Norms of structured random matrices

Radosław Adamczak[1] · Joscha Prochno[2] · Marta Strzelecka[1,3] ·
Michał Strzelecki[1]

Abstract

For $m, n \in \mathbb{N}$, let $X = (X_{ij})_{i \leq m, j \leq n}$ be a random matrix, $A = (a_{ij})_{i \leq m, j \leq n}$ a real deterministic matrix, and $X_A = (a_{ij} X_{ij})_{i \leq m, j \leq n}$ the corresponding structured random matrix. We study the expected operator norm of X_A considered as a random operator between ℓ_p^n and ℓ_q^m for $1 \leq p, q \leq \infty$. We prove optimal bounds up to logarithmic terms when the underlying random matrix X has i.i.d. Gaussian entries, independent mean-zero bounded entries, or independent mean-zero ψ_r ($r \in (0, 2]$) entries. In certain cases, we determine the precise order of the expected norm up to constants. Our results are expressed through a sum of operator norms of Hadamard products $A \circ A$ and $(A \circ A)^T$.

✉ Marta Strzelecka
  martast@mimuw.edu.pl

  Radosław Adamczak
  radamcz@mimuw.edu.pl

  Joscha Prochno
  joscha.prochno@uni-passau.de

  Michał Strzelecki
  michalst@mimuw.edu.pl

[1] Institute of Mathematics, University of Warsaw, Banacha 2, 02–097 Warsaw, Poland

[2] Faculty of Computer Science and Mathematics, University of Passau, Innstraße 33, 94032 Passau, Germany

[3] Institute of Mathematics and Scientific Computing, University of Graz, Heinrichstraße 36, 8010 Graz, Austria

Contents

1 Introduction and main results

With his work on the statistical analysis of large samples [69], Wishart initiated the systematic study of large random matrices. Ever since, random matrices have continuously entered more and more areas of mathematics and applied sciences beyond probability theory and statistics, for instance, in numerical analysis through the work of Goldstine and von Neumann [20, 65] and in quantum physics through the works of Wigner [66–68] on his famous semicircle law, which resulted in significant effort to understand spectral statistics of random matrices from an asymptotic point of view. Today, random matrix theory has grown into a vital area of probability theory and statistics, and within the last two decades, random matrices have come to play a major role in many areas of (algorithmic) computational mathematics, for instance, in questions related to sparsification methods [1, 54] and sparse approximation [57, 58],

dimension reduction [4, 12, 44], or combinatorial optimization [46, 53]. We refer the reader to [5, 6, 60] for more information.

In this paper, we are interested in the non-asymptotic theory of (large) random matrices. This theory plays a fundamental role in geometric functional analysis at least since the '70s, the connection coming in various different flavors. It is of particular importance in the geometry of Banach spaces and the theory of operator algebras [9, 10, 15, 18, 21, 30] and their applications to high-dimensional problems, for instance, in convex geometry [17, 22], compressed sensing [14, 16, 48, 63], information-based complexity [27, 28], or statistical learning theory [50, 64]. On the other hand, geometric functional analysis had and still has enduring influence on random matrix theory as is witnessed, for instance, through applications of measure concentration techniques; we refer to [15, 42] and the references cited therein. The quantity we study and focus on here concerns the expected operator norm of random matrices considered as operators between finite-dimensional ℓ_p spaces; recall that ℓ_p^n denotes the space $\mathbb{R}^n$ equipped with the (quasi-)norm $\|\cdot\|_p$, given by $\|(x_j)_{j=1}^n\|_p = (\sum_{j=1}^n |x_j|^p)^{1/p}$ for $0 < p < \infty$ and $\|(x_j)_{j=1}^n\|_\infty = \max_{j \leq n} |x_j|$ if $p = \infty$. We address the following problem: for $1 \leq p, q \leq \infty$ and $m, n \in \mathbb{N}$, determine the right order (up to constants that may depend on the parameters p and q) of

$$\mathbb{E}\|X_A \colon \ell_p^n \to \ell_q^m\|,$$

where, given a deterministic real $m \times n$ matrix $A = (a_{ij})_{i \leq m, j \leq n}$ and a random matrix $X = (X_{ij})_{i \leq m, j \leq n}$, we denote by

$$X_A := A \circ X = (a_{ij} X_{ij})_{i \leq m, j \leq n}$$

the structured random matrix; the symbol $\circ$ stands for the Hadamard product of matrices (i.e., entrywise multiplication). The bounds on the expected operator norm should be of optimal order and expressed in terms of the coefficients a_{ij}, $i \leq m$, $j \leq n$. Understanding such expressions and related quantities is important, for instance, when studying the worst-case error of optimal algorithms which are based on random information in function approximation problems [28] (see also [33]) or the quality of random information for the recovery of vectors from an ℓ_p-ellipsoid, where (the radius of) optimal information is given by Gelfand numbers of a diagonal operator [29].

In the case where the random entries of X are i.i.d. standard Gaussians (then we write G_A instead of X_A) and $1 \leq p, q \leq \infty$, we will show the following bound, which is sharp up to logarithmic terms:

$$D_1 + D_2 \lesssim \mathbb{E}\|G_A \colon \ell_p^n \to \ell_q^m\| \lesssim (\ln n)^{1/p^*}(\ln m)^{1/q}\left[\sqrt{\ln(mn)}\,D_1 + \sqrt{\ln n}\,D_2\right],$$
(1.1)

where $D_1 := \|A \circ A \colon \ell_{p/2}^n \to \ell_{q/2}^m\|^{1/2}$, $D_2 := \|(A \circ A)^T \colon \ell_{q^*/2}^m \to \ell_{p^*/2}^n\|^{1/2}$, and p^* denotes the Hölder conjugate of p defined by the relation $1/p + 1/p^* = 1$. As will be explained later, we obtain sharp estimates in certain cases and derive results similar to (1.1) for other models of randomness.

1.1 History of the problem and known results

In what follows, $A = (a_{ij})_{i,j}$ is a real deterministic matrix and $G = (g_{ij})_{i,j}$ always stands for a random matrix with i.i.d. standard Gaussian entries (usually the matrices are of size $m \times n$ unless explicitly stated otherwise). We use $C(r)$, $C(r, K)$, etc. for positive constants which may depend only on the parameters given in brackets and write $C, C', c, c', \ldots$ for positive absolute constants. The symbols $\lesssim, \lesssim_r, \lesssim_{r,K}$, etc. denote that the inequality holds up to multiplicative constants depending only on the parameters given in the subscripts; we write $a \asymp b$ if $a \lesssim b$ and $b \lesssim a$, and $\asymp_r, \asymp_{r,K}$, etc. if the constants may depend on the parameters given in the subscript.

In 1975, Bennett, Goodman, and Newman [9] proved that if X is an $m \times n$ random matrix with independent, mean-zero entries taking values in $[-1, 1]$, and $2 \le q < \infty$, then

$$\mathbb{E}\|X: \ell_2^n \to \ell_q^m\| \lesssim_q \max\{n^{1/2}, m^{1/q}\}. \tag{1.2}$$

In fact, up to constants, this estimate is best possible: for any $m \times n$ matrix X' with ± 1 entries one readily sees that $\|X': \ell_2^n \to \ell_q^m\| \ge \max\{n^{1/2}, m^{1/q}\}$; just use standard unit vectors and operator duality. Moreover, in this 'unstructured' case, where $a_{ij} = 1$ for all i, j, it is easy to extend (1.2) to the whole range of $p, q \in [1, \infty]$ (see [8, 13] or Remark 4.2 below). Also, if all entries are i.i.d. Rademacher random variables, then the bounds are two-sided, i.e., the expected operator norm is, up to constants, the same as the minimal norm for all p, q (see [8, Proposition 3.2] or [13, Satz 2]).

The case most studied in the literature is the one of the spectral norm, i.e., the $\ell_2^n \to \ell_2^m$ operator norm. Seginer [51] proved in 2000 that if $X = (X_{ij})_{i \le m, j \le n}$ is an $m \times n$ random matrix with i.i.d. mean-zero entries, then its operator norm is of the same order as the sum of expectations of the maximum Euclidean norm of rows and columns of X, i.e.,

$$\mathbb{E}\|X: \ell_2^n \to \ell_2^m\| \asymp \mathbb{E} \max_{j \le n} \|(X_{ij})_{i=1}^m\|_2 + \mathbb{E} \max_{i \le m} \|(X_{ij})_{j=1}^n\|_2. \tag{1.3}$$

Riemer and Schütt [49] proved that, up to a logarithmic factor $\ln(en)^2$, the same holds true for any random matrix with independent but not necessarily identically distributed mean-zero entries. Let us also mention that in the Gaussian setting one can use a non-commutative Khintchine bound (see, e.g., [59, Equation (4.9)]) to show that, up to a factor $\sqrt{\ln n}$, the expected spectral norm is of the order of the largest Euclidean norm of its rows and columns.

In the very same setting that was considered by Riemer and Schütt, Latała [37] had obtained a few years earlier the dimension-free estimate

$$\mathbb{E}\|X: \ell_2^n \to \ell_2^m\| \lesssim \max_{j \le n}\Big(\sum_{i=1}^m \mathbb{E}X_{ij}^2\Big)^{1/2} + \max_{i \le m}\Big(\sum_{j=1}^n \mathbb{E}X_{ij}^2\Big)^{1/2} + \Big(\sum_{i=1}^m \sum_{j=1}^n \mathbb{E}X_{ij}^4\Big)^{1/4}.$$

This bound is superior to the Riemer–Schütt bound in the case of matrices with all entries equal to 1 and is optimal for Wigner matrices. In other cases, like the one of diagonal matrices, the Riemer–Schütt bound is better.

In the case of structured Gaussian matrices, Latała, van Handel, and Youssef [40], building upon earlier work of Bandeira and van Handel [7] (which combined the moment method with combinatorial considerations) as well as results proved by van Handel in [61] (which used Slepian's lemma), obtained the precise behavior without any logarithmic terms in the dimension, namely

$$\mathbb{E}\|G_A \colon \ell_2^n \to \ell_2^m\| \asymp \mathbb{E} \max_{j \le n} \|(a_{ij}g_{ij})_{i=1}^m\|_2 + \mathbb{E} \max_{i \le m} \|(a_{ij}g_{ij})_{j=1}^n\|_2$$
$$\asymp \max_{j \le n} \|(a_{ij})_{i=1}^m\|_2 + \max_{i \le m} \|(a_{ij})_{j=1}^n\|_2 + \mathbb{E} \max_{i \le m, j \le n} |a_{ij}g_{ij}|.$$
$$(1.4)$$

Their proof is based on a clever block decomposition of the underlying matrix (see [40, Fig. 3.1]). This result finally answered in the affirmative a conjecture made by Latała more than a decade before. We also refer the reader to the survey [62] discussing in quite some detail results prior to [40] and [61]—the latter work discusses the conjectures of Latała and van Handel and shows their equivalence.

Very recently, Latała and Świątkowski [39] investigated a similar problem when the underlying random matrix has Rademacher entries. They proved a lower bound which, up to a $\ln \ln n$ factor, can be reversed for randomized $n \times n$ circulant matrices.

In [23], Guédon, Hinrichs, Litvak, and Prochno studied our main and motivating question on the order of the expected operator norm of structured random matrices considered as operators between ℓ_p^n and ℓ_q^m in the special case where $p \le 2 \le q$ and the random entries are Gaussian. In this situation, where we are not dealing with the spectral norm, the moment method cannot be employed. The approach in [23] was therefore different and based on a majorizing measure construction combining the works [24] and [25]. In [23, Theorem 1.1], the authors proved that if $1 < p \le 2 \le q < \infty$, then

$$\mathbb{E}\|G_A \colon \ell_p^n \to \ell_q^m\| \lesssim \gamma_q \max_{j \le n} \|(a_{ij})_{i=1}^m\|_q + (p^*)^{5/q} (\ln m)^{1/q} \gamma_{p^*} \max_{i \le m} \|(a_{ij})_{j=1}^n\|_{p^*}$$
$$+ (p^*)^{5/q} (\ln m)^{1/q} \gamma_q \mathbb{E} \max_{i \le m, j \le n} |a_{ij}g_{ij}|, \qquad (1.5)$$

where $\gamma_r := (\mathbb{E}|g|^r)^{1/r}$ for a standard Gaussian random variable g. Moreover, for $p = 1$ and $q \ge 2$, it was noted in [23, Remark 1.4] (see also [45, Twierdzenie 2]) that

$$\mathbb{E}\|G_A \colon \ell_1^n \to \ell_q^m\| \lesssim \sqrt{q} \max_{j \le n} \|(a_{ij})_{i=1}^m\|_q + \mathbb{E} \max_{i \le m, j \le n} |a_{ij}g_{ij}|. \qquad (1.6)$$

Later, an extension of (1.5) to the case of matrices with i.i.d. isotropic log-concave rows was obtained by Strzelecka in [55].

Trying to extend the upper bound for $\mathbb{E}\|G_A \colon \ell_p^n \to \ell_q^m\|$ to the whole range $1 \le p, q \le \infty$ one encounters two difficulties. First of all, the methods used in order to

prove (1.5) fail if $q \leq 2$ or $p \geq 2$, because the majorizing measure construction used in [23] is restricted to the case $q \geq 2$ and the assumption $1 < p \leq 2$ is required in a Hölder bound. Moreover, when $q \leq 2$ or $p \geq 2$ the result cannot hold with the right-hand side of the same form as in (1.5) (see Remark 4.2 below for counterexamples[1] to (1.5) in the cases $q \leq 2$ and $p \geq 2$). This explains the different form of expressions D_1 and D_2 in (1.1), which in the range $p \leq 2 \leq q$ reduce to the maxima of norms on the right-hand side of (1.5)—see (1.9) below.

1.2 Lower bounds and conjectures

By arguments similar to the ones used in order to prove the lower bound in (1.4), one can check that in the range considered in [23, 45] (i.e., $1 \leq p \leq 2 \leq q \leq \infty$) one has

$$
\mathbb{E}\|G_A \colon \ell_p^n \to \ell_q^m\| \gtrsim_{p,q} \max_{j \leq n} \|(a_{ij})_{i=1}^m\|_q + \max_{i \leq m} \|(a_{ij})_{j=1}^n\|_{p*}
$$
$$
+ \mathbb{E} \max_{i \leq m, j \leq n} |a_{ij} g_{ij}|. \tag{1.7}
$$

Note that for $p = 1$,

$$
\max_{i \leq m} \|(a_{ij})_{j=1}^n\|_{p*} = \max_{i \leq m, j \leq n} |a_{ij}| \leq \sqrt{\pi/2}\, \mathbb{E} \max_{i \leq m, j \leq n} |a_{ij} g_{ij}|,
$$

which explains the simplified form of (1.6).

We remark that the proof of (1.7) is based merely on the observation that the operator norm is greater than the maximum entry of the matrix and the appropriate maximum norms of its rows and columns, combined with comparison of moments for Gaussian random vectors. Another but related way to proceed, valid for all $1 \leq p, q \leq \infty$, is to exchange expectation and suprema over the ℓ_p^n and ℓ_{q*}^m balls in the definition of the operator norm. We present the details in Sect. 5.1. In particular, Proposition 5.1 and Corollary 5.2 imply[2] that, for $1 \leq p, q \leq \infty$,

$$
\mathbb{E}\|G_A \colon \ell_p^n \to \ell_q^m\| \gtrsim \|A \circ A \colon \ell_{p/2}^n \to \ell_{q/2}^m\|^{1/2} + \|(A \circ A)^T \colon \ell_{q*/2}^m \to \ell_{p*/2}^n\|^{1/2}
$$
$$
+ \mathbb{E} \max_{i \leq m, j \leq n} |a_{ij} g_{ij}|. \tag{1.8}
$$

It is an easy observation (see Lemma 2.1 below) that for $p \leq 2 \leq q$,

$$
\|A \circ A \colon \ell_{p/2}^n \to \ell_{q/2}^m\|^{1/2} = \max_{j \leq n} \|(a_{ij})_{i=1}^m\|_q,
$$
$$
\|(A \circ A)^T \colon \ell_{q*/2}^m \to \ell_{p*/2}^n\|^{1/2} = \max_{i \leq m} \|(a_{ij})_{j=1}^n\|_{p*}. \tag{1.9}
$$

---

[1] By Jensen's inequality, the expected norm of a matrix with i.i.d. Rademacher entries is less than or equal to $\sqrt{2/\pi}$ times the expected norm of the matrix with Gaussian entries, so (1.5) for $q \leq 2$ or $p \geq 2$ would imply the same (up to a constant) bound for ± 1 matrices, which does not hold in this range of (p, q) as we explain in Remark 4.2.

[2] We use here also a trivial observation that $\|G_A \colon \ell_p^n \to \ell_q^m\| \geq \max_{i,j} |a_{ij} g_{ij}|$.

Thus, in the range $1 \le p \le 2 \le q < \infty$ considered in [23, 45], the lower bounds (1.7) and (1.8) coincide.

Although it would be natural to conjecture at this point that the bound (1.8) may be reversed up to a multiplicative constant depending only on p, q, such a reverse bound turns out not to be true in the case $p \le q < 2$ (and in the dual one $2 < p \le q$) as we shall show in Sect. 5.3.

In order to conjecture the right asymptotic behavior of $\mathbb{E}\|G_A \colon \ell_p^n \to \ell_q^m\|$, one may take a look at the boundary values of p and q, i.e., $p \in \{1, \infty\}$ or $q \in \{1, \infty\}$. Note that (1.6) provides an asymptotic behavior of $\mathbb{E}\|G_A \colon \ell_p^n \to \ell_q^m\|$ on a part of this boundary (i.e., for $p = 1$ and $2 \le q \le \infty$ and in the dual case $q = \infty$ and $1 \le p \le 2$). We provide sharp results on the remaining parts of the boundary of $[1, \infty] \times [1, \infty]$ (see dense lines on the boundary of Fig. 1 below):

$$\mathbb{E}\|G_A \colon \ell_p^n \to \ell_1^m\| \asymp_p D_1 + D_2 \qquad \text{for all } 1 < p \le \infty,$$

$$\mathbb{E}\|G_A \colon \ell_\infty^n \to \ell_q^m\| \asymp_q D_1 + D_2 \qquad \text{for all } 1 \le q < \infty,$$

$$\mathbb{E}\|G_A \colon \ell_1^n \to \ell_q^m\| \asymp D_1 + \max_{j \le n}(\sqrt{\ln(j+1)}b_j^\downarrow) \qquad \text{for all } 1 \le q \le 2,$$

$$\mathbb{E}\|G_A \colon \ell_p^n \to \ell_\infty^m\| \asymp D_2 + \max_{i \le m}(\sqrt{\ln(i+1)}d_i^\downarrow) \qquad \text{for all } 2 \le p \le \infty,$$

where

$$D_1 := \|A \circ A \colon \ell_{p/2}^n \to \ell_{q/2}^m\|^{1/2}, \qquad b_j := \|(a_{ij})_{i \le m}\|_{2q/(2-q)},$$

$$D_2 := \|(A \circ A)^T \colon \ell_{q^*/2}^m \to \ell_{p^*/2}^n\|^{1/2}, \qquad d_i := \|(a_{ij})_{j \le n}\|_{2p/(p-2)},$$

and with $(x_1^\downarrow, \ldots, x_n^\downarrow)$ denoting the non-increasing rearrangement of $(|x_1|, \ldots, |x_n|)$ for a given $(x_j)_{j \le n} \in \mathbb{R}^n$. (For the precise formulation see Propositions 1.8 and 1.10, and Corollary 1.11 below.)

Moreover, in Sect. 5.1 we generalize the lower bounds from the boundary into the whole range $(p, q) \in [1, \infty] \times [1, \infty]$ (see Fig. 1 below), i.e., we prove

$$\mathbb{E}\|G_A \colon \ell_p^n \to \ell_q^m\| \gtrsim_{p,q} D_1 + D_2 + \begin{cases} \mathbb{E} \max_{i \le m, j \le n} |a_{ij}g_{ij}| & \text{if } p \le 2 \le q, \\ \max_{j \le n} \sqrt{\ln(j+1)}b_j^\downarrow & \text{if } p \le q \le 2, \\ \max_{i \le m} \sqrt{\ln(i+1)}d_i^\downarrow & \text{if } 2 \le p \le q, \\ 0 & \text{if } q < p. \end{cases}$$

(1.10)

Let us now discuss the relation between the terms appearing above. We postpone the proofs of all the following claims to Sect. 5.

In the case $p \le 2 \le q$, we have

$$D_1 + D_2 + \mathbb{E} \max_{i \le m, j \le n} |a_{ij}g_{ij}| \asymp_{p,q} D_1 + D_2 + \max_{i \le m, j \le n} \sqrt{\ln(j+1)}a'_{ij}$$

$$\asymp_{p,q} D_1 + D_2 + \max_{i \le m, j \le n} \sqrt{\ln(i+1)}a''_{ij}, \quad (1.11)$$
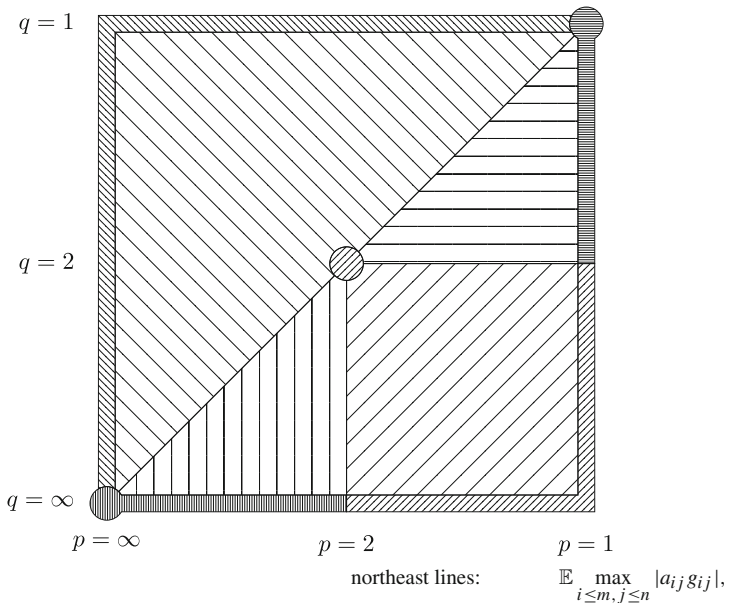
Fig. 1 The third summand in (1.10) and in Conjecture 1:

northeast lines: $\mathbb{E} \max_{i \leq m, j \leq n} |a_{ij} g_{ij}|$,

horizontal lines: $\max_{j \leq n} \sqrt{\ln(j+1)} b_j^{\downarrow}$,

vertical lines: $\max_{i \leq m} \sqrt{\ln(i+1)} d_i^{\downarrow}$,

northwest lines: 0.

Note that the horizontal axis represents $1/p$ and the vertical one $1/q$. Dense lines correspond to exact asymptotics and loosely spaced lines to upper and lower bounds matching up to logarithms

where the matrices $(a'_{ij})_{i,j}$ and $(a''_{ij})_{i,j}$ are obtained by permuting the columns and rows, respectively, of the matrix $(|a_{ij}|)_{i,j}$ in such a way that $\max_i a'_{i1} \geq \cdots \geq \max_i a'_{in}$ and $\max_j a''_{1j} \geq \cdots \geq \max_j a''_{mj}$. Therefore, in the range $1 \leq p \leq q \leq \infty$ the right-hand side of (1.10) changes continuously with p and q (for a fixed matrix A).

Obviously, $\max_{j \leq n} \sqrt{\ln(j+1)} b_j^{\downarrow} \geq \max_{i \leq m, j \leq n} \sqrt{\ln(j+1)} a'_{ij}$ and, in general, the former quantity may be of larger order than the latter one. In Sect. 5.3 we shall present a more subtle relation: for every $1 \leq p \leq q < 2$ we shall give an example showing that the right-hand side of (1.10) may be of larger order than $D_1 + D_2 + \mathbb{E} \max_{i \leq m, j \leq n} |a_{ij} g_{ij}|$. Note that by duality, i.e., the fact that

$$\|X_A : \ell_p^n \to \ell_q^m\| = \|(X_A)^T : \ell_{q*}^m \to \ell_{p*}^n\| = \|(X^T)_{A^T} : \ell_{q*}^m \to \ell_{p*}^n\|, \quad (1.12)$$

the same holds in the case $2 < p \leq q$. This suggests that the behavior of $\mathbb{E}\|G_A : \ell_p^n \to \ell_q^m\|$ is different in the regions with horizontal or vertical lines than in the region with northeast lines.

Moreover, we have

$$D_1 + D_2 \gtrsim_{p,q} \begin{cases} \max_{j \leq n} \sqrt{\ln(j+1)} b_j^{\downarrow} & \text{if } q < p \text{ and } q < 2, \\ \max_{i \leq m} \sqrt{\ln(i+1)} d_i^{\downarrow} & \text{if } q < p \text{ and } p^* < 2 \end{cases} \quad (1.13)$$

(see Sect. 5.2). Note that this is not the case for $p \leq q$, as one can easily see by considering, e.g., A equal to the identity matrix. This suggests a different (than in other regions), simplified, behavior of $\mathbb{E}\|G_A : \ell_p^n \to \ell_q^m\|$ in the region with northwest lines.

Given the discussion above, the lower bounds presented in (1.10), and the fact that they can be reversed for all $p \in [1, \infty]$, $q \in \{1, \infty\}$ (and for all $q \in [1, \infty]$, $p \in \{1, \infty\}$), it is natural to conjecture the following.

