# ON THE EXPECTATION OF OPERATOR NORMS OF RANDOM MATRICES

OLIVIER GUÉDON, AICKE HINRICHS, ALEXANDER E. LITVAK,
AND JOSCHA PROCHNO

ABSTRACT. We prove estimates for the expected value of operator norms of Gaussian random matrices with independent (but not necessarily identically distributed) and centered entries, acting as operators from $\ell_{p*}^n$ to $\ell_q^m$, $1 \leq p^* \leq 2 \leq q < \infty$.

## 1. INTRODUCTION AND MAIN RESULTS

Random matrices and their spectra are under intensive study in Statistics since the work of Wishart [28] on sample covariance matrices, in Numerical Analysis since their introduction by von Neumann and Goldstine [25] in the 1940s, and in Physics as a consequence of Wigner's work [26, 27] since the 1950s. His Semicircle Law, a fundamental theorem in the spectral theory of large random matrices describing the limit of the empirical spectral measure for what is nowadays known as Wigner matrices, is among the most celebrated results of the theory.

In Banach Space Theory and Asymptotic Geometric Analysis, random matrices appeared already in the 70s (see e.g. [2, 3, 9]). In [2], the authors obtained asymptotic bounds for the expected value of the operator norm of a random matrix $B = (b_{ij})_{i,j=1}^{m,n}$ with independent mean-zero entries with $|b_{ij}| \leq 1$ from $\ell_2^n$ to $\ell_q^m$, $2 \leq q < \infty$. To be more precise, they proved that

$$\mathbb{E} \left\| B : \ell_2^n \to \ell_q^m \right\| \leq C_q \cdot \max\left( m^{1/q}, \sqrt{n} \right),$$

where $C_q$ depends only on $q$. This was then successfully used to characterize $(p, q)$-absolutely summing operators on Hilbert spaces. Ever since, random matrices are extensively studied and methods of Banach spaces have produced numerous deep and new results. In particular, in many applications the spectral properties of a Gaussian matrix, whose entries are independent identically distributed (i.i.d.) standard Gaussian random variables, were used. Seginer proved in [22] that for an $m \times n$ random matrix with i.i.d. symmetric random variables the expectation of its spectral norm (that is, the operator norm from $\ell_2^n$ to $\ell_2^m$) is of the order of the expectation of the largest Euclidean norm of its rows and columns. He also obtained an optimal result in the case of random matrices with entries $\varepsilon_{ij} a_{ij}$, where $\varepsilon_{ij}$ are independent Rademacher random variables and $a_{ij}$ are fixed numbers. We refer the interested reader to the surveys [7, 6] and references therein.

It is natural to ask similar questions about general random matrices, in particular about Gaussian matrices whose entries are still independent centered Gaussian random variables, but with different variances. In this structured case, where we

drop the assumption of identical distributions, very little is known. It is conjectured that the expected spectral norm of such a Gaussian matrix is as in Seginer's result, that is, of the order of the expectation of the largest Euclidean norm of its rows and columns. A big step toward the solution was made by Latała in [15], who proved a bound involving fourth moments, which is of the right order $\max(\sqrt{m}, \sqrt{n})$ in the i.i.d. setting, but does not capture the right behavior in the case of, for instance, diagonal matrices. On one hand, as is mentioned in [15], in view of the classical Bai-Yin theorem, the presence of fourth moments is not surprising, on the other hand they are not needed if the conjecture is true.

Later in [20], Riemer and Schütt proved the conjecture up to a $\log n$ factor. The two results are incomparable – depending on the choice of variances, one or another gives a better bound. The Riemer-Schütt estimate was used recently in [21].

We would also like to mention that the non-commutative Khintchine inequality can be used to show that the expected spectral norm is bounded from above by the largest Euclidean norm of its rows and columns times a factor $\sqrt{\log n}$ (see e.g. (4.9) in [23]).

Another big step toward the solution was made a short while ago by Bandeira and Van Handel [1]. In particular, they proved that

$$(1.1) \qquad \mathbb{E}\left\|(a_{ij}g_{ij}) : \ell_2^n \to \ell_2^m\right\| \leq C\left(\|\!|A|\!\| + \sqrt{\log \min(n, m)} \cdot \max_{ij} |a_{ij}|\right),$$

where $\|\!|A|\!\|$ denotes the largest Euclidean norm of the rows and columns of $(a_{ij})$, $C > 0$ is a universal constant, and $g_{ij}$ are independent standard Gaussian random variables (see [1, Theorem 3.1]). Under mild structural assumptions, the bound (1.1) is already optimal. Further progress was made by Van Handel [24] who verified the conjecture up to a $\sqrt{\log \log n}$ factor. In fact, more was proved in [24]. He computed precisely the expectation of the largest Euclidean norm of the rows and columns using Gaussian concentration. And, while the moment method is at the heart of the proofs in [22] and [1], he proposed a very nice approach based on the comparison of Gaussian processes to improve the result of Latała. His approach can be also used for our setting. We comment on this in Section 4.

