln(8(k-1))}{N(1)}$. Thus

$$\begin{aligned}
\mathbb{E}\, k\text{-}\max_{1\le i\le n} |x_i \xi_i| &\ge \frac{1}{4A}\, \|(x_{k+k_0}, \dots, x_n)\|_M + \frac{1}{2}\left(1 - \tfrac{1}{\sqrt{2\pi}}\right) \max_{1\le j\le k_0} \|\,(1/x_i)_{i=1}^{k+k_0-j}\,\|^{-1}_{\frac{2e}{k_0-j+1}N} \\
&= \frac{1}{4A}\, \|(x_{k+k_0}, \dots, x_n)\|_M + \frac{1}{2}\left(1 - \tfrac{1}{\sqrt{2\pi}}\right) \max_{0\le \ell\le k_0-1} \|\,(1/x_i)_{i=1}^{k+\ell}\,\|^{-1}_{\frac{2e}{\ell+1}N}.
\end{aligned}$$

□

References

[1] B. C. Arnold, N. Narayanaswamy, *Relations, Bounds and Approximations for Order Statistics*, Lecture Notes in Statistics, 53, Berlin etc.: Springer-Verlag. viii (1989).

[2] N. Balakrishnan, A.C. Cohen, *Order Statistics and Inference*, New York, NY: Academic Press (1991).

[3] R. Baraniuk, M. Davenport, R. Devore, M. Wakin, *A simple proof of the restricted isometry property for random matrices*, Constr. Approx. 28 (2008), 253-263.

[4] Y. Benjamini, M. Leshno, *Statistical Methods for Data Mining*, Data Mining and Knowledge Discovery Handbook, Springer US (2005).

[5] E. J. Candès, J. Romberg, T. Tao, *Robust uncertaintity principles: exact signal reconstruction from highly incomplete frequency information*, IEEE Trans. Inform. Theory 52 (2006), 489-509

[6] A. Cohen, W. Dahmen, R. DeVore, *Compressed sensing and best k-term approximation*, J. Amer. Math. Soc. 22 (2009), 211-231.

[7] H. A. David, H. N. Nagaraja, *Order statistics*, 3rd ed., Wiley Series in Probability and Statistics. Chichester: John Wiley & Sons (2003).

[8] A. Dimitriyuk, Y. Gordon, *Generalizing the Johnson-Lindenstrauss lemma to k-dimensional affine subspaces*, Stud. Math. 195, No. 3, 227-241 (2009).

[9] D. Donoho, *Compressed Sensing*, IEEE Trans. Information Theory, 52 (2006), 1289–1306.

[10] Y. Gordon, *Majorization of Gaussian processes and Geometric Applications*, Prob. Th. Rel. Fields, 91, No. 2 (1992), 251–267.

[11] Y. Gordon, O. Guédon, M. Meyer, A. Pajor, *On the Euclidean sections of some Banach spaces and operator spaces*, Math. Scandinavica 91 (2002), 247–268.

[12] Y. Gordon, A. E. Litvak, S. Mendelson, A. Pajor, *Gaussian averages of interpolated bodies and applications to approximate reconstruction*, J. Approx. Theory, 149 (2007), 59–73.

[13] Y. Gordon, A. E. Litvak, C. Schütt, E. Werner, *Orlicz Norms of Sequences of Random Variables*, Ann. of Prob., 30 (2002), 1833–1853.

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

[15] Y. Gordon, A. E. Litvak, C. Schütt, E. Werner, *Minima of sequences of Gaussian random variables*, C.R. Acad. Sci. Paris, Ser 1, Math., 340 (2005), 445–448.

[16] Y. Gordon, A. E. Litvak, C. Schütt, E. Werner, *On the minimum of several random variables*, Proc. Am. Math. Soc. 134, No. 12, 3665–3675 (2006).

[17] O. Guèdon, *Gaussian Version of a Theorem of Milman and Schechtman*, Positivity 1, No.1 (1997), 1–5.

[18] G. H. Hardy, J. E. Littlewood, G. Polya, *Inequalities*, 2nd ed., Cambridge, The University Press. XII (1952).

[19] P. Hitczenko, S. Kwapień, W. V. Li, G. Schechtman, T. Schlumprecht, J. Zinn, *Hypercontractivity and comparison of moments of iterated maxima and minima of independent random variables*, Electron. J. Probab. 3, Paper 2, 26 p. (1998).

[20] M. Junge, *The optimal order for the p-th moment of sums of independent random variables with respect to symmetric norms and related combinatorial estimates*, Positivity 10 (2006), 201-230.

[21] M.A. Krasnosel'skii, Ya. B. Rutickii, *Convex Functions and Orlicz Functions*, P. Noordhoff, Groningen (1961).

[22] J. Lindenstrauss, L. Tzafriri, *Classical Banach Spaces I*, Springer-Verlag (1977).

[23] S. Mallat, O. Zeitouni, *Optimality of the Karhunen-Loeve basis in nonlinear reconstruction*, preprint.

[24] V. A. Marchenko, L. A. Pastur, *Distribution of eigenvalues in certain sets of random matrices*, Mat. Sb. (N.S.), 72 (1967), 407–535 (Russian).

[25] S. Mendelson, A. Pajor, N. Tomczak-Jaegermann, *Reconstruction and Subgaussian Operators in Asymptotic Geometric Analysis*, GAFA, Geom. and Funct. Anal., 17 (2007), 1248–1282.

[26] S. Montgomery-Smith, *Rearrangement invariant norms of symmetric sequence norms of independent sequences of random variables*, Isr. J. Math. 131 (2002), 51–60.

[27] M.M. Rao, Z.D. Ren, *Theory of Orlicz Spaces*, Marcel Dekker (1991).

[28] M. Rudelson *Lower estimates for the singular values of random matrices*, C. R. Math. Acad. Sci. Paris 342 (2006), no. 4, 247–252.

[29] S. J. Szarek, *Spaces with large distance to ℓ_∞^n and random matrices*, Amer. J. Math. 112 (1990), no. 6, 899–942.

[30] S. J. Szarek, *Condition numbers of random matrices*, J. Complexity 7 (1991), no. 2, 131–149.

Y. Gordon, Dept. of Math., Technion, Haifa 32000, Israel, `gordon@techunix.technion.ac.il`

A. E. Litvak, Dept. of Math. and Stat. Sciences, University of Alberta, Edmonton, AB, Canada T6G 2G1, `alexandr@math.ualberta.ca`

C. Schütt, Mathematisches Seminar, Christian Albrechts Universität, 24098 Kiel, Germany, `schuett@math.uni-kiel.de`

E. Werner, Dept. of Math., Case Western Reserve University, Cleveland, Ohio 44106, U.S.A. and Université de Lille 1, UFR de Mathématique, 59655 Villeneuve d'Ascq, France, `emw2@po.cwru.edu`