Conjecture 1 *For all* $1 \leq p, q \leq \infty$, *we conjecture that*

$$\mathbb{E}\|G_A : \ell_p^n \to \ell_q^m\| \asymp_{p,q} D_1 + D_2 + \begin{cases} \mathbb{E} \max_{i \leq m, j \leq n} |a_{ij} g_{ij}| & \text{if } p \leq 2 \leq q, \\ \max_{j \leq n} \sqrt{\ln(j+1)} b_j^\downarrow & \text{if } p \leq q \leq 2, \\ \max_{i \leq m} \sqrt{\ln(i+1)} d_i^\downarrow & \text{if } 2 \leq p \leq q, \\ 0 & \text{if } q < p. \end{cases}$$

(1.14)

Remark 1.1 One could pose another natural conjecture, based on the potential generalization of the first line of the bound (1.4), namely that the inequality

$$\mathbb{E}\|G_A : \ell_p^n \to \ell_q^m\| \asymp_{p,q} \mathbb{E} \max_{i \leq m} \|(a_{ij} g_{ij})_j\|_{p^*} + \mathbb{E} \max_{j \leq n} \|(a_{ij} g_{ij})_i\|_q \quad (1.15)$$

holds for all $1 \leq p, q \leq \infty$. Indeed, the lower bound is true with constant $\frac{1}{2}$, since for every deterministic matrix X one has

$$\|X : \ell_p^n \to \ell_q^m\| \geq \max\left\{ \max_{i \leq m} \|(X_{ij})_j\|_{p^*}, \max_{j \leq n} \|(X_{ij})_i\|_q \right\}.$$

However, as we prove in Sect. 5.4, this conjecture is wrong: although the right-hand sides of (1.14) and (1.15) are comparable in the range $1 \leq p \leq 2 \leq q \leq \infty$, for every pair of p, q outside this range the right-hand side of (1.15) may be of smaller order than the left-hand side.

Let us now present a conjecture concerning the boundedness of linear operators given by infinite dimensional matrices. In what follows, we say that a matrix $B = (b_{ij})_{i,j\in\mathbb{N}}$ defines a bounded operator from $\ell_p(\mathbb{N})$ to $\ell_q(\mathbb{N})$ if for all $x \in \ell_p(\mathbb{N})$ the product Bx is well defined, belongs to $\ell_q(\mathbb{N})$ and the corresponding linear operator is bounded.

Conjecture 2 *Let* $A = (a_{ij})_{i,j\in\mathbb{N}}$ *be an infinite matrix with real coefficients and let* $1 \leq p, q \leq \infty$. *We conjecture that the matrix* $G_A = (a_{ij} g_{ij})_{i,j\in\mathbb{N}}$ *defines a bounded linear operator between* $\ell_p(\mathbb{N})$ *and* $\ell_q(\mathbb{N})$ *almost surely if and only if the matrix* $A \circ A$ *defines a bounded linear operator between* $\ell_{p/2}(\mathbb{N})$ *and* $\ell_{q/2}(\mathbb{N})$, *the matrix* $(A \circ A)^T$ *defines a bounded linear operator between* $\ell_{q^*/2}(\mathbb{N})$ *and* $\ell_{p^*/2}(\mathbb{N})$, *and*

- *in the case* $p \leq 2 \leq q$, $\mathbb{E} \sup_{i,j\in\mathbb{N}} |a_{ij} g_{ij}| < \infty$,

- *in the case $p \leq q \leq 2$, $\lim_{j \to \infty} b_j = 0$, and $\sup_{j \in \mathbb{N}} \sqrt{\ln(j+1)} b_j^{\downarrow} < \infty$, where*
  $b_j = \|(a_{ij})_{i \in \mathbb{N}}\|_{2q/(2-q)}$, $j \in \mathbb{N}$,
- *in the case $2 \leq p \leq q$, $\lim_{i \to \infty} d_i = 0$, and $\sup_{i \in \mathbb{N}} \sqrt{\ln(i+1)} d_i^{\downarrow} < \infty$, where*
  $d_i := \|(a_{ij})_{j \in \mathbb{N}}\|_{2p/(p-2)}$, $i \in \mathbb{N}$,
- *(in the case $q < p$ we do not need to assume any additional conditions).*

We remark that it suffices to prove Conjecture 1 in order to confirm Conjecture 2.

Proposition 1.2 *Assume $1 \leq p, q \leq \infty$. Then (1.14) for this choice of p, q implies the assertion of Conjecture 2 for the same choice of p, q.*

We postpone the proof of this proposition to Subsection 5.5.

In this article, in addition to the cases $p = q = 2$ obtained in [40] and $p = 1, q \geq 2$ proved in [23, 45], we confirm Conjecture 1 when $p \in \{1, \infty\}$, $q \in [1, \infty]$ and when $q \in \{1, \infty\}$, $p \in [1, \infty]$. In all the other cases, we are able to prove the upper bounds only up to logarithmic (in the dimensions m, n) multiplicative factors (see Corollary 1.4 below). In particular, Proposition 1.2 implies that Conjecture 2 holds for all $p \in \{1, \infty\}$, $q \in [1, \infty]$ and for all $q \in \{1, \infty\}$, $p \in [1, \infty]$.

Note that in the structured case we work with, interpolating the results obtained for the boundary cases $p \in \{1, \infty\}$ or $q \in \{1, \infty\}$ gives a bound with polynomial (in the dimensions) multiplicative constants which are much worse than logarithmic constants from Corollary 1.4 below. However, as we shall see in Remark 4.2 below, interpolation techniques work well in the non-structured case.

1.3 Main results valid for $1 \leq p, q \leq \infty$

We start with general theorems valid for the whole range of p, q. Results which are based on methods working only for specific values of p, q, but yielding better logarithmic terms, are presented in the next subsection. A brief summary and comparison of all results can be found in Table 1.

Before stating our main results, we need to introduce additional notation. For a non-empty set $J \subset \{1, \ldots, n\}$, and $1 \leq p \leq \infty$, we define

$$B_p^J := \left\{ (x_j)_{j \in J} : \sum_{j \in J} |x_j|^p \leq 1, \quad x_j \in \mathbb{R} \right\}.$$

By ℓ_p^J we denote the space $\mathbb{R}^J := \{(x_j)_{j \in J} : x_j \in \mathbb{R}\}$ equipped with the norm

$$\|x\|_{\ell_p^J} = \left(\sum_{j \in J} |x_j|^p \right)^{1/p},$$

whose unit ball is B_p^J. Obviously, the space ℓ_p^J can be identified with a subspace of ℓ_p^n. If $A : \ell_p^n \to \ell_q^m$ is a linear operator, the notation $A : \ell_p^J \to \ell_q^I$ means that A is restricted to the space ℓ_p^J and composed with a projection onto ℓ_q^I. Moreover, for $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, $\sup_J \|x\|_{\ell_p^J} = \left(\sum_{j \leq k} |x_j^{\downarrow}|^p \right)^{1/p}$, where the supremum is

taken over all $J \subset \{1, \ldots, n\}$ with $|J| = k$, and $(x_1^\downarrow, \ldots, x_n^\downarrow)$ is the non-increasing rearrangement of $(|x_1|, \ldots, |x_n|)$.

Theorem 1.3 (Main theorem in a general version with sets I_0, J_0) *Assume that $m \leq M$, $n \leq N$, $1 \leq p, q \leq \infty$, and $G = (g_{ij})_{i \leq M, j \leq N}$ has i.i.d. standard Gaussian entries. Then*

$$
\begin{aligned}
\mathbb{E} \sup_{I_0, J_0} \|G_A \colon \ell_p^{J_0} \to \ell_q^{I_0}\| &= \mathbb{E} \sup_{I_0, J_0} \sup_{x \in B_p^{J_0}} \sup_{y \in B_{q^*}^{I_0}} \sum_{i \in I_0} \sum_{j \in J_0} y_i a_{ij} g_{ij} x_j \\
&\leq \ln(en)^{1/p^*} \ln(em)^{1/q} \Big[\big(2.4\sqrt{\ln(mn)} + 8\sqrt{\ln M} + \sqrt{2/\pi}\big) \sup_{I_0, J_0} \|A \circ A \colon \ell_{p/2}^{J_0} \to \ell_{q/2}^{I_0}\|^{1/2} \\
&\quad + \big(8\sqrt{\ln N} + 2\sqrt{2/\pi}\big) \sup_{I_0, J_0} \|(A \circ A)^T \colon \ell_{q^*/2}^{I_0} \to \ell_{p^*/2}^{J_0}\|^{1/2} \Big],
\end{aligned}
$$

where the suprema are taken over all sets $I_0 \subset \{1, \ldots, M\}$, $J_0 \subset \{1, \ldots, N\}$ such that $|I_0| = m$, $|J_0| = n$.

The above theorem gives an estimate on the largest operator norm among all sub-matrices of G_A of fixed size. Let us remark that apart from being of intrinsic interest, quantities of this type (for $p = q = 2$) have appeared in connection with the study of the restricted isometry property of random matrices with independent rows [2] or in the analysis of entropic uncertainty principles for random quantum measurements [3, 47].

Let us now give an outline of the proof of Theorem 1.3. Note that

$$
\|G_A \colon \ell_p^{J_0} \to \ell_q^{I_0}\| = \sup_{x \in B_p^{J_0}} \sup_{y \in B_{q^*}^{I_0}} \sum_{i \in I_0} \sum_{j \in J_0} y_i a_{ij} g_{ij} x_j. \tag{1.16}
$$

In the first step of our proof, we find polytopes L and K approximating (with accuracy depending logarithmically on the dimension) the unit balls in $\ell_p^{J_0}$ and $\ell_{q^*}^{I_0}$, respectively. The extreme points of the sets K and L have a special and simple structure: absolute values of their non-zero coordinates are all equal to a constant depending only on the size of the support of a given point. Since K is close to $B_{q^*}^{I_0}$ and L is close to $B_p^{J_0}$, we may consider only $x \in \text{Ext}(L)$, $y \in \text{Ext}(K)$ in (1.16). Since non-zero coordinates of $x \in \text{Ext}(L)$ and $y \in \text{Ext}(K)$, respectively, are all equal up to a sign we may use a symmetrization argument and the contraction principle to remove x and y in the sum on the right-hand side of (1.16). Thus, in the next step of the proof we only need to estimate the expected value of

$$
\sup_{I_0, J_0} \sup_{\emptyset \neq I \subset I_0} \sup_{\emptyset \neq J \subset J_0} |I|^{-1/q^*} |J|^{-1/p} \sum_{i \in I, j \in J} a_{ij} g_{ij},
$$

where I and J represent the potential supports of points in $\text{Ext}(K)$ and $\text{Ext}(L)$. To deal with this quantity, we first consider the suprema over the subsets of fixed sizes and use Slepian's lemma to compare the supremum of the Gaussian process above with the supremum of another Gaussian process, which may be bounded easily. Then

we make use of the term $|I|^{-1/q^*}|J|^{-1/p} < 1$, which allows us to go back to suprema over the sets $B_p^{J_0}$ and $B_{q^*}^{I_0}$. At the end, we use the Gaussian concentration inequality to unfix the sizes of sets I and J and complete the proof.

Applying Theorem 1.3 with $N = n$, $M = m$ immediately yields the following result, which confirms Conjecture 1 up to some logarithmic terms.

Corollary 1.4 (Main theorem – ℓ_p to ℓ_q version) *Assume that $1 \leq p, q \leq \infty$ and $G = (g_{ij})_{i \leq m, j \leq n}$ has i.i.d. standard Gaussian entries. Then,*

$$\mathbb{E}\|G_A \colon \ell_p^n \to \ell_q^m\| \lesssim (\ln n)^{1/p^*}(\ln m)^{1/q}\Big[\sqrt{\ln(mn)}\|A \circ A \colon \ell_{p/2}^n \to \ell_{q/2}^m\|^{1/2}$$
$$+ \sqrt{\ln n}\|(A \circ A)^T \colon \ell_{q^*/2}^m \to \ell_{p^*/2}^n\|^{1/2}\Big].$$

Moreover, we easily recover the same bound in the case of independent bounded entries. We state and prove a general version with sets I_0 and J_0 akin to Theorem 1.3 in Sect. 3.2.

Corollary 1.5 *Assume that $1 \leq p, q \leq \infty$ and $X = (X_{ij})_{i \leq m, j \leq n}$ has independent mean-zero entries taking values in $[-1, 1]$. Then*

$$\mathbb{E}\|X_A \colon \ell_p^n \to \ell_q^m\| \lesssim (\ln n)^{1/p^*}(\ln m)^{1/q}\Big[\sqrt{\ln(mn)}\|A \circ A \colon \ell_{p/2}^n \to \ell_{q/2}^m\|^{1/2}$$
$$+ \sqrt{\ln n}\|(A \circ A)^T \colon \ell_{q^*/2}^m \to \ell_{p^*/2}^n\|^{1/2}\Big].$$

We use the two results above to obtain their analogue in the case of ψ_r entries for $r \leq 2$; these random variables are defined by (1.17).

This class contains, among others,

- log-concave random variables (which are ψ_1),
- heavy tailed Weibull random variables (of shape parameter $r \in (0, 1)$, i.e., $\mathbb{P}(|X_{ij}| \geq t) = e^{-t^r/L}$ for $t \geq 0$),
- random variables satisfying the condition

$$\|X_{ij}\|_{2\rho} \leq \alpha \|X_{ij}\|_\rho \qquad \text{for all } \rho \geq 1.$$

These random variables are ψ_r with $r = 1/\log_2 \alpha$. They were considered recently in [38].

A general version of the following Corollary 1.6 with sets I_0 and J_0 is stated and proved in Subsection 3.2.

Corollary 1.6 *Assume that $K, L > 0$, $r \in (0, 2]$, $1 \leq p, q \leq \infty$, and $X = (X_{ij})_{i \leq m, j \leq n}$ has independent mean-zero entries satisfying*

$$\mathbb{P}(|X_{ij}| \geq t) \leq K e^{-t^r/L} \quad \text{for all } t \geq 0, i \leq m, j \leq n. \tag{1.17}$$

Then

$$\mathbb{E}\|X_A\colon \ell_p^n \to \ell_q^m\|$$
$$\lesssim_{r,K,L} (\ln n)^{1/p^*} (\ln m)^{1/q} \ln(mn)^{\frac{1}{r}-\frac{1}{2}} \left[\sqrt{\ln(mn)}\|A \circ A\colon \ell_{p/2}^n \to \ell_{q/2}^m\|^{1/2} \right.$$
$$\left. + \sqrt{\ln n}\|(A \circ A)^T\colon \ell_{q^*/2}^m \to \ell_{p^*/2}^n\|^{1/2} \right].$$

1.4 Results for particular ranges of *p, q*

We continue with results for some specific ranges of p, q, where we are able to prove estimates with better logarithmic dependence (results which follow from them by duality (1.12) are stated in Table 1 to keep the presentation short). We postpone their proofs to Sect. 4. We start with the case of Gaussian random variables. Recall that $\gamma_q = (\mathbb{E}|g|^q)^{1/q}$, where g is a standard Gaussian random variable.

Proposition 1.7 *For all* $1 \leq p \leq 2$ *and* $1 \leq q < \infty$, *we have*

$$\mathbb{E}\|G_A\colon \ell_p^n \to \ell_q^m\| \leq \gamma_q \ln(en)^{1/p^*}\|A \circ A\colon \ell_{p/2}^n \to \ell_{q/2}^m\|^{1/2}$$
$$+ 2.2 \ln(en)^{1/2+1/p^*}\|(A \circ A)^T\colon \ell_{q^*/2}^m \to \ell_{p^*/2}^n\|^{1/2}. \quad (1.18)$$

If $q = 1$ or $p = \infty$, then we are able to get a result without logarithmic terms. Recall that for a sequence $(x_j)_{j \leq n}$ we denote by $(x_j^\downarrow)_{j \leq n}$ the non-increasing rearrangement of $(|x_j|)_{j \leq n}$.

Proposition 1.8 (i) *For* $1 < p \leq \infty$, *we have*

$$\|A \circ A\colon \ell_{p/2}^n \to \ell_{1/2}^m\|^{1/2} + \|(A \circ A)^T\colon \ell_\infty^m \to \ell_{p^*/2}^n\|^{1/2} \lesssim \mathbb{E}\|G_A\colon \ell_p^n \to \ell_1^m\|$$
$$\leq \gamma_1\|A \circ A\colon \ell_{p/2}^n \to \ell_{1/2}^m\|^{1/2} + 2\gamma_{p^*}\|(A \circ A)^T\colon \ell_\infty^m \to \ell_{p^*/2}^n\|^{1/2}.$$

(ii) *Moreover,*

$$\mathbb{E}\|G_A\colon \ell_1^n \to \ell_1^m\| \asymp \|A \circ A\colon \ell_{1/2}^n \to \ell_{1/2}^m\|^{1/2} + \max_{j \leq n} \sqrt{\ln(j+1)}b_j^\downarrow,$$

where $b_j := \|(a_{ij})_{i \leq m}\|_2, \ j \leq n$.

Note that (ii) shows in particular that a blow up of the constant γ_{p^*} in the upper estimate (i) for $p \to 1$ is necessary, since the right most summands in (i) and (ii) are non-comparable.

Remark 1.9 It shall be clear from the proof that the upper bound in part (i) of Proposition 1.8 remains valid for any random matrix X (instead of G) with independent

Table 1 In several cases we prove a result of the type: $\mathbb{E}\|X_A : \ell_p^n \to \ell_q^m\| \lesssim_{p,q} L_0\big(L_1\|A \circ A : \ell_p^n \to \ell_{q/2}^m\|^{1/2} + L_2\|(A \circ A)^T : \ell_{q^*/2}^m \to \ell_{p^*/2}^n\|^{1/2}\big)$, where L_0, L_1, L_2 are some logarithmic terms depending only on n, m and $p, q \in [1, \infty]$. The table presents a summary of the results. The expression 'set K' refers to the use of Lemma 2.3. We denote $L_K := \ln(n)^{1/p^*}\ln(m)^{1/q}$.

X_{ij}	p, q	Logarithmic terms			Statement	Tools used in proof
		L_0	L_1	L_2		
Gaussian	Any	L_K	$\ln(mn)^{1/2}$	$\ln(n)^{1/2}$	Cor. 1.4	Set K + Gaussian concentration, contraction principle, Slepian's lemma
Bounded	Any	L_K	$\ln(mn)^{1/2}$	$\ln(n)^{1/2}$	Cor. 1.5	Uses the Gaussian case
$\psi_r, r \le 2$	Any	L_K	$\ln(mn)^{1/r}$	$\ln(n)^{1/2}\ln(mn)^{1/r-1/2}$	Cor. 1.6	Coupling, uses Gaussian and bounded case
Gaussian	(a) $p \le 2$	1	$\ln(n)^{1/p^*}$	$\ln(n)^{1/2+1/p^*}$	Prop. 1.7	Set K + Gaussian concentration
	(a') $q \ge 2$	1	$\ln(m)^{1/2+1/q}$	$\ln(m)^{1/q}$		Case dual to (a)
	(b) $p > q = 1$	1	1	1	Prop. 1.8	Contraction principle
	(b') $p = \infty > q$	1	1	1		Case dual to (b)
	(c) $p = 1$	Different behavior, confirming Conjecture 1			Prop. 1.10, (1.6)	Gaussian concentration, Lem. 2.11 and 2.12
	(c') $q = \infty$	Different behavior, confirming Conjecture 1			Cor. 1.11	Case dual to(c)
Bounded	(a) $p \le 2$	1	$\ln(n)^{1/p^*}$	$\ln(n)^{1/2+1/p^*}$	Cor. 1.13	Set K + Talagrand's convex concentration
	(a') $q \ge 2$	1	$\ln(m)^{1/2+1/q}$	$\ln(m)^{1/q}$		Case dual to (a)
	(b) $p \le 2 \le q$	1	$\ln(n)^{1/p^*}$	$\ln(n)^{1/q+1/p^*}$	Prop. 1.14	Set K + method à la [9]
$\psi_r, r \le 2$	(a) $p \le 2$	1	$\ln(n)^{1/p^*}$	$\ln(n)^{1/2+1/p^*}\ln(mn)^{1/r}$	Thm. 1.15	Set K + Talagrand's convex concentration, cut-off at lvl $\ln(mn)^{1/r}$
	(b) $p \le 2$	$\ln(mn)^{1/r-1/2}$	$\ln(n)^{1/p^*}$	$\ln(n)^{1/2+1/p^*}$	Cor. 1.13	Coupling, uses Gaussian and bounded cases
	(a') $q \ge 2$	1	$\ln(m)^{1/2+1/q}\ln(mn)^{1/r}$	$\ln(m)^{1/q}$		Case dual to (a)
	(b') $q \ge 2$	$\ln(mn)^{1/r-1/2}$	$\ln(m)^{1/2+1/q}$	$\ln(m)^{1/q}$		Case dual to (b)

isotropic rows (i.e., rows with mean zero and the covariance matrix equal to the identity) such that

$$\left(\mathbb{E}\Big|\sum_{i=1}^m \alpha_i X_{ij}\Big|^{p^*}\right)^{1/p^*} \lesssim_p \left(\sum_{i=1}^m \alpha_i^2\right)^{1/2} \qquad \text{for all } \alpha \in \mathbb{R}^m, \, j \le n. \quad (1.19)$$

Note that the independence and the isotropicity of rows imply that also the columns of X are isotropic (since the coordinates of every column are independent and have mean zero and variance 1). Therefore, whenever $p \ge 2$, condition (1.19) is always satisfied (because the p^*-integral norm is bounded above by the 2-integral norm, which is then equal to the right-hand side of (1.19), since the covariance matrix of each column is equal to the $m \times m$ identity matrix).

The following proposition generalizes part (ii) of Proposition 1.8 to an arbitrary $q \le 2$. We list it separately since we present a proof using different arguments. Recall that the case $p = 1$, $q \ge 2$ was established before, see (1.6).

Proposition 1.10 *If* $1 \le q \le 2$, *then*

$$\|G_A: \ell_1^n \to \ell_q^m\| \asymp \|A \circ A: \ell_{1/2}^n \to \ell_{q/2}^m\|^{1/2} + \max_{j \le n}(\sqrt{\ln(j+1)}b_j^\downarrow)$$

$$= \max_{j \le n}\|(a_{ij})_{i \le m}\|_q + \max_{j \le n}(\sqrt{\ln(j+1)}b_j^\downarrow),$$

where $b_j = \|(a_{ij})_{i \le m}\|_{2q/(2-q)}$ *for* $j \le n$.

Proposition 1.10 immediately implies its dual version.

Corollary 1.11 *If* $2 \le p \le \infty$, *then*

$$\|G_A: \ell_p^n \to \ell_\infty^m\| \asymp \|(A \circ A)^T: \ell_{1/2}^m \to \ell_{p^*/2}^n\|^{1/2} + \max_{i \le m}(\sqrt{\ln(i+1)}d_i^\downarrow)$$

$$= \max_{i \le m}\|(a_{ij})_{j \le n}\|_{p^*} + \max_{i \le m}(\sqrt{\ln(i+1)}d_i^\downarrow),$$

where $d_i = \|(a_{ij})_{j \le n}\|_{2p^*/(2-p^*)} = \|(a_{ij})_{j \le n}\|_{2p/(p-2)}$ *for* $i \le m$.

Remark 1.12 Corollary 1.11 and the dual version of (1.6) provide the exact behavior of expected norm of Gaussian operator from ℓ_p^n to ℓ_q^m not only when $q = \infty$, but also for $q \ge c_0 \ln m$, as we explain now. For all $q \ge q_0 := c_0 \ln m$ we have the following inequalities for norms on $\mathbb{R}^m$,

$$\|\cdot\|_q \ge \|\cdot\|_\infty \ge m^{-1/q_0}\|\cdot\|_{q_0} = e^{-1/c_0}\|\cdot\|_{q_0} \ge e^{-1/c_0}\|\cdot\|_q,$$

therefore,

$$\frac{1}{e^{1/c_0}}\mathbb{E}\|X_A: \ell_p^n \to \ell_q^m\| \le \mathbb{E}\|X_A: \ell_p^n \to \ell_\infty^m\| \le \mathbb{E}\|X_A: \ell_p^n \to \ell_q^m\|.$$

Similarly,

$$\|(A \circ A)^T : \ell_{q^*/2}^m \to \ell_{p^*/2}^n\| \asymp_{c_0} \|(A \circ A)^T : \ell_{1/2}^m \to \ell_{p^*/2}^n\|.$$

Proposition 1.7 implies the following estimate for matrices with independent ψ_r entries, in the same way as Corollary 1.4 implies Corollary 1.6 (see Sect. 3.2).

Corollary 1.13 *Assume that $K, L > 0$, $r \in (0, 2]$, and $X = (X_{ij})_{i \leq m, j \leq n}$ has independent mean-zero entries satisfying*

$$\mathbb{P}(|X_{ij}| \geq t) \leq K e^{-t^r/L} \quad \text{for all } t \geq 0. \tag{1.20}$$

Then, for $1 \leq p \leq 2$, $1 \leq q \leq \infty$,

$$\mathbb{E}\|X_A : \ell_p^n \to \ell_q^m\| \lesssim_{r,K,L} (\ln n)^{1/p^*} \ln(nm)^{1/r-1/2} \|A \circ A : \ell_{p/2}^n \to \ell_{q/2}^m\|^{1/2}$$
$$+ (\ln n)^{1/2+1/p^*} \ln(nm)^{1/r-1/2} \|(A \circ A)^T : \ell_{q^*/2}^m \to \ell_{p^*/2}^n\|^{1/2}.$$

By Hoeffding's inequality (i.e., Lemma 2.13) we know that matrices with independent valued in $[-1, 1]$ entries having mean zero satisfy (1.20) with $r = 2$ and $K = 2 = L$. In this special case of independent bounded random variables one can also adapt the methods of [9] to prove in the smaller range $1 \leq p \leq 2 \leq q < \infty$ the following result with explicit numerical constants and improved dependence on n (note that the second logarithmic term is better than in Corollary 1.13, where the exponent equals $1/2 + 1/p^*$).

Proposition 1.14 *Assume that $X = (X_{ij})_{i \leq m, j \leq n}$ has independent mean-zero entries taking values in $[-1, 1]$. Then, for $1 \leq p \leq 2 \leq q < \infty$,*

$$\mathbb{E}\|X_A : \ell_p^n \to \ell_q^m\| \leq C(q) \ln(en)^{1/p^*} \|A \circ A : \ell_{p/2}^n \to \ell_{q/2}^m\|^{1/2}$$
$$+ 10^{1/q} \ln(en)^{1/q+1/p^*} \|(A \circ A)^T : \ell_{q^*/2}^m \to \ell_{p^*/2}^n\|^{1/2},$$

where $C(q) := 2(q\Gamma(q/2))^{1/q} \asymp \sqrt{q}$.