The purpose of this work is to provide bounds for operator norms of such structured Gaussian random matrices considered as operators from $\ell_{p^*}^n$ to $\ell_q^m$.

In what follows, by $g_i$, $g_{ij}$, $i \geq 1$, $j \geq 1$ we always denote independent standard Gaussian random variables. Let $n, m \in \mathbb{N}$ and $A = (a_{ij})_{i,j=1}^{m,n} \in \mathbb{R}^{m \times n}$. We write $G = G_A = (a_{ij}g_{ij})_{i,j=1}^{m,n}$. For $r \geq 1$, we denote by $\gamma_r \approx \sqrt{r}$ the $L_r$-norm of a standard Gaussian random variable. The notation $f \approx h$ means that there are two absolute positive constants $c$ and $C$ (that is, independent of any parameters) such that $cf \leq h \leq Cf$ and $f \approx_{p,q} h$ means that there are two positive constants $c(p, q)$ and $C(p, q)$, which depend only on the parameters $p$ and $q$, such that $c(p, q)f \leq h \leq C(p, q)f$.

Our main result is the following theorem.

**Theorem 1.1.** *For every $1 < p^* \leq 2 \leq q < \infty$ one has*

$$
\begin{aligned}
\mathbb{E} \left\| G : \ell_{p^*}^n \to \ell_q^m \right\| &\leq \left( \mathbb{E} \left\| G : \ell_{p^*}^n \to \ell_q^m \right\|^q \right)^{1/q} \\
&\leq C \, p^{5/q} \, (\log m)^{1/q} \left[ \gamma_p \max_{i \leq m} \| (a_{ij})_{j=1}^n \|_p + \gamma_q \, \mathbb{E} \max_{\substack{i \leq m \\ j \leq n}} |a_{ij} g_{ij}| \right] \\
&\quad + 2^{1/q} \, \gamma_q \max_{j \leq n} \| (a_{ij})_{i=1}^m \|_q,
\end{aligned}
$$

*where $C$ is a positive absolute constant.*

We conjecture the following bound.

**Conjecture 1.2.** *For every $1 \leq p^* \leq 2 \leq q \leq \infty$ one has*

$$
\mathbb{E} \left\| G : \ell_{p^*}^n \to \ell_q^m \right\| \approx \max_{i \leq m} \| (a_{ij})_{j=1}^n \|_p + \max_{j \leq n} \| (a_{ij})_{i=1}^m \|_q + \mathbb{E} \max_{\substack{i \leq m \\ j \leq n}} |a_{ij} g_{ij}|.
$$

Here, as usual, $p$ is defined via the relation $1/p + 1/p^* = 1$. This conjecture extends the corresponding conjecture for the case $p = q = 2$ and $m = n$. In this case, Bandeira and Van Handel proved in [1] an estimate with $\sqrt{\log \min(m, n)} \max |a_{ij}|$ instead of $\mathbb{E} \max |a_{ij} g_{ij}|$ (see equation (1.1)), while in [24] the corresponding bound is proved with $\sqrt{\log \log n}$ in front of the right hand side.

**Remark 1.3.** The lower bound in the conjecture is almost immediate and follows from standard estimates. Thus the upper bound is the only difficulty.

**Remark 1.4.** In the case $p^* = 1$ and $q \geq 2$, a direct computation following along the lines of Lemma 3.2 below, shows that

$$
\mathbb{E} \left\| G : \ell_1^n \to \ell_q^m \right\| \lesssim \gamma_q \max_{j \leq n} \| (a_{ij})_{i=1}^m \|_q + \mathbb{E} \max_{\substack{i \leq m \\ j \leq n}} |a_{ij} g_{ij}|.
$$

**Remark 1.5.** Note that if $1 \leq p^* \leq 2 \leq q \leq \infty$, in the case of matrices of tensor structure, that is, $(a_{ij})_{i,j=1}^n = x \otimes y = (x_j \cdot y_i)_{i,j=1}^n$, with $x, y \in \mathbb{R}^n$, Chevet's theorem [4, 3] and a direct computation show that

$$
\mathbb{E} \left\| G : \ell_{p^*}^n \to \ell_q^n \right\| \approx_{p,q} \|y\|_q \|x\|_\infty + \|y\|_\infty \|x\|_p.
$$

If the matrix is diagonal, that is, $(a_{ij})_{i,j=1}^n = \text{diag}(a_{11}, \ldots, a_{nn})$, then we immediately obtain

$$
\mathbb{E} \left\| G : \ell_{p^*}^n \to \ell_q^n \right\| = \mathbb{E} \| (a_{ii} g_{ii})_{i=1}^n \|_\infty \approx \max_{i \leq n} \sqrt{\ln(i+3)} \cdot a_{ii}^* \approx \| (a_{ii})_{i=1}^n \|_{M_g},
$$

where $(a_{ii}^*)_{i \leq n}$ is the decreasing rearrangement of $(|a_{ii}|)_{i \leq n}$ and $M_g$ is the Orlicz function given by

$$
M_g(s) = \sqrt{\frac{2}{\pi}} \int_0^s e^{-\frac{1}{2t^2}} \, dt
$$

(see Lemma 2.2 below and [11, Lemma 5.2] for the Orlicz norm expression).

Slightly different estimates, but of the same flavour, can also be obtained in the case $1 \leq q \leq 2 \leq p^* \leq \infty$.

## 2. Notation and Preliminaries

By $c, C, C_1, \ldots$ we always denote positive absolute constants, whose values may change from line to line, and we write $c_p, C_p, \ldots$ if the constants depend on some parameter $p$.

Given $p \in [1, \infty]$, $p^*$ denotes its conjugate and is given by the relation $1/p + 1/p^* = 1$. For $x = (x_i)_{i \leq n} \in \mathbb{R}^n$, $\|x\|_p$ denotes its $\ell_p$-norm, that is $\|x\|_\infty = \max_{i \leq n} |x_i|$ and, for $p < \infty$,

$$\|x\|_p = \Big( \sum_{i=1}^n |x_i|^p \Big)^{1/p}.$$

The corresponding space $(\mathbb{R}^n, \|\cdot\|_p)$ is denoted by $\ell_p^n$, its unit ball by $B_p^n$.

If $E$ is a normed space, then $E^*$ denotes its dual space and $B_E$ its closed unit ball. The modulus of convexity of $E$ is defined for any $\varepsilon \in (0, 2)$ by

$$\delta_E(\varepsilon) := \inf \Big\{ 1 - \Big\| \frac{x+y}{2} \Big\|_E \ : \ \|x\|_E = 1, \ \|y\|_E = 1, \ \|x - y\|_E > \varepsilon \Big\}.$$

We say that $E$ has modulus of convexity of power type 2 if there exists a positive constant $c$ such that for all $\varepsilon \in (0, 2)$, $\delta_E(\varepsilon) \geq c\varepsilon^2$. It is well known that this property (see e.g. [8] or [18, Proposition 2.4]) is equivalent to the fact that

$$\Big\| \frac{x+y}{2} \Big\|_E^2 + \lambda^{-2} \Big\| \frac{x-y}{2} \Big\|_E^2 \leq \frac{\|x\|_E^2 + \|y\|_E^2}{2}$$

holds for all $x, y \in E$, where $\lambda > 0$ is a constant depending only on $c$. In that case, we say that $E$ has modulus of convexity of power type 2 with constant $\lambda$. We clearly have $\delta_E(\varepsilon) \geq \varepsilon^2/(2\lambda^2)$.

Recall that a Banach space $E$ is of Rademacher type $r$ for some $1 \leq r \leq 2$ if there is $C > 0$ such that for all $n \in \mathbb{N}$ and for all $x_1, \ldots, x_n \in E$,

$$\Big( \mathbb{E}_\varepsilon \Big\| \sum_{i=1}^n \varepsilon_i x_i \Big\|^2 \Big)^{1/2} \leq C \Big( \sum_{i=1}^n \|x_i\|^r \Big)^{1/r},$$

where $(\varepsilon_i)_{i=1}^\infty$ is a sequence of independent random variables defined on some probability space $(\Omega, \mathbb{P})$ such that $\mathbb{P}(\varepsilon_i = 1) = \mathbb{P}(\varepsilon_i = -1) = \frac{1}{2}$ for every $i \in \mathbb{N}$. The smallest $C$ is called type-$r$ constant of $E$, denoted by $T_r(E)$. This concept was introduced into Banach space theory by Hoffmann-Jørgensen [14] in the early 1970s and the basic theory was developed by Maurey and Pisier [17].