Finally, we have the following general result for matrices with independent ψ_r entries (cf. Corollary 1.6).

Theorem 1.15 *Let $K, L > 0$, $r \in (0, 2]$, and assume that $X = (X_{ij})_{i \leq m, j \leq n}$ has independent mean-zero entries satisfying*

$$\mathbb{P}(|X_{ij}| \geq t) \leq K e^{-t^r/L} \quad \text{for all } t \geq 0.$$

Then, for all $1 \leq p \leq 2$ and $1 \leq q < \infty$,

$$\mathbb{E}\|X_A : \ell_p^n \to \ell_q^m\| \lesssim_{r,K,L} q^{1/r}(\ln n)^{1/p^*} \|A \circ A : \ell_{p/2}^n \to \ell_{q/2}^m\|^{1/2}$$
$$+ (\ln n)^{1/2+1/p^*} \ln(mn)^{1/r} \|(A \circ A)^T : \ell_{q^*/2}^m \to \ell_{p^*/2}^n\|^{1/2}.$$

Having in mind the strategy of proof described after Theorem 1.3, let us elaborate on the idea of proof of Theorem 1.15. We shall split the matrix X into two parts $X^{(1)}$ and $X^{(2)}$ which we treat separately. In our decomposition, all entries of $X^{(1)}$ are bounded by $C \ln(mn)^{1/r}$ and the probability that $X^{(2)} \neq 0$ is very small. Then we shall deal with $X^{(2)}$ using a crude bound (Lemma 4.3) and the fact that the probability that $X^{(2)} \neq 0$ is small enough to compensate it. In order to bound the expectation of the norm of $X^{(1)}$, we require a cut-off version of Theorem 1.15 (Lemma 4.4). To obtain it, we shall replace B_p^n in the expression for the operator norm with a suitable polytope K (and leave $\sup_{y \in B_{q^*}^m}$ as it is) and then apply a Gaussian-type concentration inequality to the function $Z \mapsto F(Z) := \|Z_A x\|_q$ for $x \in \mathrm{Ext}(K)$.

1.5 Tail bounds

All the bounds for $\mathbb{E}\|X_A : \ell_p^n \to \ell_q^m\|$ provided in this work for random matrices X also yield a tail bound for $\|X_A : \ell_p^n \to \ell_q^m\|$. (It is clear from the proof of Proposition 1.16— see Sect. 3.2—that the same applies to the estimates for $\sup_{I_0, J_0} \|G_A : \ell_p^{J_0} \to \ell_q^{I_0}\|$, but we omit the details to keep the presentation clear.)

Proposition 1.16 (Tail bound) *Assume that $K, L \geq 1$, $r \in (0, 2]$, $1 \leq p, q \leq \infty$, and $\gamma \geq 1$. Fix a deterministic $m \times n$ matrix A and assume that*

$$D \geq \|A \circ A : \ell_{p/2}^n \to \ell_{q/2}^m\|^{1/2}.$$

If for all random matrices $X = (X_{ij})_{i \leq m, j \leq n}$ with independent mean-zero entries satisfying

$$\mathbb{P}(|X_{ij}| \geq t) \leq K e^{-t^r/L} \quad \text{for all } t \geq 0, \ i \leq m, \ j \leq n, \tag{1.21}$$

we have

$$\mathbb{E}\|X_A : \ell_p^n \to \ell_q^m\| \leq \gamma D, \tag{1.22}$$

then, for all random matrices with independent mean-zero entries satisfying (1.21), we also have

$$\left(\mathbb{E}\|X_A : \ell_p^n \to \ell_q^m\|^\rho\right)^{1/\rho} \lesssim_{r,K,L} \rho^{1/r} \gamma D \quad \text{for all } \rho \geq 1, \tag{1.23}$$

and, for all $t > 0$,

$$\mathbb{P}\big(\|X_A : \ell_p^n \to \ell_q^m\| \geq t\gamma D\big) \leq C(r, K, L) \exp\big(-t^r/C(r, K, L)\big). \tag{1.24}$$

Note that random variables taking values in $[-1, 1]$ satisfy condition (1.21) with $r = 2$, $K = e$, and $L = 1$. Thus, Proposition 1.16 applies also in the setting of bounded or Gaussian entries.

1.6 Organization of the paper

In Sect. 2 we gather various preliminary results we shall use in the sequel. Section 3 contains the proofs of the main results valid for all p, q (i.e., Theorem 1.3 and its corollaries) and the tail bound from Proposition 1.16. In Sect. 4 we prove the results for specific choices/ranges of p, q. In Sect. 5 we prove lower bounds on expected operator norms, showing in particular that our estimates are optimal up to logarithmic factors. We also prove other results justifying the proposed form of Conjecture 1. The last subsection of Sect. 5 is devoted to infinite dimensional Gaussian operators.

2 Preliminaries

2.1 General facts

We start with some easy lemmas which will be used repeatedly throughout the paper.

Lemma 2.1 *For any real $m \times n$ matrix $B = (b_{ij})_{i \leq m, j \leq n}$ and $0 < r \leq 1 \leq s \leq \infty$, we have*

$$\|B \colon \ell_r^n \to \ell_s^m\| = \|B \colon \ell_1^n \to \ell_s^m\| = \max_{j \leq n} \|(b_{ij})_{i=1}^m\|_s.$$

Furthermore, for a real $m \times n$ matrix $A = (a_{ij})_{i \leq m, j \leq n}$ and $1 \leq p \leq 2$, $p \leq q \leq \infty$,

$$\|A \circ A \colon \ell_{p/2}^n \to \ell_{q/2}^m\|^{1/2} = \max_{j \leq n} \|(a_{ij})_{i=1}^m\|_q.$$

Proof. Since $0 < r \leq 1$, we have conv $B_r^n = B_1^n$, where conv S denotes the convex hull of the set S. Moreover, the extreme points of B_1^n are the signed standard unit vectors, i.e., $\pm e_1, \ldots, \pm e_n$, and $z \mapsto \|z\|_s$ is a convex function (since $s \geq 1$). Thus,

$$\sup_{x \in B_r^n} \|Bx\|_s = \sup_{x \in \mathrm{conv}\, B_r^n} \|Bx\|_s = \sup_{x \in B_1^n} \|Bx\|_s = \max_{1 \leq j \leq n} \|Be_j\|_s = \max_{1 \leq j \leq n} \|(b_{ij})_{i=1}^m\|_s.$$

This immediately implies the result for the Hadamard product $A \circ A =: B$ if $1 \leq p \leq 2 \leq q \leq \infty$.

If, on the other hand, $1 \leq p \leq q \leq 2$, then by the subadditivity of the function $t \mapsto |t|^{q/2}$,

$$\|A \circ A \colon \ell_{p/2}^n \to \ell_{q/2}^m\|^{q/2} = \sup_{x \in B_{p/2}^n} \sum_{i=1}^m \left| \sum_{j=1}^n a_{ij}^2 x_j \right|^{q/2} \leq \sup_{x \in B_{p/2}^n} \sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^q |x_j|^{q/2}$$

$$= \|(|a_{ij}|^q)_{i \leq m, j \leq n} \colon \ell_{p/q}^n \to \ell_1^m\| = \max_{j \leq n} \|(a_{ij})_{i \leq m}\|_q^q,$$

where in the last equality we used the first part of the Lemma. Since we clearly have

$$\|A \circ A \colon \ell_{p/2}^n \to \ell_{q/2}^m\| \geq \max_{j \leq n} \|(a_{ij}^2)_{i \leq m}\|_{q/2} = \max_{j \leq n} \|(a_{ij})_{i \leq m}\|_q^2,$$

we thus obtain

$$\|A \circ A \colon \ell_{p/2}^n \to \ell_{q/2}^m\|^{1/2} = \max_{j \leq n} \|(a_{ij})_{i \leq m}\|_q. \quad \square$$

Definition 2.2 A set $K \subset \mathbb{R}^n$ is called unconditional, if for every $(x_j)_{j \leq n} \in K$ and every $(\varepsilon_j)_{j \leq n} \in \{-1, 1\}^n$ we have $(\varepsilon_j x_j)_{j \leq n} \in K$.

We shall use the following version of [49, Lemma 2.1].

Lemma 2.3 *Assume that $1 \leq p \leq \infty$, $n \in \mathbb{N}$, and define the convex set*

$$K := \mathrm{conv}\left\{ \frac{1}{|J|^{1/p}} \left(\varepsilon_j \mathbf{1}_{\{j \in J\}} \right)_{j=1}^n : J \subset \{1, \ldots, n\}, J \neq \emptyset, (\varepsilon_j)_{j=1}^n \in \{-1, 1\}^n \right\}.$$

Then $B_p^n \subset \ln(en)^{1/p^*} K$.

Proof Fix a vector $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$. We want to prove that $\|x\|_K \leq \ln(en)^{1/p^*} \|x\|_p$, where

$$\|x\|_K = \inf\{\lambda > 0 \colon x \in \lambda K\}$$

denotes the norm generated by K, i.e., its Minkowski gauge. Since both K and B_p^n are permutationally invariant and unconditional (see Definition 2.2), we may and will assume that $x_1 \geq \cdots \geq x_n \geq 0$. If we put $x_{n+1} := 0$, then

$$x = \sum_{j=1}^n x_j e_j = \sum_{j=1}^n (x_j - x_{j+1})(e_1 + \cdots + e_j).$$

Since $\|e_1 + \cdots + e_j\|_K = j^{1/p}$ for $1 \leq j \leq n$,[3] the triangle and Hölder inequalities yield

$$\|x\|_K \leq \sum_{j=1}^n (x_j - x_{j+1}) j^{1/p} = \sum_{j=1}^n x_j (j^{1/p} - (j-1)^{1/p})$$

$$\leq \sum_{j=1}^n x_j j^{1/p-1} \leq \|x\|_p \left(\sum_{j=1}^n \frac{1}{j} \right)^{1/p^*} \leq \|x\|_p \ln(en)^{1/p^*},$$

where we also used the elementary estimates $j^{1/p} - (j-1)^{1/p} \leq j^{\frac{1}{p}-1}$ and $\sum_{j=1}^n \frac{1}{j} \leq 1 + \int_1^n \frac{1}{t} dt = \ln(en)$. This completes the proof. $\quad \square$

---

[3] Indeed, $j^{-1/p}(e_1 + \cdots + e_j) \in K$, so $\|e_1 + \cdots + e_j\|_K \leq j^{1/p}$; on the other hand, $K \subset B_p^n$, so $\|e_1 + \cdots + e_j\|_K \geq \|e_1 + \cdots + e_j\|_p = j^{1/p}$.

Remark 2.4 The term $\ln(en)^{1/p^*}$ can be replaced by $1 + \frac{1}{p}\ln(en)^{1/p^*}$ by writing in the above proof

$$\sum_{j=1}^{n} x_j(j^{1/p} - (j-1)^{1/p}) \le x_1 + \frac{1}{p}\sum_{j=2}^{n} x_j(j-1)^{\frac{1}{p}-1} \le \|x\|_p\left(1 + \frac{1}{p}\left(\sum_{j=1}^{n-1}\frac{1}{j}\right)^{1/p^*}\right).$$

Here we used the estimates $j^{1/p} - (j-1)^{1/p} \le \frac{1}{p}(j-1)^{\frac{1}{p}-1}$ for $j > 1$ (which follows from the concavity of the function $t \mapsto t^{1/p}$) and the trivial one $x_1 \le \|x\|_p$.

Remark 2.5 The constant $(\ln n)^{1/p^*}$ in Lemma 2.3 is sharp up to a constant depending on p for every $1 \le p < \infty$ (when $p = \infty$, $K = B_p^n$ and the constant depending on p degenerates as $p \to \infty$). More precisely, we shall prove that if $B_p^n \subset C(p,n)K$, then $C(p,n) \gtrsim_p (\ln n)^{1/p^*}$. Note that $B_p^n \subset C(p,n)K$ if and only if

$$\|\cdot\|_{p^*} \le C(p,n)\|\cdot\|_K^*, \tag{2.1}$$

where $\|\cdot\|_K^*$ is norm dual to $\|\cdot\|_K$.

Let Ext K be the set of extreme points of K, and let $(y_j^\downarrow)_{j \le n}$ be the non-increasing rearrangement of $(|y_j|)_{j \le n}$. For every $y \in \mathbb{R}^n$,

$$\|y\|_K^* = \sup_{x \in K}\sum_{j=1}^{n} x_j y_j = \sup_{x \in \text{Ext }K}\sum_{j=1}^{n} x_j y_j = \sup_{J \subset [n], J \ne \emptyset}\sum_{j \in J}|y_j|\frac{1}{|J|^{1/p}}$$

$$= \sup_{k \le n}\sum_{j=1}^{k} y_j^\downarrow \frac{1}{k^{1/p}}.$$

Assume that $p^* \ne 1$ and put $y_j := j^{-1/p^*}$. We get

$$\|y\|_K^* = \sup_{k \le n}\sum_{j=1}^{k} j^{-1/p^*}\frac{1}{k^{1/p}} \asymp_p \sup_{k \le n} k^{1-\frac{1}{p^*}}\frac{1}{k^{1/p}} = 1,$$

whereas

$$\|y\|_{p^*} = \left(\sum_{j=1}^{n} j^{-1}\right)^{1/p^*} \asymp (\ln n)^{1/p^*},$$

so inequality (2.1) yields that $C(p,n) \gtrsim_p (\ln n)^{1/p^*}$.

We shall also need the following standard lemma (see, e.g., [41, Sect. 1.3]). We will use the versions with $r = 1$ and $r = 2$.

Lemma 2.6 *Let Z be a nonnegative random variable. If there exist $a \geq 0$, b, α, β, $s_0 > 0$, and $r \geq 1$ such that*

$$\mathbb{P}(Z \geq a + bs) \leq \alpha e^{-\beta s^r} \quad \text{for } s \geq s_0,$$

then

$$\mathbb{E}Z \leq a + b\left(s_0 + \alpha \frac{e^{-\beta s_0^r}}{r\beta s_0^{r-1}}\right).$$

Proof. Integration by parts yields

$$\mathbb{E}Z \leq a + bs_0 + \int_{a+bs_0}^{\infty} \mathbb{P}(Z \geq u)du = a + bs_0 + b\int_{s_0}^{\infty} \mathbb{P}(Z \geq a + bs)ds$$

$$\leq a + bs_0 + b\alpha \int_{s_0}^{\infty} e^{-\beta s^r} ds$$

$$\leq a + bs_0 + \frac{b\alpha}{r\beta s_0^{r-1}} \int_{s_0}^{\infty} r\beta s^{r-1} e^{-\beta s^r} ds = a + b\left(s_0 + \alpha \frac{e^{-\beta s_0^r}}{r\beta s_0^{r-1}}\right). \qquad \square$$

2.2 Contraction principles

Below we recall the well-known contraction principle due to Kahane and its extension by Talagrand (see, e.g., [64, Exercise 6.7.7] and [43, Theorem 4.4 and the proof of Theorem 4.12]).

Lemma 2.7 (Contraction principle) *Let $(X, \|\cdot\|)$ be a normed space, $n \in \mathbb{N}$, and $\rho \geq 1$. Assume that $x_1, \ldots, x_n \in X$ and $\alpha:=(\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n$. Then, if $\varepsilon_1, \ldots, \varepsilon_n$ are independent Rademacher random variables, we have*

$$\mathbb{E}\Big\|\sum_{i=1}^{n} \alpha_i \varepsilon_i x_i\Big\|^{\rho} \leq \|\alpha\|_{\infty}^{\rho} \mathbb{E}\Big\|\sum_{i=1}^{n} \varepsilon_i x_i\Big\|^{\rho}.$$

Lemma 2.8 (Contraction principle) *Let T be a bounded subset of $\mathbb{R}^n$. Assume that $\varphi_i : \mathbb{R} \to \mathbb{R}$ are 1-Lipschitz and $\varphi_i(0) = 0$ for $i = 1, \ldots, n$. Then, if $\varepsilon_1, \ldots, \varepsilon_n$ are independent Rademacher random variables, we have*

$$\mathbb{E}\sup_{t\in T} \sum_{i=1}^{n} \varepsilon_i \varphi_i(t_i) \leq \mathbb{E}\sup_{t\in T} \sum_{i=1}^{n} \varepsilon_i t_i.$$

2.3 Gaussian random variables

The following result is fundamental to the theory of Gaussian processes and referred to as Slepian's inequality or Slepian's lemma [52]. We use the following (slightly adapted) version taken from [11, Theorem 13.3].

Lemma 2.9 (Slepian's lemma) *Let $(X_t)_{t \in T}$ and $(Y_t)_{t \in T}$ be two Gaussian random vectors satisfying $\mathbb{E}[X_t] = \mathbb{E}[Y_t]$ for all $t \in T$. Assume that, for all $s, t \in T$, we have $\mathbb{E}[(X_s - X_t)^2] \le \mathbb{E}[(Y_s - Y_t)^2]$. Then*

$$\mathbb{E} \sup_{t \in T} X_t \le \mathbb{E} \sup_{t \in T} Y_t.$$

The next lemma is folklore. We include a short proof of an estimate with specific constants for the sake of completeness.

Lemma 2.10 *Assume that $k \ge 2$ and let g_i, $i \le k$, be standard Gaussian random variables (not necessarily independent). Then*

$$\mathbb{E} \max_{1 \le i \le k} g_i \le \sqrt{2 \ln k},$$

$$\mathbb{E} \max_{1 \le i \le k} |g_i| \le 2\sqrt{\ln k}.$$

Proof. Since the moment generating function of a Gaussian random variable is given by $\mathbb{E} e^{t g_1} = e^{t^2/2}$, it follows from Jensen's inequality that

$$
\begin{aligned}
\mathbb{E} \max_{i \le k} g_i &\le \frac{1}{t} \ln\left(\mathbb{E} \exp(t \max_{i \le k} g_i) \right) \\
&\le \frac{1}{t} \ln\left(\mathbb{E} \sum_{i=1}^{k} \exp(t g_i) \right) = \frac{1}{t} \ln\left(k e^{t^2/2} \right) = \frac{\ln k}{t} + \frac{t}{2}.
\end{aligned}
$$

By taking $t = \sqrt{2 \ln k}$, we get the first assertion. We apply this inequality with random variables $g_1, -g_1, \ldots, g_k, -g_k$ to get the second assertion, namely

$$\mathbb{E} \max_{i \le k} |g_i| = \mathbb{E} \max_{i \le k} \max\{g_i, -g_i\} \le \sqrt{2 \ln(2k)} \le \sqrt{2 \ln(k^2)} = 2\sqrt{\ln k}. \quad \square$$

The next two lemmas are taken from [61]. Recall that $b_1^{\downarrow} \ge \ldots \ge b_n^{\downarrow}$ is the non-increasing rearrangement of $(|b_j|)_{j \le n}$.

Lemma 2.11 ([61, Lemma 2.3]) *Assume that $(b_j)_{j \le n} \in \mathbb{R}^n$ and let $(X_j)_{j \le n}$ be random variables (not necessarily independent) satisfying*

$$\mathbb{P}(|X_j| > t) \le K e^{-t^2/b_j^2} \quad \text{for all } t \ge 0, \ j \le n.$$

Then

$$\mathbb{E} \max_{j \le n} |X_j| \lesssim_K \max_{j \le n} b_j^{\downarrow} \sqrt{\ln(j+1)}.$$

Lemma 2.12 ([61], Lemma 2.4]) *Assume that $(b_j)_{j \leq n} \in \mathbb{R}^n$ and let $(X_j)_{j \leq n}$ be independent random variables with $X_j \sim \mathcal{N}(0, b_j^2)$ for $j \leq n$. Then*

$$\mathbb{E} \max_{j \leq n} |X_j| \gtrsim \max_{j \leq n} b_j^{\downarrow} \sqrt{\ln(j+1)}.$$

Lemma 2.13 (Hoeffding's inequality, [32, Theorem 2]) *Assume that $(b_j)_{j \leq n} \in \mathbb{R}^n$ and let X_j, $j \leq n$, be independent mean-zero random variables such that $|X_j| \leq 1$ a.s. Then, for all $t \geq 0$,*

$$\mathbb{P}\Big(\Big|\sum_{j=1}^n b_j X_j\Big| \geq t\Big) \leq 2 \exp\Big(-\frac{t^2}{2\sum_{j=1}^n b_j^2}\Big).$$

2.4 Random variables with heavy tails

The following lemma is a special case of [34, Theorem 1].

Lemma 2.14 (Contraction principle) *Let $K, L > 0$ and assume that $(\eta_i)_{i \leq n}$ and $(\xi_i)_{i \leq n}$ are two sequences of independent symmetric random variables satisfying for every $i \leq n$ and $t \geq 0$,*

$$\mathbb{P}(|\eta_i| \geq t) \leq K \mathbb{P}(L|\xi_i| \geq t).$$

Then, for every convex function φ and every $a_1, \ldots, a_n \in \mathbb{R}$,

$$\mathbb{E}\varphi\Big(\sum_{i=1}^n a_i \eta_i\Big) \leq \mathbb{E}\varphi\Big(KL \sum_{i=1}^n a_i \xi_i\Big).$$

Lemma 2.15 ([31, Theorem 6.2]) *Assume that $Z_1, \ldots, Z_n$ are independent symmetric Weibull random variables with shape parameter $r \in (0, 1]$ and scale parameter 1, i.e., $\mathbb{P}(|Z_i| \geq t) = e^{-t^r}$ for $t \geq 0$. Then, for every $\rho \geq 2$ and $a \in \mathbb{R}^n$,*

$$\Big\|\sum_{i=1}^n a_i Z_i\Big\|_{\rho} \asymp \max\big\{\sqrt{\rho}\|a\|_2 \|Z_1\|_2, \|a\|_{\rho}\|Z_1\|_{\rho}\big\}.$$

Remark 2.16 (Moments of Weibull random variables) Note that if Z is a symmetric random variable such that $\mathbb{P}(|Z| \geq t) = e^{-t^r}$, $r \in (0, 2]$, then $Y = |Z|^r \operatorname{sgn}(Z)$ has (symmetric) exponential distribution with parameter 1, so by Stirling's formula we obtain, for all $\rho \geq 1$,

$$\|Z\|_{\rho} = \|Y\|_{\rho/r}^{1/r} = \Gamma\Big(\frac{\rho}{r} + 1\Big)^{1/\rho} \leq \Big(\frac{C}{r}\Big)^{\frac{1}{r} + \frac{1}{2\rho}} \rho^{1/r} \leq \Big(\frac{C}{r}\Big)^{\frac{1}{r} + \frac{1}{2}} \rho^{1/r},$$

with $C \geq 1$.

The three previous results easily imply the following estimate for integral norms of linear combinations of independent ψ_r random variables.

Proposition 2.17 *Let $K, L > 0$, $r \in (0, 1]$ and assume that $Z_1, \ldots, Z_n$ are independent symmetric random variables satisfying $\mathbb{P}(|Z_i| \geq t) \leq K e^{-t^r/L}$ for all $t \geq 0$ and $i \leq n$. Then, for every $\rho \geq 2$ and $a \in \mathbb{R}^n$,*

$$\Big\| \sum_{i=1}^n a_i Z_i \Big\|_\rho \lesssim (C/r)^{\frac{1}{r}+\frac{1}{2}} K L^{1/r} \max\{\sqrt{\rho}\|a\|_2, \rho^{1/r}\|a\|_\rho\}$$

$$\lesssim (C'/r)^{\frac{1}{r}+\frac{1}{2}} K L^{1/r} \max\{\sqrt{\rho}\|a\|_2, \rho^{1/r}\|a\|_\infty\}.$$

Proof The first inequality is an immediate consequence of Lemma 2.14 (applied with $\eta_i = Z_i$, independent Weibull variables ξ_i with shape parameter r and scale parameter 1, and with the convex function $\varphi : t \mapsto |t|^\rho$), Lemma 2.15, and Remark 2.16. The second inequality follows from

$$\|a\|_\rho \leq \|a\|_2^{2/\rho}\|a\|_\infty^{1-2/\rho} = \rho^{\frac{2}{\rho r}}\|\rho^{-1/r}a\|_2^{2/\rho}\|a\|_\infty^{1-2/\rho}$$

$$\leq \rho^{\frac{2}{\rho r}}\Big(\frac{2}{\rho^{1+1/r}}\|a\|_2 + \Big(1-2/\rho\Big)\|a\|_\infty\Big),$$

where in the last step we used the inequality between weighted arithmetic and geometric means. $\qquad\square$

The next lemma is standard and provides us with several equivalent formulations of the ψ_r property expressed through tail bounds, growth of moments, and the exponential moments, respectively. We provide a brief proof, since in the literature one usually finds versions for $r \geq 1$ only.