We will need the following theorem.

**Theorem 2.1.** *Let $E$ be a Banach space with modulus of convexity of power type 2 with constant $\lambda$. Let $X_1, \ldots, X_m \in E^*$ be independent random vectors, $q \geq 2$ and define*

$$B := C\lambda^4 T_2(E^*) \sqrt{\frac{\log m}{m}} \Big( \mathbb{E} \max_{i \leq m} \|X_i\|_{E^*}^q \Big)^{1/2},$$

*and*

$$\sigma := \sup_{y \in B_E} \Big( \frac{1}{m} \sum_{i=1}^m \mathbb{E} |\langle X_i, y \rangle|^q \Big)^{1/q}.$$

*Then*

$$\mathbb{E} \sup_{y \in B_E} \left| \frac{1}{m} \sum_{i=1}^{m} |\langle X_i, y \rangle|^q - \mathbb{E}|\langle X_i, y \rangle|^q \right| \leq B^2 + B \cdot \sigma^{q/2}.$$

Its proof is done following the argument "proof of condition (H)" of [13] in combination with the improvement on covering numbers established in [12, Lemma 2]. Indeed, in [12], the argument is only made in the simpler case $q = 2$, but it can be extended verbatim to the case $q \geq 2$.

We also recall known facts about Gaussian random variables. The next lemma is well-known (see e.g. Lemmas 2.3, 2.4 in [24]).

**Lemma 2.2.** *Let $a = (a_i)_{i \leq n} \in \mathbb{R}^n$ and $(a_i^*)_{i \leq n}$ be the decreasing rearrangement of $(|a_i|)_{i \leq n}$. Then*

$$\mathbb{E} \max_{i \leq n} |a_i g_i| \approx \max_{i \leq n} \sqrt{\ln(i+3)} \cdot a_i^*.$$

Note that in general the maximum of i.i.d. random variables weighted by coordinates of a vector $a$ is equivalent to a certain Orlicz norm $\|a\|_M$, where the function $M$ depends only on the distribution of random variables (see [10, Corollary 2] and Lemma 5.2 in [11]).

The following theorem is the classical Gaussian concentration inequality (see e.g. [5] or inequality (2.35) and Proposition 2.18 in [16]).

**Theorem 2.3.** *Let $n \in \mathbb{N}$ and $(Y, \|\cdot\|_Y)$ be a Banach space. Let $y_1, \ldots, y_n \in Y$ and $X = \sum_{i=1}^{n} g_i y_i$. Then, for every $t > 0$,*

$$(2.1) \qquad \mathbb{P}\Big( \big| \|X\|_Y - \mathbb{E}\|X\|_Y \big| \geq t \Big) \leq 2 \exp\left( -\frac{t^2}{2\sigma_Y(X)^2} \right),$$

*where $\sigma_Y(X) = \sup_{\|\xi\|_{Y^*}=1} \left( \sum_{i=1}^{n} |\xi(y_i)|^2 \right)^{1/2}$.*

**Remark 2.4.** Let $p \geq 2$. Let $a = (a_j)_{j \leq n} \in \mathbb{R}^n$ and $X = (a_j g_j)_{j \leq n}$. Then we clearly have

$$\sigma_{\ell_p^n}(X) = \max_{j \leq n} |a_j|.$$

Thus, Theorem 2.3 implies for $X = (a_j g_j)_{j \leq n}$

$$(2.2) \qquad \mathbb{P}\Big( \big| \|X\|_p - \mathbb{E}\|X\|_p \big| > t \Big) \leq 2 \exp\left( -\frac{t^2}{2 \max_{j \leq n} |a_j|^2} \right).$$

Note also that

$$(2.3) \qquad \mathbb{E}\|X\|_p \leq \left( \sum_{j=1}^{n} |a_j|^p \, \mathbb{E}|g_j|^p \right)^{1/p} = \gamma_p \|a\|_p.$$

## 3. Proof of the main result

We will apply Theorem 2.1 with $E = \ell_{p^*}^n$, $1 < p^* \leq 2$ and $X_1, \ldots, X_m$ being the rows of the matrix $G = (a_{ij} g_{ij})_{i,j=1}^{m,n}$. We start with two lemmas in which we estimate the quantity $\sigma$ and the expectation, appearing in that theorem.