Lemma 2.18 *Assume that $r \in (0, 2]$. Let Z be a non-negative random variable. The following conditions are equivalent:*

(i) *There exist $K_1, L_1 > 0$ such that*

$$\mathbb{P}(Z \geq t) \leq K_1 e^{-t^r/L_1} \quad \text{for all } t \geq 0.$$

(ii) *There exists K_2 such that*

$$\|Z\|_\rho \leq K_2\rho^{1/r} \quad \text{for all } \rho \geq 1.$$

(iii) *There exist $K_3, u > 0$ such that*

$$\mathbb{E}\exp(uZ^r) \leq K_3.$$

Here, (i) *implies* (ii) *with $K_2 = C(r)K_1 L_1^{1/r}$,* (ii) *implies* (iii) *with $K_3 = 1 + e^{(2er)^{-1}}$, $u = (2erK_2^r)^{-1}$, and* (iii) *implies* (i) *with $K_1 = K_3$, $L_1 = u^{-1}$.*

Proof. Property (i) implies (ii) by Lemma 2.14 (applied with $n = 1$, $\eta_1 = Z$ and an independent Weibull variable ξ_1 with parameter r) and Remark 2.16. Property (iii) implies (i) by Chebyshev's inequality:

$$\mathbb{P}(Z \geq t) = \mathbb{P}\big(\exp(uZ^r) \geq \exp(ut^r)\big) \leq K_3 \exp(-ut^r).$$

Assume now that (ii) holds and denote $k_0 = \lfloor \frac{1}{r} \rfloor$. Then, for every $k \in [1, k_0]$, we have $kr \leq 1$ and

$$\mathbb{E}Z^{kr} \leq (\mathbb{E}Z)^{kr} \leq K_2^{kr},$$

while for $k \geq k_0 + 1$, we have $kr \geq 1$ and, hence, property (ii) yields

$$\mathbb{E}Z^{kr} \leq K_2^{kr}(kr)^k.$$

Hence, by Stirling's formula we have for $u = (2er K_2^r)^{-1}$,

$$\begin{aligned}
\mathbb{E}\exp(uZ^r) &= 1 + \sum_{k=1}^{k_0} \frac{u^k \mathbb{E}Z^{kr}}{k!} + \sum_{k=k_0+1}^{\infty} \frac{u^k \mathbb{E}Z^{kr}}{k!} \\
&\leq 1 + \sum_{k=1}^{k_0} \frac{u^k K_2^{kr}}{k!} + \sum_{k=k_0+1}^{\infty} \frac{u^k K_2^{kr}(kr)^k}{(k/e)^k} \\
&= 1 + \sum_{k=1}^{k_0} \frac{u^k K_2^{kr}}{k!} + \sum_{k=k_0+1}^{\infty} 2^{-k} \leq e^{uK_2^r} + 1. \qquad \square
\end{aligned}$$

The next lemma states that a linear combination of independent ψ_r random variables is a ψ_r random variable.

Lemma 2.19 *Assume that $u > 0$, $r \in (0, 2]$, and let $(Z_i)_{i \leq k}$ be independent symmetric random variables satisfying $\mathbb{P}(|Z_i| \geq t) \leq K e^{-t^r/L}$ for all $t \geq 0$. Then for every $a \in \mathbb{R}^k$ the random variable $Y := \|a\|_2^{-1} \sum_{i=1}^{k} a_i Z_i$ satisfies, for all $t \geq 0$,*

$$\mathbb{P}(|Y| \geq t) \leq K' e^{-t^r/L'},$$

where K', L' depend only on K, L, and r.

Proof The case $r \geq 1$ is standard (see, e.g., [14, Theorem 1.2.5]), therefore we skip a proof in this case (however, in order to prove the lemma in the case $r \geq 1$ it suffices to use the result of Gluskin and Kwapień [19] (together with Lemma 2.14) instead of Lemma 2.15 in the proof below).

Assume that $r \in (0, 1]$ and recall that $Y = \|a\|_2^{-1} \sum_{i=1}^{k} a_i Z_i$. By Proposition 2.17,

$$\|Y\|_\rho \lesssim_{K,L,r} \max\{\sqrt{\rho}, \rho^{1/r}\} = \rho^{1/r} \qquad \text{for all } \rho \geq 1.$$

Hence, Lemma 2.18 yields the assertion. $\qquad \square$

Lemma 2.20 *Assume that $r \in (0, 2]$, $\frac{1}{s} := \frac{1}{r} - \frac{1}{2}$, Y is a non-negative random variable such that $\mathbb{P}(Y \geq t) = e^{-t^s}$ for all $t \geq 0$, and $g \sim \mathcal{N}(0, 1)$ is independent of Y. Then, for every $t \geq 0$,*

$$\mathbb{P}(|g|Y \geq t) \geq ce^{-4t^r},$$

where $c := \sqrt{2/\pi}e^{-2}$.

Proof In the case $r = 2$ we have $s = \infty$ and then $Y = 1$ almost surely and the assertion is trivial. Assume now that $r < 2$. By our assumptions $r = \frac{2s}{2+s}$. Let $x_0 := (2t^s)^{1/(2+s)}$. Note that $x \geq x_0$ is equivalent to $\frac{t^s}{x^s} \leq \frac{x^2}{2}$. Thus,

$$\mathbb{P}(|g|Y \geq t) = \mathbb{E}e^{-\frac{t^s}{|g|^s}} = \sqrt{\frac{2}{\pi}} \int_0^\infty e^{-\frac{t^s}{x^s} - \frac{x^2}{2}} dx \geq \sqrt{\frac{2}{\pi}} \int_{x_0}^{x_0+1} e^{-\frac{t^s}{x^s} - \frac{x^2}{2}} dx$$

$$\geq \sqrt{\frac{2}{\pi}} \int_{x_0}^{x_0+1} e^{-x^2} dx \geq \sqrt{\frac{2}{\pi}} e^{-(x_0+1)^2} \geq \sqrt{\frac{2}{\pi}} e^{-2(x_0^2+1)}$$

$$= ce^{-2x_0^2} \geq ce^{-4t^{2s/(2+s)}} = ce^{-4t^r},$$

where we used $2^{2/(2+s)} \leq 2$ and chose $c := \sqrt{2/\pi}e^{-2}$. □

Lemma 2.21 *Assume that $K, L > 0$, $r \in (0, 2]$ and that Z is a random variable satisfying $\mathbb{P}(|Z| \geq t) \leq Ke^{-t^r/L}$ for all $t \geq 0$. Let Y, g, and $c = \sqrt{2/\pi}e^{-2}$ be as in Lemma 2.20. Then there exist random variables $U \sim |Z|$ and $V \sim |g|Y$ such that*

$$U \leq (8L)^{1/r}\left(\left(\frac{\ln(K/c)}{4}\right)^{1/r} + V\right) \quad a.s.$$

Proof For $t = 0$ we have $1 = \mathbb{P}(|Z| \geq 0) \leq K$, so $K \geq 1$, and thus $\ln(K/c) = \ln(Ke^2\sqrt{\pi/2}) > 0$. We use our assumptions, the inequality $(a + b)^r \geq (a^r + b^r)/2$, and Lemma 2.20 to obtain for any $t \geq 0$,

$$\mathbb{P}\left((8L)^{-1/r}|Z| \geq t + \left(\ln(K/c)/4\right)^{1/r}\right) \leq K\exp\left(-8\left[t + \left(\ln(K/c)/4\right)^{1/r}\right]^r\right)$$

$$\leq K\exp\left(-4(t^r + \ln(K/c)/4)\right) = ce^{-4t^r}$$

$$\leq \mathbb{P}(|g|Y \geq t).$$

Consider the version U of $|Z|$ and the version V of $|g|Y$ defined on the (common) probability space $(0, 1)$ equipped with Lebesgue measure, constructed as the (generalised) inverses of cumulative distribution functions of $|Z|$ and $|g|Y$, respectively. Then $(8L)^{-1/r}U - \left(\ln(K/c)/4\right)^{1/r} \leq V$, which implies the assertion. □

Lemma 2.22 *Let $K, L > 0$, $r \in (0, 2]$ and $k \geq 3$, and assume that $(Z_i)_{i \leq k}$, are random variables satisfying $\mathbb{P}(|Z_i| \geq t) \leq Ke^{-t^r/L}$ for all $t \geq 0$. Then*

$$\mathbb{P}\left(\max_{i \leq k} |Z_i| \geq (vL \ln k)^{1/r}\right) \leq Kk^{-v+1} \leq eKe^{-v} \quad \text{for every } v \geq 1$$

and

$$\mathbb{E} \max_{i \leq k} |Z_i| \lesssim \left(LK^r r^{-1} \ln k\right)^{1/r} \lesssim_{r,K,L} (\ln k)^{1/r}.$$

Proof. By a union bound and the assumptions we get, for every $v \geq 1$,

$$\mathbb{P}\left(\max_{i \leq k} |Z_i| \geq (vL \ln k)^{1/r}\right) \leq \sum_{i=1}^{k} \mathbb{P}\left(|Z_i| \geq (vL \ln k)^{1/r}\right) \leq k \cdot K e^{-v \ln k}$$

$$= K e^{-(v-1) \ln k} = Kk^{-v+1} \leq eK e^{-v},$$

where we used $k \geq 3$ in the last step. We integrate by parts, change the variables, and use the above bound to obtain the second part of the assertion, i.e.,

$$\mathbb{E} \max_{i \leq k} |Z_i| = \int_0^\infty \mathbb{P}\left(\max_{i \leq k} |Z_i| \geq u\right) du \leq (L \ln k)^{1/r} + \int_{(L \ln k)^{1/r}}^\infty \mathbb{P}\left(\max_{i \leq k} |Z_i| \geq u\right) du$$

$$= (L \ln k)^{1/r} + \frac{(L \ln k)^{1/r}}{r} \int_1^\infty v^{\frac{1}{r}-1} \mathbb{P}\left(\max_{i \leq k} |Z_i| \geq (vL \ln k)^{1/r}\right) dv$$

$$\leq (L \ln k)^{1/r} \left(1 + \frac{eK}{r} \int_1^\infty v^{\frac{1}{r}-1} e^{-v} dv\right)$$

$$\leq (L \ln k)^{1/r} \left(1 + eK \, \Gamma\left(\frac{1}{r}+1\right)\right). \qquad \square$$

3 Proofs of the main results

After the preparation in the previous section, we shall now present the proofs of our main results.

3.1 General bound via Slepian's lemma

In order to obtain Theorem 1.3 we first prove its weaker version, for $p = \infty$ and $q = 1$ only. After that we shall use the polytope K from Lemma 2.3 and the Gaussian concentration to see how Proposition 3.1 implies the general bound. The proof of this proposition relies on the symmetrization together with the contraction principle, which allow us to get rid of y_i and x_j, and make use of Slepian's lemma.

Proposition 3.1 *Assume that* $G = (g_{ij})_{i \leq m, j \leq n}$ *has i.i.d. standard Gaussian entries and* $k \leq m, l \leq n$. *Then*

$$\mathbb{E} \sup_{I,J} \sup_{y \in B_\infty^m} \sup_{x \in B_\infty^n} \sum_{i \in I, j \in J} y_i a_{ij} g_{ij} x_j \leq \left(8\sqrt{\ln m} + \sqrt{2/\pi}\right) \sup_{I,J} \sum_{i \in I} \sqrt{\sum_{j \in J} a_{ij}^2}$$

$$+ \left(8\sqrt{\ln n} + 2\sqrt{2/\pi}\right) \sup_{I,J} \sum_{j \in J} \sqrt{\sum_{i \in I} a_{ij}^2},$$

where the suprema are taken over all sets $I \subset \{1, \ldots, m\}$, $J \subset \{1, \ldots, n\}$ such that $|I| = k$, $|J| = l$.

Proof Throughout the proof, $k \leq m$ and $l \leq n$ are fixed and the suprema are taken over all index sets satisfying $I \subset \{1, \ldots, m\}$, $|I| = k$ and $J \subset \{1, \ldots, n\}$, $|J| = l$.

Let us denote by $(\widetilde{g}_{ij})_{i \leq m, j \leq n}$ an independent copy of $(g_{ij})_{i \leq m, j \leq n}$. Using the duality $(\ell_1^m)^* = \ell_\infty^m$, centering the expression, noticing that $\sum_{j \in J} a_{ij} \widetilde{g}_{ij} x_j$ is a Gaussian $\sqrt{\sum_{j \in J} a_{ij}^2 x_j^2}$, and using Jensen's inequality, we see that

$$
\begin{aligned}
\mathbb{E} \sup_{I,J} \sup_{x \in B_\infty^n} \sup_{y \in B_\infty^m} \sum_{i \in I, j \in J} y_i a_{ij} g_{ij} x_j &= \mathbb{E} \sup_{I,J} \sup_{x \in B_\infty^n} \sum_{i \in I} \left| \sum_{j \in J} a_{ij} g_{ij} x_j \right| \\
&\leq \mathbb{E} \sup_{I,J} \sup_{x \in B_\infty^n} \sum_{i \in I} \left(\left| \sum_{j \in J} a_{ij} g_{ij} x_j \right| - \mathbb{E} \left| \sum_{j \in J} a_{ij} \widetilde{g}_{ij} x_j \right| \right) \\
&\quad + \sup_{I,J} \sup_{x \in B_\infty^n} \sum_{i \in I} \mathbb{E} \left| \sum_{j \in J} a_{ij} \widetilde{g}_{ij} x_j \right| \\
&= \mathbb{E} \sup_{I,J} \sup_{x \in B_\infty^n} \sum_{i \in I} \left(\left| \sum_{j \in J} a_{ij} g_{ij} x_j \right| - \mathbb{E} \left| \sum_{j \in J} a_{ij} \widetilde{g}_{ij} x_j \right| \right) \\
&\quad + \sup_{I,J} \sup_{x \in B_\infty^n} \sum_{i \in I} \sqrt{\sum_{j \in J} a_{ij}^2 x_j^2} \, \mathbb{E}|g| \\
&\leq \mathbb{E} \sup_{I,J} \sup_{x \in B_\infty^n} \sum_{i \in I} \left(\left| \sum_{j \in J} a_{ij} g_{ij} x_j \right| - \left| \sum_{j \in J} a_{ij} \widetilde{g}_{ij} x_j \right| \right) + \sqrt{\frac{2}{\pi}} \sup_{I,J} \sum_{i \in I} \sqrt{\sum_{j \in J} a_{ij}^2}.
\end{aligned}
$$
(3.1)

To estimate the expected value on the right-hand side, we use a symmetrization trick together with the contraction principle (Lemma 2.8). Let $(\varepsilon_i)_{i \leq m}$ be a sequence of independent Rademacher random variables independent of all others. Since the random vectors

$$
Z_i = \left(\mathbf{1}_{\{i \in I\}} \left(\left| \sum_{j \in J} a_{ij} g_{ij} x_j \right| - \left| \sum_{j \in J} a_{ij} \widetilde{g}_{ij} x_j \right| \right) \right)_{I \subset [m], J \subset [n], x \in B_\infty^n}
$$

(where $i \leq m$) are independent and symmetric, $(Z_i)_{i \leq m}$ has the same distribution as $(\varepsilon_i Z_i)_{i \leq m}$. Therefore,

$$
\begin{aligned}
\mathbb{E} \sup_{I,J} \sup_{x \in B_\infty^n} \sum_{i \in I} &\left(\left| \sum_{j \in J} a_{ij} g_{ij} x_j \right| - \left| \sum_{j \in J} a_{ij} \widetilde{g}_{ij} x_j \right| \right) \\
&= \mathbb{E} \sup_{I,J} \sup_{x \in B_\infty^n} \sum_{i \in I} \varepsilon_i \left(\left| \sum_{j \in J} a_{ij} g_{ij} x_j \right| - \left| \sum_{j \in J} a_{ij} \widetilde{g}_{ij} x_j \right| \right) \\
&\leq 2 \mathbb{E} \sup_{I,J} \sup_{x \in B_\infty^n} \sum_{i \in I} \varepsilon_i \left| \sum_{j \in J} a_{ij} g_{ij} x_j \right|
\end{aligned}
$$

$$= 2\mathbb{E} \sup_{I,J} \sup_{x \in B_\infty^n} \sum_{i=1}^m \varepsilon_i \left| \sum_{j \in J} a_{ij} g_{ij} x_j \mathbf{1}_{\{i \in I\}} \right|. \tag{3.2}$$

Applying (conditionally, with the values of g_{ij}'s fixed) the contraction principle (i.e., Lemma 2.8) with the set

$$T = \left\{ \left(\sum_{j \in J} a_{ij} g_{ij} x_j \mathbf{1}_{\{i \in I\}} \right)_{i \le m} : I \subset [m], |I| = k, J \subset [n], |J| = l, x \in B_\infty^n \right\}$$

and the function $u \mapsto |u|$ (which is 1-Lipschitz and takes the value 0 at the origin), we get

$$\mathbb{E} \sup_{I,J} \sup_{x \in B_\infty^n} \sum_{i=1}^m \varepsilon_i \left| \sum_{j \in J} a_{ij} g_{ij} x_j \mathbf{1}_{\{i \in I\}} \right| \le \mathbb{E} \sup_{I,J} \sup_{x \in B_\infty^n} \sum_{i=1}^m \varepsilon_i \sum_{j \in J} a_{ij} g_{ij} x_j \mathbf{1}_{\{i \in I\}}$$

$$= \mathbb{E} \sup_{I,J} \sup_{x \in B_\infty^n} \sum_{j \in J} \sum_{i \in I} a_{ij} \varepsilon_i g_{ij} x_j = \mathbb{E} \sup_{I,J} \sup_{x \in B_\infty^n} \sum_{j \in J} \sum_{i \in I} a_{ij} g_{ij} x_j. \tag{3.3}$$

By proceeding similarly as in (3.1), we obtain

$$\mathbb{E} \sup_{I,J} \sup_{x \in B_\infty^n} \sum_{j \in J} \sum_{i \in I} a_{ij} g_{ij} x_j = \mathbb{E} \sup_{I,J} \sum_{j \in J} \left| \sum_{i \in I} a_{ij} g_{ij} \right|$$

$$\le \mathbb{E} \sup_{I,J} \sum_{j \in J} \left(\left| \sum_{i \in I} a_{ij} g_{ij} \right| - \mathbb{E} \left| \sum_{i \in I} a_{ij} \widetilde{g}_{ij} \right| \right) + \sqrt{\frac{2}{\pi}} \sup_{I,J} \sum_{j \in J} \sqrt{\sum_{i \in I} a_{ij}^2}. \tag{3.4}$$

Observe that using symmetrization and the contraction principle similarly as in (3.2) and (3.3), we can estimate the first summand on right-hand side of (3.4) as follows,

$$\mathbb{E} \sup_{I,J} \sum_{j \in J} \left(\left| \sum_{i \in I} a_{ij} g_{ij} \right| - \mathbb{E} \left| \sum_{i \in I} a_{ij} \widetilde{g}_{ij} \right| \right) \le 2\mathbb{E} \sup_{I,J} \sum_{i \in I} \sum_{j \in J} a_{ij} g_{ij}. \tag{3.5}$$

Altogether, the inequalities in (3.1) – (3.5) yield that

$$\mathbb{E} \sup_{I,J} \sup_{y \in B_\infty^m} \sup_{x \in B_\infty^n} \sum_{i \in I, j \in J} y_i a_{ij} g_{ij} x_j \le 4\mathbb{E} \sup_{I,J} \sum_{i \in I} \sum_{j \in J} a_{ij} g_{ij} + 2\sqrt{\frac{2}{\pi}} \sup_{I,J} \sum_{j \in J} \sqrt{\sum_{i \in I} a_{ij}^2}$$

$$+ \sqrt{\frac{2}{\pi}} \sup_{I,J} \sum_{i \in I} \sqrt{\sum_{j \in J} a_{ij}^2}. \tag{3.6}$$

We shall now estimate the first summand on the right-hand side of (3.6) using Slepian's lemma (i.e., Lemma 2.9). Denote

$$X_{I,J} := \sum_{i \in I} \sum_{j \in J} a_{ij} g_{ij},$$

$$Y_{I,J} := \sum_{i \in I} g_i \sqrt{\sum_{j \in J} a_{ij}^2} + \sum_{j \in J} \widetilde{g}_j \sqrt{\sum_{i \in I} a_{ij}^2},$$

where $g_i, i = 1, \ldots, m, \widetilde{g}_j, j = 1, \ldots, n$ are independent standard Gaussian variables. The random variables $X_{I,J}, Y_{I,J}$ clearly have zero mean. Thus, we only need to calculate and compare $\mathbb{E}(X_{I,J} - X_{\widetilde{I},\widetilde{J}})^2$ and $\mathbb{E}(Y_{I,J} - Y_{\widetilde{I},\widetilde{J}})^2$. In the calculations below it will be evident over which sets the index i (resp. j) runs, so in order to shorten the notation and improve readability, we use the notational convention

$$\sum_I := \sum_{i \in I}, \quad \sum_{\widetilde{J}} := \sum_{j \in \widetilde{J}}, \quad \sum_{I \cap \widetilde{I}, J \setminus \widetilde{J}} := \sum_{i \in I \cap \widetilde{I}, j \in J \setminus \widetilde{J}}, \quad \text{etc.}$$

By independence,

$$\mathbb{E}(X_{I,J} - X_{\widetilde{I},\widetilde{J}})^2 = \sum_{I,J} a_{ij}^2 + \sum_{\widetilde{I},\widetilde{J}} a_{ij}^2 - 2 \sum_{I \cap \widetilde{I}, J \cap \widetilde{J}} a_{ij}^2$$

$$= \sum_{I,J} a_{ij}^2 + \sum_{\widetilde{I},\widetilde{J}} a_{ij}^2 - \sum_{I \cap \widetilde{I}, J} a_{ij}^2 - \sum_{I \cap \widetilde{I}, \widetilde{J}} a_{ij}^2 + \sum_{I \cap \widetilde{I}, J \setminus \widetilde{J}} a_{ij}^2 + \sum_{I \cap \widetilde{I}, \widetilde{J} \setminus J} a_{ij}^2.$$

By independence and the inequality $2\sqrt{ab} \leq a + b$ (valid for $a, b \geq 0$),

$$\mathbb{E}(Y_{I,J} - Y_{\widetilde{I},\widetilde{J}})^2 = 2 \sum_{I,J} a_{ij}^2 + 2 \sum_{\widetilde{I},\widetilde{J}} a_{ij}^2$$

$$- 2 \sum_{I \cap \widetilde{I}} \sqrt{\sum_J a_{ij}^2} \sqrt{\sum_{\widetilde{J}} a_{ij}^2} - 2 \sum_{J \cap \widetilde{J}} \sqrt{\sum_I a_{ij}^2} \sqrt{\sum_{\widetilde{I}} a_{ij}^2}$$

$$\geq 2 \sum_{I,J} a_{ij}^2 + 2 \sum_{\widetilde{I},\widetilde{J}} a_{ij}^2 - \sum_{I \cap \widetilde{I}, J} a_{ij}^2 - \sum_{I \cap \widetilde{I}, \widetilde{J}} a_{ij}^2 - \sum_{I, J \cap \widetilde{J}} a_{ij}^2 - \sum_{\widetilde{I}, J \cap \widetilde{J}} a_{ij}^2$$

$$= \sum_{I,J} a_{ij}^2 + \sum_{\widetilde{I},\widetilde{J}} a_{ij}^2 - \sum_{I \cap \widetilde{I}, J} a_{ij}^2 - \sum_{I \cap \widetilde{I}, \widetilde{J}} a_{ij}^2 + \sum_{I, J \setminus \widetilde{J}} a_{ij}^2 + \sum_{\widetilde{I}, \widetilde{J} \setminus J} a_{ij}^2.$$

Thus, we clearly have

$$\mathbb{E}(X_{I,J} - X_{\widetilde{I},\widetilde{J}})^2 \leq \mathbb{E}(Y_{I,J} - Y_{\widetilde{I},\widetilde{J}})^2$$

(cf. Remark 3.2 below). Hence, by Slepian's lemma (Lemma 2.9) and Lemma 2.10 on the expected maxima of standard Gaussian random variables,

$$\mathbb{E} \sup_{I,J} \sum_{i \in I} \sum_{j \in J} a_{ij} g_{ij} \leq \mathbb{E} \sup_{I,J} \left[\sum_{i \in I} g_i \sqrt{\sum_{j \in J} a_{ij}^2} + \sum_{j \in J} \widetilde{g}_j \sqrt{\sum_{i \in I} a_{ij}^2} \right]$$

$$\leq \mathbb{E} \sup_{I,J} \sum_{i \in I} g_i \sqrt{\sum_{j \in J} a_{ij}^2} + \mathbb{E} \sup_{I,J} \sum_{j \in J} \widetilde{g}_j \sqrt{\sum_{i \in I} a_{ij}^2}$$

$$\leq \mathbb{E} \sup_{i \leq m} |g_i| \sup_{I,J} \sum_{i \in I} \sqrt{\sum_{j \in J} a_{ij}^2} + \mathbb{E} \sup_{j \leq n} |\widetilde{g}_j| \sup_{I,J} \sum_{j \in J} \sqrt{\sum_{i \in I} a_{ij}^2}$$

$$\leq 2\sqrt{\ln m} \sup_{I,J} \sum_{i \in I} \sqrt{\sum_{j \in J} a_{ij}^2} + 2\sqrt{\ln n} \sup_{I,J} \sum_{j \in J} \sqrt{\sum_{i \in I} a_{ij}^2}.$$

Recalling the estimate (3.6), we arrive at

$$\mathbb{E} \sup_{I,J} \sup_{y \in B_\infty^m} \sup_{x \in B_\infty^n} \sum_{i \in I, j \in J} y_i a_{ij} g_{ij} x_j \leq \left(8\sqrt{\ln m} + \sqrt{2/\pi}\right) \sup_{I,J} \sum_{i \in I} \sqrt{\sum_{j \in J} a_{ij}^2}$$

$$+ \left(8\sqrt{\ln n} + 2\sqrt{2/\pi}\right) \sup_{I,J} \sum_{j \in J} \sqrt{\sum_{i \in I} a_{ij}^2},$$

which completes the proof of Proposition 3.1. $\qquad \square$

Remark 3.2 In the above proof, we also have

$$\mathbb{E}(X_{I,J} - X_{\widetilde{I},\widetilde{J}})^2 = \sum_{I,J} a_{ij}^2 + \sum_{\widetilde{I},\widetilde{J}} a_{ij}^2 - \sum_{I \cap \widetilde{I}, J \cap \widetilde{J}} a_{ij}^2 - \sum_{J \cap \widetilde{J}, I \cap \widetilde{I}} a_{ij}^2$$

$$\geq \sum_{I,J} a_{ij}^2 + \sum_{\widetilde{I},\widetilde{J}} a_{ij}^2 - \sum_{I \cap \widetilde{I}} \sqrt{\sum_J a_{ij}^2} \sqrt{\sum_{\widetilde{J}} a_{ij}^2} - \sum_{J \cap \widetilde{J}} \sqrt{\sum_I a_{ij}^2} \sqrt{\sum_{\widetilde{I}} a_{ij}^2}$$

$$= \frac{1}{2} \mathbb{E}(Y_{I,J} - Y_{\widetilde{I},\widetilde{J}})^2.$$

Therefore, by Slepian's lemma (Lemma 2.9) we may reverse the estimate from the proof as follows:

$$\mathbb{E} \sup_{I,J} \sup_{y \in B_\infty^m} \sup_{x \in B_\infty^n} \sum_{i \in I, j \in J} y_i a_{ij} g_{ij} x_j \geq \frac{1}{\sqrt{2}} \mathbb{E} \sup_{I,J} \left[\sum_{i \in I} g_i \sqrt{\sum_{j \in J} a_{ij}^2} + \sum_{j \in J} \widetilde{g}_j \sqrt{\sum_{i \in I} a_{ij}^2} \right].$$

Proof of Theorem 1.3 Recall that $\sup_{I_0, J_0}$ stands for the supremum taken over all sets $I_0 \subset [M] := \{1, \ldots, M\}$, $J_0 \subset [N] := \{1, \ldots, N\}$ with $|I_0| = m$, $|J_0| = n$. Given such sets I_0, J_0, we introduce the sets

$$K = K(I_0) := \text{conv}\left\{ \frac{1}{|I|^{1/q^*}} \left(\varepsilon_i \mathbf{1}_{\{i \in I\}}\right)_{i \in I_0} : I \subset I_0, I \neq \emptyset, (\varepsilon_i)_{i \in I_0} \in \{-1, 1\}^{I_0} \right\},$$

$$L = L(J_0) := \text{conv}\left\{ \frac{1}{|J|^{1/p}} \left(\eta_j \mathbf{1}_{\{j \in J\}}\right)_{j \in J_0} : J \subset J_0, J \neq \emptyset, (\eta_j)_{j \in J_0} \in \{-1, 1\}^{J_0} \right\}.$$

Then, by Lemma 2.3, $B_{q^*}^{I_0} \subset \ln(em)^{1/q} K$ and $B_p^{J_0} \subset \ln(en)^{1/p^*} L$. Therefore,

$$\mathbb{E} \sup_{I_0, J_0} \sup_{x \in B_p^{J_0}} \sup_{y \in B_{q^*}^{I_0}} \sum_{i \in I_0} \sum_{j \in J_0} y_i a_{ij} g_{ij} x_j$$

$$\leq \ln(em)^{1/q} \ln(en)^{1/p^*}$$

$$\cdot \mathbb{E} \sup_{I_0, J_0} \sup_{I \subset I_0, J \subset J_0} \sup\left\{ \frac{1}{|I|^{1/q^*}|J|^{1/p}} \sum_{i \in I} \sum_{j \in J} \varepsilon_i a_{ij} g_{ij} \eta_j \;:\; \varepsilon_i, \eta_j \in \{-1, 1\} \right\}$$

$$= \ln(em)^{1/q} \ln(en)^{1/p^*}$$

$$\cdot \mathbb{E} \max_{k \leq m, l \leq n} \frac{1}{k^{1/q^*}l^{1/p}} \sup_{I \subset [M], |I| = k} \sup_{J \subset [N], |J| = l} \sup_{x \in B_\infty^N} \sup_{y \in B_\infty^M} \sum_{i \in I} \sum_{j \in J} y_i a_{ij} g_{ij} x_j$$

$$= \ln(em)^{1/q} \ln(en)^{1/p^*} \mathbb{E} \max_{k \leq m, l \leq n} Z_{k,l}, \tag{3.7}$$

where we denoted

$$Z_{k,l} := \frac{1}{k^{1/q^*}l^{1/p}} \sup_{I, J} \sup_{x \in B_\infty^N} \sup_{y \in B_\infty^M} \sum_{i \in I} \sum_{j \in J} y_i a_{ij} g_{ij} x_j,$$

with the suprema here (and later on in this proof) being always taken over all sets $I \subset [M], |I| = k$ and $J \subset [N], |J| = l$.

By Proposition 3.1, we only know that for all $k \leq m$ and $l \leq n$,

$$\mathbb{E} Z_{k,l} \leq \left(8\sqrt{\ln M} + \sqrt{2/\pi} \right) \frac{1}{k^{1/q^*}l^{1/p}} \sup_{I, J} \sum_{i \in I} \sqrt{\sum_{j \in J} a_{ij}^2}$$

$$+ \left(8\sqrt{\ln N} + 2\sqrt{2/\pi} \right) \frac{1}{k^{1/q^*}l^{1/p}} \sup_{I, J} \sum_{j \in J} \sqrt{\sum_{i \in I} a_{ij}^2}, \tag{3.8}$$

but we shall use the Gaussian concentration and the union bound to obtain an estimate for $\mathbb{E} \max_{k \leq m, l \leq n} Z_{k,l}$.

Note first that $(k^{-1/q^*} \mathbf{1}_{\{i \in I\}})_{i \in I_0} \in K(I_0) \subset B_{q^*}^{I_0}$ and $(l^{-1/p} \mathbf{1}_{\{j \in J\}})_{j \in J_0} \in L(J_0) \subset B_p^{J_0}$, provided that $|I| = k, |J| = l, I \subset I_0, J \subset J_0$. Therefore,

$$\frac{1}{k^{1/q^*}l^{1/p}} \sup_{I, J} \sum_{i \in I} \sqrt{\sum_{j \in J} a_{ij}^2} \leq \sup_{I_0, J_0} \sup_{x \in B_p^{J_0}} \sup_{y \in B_{q^*}^{I_0}} \sum_{i \in I_0} y_i \sqrt{\sum_{j \in J_0} a_{ij}^2 x_j^2}$$

$$= \sup_{I_0, J_0} \sup_{z \in B_{p/2}^{J_0}} \left(\sum_{i \in I_0} \left(\sum_{j \in J_0} a_{ij}^2 z_j \right)^{q/2} \right)^{1/q}$$

$$= \sup_{I_0, J_0} \| A \circ A \colon \ell_{p/2}^{J_0} \to \ell_{q/2}^{I_0} \|^{1/2}$$

and, similarly,

$$\frac{1}{k^{1/q^*}l^{1/p}} \sup_{I, J} \sum_{j \in J} \sqrt{\sum_{i \in I} a_{ij}^2} \leq \sup_{I_0, J_0} \| (A \circ A)^T \colon \ell_{q^*/2}^{I_0} \to \ell_{p^*/2}^{J_0} \|^{1/2}.$$

This together with the estimate in (3.8) gives

$$\mathbb{E} Z_{k,l} \leq \left(8\sqrt{\ln M} + \sqrt{2/\pi}\right) \sup_{I_0, J_0} \| A \circ A : \ell_{p/2}^{J_0} \to \ell_{q/2}^{I_0} \|^{1/2}$$
$$+ \left(8\sqrt{\ln N} + 2\sqrt{2/\pi}\right) \sup_{I_0, J_0} \| (A \circ A)^T : \ell_{q^*/2}^{I_0} \to \ell_{p^*/2}^{J_0} \|^{1/2}. \qquad (3.9)$$

Note that by the Cauchy–Schwarz inequality, the function

$$z \mapsto \frac{1}{k^{1/q^*} l^{1/p}} \sup_{I,J} \sup_{x \in B_\infty^N} \sup_{y \in B_\infty^M} \sum_{i \in I} \sum_{j \in J} y_i a_{ij} z_{ij} x_j$$

is D-Lipschitz with

$$D \leq \frac{1}{k^{1/q^*} l^{1/p}} \sup_{I,J} \sqrt{\sum_{j \in J} \sum_{i \in I} a_{ij}^2} \leq \sup_{I,J} \sqrt{\sup_{x \in B_{p/2}^N} \sup_{y \in B_{q^*/2}^M} \sum_{i \in I} \sum_{j \in J} y_i a_{ij}^2 x_j}$$

$$\leq \sup_{I_0, J_0} \sqrt{\sup_{x \in B_{p/2}^N} \sup_{y \in B_{q^*/2}^M} \sum_{i \in I_0} \sum_{j \in J_0} y_i a_{ij}^2 x_j},$$

where in the last inequality we used the fact that $k \leq m$ and $l \leq n$. In order to estimate the right-hand side of the latter inequality, we consider the following two cases:

Case 1. If $q^* \geq 2$, then $(q^*/2)^* = q/(2-q) \geq q/2$ and $\| \cdot \|_{q/(2-q)} \leq \| \cdot \|_{q/2}$. Consequently,

$$\sup_{x \in B_{p/2}^N, y \in B_{q^*/2}^M} \sum_{i \in I_0} \sum_{j \in J_0} y_i a_{ij}^2 x_j = \| A \circ A : \ell_{p/2}^{J_0} \to \ell_{q/(2-q)}^{I_0} \|$$

$$\leq \| A \circ A : \ell_{p/2}^{J_0} \to \ell_{q/2}^{I_0} \|. \qquad (3.10)$$

Case 2. If $q^* \leq 2$, then $B_{q^*/2}^M \subset B_1^M$ and $\| \cdot \|_\infty \leq \| \cdot \|_{q/2}$. Thus,

$$\sup_{x \in B_{p/2}^N, y \in B_{q^*/2}^M} \sum_{i \in I_0} \sum_{j \in J_0} y_i a_{ij}^2 x_j \leq \sup_{u \in B_{p/2}^N, v \in B_1^M} \sum_{i \in I_0} \sum_{j \in J_0} v_i a_{ij}^2 u_j$$

$$= \| A \circ A : \ell_{p/2}^{J_0} \to \ell_\infty^{I_0} \| \leq \| A \circ A : \ell_{p/2}^{J_0} \to \ell_{q/2}^{I_0} \|. \qquad (3.11)$$

In both cases we have

$$D \leq \sup_{I_0, J_0} \| A \circ A : \ell_{p/2}^{J_0} \to \ell_{q/2}^{I_0} \|^{1/2},$$

so the Gaussian concentration inequality (see, e.g., [41, Chapter 5.1]) implies that for all $u \geq 0$, $k \leq m$, and $l \leq n$,

$$\mathbb{P}(Z_{k,l} \geq \mathbb{E}Z_{k,l} + u) \leq \exp\Big(-\frac{u^2}{2 \sup_{I_0,J_0} \|A \circ A : \ell_{p/2}^{J_0} \to \ell_{q/2}^{I_0}\|}\Big),$$

so

$$\mathbb{P}\Big(Z_{k,l} \geq \max_{k \leq m, l \leq n} \mathbb{E}Z_{k,l} + \sqrt{2\ln(mn)}u \sup_{I_0,J_0} \|A \circ A : \ell_{p/2}^{J_0} \to \ell_{q/2}^{I_0}\|^{1/2}\Big)$$
$$\leq \exp(-u^2 \ln(mn)).$$

This, together with the union bound, implies that for $u \geq \sqrt{2}$, we have

$$\mathbb{P}\Big(\max_{k \leq m, l \leq n} Z_{k,l} \geq \max_{k \leq m, l \leq n} \mathbb{E}Z_{k,l} + \sqrt{2\ln(mn)}u \sup_{I_0,J_0} \|A \circ A : \ell_{p/2}^{J_0} \to \ell_{q/2}^{I_0}\|^{1/2}\Big)$$
$$\leq mn e^{-u^2 \ln(mn)} = \exp\Big(-(u^2 - 1)\ln(mn)\Big) \leq e^{-u^2/2}.$$

Hence, by Lemma 2.6 and the estimate in (3.9),

$$\mathbb{E}\max_{k \leq m, l \leq n} Z_{k,l} \leq \max_{k \leq m, l \leq n} \mathbb{E}Z_{k,l}$$
$$+ \sqrt{2\ln(mn)}\Big(\sqrt{2} + \frac{1}{e\sqrt{2}}\Big) \sup_{I_0,J_0} \|A \circ A : \ell_{p/2}^{J_0} \to \ell_{q/2}^{I_0}\|^{1/2}$$
$$\leq \big(2.4\sqrt{\ln(mn)} + 8\sqrt{\ln M} + \sqrt{2/\pi}\big) \sup_{I_0,J_0} \|A \circ A : \ell_{p/2}^{J_0} \to \ell_{q/2}^{I_0}\|^{1/2}$$
$$+ \big(8\sqrt{\ln N} + 2\sqrt{2/\pi}\big) \sup_{I_0,J_0} \|(A \circ A)^T : \ell_{q^*/2}^{I_0} \to \ell_{p^*/2}^{J_0}\|^{1/2}.$$

Recalling (3.7) yields the assertion.   □

3.2 Coupling

In this subsection we use contraction principles and the coupling described in Lemma 2.21 to prove Corollaries 1.5 and 1.6, and Proposition 1.16. Below we state more general versions of the corollaries akin to Theorem 1.3 (the versions from the introduction follow by setting $M = m$, $N = n$).

Theorem 3.3 (General version of Corollary 1.5) *Assume that $m \leq M$, $n \leq N$, $1 \leq p, q \leq \infty$, and $X = (X_{ij})_{i \leq M, j \leq N}$ has independent mean-zero entries taking values in $[-1, 1]$. Then*

$$\mathbb{E}\sup_{I,J} \|X_A : \ell_p^J \to \ell_q^I\| = \mathbb{E}\sup_{I,J} \sup_{x \in B_p^J} \sup_{y \in B_{q^*}^I} \sum_{i \in I} \sum_{j \in J} y_i a_{ij} X_{ij} x_j$$

$$\leq \ln(en)^{1/p^*} \ln(em)^{1/q}$$
$$\cdot \left[\left(2.4\sqrt{2\pi}\sqrt{\ln(mn)} + 8\sqrt{2\pi}\sqrt{\ln M} + 2 \right) \sup_{I,J} \| A \circ A : \ell_{p/2}^J \to \ell_{q/2}^I \|^{1/2} \right.$$
$$\left. + \left(8\sqrt{2\pi}\sqrt{\ln N} + 4 \right) \sup_{I,J} \| (A \circ A)^T : \ell_{q^*/2}^I \to \ell_{p^*/2}^J \|^{1/2} \right],$$

where the suprema are taken over all sets $I \subset \{1, \ldots, M\}$, $J \subset \{1, \ldots, N\}$ *such that* $|I| = m$, $|J| = n$.

Remark 3.4 (*Symmetrization of entries of a random matrix*) Let $\widetilde{Z}$ be an independent copy of a random matrix Z with mean 0 entries. Then for any norm $\| \cdot \|$, including the operator norm from ℓ_p^n to ℓ_q^m, we have by Jensen's inequality

$$\mathbb{E}\|Z\| = \mathbb{E}\|Z - \mathbb{E}\widetilde{Z}\| \leq \mathbb{E}\|Z - \widetilde{Z}\| \leq \mathbb{E}\|Z\| + \mathbb{E}\|\widetilde{Z}\| = 2\mathbb{E}\|Z\|.$$

Therefore, in many cases we may simply assume that we deal with matrices with symmetric (not only mean 0) entries. For example, in the setting of Theorem 3.3, the entries of $X - \widetilde{X}$ are symmetric and take values in $[-2, 2]$, so it suffices to prove the assertion of this theorem (with a two times smaller constant on the right-hand side) under the additional assumption that the entries of the given random matrix are symmetric.

Proof of Theorem 3.3 By Remark 3.4 we may and do assume that the entries of X are symmetric—in this case we need to prove the assertion with a two times smaller constant.

Since the entries of X are independent and symmetric, X has the same distribution as $(\varepsilon_{ij}|X_{ij}|)_{i,j}$, where $(\varepsilon_{ij})_{i \leq M, j \leq N}$ is a random matrix with i.i.d. Rademacher entries, independent of all other random variables. Thus, the contraction principle (see Lemma 2.7) applied conditionally yields (below the suprema are taken over all sets $I \subset \{1, \ldots, M\}$, $J \subset \{1, \ldots, N\}$ such that $|I| = m$, $|J| = n$, *and over all* $x \in B_p^J$, $y \in B_{q^*}^I$, *and the sums run over all* $i \in I$ *and* $j \in J$)

$$\mathbb{E}\sup_{I,J} \sum y_i a_{ij} X_{ij} x_j = \mathbb{E}\sup_{I,J} \sum y_i a_{ij} \varepsilon_{ij} |X_{ij}| x_j \leq \mathbb{E}\sup_{I,J} \sum y_i a_{ij} \varepsilon_{ij} x_j$$
$$= \sqrt{\frac{\pi}{2}} \mathbb{E}\sup_{I,J} \sum y_i a_{ij} \varepsilon_{ij} \mathbb{E}|g_{ij}| x_j \leq \sqrt{\frac{\pi}{2}} \mathbb{E}\sup_{I,J} \sum y_i a_{ij} \varepsilon_{ij} |g_{ij}| x_j$$
$$= \sqrt{\frac{\pi}{2}} \mathbb{E}\sup_{I,J} \sum y_i a_{ij} g_{ij} x_j,$$

and the assertion follows from Theorem 1.3. □

Theorem 3.5 (General version of Corollary 1.6) *Assume that* $K, L > 0$, $r \in (0, 2]$, $m \leq M$, $n \leq N$, $1 \leq p, q \leq \infty$, *and* $X = (X_{ij})_{i \leq M, j \leq N}$ *has independent mean-zero entries satisfying*

$$\mathbb{P}(|X_{ij}| \geq t) \leq Ke^{-t^r/L} \quad \text{for all } t \geq 0, \, i \leq M, \, j \leq N.$$

Then

$$\mathbb{E}\sup_{I,J}\|X_A : \ell_p^J \to \ell_q^I\| = \mathbb{E}\sup_{I,J}\sup_{x\in B_p^J}\sup_{y\in B_{q*}^I}\sum_{i\in I}\sum_{j\in J}y_i a_{ij}X_{ij}x_j$$

$$\lesssim_{r,K,L}(\ln n)^{1/p^*}(\ln m)^{1/q}\ln(MN)^{\frac{1}{r}-\frac{1}{2}}$$

$$\cdot\left[\left(\sqrt{\ln(mn)}+\sqrt{\ln M}\right)\sup_{I,J}\|A\circ A : \ell_{p/2}^J \to \ell_{q/2}^I\|^{1/2}\right.$$

$$\left.+\sqrt{\ln N}\sup_{I,J}\|(A\circ A)^T : \ell_{q^*/2}^I \to \ell_{p^*/2}^J\|^{1/2}\right],$$

where the suprema are taken over all sets $I \subset \{1,\ldots,M\}$, $J \subset \{1,\ldots,N\}$ such that $|I| = m$, $|J| = n$.

Proof Let $\widetilde{X}$ be an independent copy of X. Then

$$\mathbb{P}(|X_{ij}-\widetilde{X}_{ij}|\geq t)\leq\mathbb{P}(|X_{ij}|\geq t/2\text{ or }|\widetilde{X}_{ij}|\geq t/2)$$

$$\leq 2\mathbb{P}(|X_{ij}|\geq t/2)\leq 2Ke^{-t^r/(2^r L)}.$$

This means that the symmetric matrix $X - \widetilde{X}$ satisfies the assumptions of Theorem 3.5. Hence, due to Remark 3.4, we may and do assume that the entries of X are symmetric.

Take the unique positive parameter s satisfying $\frac{1}{r} = \frac{1}{2} + \frac{1}{s}$. For $i \leq M$, $j \leq N$, let g_{ij} be i.i.d. standard Gaussian variables, independent of other variables, and let Y_{ij} be i.i.d. non-negative Weibull random variables with shape parameter s scale parameter 1 (i.e., $\mathbb{P}(Y_{ij}\geq t)=e^{-t^s}$ for $t\geq 0$), independent of other variables. (In the case $r=2$, we have $s=\infty$ and then $Y_{ij}=1$ almost surely.) Take

$$(U_{ij})_{i\leq M,j\leq N}\overset{d}{\sim}(|X_{ij}|)_{i\leq M,j\leq N},\qquad(V_{ij})_{i\leq M,j\leq N}\overset{d}{\sim}(|g_{ij}|Y_{ij})_{i\leq M,j\leq N}$$

as in Lemma 2.21 (we pick a pair (U_{ij}, V_{ij}) separately for every (i, j), and then take such a version of each pair that the system of MN random pairs (U_{ij}, V_{ij}) is independent).

Let $(\varepsilon_{ij})_{i\leq M,j\leq N}$ be a random matrix with i.i.d. Rademacher entries, independent of all other random variables. Since the entries of X are symmetric and independent, X has the same distribution as $(\varepsilon_{ij}|X_{ij}|)_{ij}$. By Lemma 2.21 we know that

$$U_{ij}\leq(8L)^{1/r}\left(\left(\frac{\ln(K/c)}{4}\right)^{1/r}+V_{ij}\right)\lesssim_{r,K,L}1+V_{ij}\qquad\text{a.s.}$$

We use the contraction principle conditionally for $\mathbb{E}_\varepsilon$, i.e., for U_{ij}'s and V_{ij}'s fixed. More precisely, we apply Lemma 2.7 to the space $\mathbf{X}$ of all $M \times N$ matrices with real coefficients, equipped with the norm

$$\|(M_{ij})_{i\leq M,j\leq N}\|:=\sup_{I,J}\|(M_{ij})_{i\in I,j\in J}:\ell_p^I\to\ell_q^J\|=\sup_{I,J}\sum y_i M_{ij}x_j$$

(where the first supremum is taken over all sets $I \subset \{1, \dots, M\}$, $J \subset \{1, \dots, N\}$ such that $|I| = m$, $|J| = n$; recall that the second supremum is taken over all sets I, J as in the first supremum, and over all $x \in B_p^J$, $y \in B_{q*}^I$, and the sum runs over all $i \in I$ and $j \in J$); note that we identify $\mathbf{X}$ with $\mathbb{R}^{MN}$ (and MN plays the role of n from Lemma 2.7). We apply the contraction principle of Lemma 2.7 (conditionally, with the values of U_{ij}'s and V_{ij}'s fixed) with coefficients $\alpha_{ij} := \frac{U_{ij}}{C(r,K,L)(1+V_{ij})}$ and points $\mathbf{x}_{ij} := \left(a_{kl} C(r, K, L)(1 + V_{kl}) \mathbf{1}_{\{(k,l)=(i,j)\}}\right)_{kl} \in \mathbf{X}$ to get

$$\mathbb{E} \sup_{I,J} \sum y_i a_{ij} X_{ij} x_j = \mathbb{E} \sup_{I,J} \sum y_i a_{ij} \varepsilon_{ij} |X_{ij}| x_j = \mathbb{E} \sup_{I,J} \sum y_i a_{ij} \varepsilon_{ij} U_{ij} x_j$$

$$\overset{\text{Lemma 2.7}}{\lesssim_{r,K,L}} \mathbb{E} \sup_{I,J} \sum y_i a_{ij} \varepsilon_{ij} x_j + \mathbb{E} \sup_{I,J} \sum y_i a_{ij} \varepsilon_{ij} V_{ij} x_j.$$
(3.12)

We may estimate the first term using Theorem 3.3 applied to the matrix $(\varepsilon_{ij})_{i \le M, j \le N}$ as follows,

$$\mathbb{E} \sup_{I,J} \sum y_i a_{ij} \varepsilon_{ij} x_j \lesssim \ln(en)^{1/p^*} \ln(em)^{1/q}$$

$$\cdot \left[\left(\sqrt{\ln(mn)} + \sqrt{\ln M}\right) \sup_{I,J} \| A \circ A : \ell_{p/2}^J \to \ell_{q/2}^I \|^{1/2} \right.$$

$$\left. + \sqrt{\ln N} \sup_{I,J} \| (A \circ A)^T : \ell_{q*/2}^I \to \ell_{p*/2}^J \|^{1/2} \right]. \quad (3.13)$$

Recall that $(\varepsilon_{ij} V_{ij})_{i \le M, j \le N} \overset{d}{\sim} (\varepsilon_{ij} g_{ij} Y_{ij})_{i \le M, j \le N}$ and that $Y_{ij} \ge 0$ almost surely. Next we again use the contraction principle (applied conditionally for $\mathbb{E}_\varepsilon$, i.e. for fixed Y_{ij}'s and g_{ij}'s) and get

$$\mathbb{E} \sup_{I,J} \sum y_i a_{ij} \varepsilon_{ij} V_{ij} x_j = \mathbb{E} \sup_{I,J} \sum y_i a_{ij} \varepsilon_{ij} g_{ij} Y_{ij} x_j$$

$$\le \mathbb{E}_Y \max_{i \le M, j \le N} |Y_{ij}| \, \mathbb{E}_{\varepsilon,g} \sup_{I,J} \sum y_i a_{ij} \varepsilon_{ij} g_{ij} x_j. \quad (3.14)$$

Moreover, Theorem 1.3 and Lemma 2.22 (applied with $r = s$, $k = MN$, $Z_{ij} = Y_{ij}$, and $K = 1 = L$), imply

$$\mathbb{E}_Y \max_{i \le M, j \le N} |Y_{ij}| \, \mathbb{E}_{\varepsilon,g} \sup_{I,J} \sum y_i a_{ij} \varepsilon_{ij} g_{ij} x_j$$

$$\lesssim_r \ln(MN)^{1/s} \, \mathbb{E} \sup_{I,J} \sum y_i a_{ij} g_{ij} x_j$$

$$\lesssim \ln(MN)^{\frac{1}{r} - \frac{1}{2}} (\ln n)^{1/p^*} (\ln m)^{1/q}$$

$$\cdot \left[\left(\sqrt{\ln(mn)} + \sqrt{\ln M} \right) \sup_{I,J} \|A \circ A : \ell_{p/2}^J \to \ell_{q/2}^I \|^{1/2} \right.$$

$$\left. + \sqrt{\ln N} \sup_{I,J} \|(A \circ A)^T : \ell_{q*/2}^I \to \ell_{p*/2}^J \|^{1/2} \right]. \tag{3.15}$$

Combining the estimates in (3.12)–(3.15) yields the assertion. □

Finally, we prove that these estimates of the operator norms translate into tail bounds.