**Lemma 3.1.** *Let $m, n \in \mathbb{N}$, $1 < p^* \leq 2 \leq q$, and for $i \leq m$ let $X_i = (a_{ij}g_{ij})_{j=1}^n$. Then*

$$\sigma = \sup_{y \in B_{p^*}^n} \left( \frac{1}{m} \sum_{i=1}^m \mathbb{E}|\langle X_i, y \rangle|^q \right)^{1/q} = \frac{\gamma_q}{m^{1/q}} \max_{j \leq n} \|(a_{ij})_{i=1}^m\|_q.$$

*Proof.* For every $i \leq m$, $\langle X_i, y \rangle = \sum_{j=1}^n a_{ij} y_j g_{ij}$, is a Gaussian random variable with variance $\|(a_{ij}y_j)_{j=1}^n\|_2$. Hence,

$$\sigma^q = \sup_{y \in B_{p^*}^n} \frac{1}{m} \sum_{i=1}^m \mathbb{E}|\langle X_i, y \rangle|^q = \frac{\gamma_q^q}{m} \sup_{y \in B_{p^*}^n} \sum_{i=1}^m \left( \sum_{j=1}^n |a_{ij}y_j|^2 \right)^{q/2}.$$

Since $p^* \leq 2 \leq q$, the function

$$\phi(z) = \sum_{i=1}^m \left( \sum_{j=1}^n |a_{ij}|^2 |z_j|^{2/p^*} \right)^{q/2}$$

is a convex function on the simplex $S = \{z \in \mathbb{R}^n \mid \sum_{j=1}^n \leq 1, \forall j : z_j \geq 0\}$. Therefore, it attains its maximum on extreme points, that is, on vectors of the canonical unit basis of $\mathbb{R}^n$, $e_1, \ldots, e_n$. Thus,

$$\sup_{y \in B_{p^*}^n} \sum_{i=1}^m \left( \sum_{j=1}^n |a_{ij}y_j|^2 \right)^{q/2} = \sup_{z \in S} \phi(z) = \sup_{k \leq n} \phi(e_k) = \max_{j \leq n} \|(a_{ij})_{i=1}^m\|_q^q,$$

which completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Now we estimate the expectation in Theorem 2.1. The proof is based on the Gaussian concentration, Theorem 2.3, and is similar to Theorem 2.1 and Remark 2.2 in [24].

**Lemma 3.2.** *Let $m, n \in \mathbb{N}$, $1 < p^* \leq 2 \leq q$, and for $i \leq m$ let $X_i = (a_{ij}g_{ij})_{j=1}^n$. Then*

$$\left( \mathbb{E} \max_{i \leq m} \|X_i\|_p^q \right)^{1/q} \leq \max_{i \leq m} \mathbb{E}\|X_i\|_p + C \gamma_q \mathbb{E} \max_{\substack{i \leq m \\ j \leq n}} |a_{ij}g_{ij}|$$

$$\leq \gamma_p \max_{i \leq m} \|(a_{ij})_{j=1}^n\|_p + C \gamma_q \mathbb{E} \max_{\substack{i \leq m \\ j \leq n}} |a_{ij}g_{ij}|,$$

*where $C$ is a positive absolute constant.*

*Proof.* We have

$$\left( \mathbb{E} \max_{i \leq m} \|X_i\|_p^q \right)^{1/q} \leq \left\| \max_{i \leq m} \left| \|X_i\|_p - \mathbb{E}\|X_i\|_p \right| + \max_{i \leq m} \mathbb{E}\|X_i\|_p \right\|_{L_q}$$

$$\leq \left( \mathbb{E} \max_{i \leq m} \left| \|X_i\|_p - \mathbb{E}\|X_i\|_p \right|^q \right)^{1/q} + \max_{i \leq m} \mathbb{E}\|X_i\|_p.$$

For all $i \leq m$ and $t > 0$ by (2.2) we have

$$(3.1) \qquad \mathbb{P}\left( \left| \|X_i\|_p - \mathbb{E}\|X_i\|_p \right| > t \right) \leq 2 \exp\left( -\frac{t^2}{2 \max_{j \leq n} |a_{ij}|^2} \right).$$