Proof of Proposition 1.16 Since (1.23) implies (1.24) (by Lemma 2.18), it suffices to prove inequality (1.23). By the symmetrization argument similar to the one from the first paragraph of the proof of Theorem 3.5, we may nad will assume that X has independent and *symmetric* entries satisfying (1.21). By assumption (1.21), and the inequality $2(a + b)^r \geq a^r + b^r$ we have for every $t \geq 0$,

$$\mathbb{P}\big((2L)^{-1/r}|X_{ij}| \geq t + (\ln K)^{1/r}\big)$$
$$\leq K \exp\left(-2\big(t + (\ln K)^{1/r}\big)^r\right) \leq K \exp\left(-t^r - \ln K\right) = e^{-t^r},$$

so (as in the proof of Lemma 2.21) there exists a random matrix $(Y_{ij})_{i \leq m, j \leq n}$ with i.i.d. entries with the symmetric Weibull distribution with shape parameter r and scale parameter 1 (i.e., $\mathbb{P}(|Y_{ij}| \geq t) = e^{-t^r}$ for $t \geq 0$) satisfying

$$|X_{ij}| \leq (2L)^{1/r}\big((\ln K)^{1/r} + |Y_{ij}|\big) \lesssim_{r,K,L} 1 + Y_{ij} \qquad \text{a.s.} \tag{3.16}$$

Let $(\varepsilon_{ij})_{i \leq m, j \leq n}$ be a matrix of independent Rademacher random variables independent of all others, and let $\|\cdot\|$ denote the operator norm from ℓ_p^n to ℓ_q^m. Let E_{ij} be a matrix with 1 at the intersection of ith row and jth column and with other entries 0. The contraction principle (i.e., Lemma 2.7) applied conditionally, (3.16), and the triangle inequality yield for any $\rho \geq 1$,

$$\left(\mathbb{E} \Big\| \sum_{i=1}^m \sum_{j=1}^n X_{ij} a_{ij} E_{ij} \Big\|^\rho \right)^{1/\rho} \leq \left(\mathbb{E} \Big\| \sum_{i,j} \varepsilon_{ij} |X_{ij}| a_{ij} E_{ij} \Big\|^\rho \right)^{1/\rho}$$

$$\lesssim_{r,K,L} \left(\mathbb{E} \Big\| \sum_{i,j} \varepsilon_{ij} a_{ij} E_{ij} \Big\|^\rho \right)^{1/\rho} + \left(\mathbb{E} \Big\| \sum_{i,j} \varepsilon_{ij} |Y_{ij}| a_{ij} E_{ij} \Big\|^\rho \right)^{1/\rho}$$

$$= \left(\mathbb{E} \Big\| \sum_{i,j} \varepsilon_{ij} a_{ij} E_{ij} \Big\|^\rho \right)^{1/\rho} + \left(\mathbb{E} \Big\| \sum_{i,j} Y_{ij} a_{ij} E_{ij} \Big\|^\rho \right)^{1/\rho}.$$

Therefore, it suffices to prove (1.23) for random matrices $(Y_{ij})_{ij}$ and $(\varepsilon_{ij})_{ij}$ instead of X.

Since by assumption $K, L \geq 1$, both random matrices $(Y_{ij})_{ij}$ and $(\varepsilon_{ij})_{ij}$ satisfy (1.21), so for them inequality (1.22) holds. By the comparison of weak and strong

moments [38, Theorem 1.1] (note that the random variables Y_{ij} satisfy the assumption $\|Y_{ij}\|_{2s} \leq \alpha \|Y_{ij}\|_s$ for all $s \geq 2$ with $\alpha = 2^{1/r}$ by [38, Remark 1.5]), we have

$$\left(\mathbb{E} \Big\| \sum_{i,j} Y_{ij} a_{ij} E_{ij} \Big\|^{\rho} \right)^{1/\rho} = \left(\mathbb{E} \sup_{x \in B_p^n,\ y \in B_{q*}^m} \Big| \sum_{i,j} y_i Y_{ij} a_{ij} x_j \Big|^{\rho} \right)^{1/\rho}$$

$$\lesssim_r \mathbb{E} \sup_{x \in B_p^n,\ y \in B_{q*}^m} \sum_{i,j} y_i Y_{ij} a_{ij} x_j + \sup_{x \in B_p^n,\ y \in B_{q*}^m} \left(\mathbb{E} \Big| \sum_{i,j} y_i Y_{ij} a_{ij} x_j \Big|^{\rho} \right)^{1/\rho} \quad (3.17)$$

Because of inequality (1.22), the first summand on the right-hand side may be estimated by γD. Lemma 2.19 and the implication (i) $\Longrightarrow$ (ii) from Lemma 2.18 yield

$$\left(\mathbb{E} \Big| \sum_{i,j} y_i Y_{ij} a_{ij} x_j \Big|^{\rho} \right)^{1/\rho} \lesssim_{r,K,L} \rho^{1/r} \sqrt{\sum_{i,j} y_i^2 a_{ij}^2 x_j^2}.$$

Moreover, by (3.10) and (3.11) (used with $m = M$ and $n = N$) and our assumption that $\|A \circ A : \ell_{p/2}^n \to \ell_{q/2}^m\|^{1/2} \leq D$,

$$\sup_{x \in B_p^n,\ y \in B_{q*}^m} \sqrt{\sum_{i,j} y_i^2 a_{ij}^2 x_j^2} \leq D,$$

so the second summand on the right-hand side of (3.17) is bounded above (up to a multiplicative constant depending only on r, K, and L) by $\rho^{1/r} D$. Thus, (1.23) indeed holds for the random matrix $(Y_{ij})_{ij}$ instead of X. A similar reasoning shows that the same inequality holds also for the random matrix $(\varepsilon_{ij})_{ij}$ (one may also simply use the Khintchine–Kahane inequality and assumption (1.22)). □

4 Proofs of further results

4.1 Gaussian random variables

Proof of Proposition 1.7 Fix $1 \leq p \leq 2$ and $1 \leq q \leq \infty$. Let K be the set defined in Lemma 2.3 for which $B_p^n \subset \ln(en)^{1/p^*} K$. Then

$$\|G_A : \ell_p^n \to \ell_q^m\| = \sup_{x \in B_p^n} \|G_A x\|_q \leq \ln(en)^{1/p^*} \sup_{x \in \mathrm{Ext}(K)} \|G_A x\|_q, \quad (4.1)$$

where $\mathrm{Ext}(K)$ is the set of extreme points of K. We shall now estimate the expected value of the right-hand side of (4.1).

To this end, we first consider a fixed $x = (x_j)_{j=1}^n \in \mathrm{Ext}(K)$. Then there exists a non-empty index set $J \subset \{1, \ldots, n\}$ of cardinality $k \leq n$ such that $x_j = \frac{\pm 1}{k^{1/p}}$ for

$j \in J$ and $x_j = 0$ for $j \notin J$. We have

$$\|G_A x\|_q = \left\| \left(\sum_{j=1}^{n} a_{ij} g_{ij} x_j \right)_{i=1}^{m} \right\|_q = \left(\sum_{i=1}^{m} \left| \sum_{j=1}^{n} a_{ij} g_{ij} x_j \right|^q \right)^{1/q}. \tag{4.2}$$

Let us estimate the Lipschitz constant of the function

$$z = (z_{ij})_{ij} \mapsto \left\| \left(\sum_{j=1}^{n} a_{ij} z_{ij} x_j \right)_{i=1}^{m} \right\|_q = \sup_{y \in B_{q^*}^m} \sum_{i=1}^{m} \sum_{j=1}^{n} y_i a_{ij} z_{ij} x_j. \tag{4.3}$$

It follows from the Cauchy–Schwarz inequality (used in $\mathbb{R}^{m \times n}$) that

$$\sup_{y \in B_{q^*}^m} \sum_{i=1}^{m} \sum_{j=1}^{n} y_i a_{ij} z_{ij} x_j \leq \|z\|_2 \sqrt{\sup_{y \in B_{q^*}^m} \sum_{i=1}^{m} \sum_{j=1}^{n} y_i^2 a_{ij}^2 x_j^2}$$

$$= \|z\|_2 \frac{1}{k^{1/p}} \sqrt{\sup_{y \in B_{q^*/2}^m} \sum_{i=1}^{m} \sum_{j \in J} y_i a_{ij}^2} = \|z\|_2 \frac{b_J}{k^{1/p}}, \tag{4.4}$$

where we put

$$b_J := \sqrt{\sup_{y \in B_{q^*/2}^m} \sum_{i=1}^{m} \sum_{j \in J} y_i a_{ij}^2}.$$

This shows that the function defined by (4.3) is $\frac{b_J}{k^{1/p}}$-Lipschitz continuous. Therefore, by the Gaussian concentration inequality (see, e.g., [41, Chapter 5.1]), for any $u \geq 0$,

$$\mathbb{P}\left(\|G_A x\|_q \geq \mathbb{E}\|G_A x\|_q + u \right) \leq \exp\left(-\frac{k^{2/p} u^2}{2 b_J^2} \right). \tag{4.5}$$

We shall transform this inequality into a form which is more convenient to work with. We want to estimate $\mathbb{E}\|G_A x\|_q$ independently of x and get rid of the dependence on J and p on the right-hand side. By (4.2) and the fact that $x \in \text{Ext}(K) \subset B_p^n$, we obtain

$$\mathbb{E}\|G_A x\|_q \leq \left(\mathbb{E}\|G_A x\|_q^q \right)^{1/q} = \gamma_q \left(\sum_{i=1}^{m} \left| \sum_{j=1}^{n} a_{ij}^2 x_j^2 \right|^{q/2} \right)^{1/q}$$

$$\leq \gamma_q \sup_{z \in B_p^n} \left(\sum_{i=1}^{m} \left| \sum_{j=1}^{n} a_{ij}^2 z_j^2 \right|^{q/2} \right)^{1/q} = \gamma_q \|A \circ A : \ell_{p/2}^n \to \ell_{q/2}^m \|^{1/2} =: a.$$

We use the definition of b_J, then interchange the sums, use the triangle inequality, and then the inequality between the arithmetic mean and the power mean of order

$p^*/2 \geq 1$ (recall that $|J| = k$ and $p \leq 2$) to obtain

$$k^{2/p^*-1}b_J^2 = k^{2/p^*-1} \sup_{y \in B_{q^*/2}^m} \sum_{i=1}^m \sum_{j \in J} a_{ij}^2 y_i = k^{2/p^*-1} \sup_{y \in B_{q^*/2}^m} \sum_{j \in J} \left| \sum_{i=1}^m a_{ij}^2 y_i \right|$$

$$\leq \sup_{y \in B_{q^*/2}^m} \left(\sum_{j \in J} \left| \sum_{i=1}^m a_{ij}^2 y_i \right|^{p^*/2} \right)^{2/p^*} \leq \sup_{y \in B_{q^*/2}^m} \left(\sum_{j=1}^n \left| \sum_{i=1}^m a_{ij}^2 y_i \right|^{p^*/2} \right)^{2/p^*}$$

$$= \|(A \circ A)^T : \ell_{q^*/2}^m \to \ell_{p^*/2}^n\| =: b^2. \tag{4.6}$$

The two inequalities above, together with inequality (4.5) (applied with $u = k^{1/p^*-1/2}b_J\sqrt{2\ln(en)}s$), imply that

$$\mathbb{P}\big(\|G_A x\|_q \geq a + b\sqrt{2\ln(en)}\,s\big) \leq \exp\big(-k^{2/p+2/p^*-1}\ln(en)s^2\big)$$
$$= \exp\big(-k\ln(en)s^2\big) \tag{4.7}$$

holds for any $s \geq 0$ and all $x \in \text{Ext}(K)$ with support of cardinality k.

For any $k \leq n$, there are $2^k \binom{n}{k} \leq 2^k n^k \leq \exp(k\ln(en))$ vectors in $\text{Ext}(K)$ with support of cardinality k. Therefore, using a union bound together with (4.7), we see that, for all $s \geq \sqrt{2}$,

$$\mathbb{P}\big(\sup_{x \in \text{Ext } K} \|G_A x\|_q \geq a + b\sqrt{2\ln(en)}s \big) \leq \sum_{k=1}^n \exp(-k\ln(en)(s^2 - 1))$$

$$\leq n \exp(-\ln(en)(s^2 - 1)) = n(en)^{-s^2+1} \leq e^{-s^2+1}.$$

Hence, by Lemma 2.6 (applied with $s_0 := \sqrt{2}$, $\alpha := e$, $\beta := 1$, and $r := 2$),

$$\mathbb{E} \sup_{x \in \text{Ext } K} \|G_A x\|_q \leq a + b\sqrt{2\ln(en)}\Big(\sqrt{2} + e\frac{e^{-2}}{2\sqrt{2}}\Big) \leq a + 2.2b\sqrt{\ln(en)}.$$

Recalling (4.1) and the definitions of a and b yields the assertion. □

We now turn to the special case $q = 1$.

Proof of Proposition 1.8 Since the first part of this proof works for general $q \geq 1$, we do not restrict our attention to $q = 1$ for now. First of all,

$$\mathbb{E}\|G_A : \ell_p^n \to \ell_q^m\| \leq \big(\mathbb{E}\|G_A : \ell_p^n \to \ell_q^m\|^q\big)^{1/q} = \Big(\mathbb{E} \sup_{x \in B_p^n} \sum_{i=1}^m |\langle X_i, x \rangle|^q\Big)^{1/q},$$

where $X_i = (a_{ij}g_{ij})_{j=1}^n$ is the i-th row of the matrix G_A. Centering this expression gives

$$\mathbb{E} \sup_{x \in B_p^n} \sum_{i=1}^m |\langle X_i, x \rangle|^q \leq \mathbb{E} \sup_{x \in B_p^n} \Big[\sum_{i=1}^m |\langle X_i, x \rangle|^q - \mathbb{E}|\langle X_i, x \rangle|^q \Big]$$

$$+ \sup_{x \in B_p^n} \sum_{i=1}^m \mathbb{E}|\langle X_i, x \rangle|^q. \tag{4.8}$$

We first take care of the second term on the right-hand side of (4.8). We have

$$\sup_{x \in B_p^n} \sum_{i=1}^m \mathbb{E}|\langle X_i, x \rangle|^q = \gamma_q^q \sup_{x \in B_p^n} \sum_{i=1}^m \Big(\sum_{j=1}^n a_{ij}^2 x_j^2 \Big)^{q/2}$$

$$= \gamma_q^q \sup_{z \in B_{p/2}^n} \Big\| \Big(\sum_{j=1}^n a_{ij}^2 z_j \Big)_{i \le m} \Big\|_{q/2}^{q/2} = \gamma_q^q \| A \circ A \colon \ell_{p/2}^n \to \ell_{q/2}^m \|^{q/2}. \tag{4.9}$$

In order to deal with the first term on the right-hand side of (4.8), we use a symmetrization trick together with the contraction principle. The latter is the reason that we need to work with $q = 1$ here. We start with the symmetrization. Denoting by $\widetilde{X}_1, \ldots, \widetilde{X}_n$ independent copies of $X_1, \ldots, X_n$ and by $(\varepsilon_i)_{i=1}^m$ a sequence of Rademacher random variables independent of all others, we obtain by Jensen's and the triangle inequalities that

$$\mathbb{E} \sup_{x \in B_p^n} \Big[\sum_{i=1}^m |\langle X_i, x \rangle|^q - \mathbb{E}|\langle X_i, x \rangle|^q \Big] = \mathbb{E} \sup_{x \in B_p^n} \Big[\sum_{i=1}^m |\langle X_i, x \rangle|^q - \mathbb{E}|\langle \widetilde{X}_i, x \rangle|^q \Big]$$

$$\le \mathbb{E} \sup_{x \in B_p^n} \Big[\sum_{i=1}^m |\langle X_i, x \rangle|^q - |\langle \widetilde{X}_i, x \rangle|^q \Big]$$

$$= \mathbb{E} \sup_{x \in B_p^n} \Big[\sum_{i=1}^m \varepsilon_i (|\langle X_i, x \rangle|^q - |\langle \widetilde{X}_i, x \rangle|^q) \Big] \le 2 \cdot \mathbb{E} \sup_{x \in B_p^n} \sum_{i=1}^m \varepsilon_i |\langle X_i, x \rangle|^q. \tag{4.10}$$

If $q = 1$, we may use the contraction principle (i.e., Lemma 2.8 applied with functions $\varphi_i(t) = |t|$) conditionally to obtain

$$\mathbb{E} \sup_{x \in B_p^n} \sum_{i=1}^m \varepsilon_i |\langle X_i, x \rangle| \le \mathbb{E} \sup_{x \in B_p^n} \sum_{i=1}^m \varepsilon_i \langle X_i, x \rangle$$

$$= \mathbb{E} \sup_{x \in B_p^n} \sum_{j=1}^n x_j \sum_{i=1}^m a_{ij} \cdot \varepsilon_i g_{ij} = \mathbb{E} \sup_{x \in B_p^n} \sum_{j=1}^n x_j \sum_{i=1}^m a_{ij} g_{ij}. \tag{4.11}$$

For $p > 1$, we have

$$\mathbb{E} \sup_{x \in B_p^n} \sum_{j=1}^n x_j \sum_{i=1}^m a_{ij} g_{ij} = \mathbb{E} \Big(\sum_{j=1}^n \Big| \sum_{i=1}^m a_{ij} g_{ij} \Big|^{p^*} \Big)^{1/p^*}$$

$$\leq \Big(\sum_{j=1}^{n} \mathbb{E} \Big| \sum_{i=1}^{m} a_{ij} g_{ij} \Big|^{p^*} \Big)^{1/p^*} = \gamma_{p^*} \Big(\sum_{j=1}^{n} \Big(\sum_{i=1}^{m} a_{ij}^2 \Big)^{p^*/2} \Big)^{1/p^*}.$$

$$(4.12)$$

Moreover, we have

$$\Big(\sum_{j=1}^{n} \Big(\sum_{i=1}^{m} a_{ij}^2 \Big)^{p^*/2} \Big)^{1/p^*} = \sup_{\delta \in \{-1,1\}^m} \Big\| \Big(\sum_{i=1}^{m} a_{ij}^2 \delta_i \Big)_{j \leq n} \Big\|_{p^*/2}^{1/2}$$

$$= \big\| (A \circ A)^T : \ell_\infty^m \to \ell_{p^*/2}^n \big\|^{1/2}. \qquad (4.13)$$

Inequalities (4.10)–(4.13) give the estimate of the first term on the right-hand side of (4.8). This ends the proof of the upper bound for $p > 1$.

If $p = 1$, then letting $g_1, \ldots, g_n$ be i.i.d. standard Gaussian random variables, we have

$$\mathbb{E} \sup_{x \in B_p^n} \sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{ij} g_{ij} = \mathbb{E} \max_{j \leq n} \Big| \sum_{i=1}^{m} a_{ij} g_{ij} \Big|$$

$$= \mathbb{E} \max_{j \leq n} g_j b_j \asymp \max_{j \leq n} (\sqrt{\ln(j+1)} b_j^{\downarrow}), \qquad (4.14)$$

where the last step follows from Lemmas 2.11 and 2.12 with $b_j := \|(a_{ij})_{i \leq m}\|_2, \; j \leq n$. Putting together (4.8)–(4.11) and (4.14) completes the proof of the upper bound in the case $p = 1$.

The lower bound in the case $p > 1$ follows from Proposition 5.1 and Corollary 5.2 below. In the case $p = 1$, we use Proposition 5.1, note that

$$\mathbb{E} \|G_A : \ell_p^n \to \ell_1^m\| \geq \mathbb{E} \sup_{x \in B_p^n} \sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{ij} g_{ij},$$

and use (4.14) to obtain a lower bound. □

Now we deal with another special case, the one where $p = 1$.

Proof of Proposition 1.10 Recall that we deal with the range $p = 1 \leq q \leq 2$. Using the structure of extreme points of B_1^n we get

$$\mathbb{E} \|G_A : \ell_1^n \to \ell_q^m\| = \mathbb{E} \max_{j \leq n} \|(a_{ij} g_{ij})_{i \leq m}\|_q.$$

Denote $Z_j = \|(a_{ij} g_{ij})_{i \leq m}\|_q$. By well-known tail estimates of norms of Gaussian variables with values in Banach spaces (see, e.g., [36, Corollary 1] for a more general formulation) we get for all $t > 0$,

$$\mathbb{P}\big(Z_j \geq C(\mathbb{E} Z_j + \sqrt{t} b_j) \big) \leq e^{-t}, \qquad (4.15)$$

$$\mathbb{P}\big(Z_j \geq c(\mathbb{E}Z_j + \sqrt{t}b_j)\big) \geq \min(c, e^{-t}), \tag{4.16}$$

where c, C are universal positive constants, and

$$b_j^2 = \|(a_{ij}^2)_{i \leq m}\|_{q/(2-q)} = \|(a_{ij}^2)_{i \leq m}\|_{(q^*/2)^*} = \sup_{x \in B_{q^*}^m} \sum_{i=1}^m a_{ij}^2 x_i^2.$$

Inequality (4.15) shows in particular that the random variables $(Z_j - C\mathbb{E}Z_j)_+$ satisfy

$$\mathbb{P}((Z_j - C\mathbb{E}Z_j)_+ \geq t) \leq \exp\Big(-\frac{t^2}{C^2 b_j^2}\Big)$$

for all $t > 0$, thus by Lemma 2.11 we get

$$\mathbb{E}\|G_A \colon \ell_1^n \to \ell_q^m\| = \mathbb{E}\max_{j \leq n} Z_j \leq C \max_{j \leq n} \mathbb{E}Z_j + \max_{j \leq n}(Z_j - C\mathbb{E}Z_j)_+$$
$$\lesssim \Big(\max_{j \leq n} \mathbb{E}Z_j + \max_{j \leq n}(\sqrt{\ln(j+1)}b_j^{\downarrow})\Big),$$

which together with the observation (following from Lemma 2.1 and the fact that $1 = p \leq q \leq 2$) that

$$\mathbb{E}Z_j \leq \Big(\mathbb{E}\sum_{i=1}^m |a_{ij}|^q |g_{ij}|^q\Big)^{1/q} = \gamma_q \|(a_{ij})_{i \leq m}\|_q = \gamma_q \|A \circ A \colon \ell_{1/2}^n \to \ell_{q/2}^m\|^{1/2},$$

proves the upper estimate of the proposition.

Using comparison of moments of norms of Gaussian random vectors, we also get

$$\mathbb{E}\|G_A \colon \ell_1^n \to \ell_q^m\| \geq \max_{j \leq n} \mathbb{E}Z_j \gtrsim \max_{j \leq n}(\mathbb{E}Z_j^q)^{1/q}$$
$$= \gamma_q \|(a_{ij})_{i \leq m}\|_q = \gamma_q \|A \circ A \colon \ell_{1/2}^n \to \ell_{q/2}^m\|^{1/2}, \tag{4.17}$$

so to end the proof it is enough to show that

$$\mathbb{E}\|G_A \colon \ell_1^n \to \ell_q^m\| \geq \max_{j \leq n}(\sqrt{\ln(j+1)}b_j^{\downarrow}). \tag{4.18}$$

This will follow by a straightforward adaptation of the argument from the proof of Lemma 2.12. We may and do assume that the sequence $(b_j)_{j \leq n}$ is non-increasing in j. By (4.16) we have for any $j \leq n$ and $k \geq 1$,

$$\mathbb{P}(Z_j \geq c\sqrt{\ln(k+1)}b_j) \geq \frac{c'}{k}.$$

Thus, since $b_j \geq b_k$ for all $j \leq k$, we have for any $k \leq n$,

$$\mathbb{P}(\max_{j \leq n} Z_j \geq \sqrt{\ln(k+1)}b_k) \geq \mathbb{P}(\exists_{j \leq k} \ Z_j \geq \sqrt{\ln(k+1)}b_j)$$

$$\geq 1 - (1 - c'/k)^k \geq 1 - e^{-c'} > 0.$$

Thus,

$$\mathbb{E}\|G_A: \ell_1^n \to \ell_q^m\| = \mathbb{E}\max_{j \leq n} Z_j \gtrsim \sqrt{\ln(k+1)}b_k.$$

Taking maximum over $k \leq n$ gives (4.18) and ends the proof. $\qquad\square$

4.2 Bounded random variables

Here we show how one can adapt the methods of [9] to prove Proposition 1.14, i.e., a version of Corollary 1.13 in the special case of bounded random variables with better logarithmic terms and with explicit numerical constants. Following [9], we start with a lemma.

Lemma 4.1 *Assume that X is as in Proposition 1.14. Let $(b_j)_{j \leq n} \in \mathbb{R}^n$ and suppose that t_0 is such that $\left|\sum_{j=1}^n b_j X_{ij}\right| \leq t_0$ almost surely. Then, for all $q \geq 2$ and $0 \leq t \leq t_0^{2-q}(4\sum_{j=1}^n b_j^2)^{-1}$,*

$$\mathbb{E}\exp\left(t\Big|\sum_{j=1}^n b_j X_{ij}\Big|^q\right) \leq 1 + C^q(q)\, t\, \Big(\sum_{j=1}^n b_j^2\Big)^{q/2}, \qquad (4.19)$$

where $C(q):=2(q\Gamma(q/2))^{1/q} \asymp \sqrt{q}$.

Proof. Without loss of generality we may and do assume that $\sum_{j=1}^n b_j^2 = 1$.

Since $q \geq 2$, for $s \in [0, t_0]$ and $t \in [0, \frac{1}{4}t_0^{2-q}]$ we have $ts^q - s^2/2 \leq -s^2/4$. Thus, integration by parts, our assumption $0 \leq \left|\sum_{j=1}^n b_j X_{ij}\right| \leq t_0$ a.s., and Hoeffding's inequality (i.e., Lemma 2.13) yield

$$\mathbb{E}\exp\left(t\Big|\sum_{j=1}^n b_j X_{ij}\Big|^q\right) = 1 + qt\int_0^{t_0} s^{q-1}\exp(ts^q)\mathbb{P}\Big(\Big|\sum_{j=1}^n b_j X_{ij}\Big| \geq s\Big)ds$$

$$\leq 1 + 2qt\int_0^{t_0} s^{q-1}\exp(ts^q - s^2/2)ds$$

$$\leq 1 + 2qt\int_0^\infty s^{q-1}\exp(-s^2/4)ds$$

$$= 1 + t2^q q\Gamma(q/2). \qquad\square$$

Proof of Proposition 1.14 We start with a bunch of reductions. Set

$$a := \|A \circ A : \ell_{p/2}^n \to \ell_{q/2}^m\|^{1/2} = \max_{j \le n} \|(a_{ij})_{i=1}^m\|_q,$$

$$b := \|(A \circ A)^T : \ell_{q^*/2}^m \to \ell_{p^*/2}^n\|^{1/2} = \max_{i \le m} \|(a_{ij})_{j=1}^n\|_{p^*}.$$

(The equalities follow from Lemma 2.1, since $p/2 \le 1 \le q/2$ and $q^*/2 \le 1 \le p^*/2$). Let K be the set defined in Lemma 2.3, so that $B_p^n \subset \ln(en)^{1/p^*} K$. Then

$$\|X_A : \ell_p^n \to \ell_q^m\| = \sup_{x \in B_p^n} \|X_A x\|_q \le \ln(en)^{1/p^*} \sup_{x \in \mathrm{Ext}(K)} \|X_A x\|_q, \quad (4.20)$$

where $\mathrm{Ext}(K)$ is the set of extreme points of K.