By permuting the rows of $(a_{ij})_{i,j=1}^{m,n}$, we can assume that

$$\max_{j \leq n} |a_{1j}| \geq \cdots \geq \max_{j \leq n} |a_{mj}|.$$

For each $i \leq m$, choose $j(i) \leq n$ such that $|a_{ij(i)}| = \max_{j \leq n} |a_{ij}|$. Clearly,

$$\max_{\substack{i \leq m \\ j \leq n}} |a_{ij} g_{ij}| \geq \max_{i \leq m} |a_{ij(i)}| \cdot |g_{ij(i)}|$$

and hence, by independence of $g_{ij}$'s and Lemma 2.2,

$$b := \mathbb{E} \max_{\substack{i \leq m \\ j \leq n}} |a_{ij} g_{ij}| \geq \mathbb{E} \max_{i \leq m} |a_{ij(i)}| \cdot |g_i| \geq c \max_{i \leq m} \sqrt{\log(i+3)} \cdot |a_{ij(i)}|,$$

where the latter inequality follows since $|a_{1j(1)}| \geq \cdots \geq |a_{nj(n)}|$. Thus, for $i \leq m$,

$$\max_{j \leq n} |a_{ij}|^2 = a_{ij(i)}^2 \leq \frac{b^2}{c \log(i+3)}.$$

By (3.1) we observe for every $t > 0$,

$$\mathbb{P}\left( \max_{i \leq m} \big| \|X_i\|_p - \mathbb{E}\|X_i\|_p \big| > t \right) \leq 2 \sum_{i=1}^{m} \exp\left( -\frac{ct^2 \log(i+3)}{2b^2} \right)$$

$$= 2 \sum_{i=1}^{m} \left( \frac{1}{i+3} \right)^{ct^2/2b^2} \leq 2 \int_3^\infty x^{-ct^2/2b^2} \, dx$$

$$\leq 6 \cdot 3^{-ct^2/2b^2},$$

whenever $ct^2/b^2 \geq 4$. Integrating the tail inequality proves that

$$\left( \mathbb{E} \max_{i \leq m} \big| \|X_i\|_p - \mathbb{E}\|X_i\|_p \big|^q \right)^{1/q} \leq C_1 \sqrt{q}\, b \leq C_2 \, \gamma_q \, \mathbb{E} \max_{\substack{i \leq m \\ j \leq n}} |a_{ij} g_{ij}|.$$

By the triangle inequality, we obtain the first desired inequality, the second one follows by (2.3). $\qquad \square$

We are now ready to present the proof of the main theorem.

*Proof of Theorem 1.1.* First observe that

$$\mathbb{E} \left\| G : \ell_{p^*}^n \to \ell_q^m \right\| \leq \left( \mathbb{E} \left\| G : \ell_{p^*}^n \to \ell_q^m \right\|^q \right)^{1/q} = \left( \mathbb{E} \sup_{y \in B_{p^*}^n} \sum_{i=1}^{m} |\langle X_i, y \rangle|^q \right)^{1/q}.$$

We have

$$\mathbb{E} \sup_{y \in B_{p^*}^n} \sum_{i=1}^{m} |\langle X_i, y \rangle|^q \leq \mathbb{E} \sup_{y \in B_{p^*}^n} \left[ \sum_{i=1}^{m} |\langle X_i, y \rangle|^q - \mathbb{E}|\langle X_i, y \rangle|^q \right] + \sup_{y \in B_{p^*}^n} \sum_{i=1}^{m} \mathbb{E}|\langle X_i, y \rangle|^q$$

$$= m \cdot \mathbb{E} \sup_{y \in B_{p^*}^n} \left[ \frac{1}{m} \sum_{i=1}^{m} |\langle X_i, y \rangle|^q - \mathbb{E}|\langle X_i, y \rangle|^q \right] + m \cdot \sigma^q.$$

Hence, Theorem 2.1 applied with $E = \ell_{p^*}^n$ implies

$$\mathbb{E} \left\| G : \ell_{p^*}^n \to \ell_q^m \right\|^q \leq m \cdot \left[ B^2 + B\sigma^{q/2} \right] + m \cdot \sigma^q \leq 2m \left( B^2 + \sigma^q \right),$$

where $B$ and $\sigma$ are defined in that theorem. Therefore,

$$\left( \mathbb{E} \left\| G : \ell_{p^*}^n \to \ell_q^m \right\|^q \right)^{1/q} \leq 2^{1/q} m^{1/q} \left( B^{2/q} + \sigma \right).$$