Consider first a fixed $x = (x_j)_{j=1}^n \in \mathrm{Ext}(K) \subset B_p^n$. We have

$$\|X_A x\|_q^q = \sum_{i=1}^m \left| \sum_{j=1}^n a_{ij} X_{ij} x_j \right|^q. \quad (4.21)$$

Denote

$$t_0 := b = \max_{i \le m} \|(a_{ij})_{j=1}^n\|_{p^*},$$

$$t := \frac{t_0^{2-q}}{4 \max_{i \le m} \|(a_{ij} x_j)_{j=1}^n\|_2^2}.$$

Then, by the boundedness of X_{ij} and by Hölder's inequality, for every $i \le m$,

$$\left| \sum_{j=1}^n a_{ij} x_j X_{ij} \right| \le \sum_{j=1}^n |a_{ij}| |x_j| \le \|(a_{ij})_{j=1}^n\|_{p^*} \|(x_j)_{j=1}^n\|_p \le t_0.$$

We can now apply, for every $i \le m$, Lemma 4.1 (with t and t_0 as above and with coefficients $b_j = a_{ij} x_j$). Since the random variables $\left| \sum_{j=1}^n a_{ij} x_j X_{ij} \right|$, $i \le m$, are independent, using Lemma 4.1 yields

$$\mathbb{E} \exp\left(t \sum_{i=1}^m \left| \sum_{j=1}^n a_{ij} x_j X_{ij} \right|^q \right) = \prod_{i=1}^m \left[\mathbb{E} \exp\left(t \left| \sum_{j=1}^n a_{ij} x_j X_{ij} \right|^q \right) \right]$$

$$\le \prod_{i=1}^m \left(1 + C^q(q) \, t \left(\sum_{j=1}^n a_{ij}^2 x_j^2 \right)^{q/2} \right)$$

$$\le \exp\left(C^q(q) \, t \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij}^2 x_j^2 \right)^{q/2} \right) \le \exp\left(C^q(q) \, t a^q \right),$$

where in the last step we used the definition of $a = \|A \circ A : \ell^n_{p/2} \to \ell^m_{q/2}\|^{1/2}$ (and the fact that $x \in B^n_p$). By Chebyshev's inequality and (4.21), we have, for every $s \geq 0$,

$$\mathbb{P}\Big(t\|X_A x\|^q_q \geq \ln\big[\mathbb{E}\exp\big(t\sum_{i=1}^{m}|\sum_{j=1}^{n}a_{ij}x_j X_{ij}|^q\big)\big] + sk\Big) \leq e^{-sk}.$$

Combining this with the previous estimate yields, for every $s \geq 0$,

$$\mathbb{P}\Big(\|X_A x\|^q_q \geq C^q(q)\, a^q + \frac{sk}{t}\Big) \leq e^{-sk}.$$

Recall that $x \in \mathrm{Ext}(K)$. Thus, there exists an index set $J \subset \{1, \ldots, n\}$ of cardinality $k \leq n$, such that $x_j = \frac{\pm 1}{k^{1/p}}$ for $j \in J$ and $x_j = 0$ for $j \notin J$. We use the definition of t and the inequality between the arithmetic mean and the power mean of order $p^*/2 \geq 1$ (recall that $|J| = k$ and $p \leq 2$) to get

$$\frac{1}{4t} = b^{q-2}\max_{i \leq m}\|(a_{ij}x_j)^n_{j=1}\|^2_2 = b^{q-2}k^{-2/p}\max_{i \leq m}\sum_{j \in J}a^2_{ij}$$

$$\leq b^{q-2}k^{-2/p+1-2/p^*}\max_{i \leq m}\big(\sum_{j \in J}|a_{ij}|^{p^*}\big)^{2/p^*} = b^q k^{-1}.$$

Putting everything together, we obtain

$$\mathbb{P}\big(\|X_A x\|^q_q \geq C^q(q)\, a^q + 4b^q s\big) \leq e^{-sk} \tag{4.22}$$

for all $s \geq 0$ and all $x \in \mathrm{Ext}(K)$ with support of cardinality k.

For any $k \leq n$, there are $2^k\binom{n}{k} \leq 2^k n^k \leq \exp(k\ln(en))$ vectors in $\mathrm{Ext}(K)$ with support of cardinality k. Thus, using the union bound and (4.22), we see that, for all $s \geq 2$,

$$\mathbb{P}\big(\sup_{x \in \mathrm{Ext}\,K}\|X_A x\|^q_q \geq C^q(q)\, a^q + 4b^q\ln(en)s\big) \leq \sum_{k=1}^{n}\exp(-k\ln(en)(s-1))$$

$$\leq n\exp(-\ln(en)(s-1)) = n(en)^{-s+1} \leq e^{-s+1}.$$

Hence, by Lemma 2.6,

$$\mathbb{E}\sup_{x \in \mathrm{Ext}\,K}\|X_A x\|_q \leq \big(\mathbb{E}\sup_{x \in \mathrm{Ext}\,K}\|X_A x\|^q_q\big)^{1/q} \leq \big(C^q(q)\, a^q + 4b^q\ln(en)(2 + e \cdot e^{-2})\big)^{1/q}$$

$$\leq C(q)a + 10^{1/q}\ln(en)^{1/q}b.$$

Recalling (4.20) and the definitions of a, b, and $C(q)$ yields the assertion. $\qquad\square$

Remark 4.2 In the unstructured case, for X_{ij} which are independent, mean-zero, and take values in $[-1, 1]$, it is easy to extend (1.2) to the whole range of $p, q \in [1, \infty]$ (see [8, 13]). Indeed, for $p \geq 2$ and $q \geq 2$,

$$\mathbb{E}\|X\colon \ell_p^n \to \ell_q^m\| \leq \|\ell_p^n \hookrightarrow \ell_2^n\| \cdot \mathbb{E}\|X\colon \ell_2^n \to \ell_q^m\|$$
$$\lesssim_q n^{1/2-1/p} \cdot \max\{n^{1/2}, m^{1/q}\} = \max\{n^{1-1/p}, n^{1/2-1/p}m^{1/q}\}.$$

Thus, for $p \geq 2$ and $1 \leq q \leq 2$,

$$\mathbb{E}\|X\colon \ell_p^n \to \ell_q^m\| \leq \mathbb{E}\|X\colon \ell_p^n \to \ell_2^m\| \cdot \|\ell_2^m \hookrightarrow \ell_q^m\|$$
$$\lesssim_q \max\{n^{1-1/p}, n^{1/2-1/p}m^{1/2}\} \cdot m^{1/q-1/2}$$
$$= \max\{n^{1-1/p}m^{1/q-1/2}, n^{1/2-1/p}m^{1/q}\}.$$

Suppose now that $1 \leq p \leq 2 \leq q \leq \infty$ and $1/p + 1/q \leq 1$ (i.e., $q \geq p^*$). Choose $\theta \in [0, 1]$ and $r \geq 2$ so that $\frac{1}{p} = \frac{\theta}{2} + \frac{1-\theta}{1}$ and $\frac{1}{q} = \frac{\theta}{r} + \frac{1-\theta}{\infty}$, i.e., $\theta = 2/p^*$ and $r = 2q/p^*$. Using the Riesz–Thorin interpolation theorem, the fact that $\|X\colon \ell_1^n \to \ell_\infty^m\| \leq 1$ (since the entries take values in $[-1, 1]$), and Jensen's inequality, we arrive at

$$\mathbb{E}\|X\colon \ell_p^n \to \ell_q^m\| \leq \mathbb{E}\|X\colon \ell_2^n \to \ell_r^m\|^\theta \|X\colon \ell_1^n \to \ell_\infty^m\|^{1-\theta}$$
$$\leq \mathbb{E}\|X\colon \ell_2^n \to \ell_r^m\|^\theta \leq \left(\mathbb{E}\|X\colon \ell_2^n \to \ell_r^m\|\right)^\theta$$
$$\leq \max\{n^{1/2}, m^{1/r}\}^\theta = \max\{n^{1/p^*}, m^{1/q}\}.$$

The estimates in the remaining ranges of p, q follow by duality (1.12). Moreover, up to constants, all these estimates are optimal, as they can be reversed for matrices with ± 1 entries (see [8, Proposition 3.2] or [13, Satz 2]).

4.3 ψ_r random variables

In this section, we prove Theorem 1.15. To this end we shall split the matrix X into two parts $X^{(1)}$ and $X^{(2)}$ such that all entries of $X^{(1)}$ are bounded by $C \ln(mn)^{1/r}$. Then, we shall deal with $X^{(2)}$ using the following crude bound and the fact that the probability that $X^{(2)} \neq 0$ is very small. In order to bound the expectation of the norm of $X^{(1)}$ we need a cut-off version of Theorem 1.15 – see Lemma 4.4 below.

Lemma 4.3 *Let $r \in (0, 2]$. Assume that $X = (X_{ij})_{i \leq m, j \leq n}$ satisfies the assumptions of Theorem 1.15. Then*

$$\left(\mathbb{E}\|X_A\colon \ell_p^n \to \ell_q^m\|^2\right)^{1/2} \lesssim_{r,K,L} (m+n)^{1/r}\|A \circ A\colon \ell_{p/2}^n \to \ell_{q/2}^m\|^{1/2}.$$

Proof By a standard volumetric estimate (see, e.g., [64, Corollary 4.2.13]), we know that there exists (in the metric $\|\cdot\|_p$) a $1/2$-net S in B_p^n of size at most 5^n. In other

words, for any $x \in B_p^n$ there exists $y \in S$ such that $x - y \in \frac{1}{2} B_p^n$. Thus, for any $z \in \mathbb{R}^n$,

$$\sup_{x \in B_p^n} \sum_{j=1}^n x_j z_j \leq \sup_{x \in B_p^n} \min_{y \in S} \sum_{j=1}^n (x_j - y_j) z_j + \sup_{y \in S} \sum_{j=1}^n y_j z_j$$

$$\leq \sup_{u \in \frac{1}{2} B_p^n} \sum_{j=1}^n u_j z_j + \sup_{y \in S} \sum_{j=1}^n y_j z_j = \frac{1}{2} \sup_{x \in B_p^n} \sum_{j=1}^n x_j z_j + \sup_{y \in S} \sum_{j=1}^n y_j z_j.$$

Hence,

$$\sup_{x \in B_p^n} \sum_{j=1}^n x_j z_j \leq 2 \sup_{y \in S} \sum_{j=1}^n y_j z_j. \tag{4.23}$$

Likewise, if we denote by T the $1/2$-net in B_{q*}^m (in the metric $\| \cdot \|_{q*}$) of size at most 5^m, then

$$\sup_{x \in B_{q*}^m} \sum_{i=1}^m x_i z_i \leq 2 \sup_{y \in T} \sum_{i=1}^m y_i z_i. \tag{4.24}$$

Combining these two estimates, we see that

$$\left(\mathbb{E} \| X_A : \ell_p^n \to \ell_q^m \|^2 \right)^{1/2} = \left(\mathbb{E} \sup_{x \in B_p^n, y \in B_{q*}^m} \left(\sum_{i=1}^m \sum_{j=1}^n y_i a_{ij} X_{ij} x_j \right)^2 \right)^{1/2}$$

$$\leq 4 \left(\mathbb{E} \sup_{x \in S, y \in T} \left(\sum_{i=1}^m \sum_{j=1}^n y_i a_{ij} X_{ij} x_j \right)^2 \right)^{1/2}. \tag{4.25}$$

Lemma 2.19 implies that for any $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, the random variable

$$Z(x, y) := \left(\sum_{i,j} y_i^2 a_{ij}^2 x_j^2 \right)^{-1/2} \sum_{i=1}^m \sum_{j=1}^n y_i a_{ij} X_{ij} x_j$$

satisfies condition (i) in Lemma 2.18. Thus, Lemma 2.18 implies that

$$\mathbb{E} \exp \left(c(r, K, L) \left(\sum_{i,j} y_i^2 a_{ij}^2 x_j^2 \right)^{-r/2} \left(\sum_{i=1}^m \sum_{j=1}^n y_i a_{ij} X_{ij} x_j \right)^r \right) \leq C(r, K, L), \tag{4.26}$$

where $c(r, K, L) \in (0, \infty)$ and $C(r, K, L) \in (0, \infty)$ depend only on r, K, and L.

The function $z \mapsto e^{z^{r/2}}$ is convex on $[(2r^{-1} - 1)^{2/r}, \infty)$. Therefore, by Jensen's inequality, for any $u > 0$ and any nonnegative random variable Z,

$$\exp \left(u (\mathbb{E} Z^2)^{r/2} \right) \leq \exp \left((u^{2/r} \mathbb{E} Z^2 + (2r^{-1} - 1)^{2/r})^{r/2} \right)$$

$$\leq \mathbb{E} \exp\big((u^{2/r} Z^2 + (2r^{-1} - 1)^{2/r})^{r/2}\big)$$
$$\leq \mathbb{E} \exp\big(u Z^r + (2r^{-1} - 1)\big) \leq e^{2/r} \mathbb{E} \exp(u Z^r).$$

Hence,

$$(\mathbb{E} Z^2)^{1/2} \leq u^{-1/r} \Big(\ln\big(e^{2/r} \mathbb{E} \exp(u Z^r)\big)\Big)^{1/r}.$$

Thus, when

$$u := c(r, K, L) \left(\max_{x \in S, y \in T} \sum_{i,j} y_i^2 a_{ij}^2 x_j^2 \right)^{-r/2},$$

we get by (4.26), (3.10), and (3.11),

$$\Big(\mathbb{E} \sup_{x \in S, y \in T} \big(\sum_{i=1}^m \sum_{j=1}^n y_i a_{ij} X_{ij} x_j\big)^2\Big)^{1/2}$$
$$\leq u^{-1/r} \ln^{1/r} \Big(e^{2/r} \mathbb{E} \exp\big(c(r, K, L) \sup_{x \in S, y \in T} Z(x, y)^r\big)\Big)$$
$$\leq u^{-1/r} \ln^{1/r} \Big(e^{2/r} \mathbb{E} \sum_{x \in S, y \in T} \exp\big(c(r, K, L) Z(x, y)^r\big)\Big)$$
$$\leq u^{-1/r} \ln^{1/r} \Big(e^{2/r} |S||T| C(r, K, L)\Big)$$
$$\leq \frac{1}{c(r, K, L)} \max_{x \in S, y \in T} \big(\sum_{i,j} y_i^2 a_{ij}^2 x_j^2\big)^{1/2} \ln^{1/r} \Big(e^{2/r} 5^m 5^n C(r, K, L)\Big)$$
$$\overset{(3.10), (3.11)}{\lesssim_{r,K,L}} \|A \circ A \colon \ell_{p/2}^n \to \ell_{q/2}^m\|^{1/2} \big(m + n + \widetilde{C}(r, K, L)\big)^{1/r},$$

where in the last two inequalities we also used inequalities $|S| \leq 5^n$ and $|T| \leq 5^m$, and the inclusions $S \subset B_p^n$, $T \subset B_{q*}^m$. Recalling (4.25) completes the proof. □

The following cut-off version of Theorem 1.15 can be proved similarly as Proposition 1.7.

Lemma 4.4 *Let $K, L, M > 0$ and $r \in (0, 2]$. Assume $X = (X_{ij})_{i \leq m, j \leq n}$ is a random matrix with independent symmetric entries taking values in $[-M, M]$ and satisfying the condition*

$$\mathbb{P}\big(|X_{ij}| \geq t\big) \leq K e^{-t^r/L} \quad \text{for all } t \geq 0. \tag{4.27}$$

Then, for $1 \leq p \leq 2$ and $1 \leq q < \infty$, we have

$$\mathbb{E}\|X_A \colon \ell_p^n \to \ell_q^m\| \lesssim q^{1/r} C(r, K, L) \ln(en)^{1/p^*} \|A \circ A \colon \ell_{p/2}^n \to \ell_{q/2}^m\|^{1/2}$$

$$+ M \ln(en)^{1/2+1/p^*} \|(A \circ A)^T : \ell_{q^*/2}^m \to \ell_{p^*/2}^n\|^{1/2}.$$

Proof Fix $1 \leq p \leq 2$ and $1 \leq q \leq \infty$. Let K be the set defined in Lemma 2.3 so that $B_p^n \subset \ln(en)^{1/p^*} K$. Then

$$\|X_A : \ell_p^n \to \ell_q^m\| = \sup_{x \in B_p^n} \|X_A x\|_q \leq \ln(en)^{1/p^*} \sup_{x \in \text{Ext}(K)} \|X_A x\|_q, \quad (4.28)$$

where $\text{Ext}(K)$ is the set of extreme points of K. We shall now estimate the expected value of the right-hand side of (4.28).

To this end, we consider a fixed $x = (x_j)_{j=1}^n \in \text{Ext}(K)$. This means that there exists a non-empty index set $J \subset \{1, \ldots, n\}$ of cardinality $k \leq n$ such that $x_j = \frac{\pm 1}{k^{1/p}}$ for $j \in J$ and $x_j = 0$ for $j \notin J$. We know from (4.4) that the Lipschitz constant of the convex function

$$z = (z_{ij})_{ij} \mapsto \left\| \left(\sum_{j=1}^n a_{ij} z_{ij} x_j \right)_i \right\|_q = \sup_{y \in B_{q^*}^m} \sum_{i=1}^m \sum_{j=1}^n y_i a_{ij} z_{ij} x_j$$

is less than or equal to

$$\frac{1}{k^{1/p}} \sqrt{\sup_{y \in B_{q^*/2}^m} \sum_{i=1}^m \sum_{j \in J} y_i a_{ij}^2} =: \frac{b_J}{k^{1/p}}.$$

Thus, Talagrand's concentration for convex functions and random vectors with independent bounded coordinates (see [56, Theorem 6.6 and Eq. (6.18)]), together with the inequality $\text{Med}(|Z|) \leq 2\mathbb{E}|Z|$, implies

$$\mathbb{P}(\|X_A x\|_q \geq 2\mathbb{E}\|X_A x\|_q + t) \leq 4 \exp\left(-\frac{k^{2/p} t^2}{16 M^2 b_J^2}\right) \quad \text{for all } t \geq 0. \quad (4.29)$$

Similar to the proof in the Gaussian case (i.e., proof of Proposition 1.7), we shall transform this into a more convenient form by getting rid of b_J and estimating $\mathbb{E}\|X_A x\|_q$. Let us denote, for each $i \in \{1, \ldots, m\}$,

$$Z_i := \sum_{j=1}^n a_{ij} X_{ij} x_j.$$

From our assumption (4.27) as well as Lemmas 2.19 and 2.18, we obtain that $(\mathbb{E}|Z_i|^q)^{1/q} \lesssim_{r,K,L} q^{1/r} \sqrt{\sum_{j=1}^n a_{ij}^2 x_j^2}$. Hence,

$$\mathbb{E}\|X_A x\|_q \leq (\mathbb{E}\|X_A x\|_q^q)^{1/q} = \left(\sum_{i=1}^m \mathbb{E}|(X_A x)_i|^q\right)^{1/q} \lesssim_{r,K,L} q^{1/r} \left(\sum_{i=1}^m \left(\sum_{j=1}^n a_{ij}^2 x_j^2\right)^{q/2}\right)^{1/q}$$

$$\leq q^{1/r} \sup_{z \in B_p^n} \Big(\sum_{i=1}^{m} \Big| \sum_{j=1}^{n} a_{ij}^2 z_j^2 \Big|^{q/2} \Big)^{1/q} = q^{1/r} \| A \circ A : \ell_{p/2}^n \to \ell_{q/2}^m \|^{1/2} =: q^{1/r} a.$$

From (4.6), we see that

$$k^{2/p^*-1} b_J^2 \leq \| (A \circ A)^T : \ell_{q^*/2}^m \to \ell_{p^*/2}^n \| =: b^2.$$

The above two inequalities together with estimate (4.29) (applied with $t = 4k^{\frac{1}{p^*} - \frac{1}{2}} b_J M \sqrt{\ln(en)s}$), imply that

$$\mathbb{P}\big(\| X_A x \|_q \geq C(r, K, L) q^{1/r} a + 4bM \sqrt{\ln(en)} s \big) \leq 4 \exp\big(-k \ln(en) s^2 \big) \quad (4.30)$$

for every $s \geq 0$ and any $x \in \mathrm{Ext}(K)$ with support of cardinality k.

For any $k \leq n$, there are $2^k \binom{n}{k} \leq 2^k n^k \leq \exp(k \ln(en))$ vectors in $\mathrm{Ext}(K)$ with support of cardinality k. Thus, using the union bound and (4.30), we see that for $s \geq \sqrt{2}$,

$$\mathbb{P}\big(\sup_{x \in \mathrm{Ext}\, K} \| X_A x \|_q \geq C(r, K, L) q^{1/r} a + 4bM \sqrt{\ln(en)} s \big)$$

$$\leq 4 \sum_{k=1}^{n} \exp(-k \ln(en)(s^2 - 1))$$

$$\leq 4n \exp(-\ln(en)(s^2 - 1)) = 4n(en)^{-s^2+1} \leq 4e^{-s^2+1}.$$

Hence, by Lemma 2.6,

$$\mathbb{E} \sup_{x \in \mathrm{Ext}\, K} \| X_A x \|_q \leq C(r, K, L) q^{1/r} a + 4bM \sqrt{\ln(en)} \Big(\sqrt{2} + 4e \frac{e^{-2}}{2\sqrt{2}} \Big).$$

Recalling (4.28) and the definitions of a and b yields the assertion. $\qquad\square$

Proof of Theorem 1.15 By a symmetrization argument (as in the first paragraph of the proof of Theorem 3.5), we may and do assume that all the entries X_{ij} are symmetric. Set $M = (4L \ln(mn)/r)^{1/r}$. Denote $\widehat{X}_{ij} = X_{ij} \mathbf{1}_{\{|X_{ij}| \leq M\}}$ and let $\widehat{X}$ be the $m \times n$ matrix with entries $\widehat{X}_{ij}$. We have

$$\mathbb{E} \| X_A : \ell_p^n \to \ell_q^m \| = \mathbb{E} \| X_A : \ell_p^n \to \ell_q^m \| \mathbf{1}_{\{\max_{k,l} |X_{kl}| \leq M\}}$$
$$+ \mathbb{E} \| X_A : \ell_p^n \to \ell_q^m \| \mathbf{1}_{\{\max_{k,l} |X_{kl}| > M\}}.$$

The random matrix $\widehat{X}$ satisfies the assumptions of Lemma 4.4. Thus, the first summand above can be estimated as follows:

$$\mathbb{E} \| X_A : \ell_p^n \to \ell_q^m \| \mathbf{1}_{\{\max_{k,l} |X_{kl}| \leq M\}}$$

$$= \mathbb{E} \sup_{y \in B_{q^*}^m, x \in B_p^n} \Big\{ \sum_{i=1}^{m} \sum_{j=1}^{n} y_i a_{ij} X_{ij} x_j \Big\} \cdot \mathbf{1}_{\{\max_{k,l} |X_{kl}| \leq M\}}$$

$$= \mathbb{E} \sup_{y \in B_{q^*}^m, x \in B_p^n} \Big\{ \sum_{i=1}^{m} \sum_{j=1}^{n} y_i a_{ij} X_{ij} \mathbf{1}_{\{|X_{ij}| \leq M\}} x_j \Big\} \cdot \mathbf{1}_{\{\max_{k,l} |X_{kl}| \leq M\}}$$

$$= \mathbb{E} \| \widehat{X}_A \colon \ell_p^n \to \ell_q^m \| \mathbf{1}_{\{\max_{k,l} |X_{kl}| \leq M\}} \leq \mathbb{E} \| \widehat{X}_A \colon \ell_p^n \to \ell_q^m \|$$

$$\lesssim_{r,K,L} \; q^{1/r} \ln(en)^{1/p^*} \|A \circ A \colon \ell_{p/2}^n \to \ell_{q/2}^m \|^{1/2}$$
$$+ M \ln(en)^{1/2 + 1/p^*} \|(A \circ A)^T \colon \ell_{q^*/2}^m \to \ell_{p^*/2}^n \|^{1/2}$$

$$\lesssim_{r,K,L} \; q^{1/r} \ln(en)^{1/p^*} \|A \circ A \colon \ell_{p/2}^n \to \ell_{q/2}^m \|^{1/2}$$
$$+ \ln(mn)^{1/r} \ln(en)^{1/2 + 1/p^*} \|(A \circ A)^T \colon \ell_{q^*/2}^m \to \ell_{p^*/2}^n \|^{1/2}.$$

For the second summand we write, using the Cauchy–Schwarz inequality and then Lemmas 4.3 and 2.22 (with $k = mn$ and $v = 4/r$; recall that $M = (4\,L \ln(mn)/r)^{1/r}$),

$$\mathbb{E} \|X_A \colon \ell_p^n \to \ell_q^m \| \mathbf{1}_{\{\max_{k,l} |X_{kl}| > M\}}$$
$$\leq \big(\mathbb{E} \|X_A \colon \ell_p^n \to \ell_q^m \|^2 \big)^{1/2} \mathbb{P}(\max_{k \leq m, l \leq n} |X_{kl}| > M)^{1/2}$$
$$\lesssim_{r,K,L} (m+n)^{1/r} \|A \circ A \colon \ell_{p/2}^n \to \ell_{q/2}^m \|^{1/2} \cdot (mn)^{-2/r + 1/2}$$
$$\lesssim_r \|A \circ A \colon \ell_{p/2}^n \to \ell_{q/2}^m \|^{1/2}.$$

Combining the above three inequalities ends the proof. $\qquad \square$

5 Lower bounds and further discussion of conjectures

5.1 Lower bounds

Let us first provide lower bounds showing that the upper bounds obtained above are indeed sharp (up to logarithms).