Now, recall that $T_2(\ell_p^n) \approx \sqrt{p}$ and that $B_{p^*}^n$ has modulus of convexity of power type 2 with $\lambda^{-2} \approx 1/p$ (see, e.g., [19, Theorem 5.3]). Therefore,

$$B^{2/q} = C^{2/q} \lambda^{8/q} T_2^{2/q}(\ell_p^n) \left( \frac{\log m}{m} \right)^{1/q} \left( \mathbb{E} \max_{i \leq m} \|X_i\|_p^q \right)^{1/q}$$

$$= C^{2/q} p^{5/q} (\log m)^{1/q} m^{-1/q} \left( \mathbb{E} \max_{i \leq m} \|X_i\|_p^q \right)^{1/q}.$$

Applying Lemma 3.1, we obtain

$$\left( \mathbb{E} \left\| G : \ell_{p^*}^n \to \ell_q^m \right\|^q \right)^{1/q}$$

$$\leq (2C^2)^{1/q} \cdot p^{5/q} \cdot (\log m)^{1/q} \left( \mathbb{E} \max_{i \leq m} \|X_i\|_p^q \right)^{1/q} + 2^{1/q} \gamma_q \cdot \max_{j \leq n} \|(a_{ij})_{i=1}^m\|_q.$$

The desired bound follows now from Lemma 3.2. □

**Remark 3.3.** This proof can be extended to the case of random matrices whose rows are centered independent vectors with multivariate Gaussian distributions. We leave the details to the interested reader.

## 4. Concluding Remarks

In this section, we briefly outline what can be obtained using the approach of [24]. We use a standard trick to pass to a symmetric matrix. The matrix $G_A$ being given, define $S$ as

$$S = \frac{1}{2} \begin{pmatrix} 0 & G_A^T \\ G_A & 0 \end{pmatrix}.$$

Then, $S$ is a random symmetric matrix and

$$\sup_w \langle Sw, w \rangle = \sup_{u \in B_{p^*}^n} \sup_{v \in B_{q^*}^m} \langle G_A u, v \rangle = \left\| G_A : \ell_{p^*}^n \to \ell_q^m \right\|,$$

where the supremum in $w$ is taken over all vectors of the form $(u, v)^T$ with $u \in B_{p^*}^n$ and $v \in B_{q^*}^m$. Repeating verbatim the proof of Theorem 4.1 in [24] one gets

$$\mathbb{E} \left\| G_A : \ell_{p^*}^n \to \ell_q^m \right\| \quad \lesssim_{p,q} \mathbb{E} \max_{i \leq m} \left( \sum_{j=1}^n |g_j|^p |a_{ij}|^p \right)^{1/p}$$

$$+ \mathbb{E} \max_{j \leq n} \left( \sum_{i=1}^m |g_i|^q |a_{ij}|^q \right)^{1/q} + \mathbb{E} \max_i Y_i,$$

where $Y \sim N(0, A^-)$ and $A^-$ is a positive definite matrix whose diagonal elements are bounded by

$$\max \left( \max_{i \leq m} \sqrt{\sum_j a_{ij}^4}, \max_{j \leq n} \sqrt{\sum_i a_{ij}^4} \right).$$

However, the bounds obtained here and in Theorem 1.1 are incomparable. Depending on the situation one may be better than the other.

We would also like to thank our colleague R. Adamczak for helpful comments. We are grateful to the anonymous referee for many useful comments and remarks helping us to improve the presentation as well as for showing the argument outlined in the last section.

## References

[1] A. Bandeira and R. van Handel. Sharp nonasymptotic bounds on the norm of random matrices with independent entries. *Ann. Probab.*, to appear.

[2] G. Bennett, V. Goodman, and C.M. Newman. Norms of random matrices. *Pacific J. Math.*, 59(2):359–365, 1975.

[3] Y. Benyamini and Y. Gordon. Random factorization of operators between banach spaces. *J. Analyse Math.*, 39(45-74), 1981.

[4] S. Chevet. Séries de variables aléatoires gaussiennes à valeurs dans $E \hat{\otimes}_\varepsilon F$. Application aux produits d'espaces de Wiener abstraits. In *Séminaire sur la Géométrie des Espaces de Banach (1977–1978)*, pages Exp. No. 19, 15. École Polytech., Palaiseau, 1978.

[5] B. S. Cirel'son, I. A. Ibragimov, and V. N. Sudakov. Norms of Gaussian sample functions. In *Proceedings of the Third Japan-USSR Symposium on Probability Theory (Tashkent, 1975)*, pages 20–41. Lecture Notes in Math., Vol. 550. Springer, Berlin, 1976.

[6] K.R. Davidson and S.J. Szarek. *Addenda and corrigenda to: "Local operator theory, random matrices and Banach spaces"*. Handbook of the geometry of Banach spaces, Vol. 2. North-Holland, North-Holland, Amsterdam, 2003.

[7] K.R. Davidson and S.J. Szarek. *Local operator theory, random matrices and Banach spaces*. Handbook of the geometry of Banach spaces, Vol. 1. North-Holland, North-Holland, Amsterdam, 2003.

[8] T. Figiel. On the moduli of convexity and smoothness. *Studia Math.*, 56(2):121–155, 1976.

[9] Y. Gordon. Some inequalities for gaussian processes and applications. *Israel J. Math.*, 50:265–289, 1985.

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

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

[12] O. Guédon, S. Mendelson, A. Pajor, and N. Tomczak-Jaegermann. Majorizing measures and proportional subsets of bounded orthonormal systems. *Revista Matemtica Iberoamericana*, 24(3):1075–1095, 2008.

[13] O. Guédon and M. Rudelson. Moments of random vectors via majorizing measures. *Adv. Math.*, 208(2):798 – 823, 2007.

[14] J. Hoffmann-Jørgensen. Sums of independent Banach space valued random variables. *Studia Math.*, 52:159–186, 1974.

[15] R. Latała. Some estimates of norms of random matrices. *Proc. Amer. Math. Soc.*, 133(5):1273–1282 (electronic), 2005.

[16] M. Ledoux. *The concentration of measure phenomenon*. Mathematical Surveys and Monographs, 89. American Mathematical Society, Providence, RI, 2001.

[17] B. Maurey and G. Pisier. Séries de variables aléatoires vectorielles indépendantes et propriétés géométriques des espaces de Banach. *Studia Math.*, 58(1):45–90, 1976.

[18] G. Pisier. Martingales with values in uniformly convex spaces. *Israel J. Math.*, 20(3-4):326–350, 1975.

[19] G. Pisier and Q. Xu. Non-commutative $L^p$-spaces. In *Handbook of the geometry of Banach spaces, Vol. 2*, pages 1459–1517. North-Holland, Amsterdam, 2003.

[20] S. Riemer and C. Schütt. On the expectation of the norm of random matrices with non-identically distributed entries. *Electron. J. Probab.*, 18:no. 29, 1–13, 2013.

[21] M. Rudelson and O. Zeitouni. Singular values of gaussian matrices and permanent estimators. *Random Structures and Algorithms*, to appear.

[22] Y. Seginer. The expected norm of random matrices. *Combin. Probab. Comput.*, 9:149–166, 2000.

[23] Joel A. Tropp. User-friendly tail bounds for sums of random matrices. *Foundations of Computational Mathematics*, 12(4):389–434, 2012.

[24] R. Van Handel. On the spectral norm of gaussian random matrices. *Trans. Amer. Math. Soc.*, to appear.

[25] J. von Neumann and H. H. Goldstine. Numerical inverting of matrices of high order. *Bull. Amer. Math. Soc.*, 53(11):1021–1099, 11 1947.

[26] E. P. Wigner. Characteristic Vectors of Bordered Matrices With Infinite Dimensions. *The Annals of Mathematics*, 62(3):548–564, 1955.

[27] E. P. Wigner. On the Distribution of the Roots of Certain Symmetric Matrices. *The Annals of Mathematics*, 67(2):325–327, March 1958.

[28] J. Wishart. The generalised product moment distribution in samples from a normal multi-variate population. *Biometrika*, 20A(1/2):32–52, July 1928.

(Olivier Guédon) Laboratoire d'Analyse et de Mathématiques Appliquées, Université Paris-Est, 77454 Marne-la-Vallée, Cedex 2, France
  *E-mail address*: `olivier.guedon@u-pem.fr`

(Alexander E. Litvak) Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, AB, T6G 2G1, Canada
  *E-mail address*: `alitvak@ualberta.ca`

(Aicke Hinrichs) Institut für Analysis, Johannes Kepler Universität Linz, Altenbergerstrasse 69, 4040 Linz, Austria
  *E-mail address*: `aicke.hinrichs@jku.at`

(Joscha Prochno) Department of Mathematics, University of Hull, Cottingham Road, Hull, HU6 7RX, United Kingdom
  *E-mail address*: `j.prochno@hull.ac.uk`