Proposition 5.1 *Let* $X = (X_{ij})_{i \leq m, j \leq n}$ *be a random matrix with independent mean-zero entries satisfying* $\mathbb{E}|X_{ij}| \geq c$ *for some* $c \in (0, \infty)$. *Then, for all* $1 \leq p, q \leq \infty$,

$$\mathbb{E} \|X_A \colon \ell_p^n \to \ell_q^m \| \geq \frac{c}{2\sqrt{2}} \|A \circ A \colon \ell_{p/2}^n \to \ell_{q/2}^m \|^{1/2}.$$

Using duality (1.12) we immediately obtain the following corollary.

Corollary 5.2 *Let* $X = (X_{ij})_{i \leq m, j \leq n}$ *be as in Proposition 5.1. Then, for all* $1 \leq p, q \leq \infty$,

$$\mathbb{E} \|X_A \colon \ell_p^n \to \ell_q^m \| \geq \frac{c}{2\sqrt{2}} \|(A \circ A)^T \colon \ell_{q^*/2}^m \to \ell_{p^*/2}^n \|^{1/2}.$$

Proof of Proposition 5.1 Let $\| \cdot \|$ denote the operator norm from ℓ_p^n to ℓ_q^m. For $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$, let us denote by E_{ij} the $m \times n$ matrix with entry

1 at the intersection of ith row and jth column and with all other entries 0. By the symmetrization trick described in Remark 3.4, it suffices to consider matrices X with symmetric entries and prove the assertion with a twice better constant $c/\sqrt{2}$ (note that, also by Remark 3.4, the lower bound for the absolute first moment of the symmetrized entries does not change and is still equal to c).

If X has symmetric independent entries, it has the same distribution as $(\varepsilon_{ij}|X_{ij}|)_{ij}$, where ε_{ij}, $i \le m$, $j \le n$, are i.i.d. Rademacher random variables, independent of all other random variables. Hence, by Jensen's inequality and the contraction principle (Lemma 2.7 applied with $\alpha_{ij} = 1/\mathbb{E}|X_{ij}| \le 1/c$ and $x_{ij} = a_{ij}\mathbb{E}|X_{ij}|E_{ij})$, we get

$$\mathbb{E}\left\|\sum_{i=1}^{m}\sum_{j=1}^{n} X_{ij}a_{ij}E_{ij}\right\| = \mathbb{E}\left\|\sum_{i,j}\varepsilon_{ij}|X_{ij}|a_{ij}E_{ij}\right\| \ge \mathbb{E}\left\|\sum_{i,j}\varepsilon_{ij}\mathbb{E}|X_{ij}|a_{ij}E_{ij}\right\|$$
$$\ge c\,\mathbb{E}\left\|\sum_{i,j}\varepsilon_{ij}a_{ij}E_{ij}\right\|. \tag{5.1}$$

Thus, it suffices to estimate from below $\mathbb{E}\|\sum_{i,j}\varepsilon_{ij}a_{ij}E_{ij}\|$.

Since the ℓ_q norm is unconditional, we obtain from the inequalities of Jensen and Khintchine (see [26]) that

$$\mathbb{E}\left\|\sum_{i=1}^{m}\sum_{j=1}^{n}\varepsilon_{ij}a_{ij}E_{ij}\right\| = \mathbb{E}\sup_{x\in B_p^n}\left\|\left(\sum_{j=1}^{n}a_{ij}\varepsilon_{ij}x_j\right)_{i=1}^{m}\right\|_q = \mathbb{E}\sup_{x\in B_p^n}\left\|\left(\left|\sum_{j=1}^{n}a_{ij}\varepsilon_{ij}x_j\right|\right)_{i=1}^{m}\right\|_q$$

$$\ge \sup_{x\in B_p^n}\left\|\left(\mathbb{E}\left|\sum_{j=1}^{n}a_{ij}\varepsilon_{ij}x_j\right|\right)_{i=1}^{m}\right\|_q$$

$$\underset{\substack{\text{Khintchine's} \\ \text{inequality}}}{\ge} \frac{1}{\sqrt{2}}\sup_{x\in B_p^n}\left\|\left(\left(\sum_{j=1}^{n}a_{ij}^2x_j^2\right)^{1/2}\right)_{i=1}^{m}\right\|_q$$

$$= \frac{1}{\sqrt{2}}\sup_{z\in B_{p/2}^n}\left\|\left(\sum_{j=1}^{n}a_{ij}^2z_j\right)_{i=1}^{m}\right\|_{q/2}^{1/2}$$

$$= \frac{1}{\sqrt{2}}\|A\circ A:\ell_{p/2}^n \to \ell_{q/2}^m\|^{1/2}.$$

This together with the estimate in (5.1) yields the assertion.   □

Since $\|G_A:\ell_p^n \to \ell_q^m\| \ge \max_{i,j}|a_{ij}g_{ij}|$, it suffices to prove the following proposition in order to provide the lower bound in Conjecture 1.

Proposition 5.3 *For the $m \times n$ Gaussian matrix G_A, we have*

$$\mathbb{E}\|G_A:\ell_p^n \to \ell_q^m\| \gtrsim_{p,q} \begin{cases} \max_{j\le n}\sqrt{\ln(j+1)}b_j^{\downarrow} & \text{if } p \le q \le 2, \\ \max_{i\le m}\sqrt{\ln(i+1)}d_i^{\downarrow} & \text{if } 2 \le p \le q, \\ 0 & \text{otherwise}, \end{cases} \tag{5.2}$$

where $b_j = \|(a_{ij})_{i\leq m}\|_{2q/(2-q)}$ and $d_i = \|(a_{ij})_{j\leq n}\|_{2p/(p-2)}$.

Proof Since $B_1^n \subset B_p^n$ for $p \geq 1$ and the b_j's do not depend on p, it suffices to prove the first part of the assertion (in the range $p \leq q \leq 2$) only in the case $p = 1 \leq q \leq 2$. In this case (5.2) follows by Propostion 1.10.

The assertion in the range $2 \leq p \leq q$ follows by duality (1.12). $\qquad\square$

5.2 The proof of Inequalities (1.13) and (1.11)

Let us now show that in the case $q < p$, the third term on the right-hand side in Conjecture 1 is not needed. To this end it suffices to prove (1.13) only in the case $q < 2$, since the case $p > 2$ follows by duality (1.12).

Proposition 5.4 *Whenever $1 \leq q < p \leq \infty$ and $q < 2$, we have*

$$D_2 = \|(A \circ A)^T : \ell_{q^*/2}^m \to \ell_{p^*/2}^n\|^{1/2} \gtrsim_{p,q} \max_{j\leq n} \sqrt{\ln(j+1)}\, b_j^\downarrow, \qquad (5.3)$$

where $b_j = \|(a_{ij})_{i\leq m}\|_{2q/(2-q)}$.

Proof Since the right-hand side of (5.3) does not depend on p, and the left-hand side is non-decreasing with p, we may consider only the case $1 \leq q < p \leq 2$. By permuting the columns of A we may and do assume without loss of generality that the sequence $(b_j)_j$ is non-increasing.

Fix $j_0 \leq n$. Let r be the midpoint of the non-empty interval $(\frac{2-p}{p}, \frac{2-q}{q})$. Take $x = (x_j)_{j\leq n}$ with $x_j = \frac{1}{j^r}$. Since $rp/(2-p) > 1$, we have

$$\sum_{j=1}^n x_j^{p/(2-p)} \leq \sum_{j=1}^\infty \frac{1}{j^{rp/(2-p)}} = C(p,q) < \infty,$$

so $x \in C'(p,q)B_{p/(2-p)}^n = C'(p,q)B_{(p^*/2)^*}^n$. Therefore, the inequality $(q^*/2)^* = q/(2-q) \geq 1$ and the facts that $b_j \geq b_{j_0}$ for all $j \leq j_0$, and that $r < (2-q)/q$ imply

$$\begin{aligned}
D_2^2 &= \sup_{z\in B_{(p^*/2)^*}^n} \left(\sum_{i=1}^m \Big(\sum_{j=1}^n a_{ij}^2 z_j\Big)^{(q^*/2)^*}\right)^{1/(q^*/2)^*}\\
&\gtrsim_{p,q} \left(\sum_{i=1}^m \Big(\sum_{j=1}^{j_0} a_{ij}^2 j^{-r}\Big)^{q/(2-q)}\right)^{(2-q)/q}\\
&\geq \left(\sum_{i=1}^m \sum_{j=1}^{j_0} a_{ij}^{2q/(2-q)} j^{-rq/(2-q)}\right)^{(2-q)/q}\\
&= \left(\sum_{j=1}^{j_0} b_j^{2q/(q-2)} j^{-rq/(2-q)}\right)^{(2-q)/q}
\end{aligned}$$

$$\geq b_{j_0}^2 j_0^{-r+(2-q)/q} \gtrsim_{p,q} b_{j_0}^2 \ln(j_0 + 1).$$

Taking the maximum over all $j_0 \leq n$ completes the proof. $\qquad\square$

Now we turn to the proof of (1.11). Note that it suffices to prove only the first two-sided inequality in (1.11), since the second one follows from it by duality (1.12).

Proposition 5.5 *For all $1 \leq p, q \leq \infty$, we have*

$$\|A \circ A: \ell_{p/2}^n \to \ell_{q/2}^m\|^{1/2} + \mathbb{E} \max_{i \leq m, j \leq n} |a_{ij} g_{ij}|$$

$$\asymp_q \|A \circ A: \ell_{p/2}^n \to \ell_{q/2}^m\|^{1/2} + \max_{i \leq m, j \leq n} \sqrt{\ln(j+1)} a'_{ij}, \qquad (5.4)$$

where the matrix $(a'_{ij})_{i,j}$ is obtained by permuting the columns of the matrix $(|a_{ij}|)_{i,j}$ in such a way that $\max_i a'_{i1} \geq \cdots \geq \max_i a'_{in}$.

Proof By permuting the columns of the matrix A, we can assume that the sequence $(\max_{i \leq m} |a_{ij}|)_{j=1}^n$ is non-increasing. We have

$$\mathbb{E} \max_{i \leq m, j \leq n} |a_{ij} g_{ij}| \leq \mathbb{E} \max_{j \leq n} \left(\max_{i \leq m} |a_{ij} g_{ij}| - \mathbb{E} \max_{i \leq m} |a_{ij} g_{ij}| \right)$$

$$+ \max_{j \leq n} \mathbb{E} \max_{i \leq m} |a_{ij} g_{ij}|. \qquad (5.5)$$

The function $y \mapsto \max_{i \leq m} |a_{ij} y_i|$ is $\max_{i \leq m} |a_{ij}|$-Lipschitz with respect to the Euclidean norm on $\mathbb{R}^m$, so by Gaussian concentration (see, e.g., [41, Chapter 5.1]),

$$\mathbb{P}\left(\max_{i \leq m} |a_{ij} g_{ij}| - \mathbb{E} \max_{i \leq m} |a_{ij} g_{ij}| \geq t\right) \leq \exp\left(-\frac{t^2}{2 \max_{i \leq m} |a_{ij}|}\right)$$

for all $t \geq 0$, $j \leq n$. Thus, Lemma 2.11 and inequality (5.5) imply

$$\mathbb{E} \max_{i \leq m, j \leq n} |a_{ij} g_{ij}| \lesssim \max_{j \leq n} \left(\sqrt{\ln(j+1)} \max_{i \leq m} |a_{ij}| \right) + \max_{j \leq n} \mathbb{E} \max_{i \leq m} |a_{ij} g_{ij}|. \qquad (5.6)$$

We have

$$\max_{j \leq n} \mathbb{E} \max_{i \leq m} |a_{ij} g_{ij}| \leq \max_{j \leq n} \mathbb{E} \left(\sum_{i=1}^m |a_{ij} g_{ij}|^q \right)^{1/q} \leq \gamma_q \max_{j \leq n} \|(a_{ij})_i\|_q$$

$$= \gamma_q \max_{j \leq n} \|(a_{ij}^2)_i\|_{q/2}^{1/2} \leq \gamma_q \|A \circ A: \ell_{p/2}^n \to \ell_{q/2}^m\|^{1/2},$$

which, together with (5.6), provides the asserted upper bound.

On the other hand, if $(a_l^\downarrow)_{l \leq mn}$ denotes the non-increasing rearrangement of the sequence of all absolute values of entries of A, then Lemma 2.12 implies

$$\mathbb{E} \max_{j \leq n} \max_{i \leq m} |a_{ij} g_{ij}| \gtrsim \max_{l \leq mn} \sqrt{\ln(l+1)} a_l^\downarrow \geq \max_{j \leq n} \sqrt{\ln(j+1)} a_j^\downarrow$$

$$\geq \max_{j \leq n} \left(\sqrt{\ln(j+1)} \max_{i \leq m} a'_{ij} \right),$$

which provides the asserted lower bound. □

Note that the above proof shows in fact that

$$\max_{j \leq n} \|(a_{ij})_i\|_q + \mathbb{E} \max_{i \leq m, j \leq n} |a_{ij} g_{ij}|$$

$$\asymp_q \max_{j \leq n} \|(a_{ij})_i\|_q + \max_{i \leq m, j \leq n} \sqrt{\ln(j+1)} a'_{ij},$$

so

$$\max_{j \leq n} \|(a_{ij})_i\|_q + \max_{i \leq m} \|(a_{ij})_j\|_{p*} + \max_{j \leq n, i \leq m} \sqrt{\ln(i+1)} a''_{ij}$$

$$\asymp_q \max_{j \leq n} \|(a_{ij})_i\|_q + \max_{i \leq m} \|(a_{ij})_j\|_{p*} + \max_{i \leq m, j \leq n} \sqrt{\ln(j+1)} a'_{ij}, \quad (5.7)$$

where the matrix $(a''_{ij})_{i,j}$ is obtained by permuting the rows of the matrix $(|a_{ij}|)_{i,j}$ in such a way that $\max_j a''_{1j} \geq \cdots \geq \max_j a''_{mj}$.

5.3 Counterexample to a seemingly natural conjecture

In this subsection we provide an example showing that for any $p \leq q < 2$ the bound

$$\mathbb{E}\|G_A : \ell_p^n \to \ell_q^m\| \lesssim_{p,q} \|A \circ A : \ell_{p/2}^n \to \ell_{q/2}^m\|^{1/2} + \|(A \circ A)^T : \ell_{q*/2}^m \to \ell_{p*/2}^n\|^{1/2}$$

$$+ \mathbb{E} \max_{i \leq m, j \leq n} |a_{ij} g_{ij}|. \quad (5.8)$$

cannot hold. By duality (1.12), it also cannot hold for any $2 < p \leq q$. This explains that Conjecture 1 cannot be simplified into a form like on the right-hand side of (1.8).

Let $p \leq q < 2, k, N \in \mathbb{N}$, and let $A_1, \ldots, A_N$ be $k \times k$ matrices with all entries equal to one. Consider a block matrix

$$A = \begin{pmatrix} A_1 & & & 0 \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_N \end{pmatrix}$$

of size $kN \times kN$, with blocks $A_1, \ldots A_N$ on the diagonal and with all other entries equal to 0.

Note that since $p \leq q \leq 2$,

$$\|A \circ A : \ell_{p/2}^{kN} \to \ell_{q/2}^{kN}\| = \max_{l \leq N} \|A_l \circ A_l : \ell_{p/2}^k \to \ell_{q/2}^k\| = \|A_1 \circ A_1 : \ell_{p/2}^k \to \ell_{q/2}^k\|$$

$$= \sup_{x \in B_{p/2}^k} \Big(\sum_{i=1}^k \Big| \sum_{j=1}^k x_i \Big|^{q/2} \Big)^{2/q} = \sup_{x \in B_{p/2}^k} k^{2/q} \Big| \sum_{i=1}^k x_i \Big| = k^{2/q},$$

and similarly, since $2 \le q^* \le p^*$,

$$\|(A \circ A)^T : \ell_{q^*/2}^{kN} \to \ell_{p^*/2}^{kN}\| = \|(A_1 \circ A_1)^T : \ell_{q^*/2}^k \to \ell_{p^*/2}^k\| = k^{2/p^*+1-2/q^*}.$$

The two bounds above and Lemma 2.10 imply that the right-hand side of (5.8) is bounded from above by

$$C\Big(k^{1/q} + k^{1/p^*+1/2-1/q^*} + \sqrt{\ln(kN)}\Big). \tag{5.9}$$

On the other hand, since for all $j \le kN$, $\|(a_{ij})_i\|_{2q/(2-q)} = k^{(2-q)/(2q)}$, we obtain from the lower bound (5.2) that

$$\mathbb{E}\|G_A : \ell_p^{kN} \to \ell_q^{kN}\| \gtrsim \sqrt{\ln(kN)}k^{(2-q)/(2q)}. \tag{5.10}$$

If we take $N \asymp e^{e^k}$, then (5.10) is of larger order than (5.9) as $k \to \infty$, so (5.8) cannot hold.

5.4 Discussion of another natural conjecture

In this subsection we prove all the assertions of Remark 1.1. We begin by showing that for every $1 \le p \le 2 \le q \le \infty$,

$$D_1 + D_2 + \mathbb{E}\max_{i,j}|a_{ij}g_{ij}| \asymp_{p,q} \mathbb{E}\max_{i \le m}\|(a_{ij}g_{ij})_j\|_{p^*} + \mathbb{E}\max_{j \le n}\|(a_{ij}g_{ij})_i\|_q, \tag{5.11}$$

and, in the case $p, q \ge 2$,

$$\mathbb{E}\max_{i \le m}\|(a_{ij}g_{ij})_j\|_{p^*} + \mathbb{E}\max_{j \le n}\|(a_{ij}g_{ij})_i\|_q$$
$$\lesssim_{p,q} \max_{i \le m}\|(a_{ij})_j\|_{p^*} + \max_{j \le n}\|(a_{ij})_i\|_q + \max_{i \le m}\sqrt{\ln(i+1)}d_i^{\downarrow}, \tag{5.12}$$

where $D_1 = \|A \circ A : \ell_{p/2}^n \to \ell_{q/2}^m\|^{1/2}$, $D_2 = \|(A \circ A)^T : \ell_{q^*/2}^m \to \ell_{p^*/2}^n\|^{1/2}$, and $d_i = \|(a_{ij})_{j \le n}\|_{2p/(p-2)}$. In other words, (5.11) shows that Conjecture 1 is equivalent to (1.15) as long as $1 \le p \le 2 \le q \le \infty$.

Proof of (5.11) and (5.12) Fix $i \le m$ and let $f(x) = \|(a_{ij}x_j)_j\|_{p^*}$ for $x \in \mathbb{R}^n$. For $p \ge 2$ we have $p^*(2/p^*)^* = 2p/(p-2)$. Thus f is Lipschitz continuous with constant L_i equal to

$$\sup_{x \in B_2^n} \Big(\sum_{j=1}^n |a_{ij}x_j|^{p^*} \Big)^{1/p^*} = \sup_{y \in B_{2/p^*}^n} \Big(\sum_{j=1}^n |a_{ij}|^{p^*}y_j \Big)^{1/p^*} = \begin{cases} \max_{j \le n}|a_{ij}| & \text{if } p \le 2, \\ \|(a_{ij})_j\|_{2p/(p-2)} & \text{if } p \ge 2. \end{cases}$$

Therefore, the Gaussian concentration inequality (see, e.g., [41, Chapter 5.1]) implies that for every $t \geq 0$ and every $i \leq m$,

$$\mathbb{P}\Big(\|(a_{ij}g_{ij})_j\|_{p^*} - \mathbb{E}\|(a_{ij}g_{ij})_j\|_{p^*} \geq t \Big) \leq e^{-t^2/2L_i^2},$$

so by Lemma 2.11 we get

$$\mathbb{E}\max_{i \leq m}\Big(\|(a_{ij}g_{ij})_j\|_{p^*} - \mathbb{E}\|(a_{ij}g_{ij})_j\|_{p^*} \Big)$$
$$\lesssim \begin{cases} \max_{i \leq m}\max_{j \leq n}\sqrt{\ln(i+1)}a''_{ij} & \text{if } p \leq 2, \\ \max_{i \leq m}\sqrt{\ln(i+1)}d_i^{\downarrow} & \text{if } p \geq 2, \end{cases} \tag{5.13}$$

where the matrix $(a''_{ij})_{i,j}$ is obtained by permuting the rows of the matrix $(|a_{ij}|)_{i,j}$ in such a way that $\max_j a''_{1j} \geq \cdots \geq \max_j a''_{mj}$.

Moreover, by Jensen's inequality,

$$\mathbb{E}\|(a_{ij}g_{ij})_j\|_{p^*} \leq \big(\mathbb{E}\|(a_{ij}g_{ij})_j\|_{p^*}^{p^*}\big)^{1/p^*} = \Big(\mathbb{E}\sum_{j=1}^n |a_{ij}g_{ij}|^{p^*}\Big)^{1/p^*} = \gamma_{p^*}\|(a_{ij})_j\|_{p^*}.$$

This together with the triangle inequality and (5.13) implies

$$\mathbb{E}\max_{i \leq m}\|(a_{ij}g_{ij})_j\|_{p^*}$$
$$\lesssim_p \max_{i \leq m}\|(a_{ij})_j\|_{p^*} + \begin{cases} \max_{i \leq m}\max_{j \leq n}\sqrt{\ln(i+1)}a''_{ij} & \text{if } p \leq 2, \\ \max_{i \leq m}\sqrt{\ln(i+1)}d_i^{\downarrow} & \text{if } p \geq 2, \end{cases}$$

and, by duality,

$$\mathbb{E}\max_{j \leq n}\|(a_{ij}g_{ij})_i\|_q$$
$$\lesssim_q \max_{j \leq n}\|(a_{ij})_i\|_q + \begin{cases} \max_{j \leq n}\max_{i \leq m}\sqrt{\ln(j+1)}a'_{ij} & \text{if } q \geq 2, \\ \max_{j \leq n}\sqrt{\ln(j+1)}b_j^{\downarrow} & \text{if } q \leq 2, \end{cases}$$

where $b_j = \|(a_{ij})_i\|_{2q/(2-q)}$, and the matrix $(a'_{ij})_{i,j}$ is obtained by permuting the columns of the matrix $(|a_{ij}|)_{i,j}$ in such a way that $\max_i a'_{i1} \geq \cdots \geq \max_i a'_{in}$. This, together with Lemma 2.1 and (5.4) yields in the case $p \leq 2 \leq q$,

$$\mathbb{E}\max_{i \leq m}\|(a_{ij}g_{ij})_j\|_{p^*} + \mathbb{E}\max_{j \leq n}\|(a_{ij}g_{ij})_i\|_q \lesssim_{p,q} D_1 + D_2 + \mathbb{E}\max_{i,j}|a_{ij}g_{ij}|,$$

what implies the lower bound of (5.11). In the case $2 < p, q$ we additionally use (5.7) and the simple observation that

$$\max_{i \leq m}\max_{j \leq n}\sqrt{\ln(i+1)}a''_{ij} \leq \max_{i \leq m}\sqrt{\ln(i+1)}d_i^{\downarrow}$$

to get (5.12).

Now we move to the proof of the upper bound of (5.11) in the case $p \leq 2 \leq q$. Since the ℓ_{p*}^n norm is unconditional, we have by Jensen's inequality and Lemma 2.1

$$\mathbb{E} \max_{i \leq m} \|(a_{ij}g_{ij})_j\|_{p*} = \mathbb{E} \max_{i \leq m} \|(|a_{ij}g_{ij}|)_j\|_{p*} \geq \max_{i \leq m} \|(|a_{ij}|\mathbb{E}|g_{ij}|)_j\|_{p*}$$
$$= \sqrt{2/\pi} \max_{i \leq m} \|(|a_{ij}|)_j\|_{p*} = \sqrt{2/\pi} D_2,$$

and dually

$$\mathbb{E} \max_{j \leq n} \|(a_{ij}g_{ij})_i\|_q \geq \sqrt{2/\pi} D_1.$$

Moreover, since $\|\cdot\|_q \geq \|\cdot\|_\infty$,

$$\mathbb{E} \max_{j \leq n} \|(a_{ij}g_{ij})_i\|_q \geq \mathbb{E} \max_j \max_i |a_{ij}g_{ij}|,$$

which finishes the proof of the upper bound of (5.11).   □

Next, for every pair $(p, q) \in [1, \infty]^2$ which does not satisfy the condition $1 \leq p \leq 2 \leq q \leq \infty$ we shall give examples of $m, n \in \mathbb{N}$, and $m \times n$ matrices A, for which

$$\mathbb{E}\|G_A : \ell_p^n \to \ell_q^m\| \gg \mathbb{E} \max_{i \leq m} \|(a_{ij}g_{ij})_j\|_{p*} + \mathbb{E} \max_{j \leq n} \|(a_{ij}g_{ij})_i\|_q \qquad (5.14)$$

when $m, n \to \infty$. This shows that the natural conjecture (1.15) is wrong outside the range $1 \leq p \leq 2 \leq q \leq \infty$. The case $p = 2 = q$, when (1.15) is valid (cf. (1.4)), is in a sense a boundary case, for which (1.15) (i.e., a natural generalization of (1.4)) may hold.

Example 5.6 (for (5.14) in the case $q < p$.) Let $m = n$, and $A = \mathrm{Id}_n$. Then by Lemmas 2.10 and 2.12 we have

$$\mathbb{E} \max_{i \leq m} \|(a_{ij}g_{ij})_j\|_{p*} + \mathbb{E} \max_{j \leq n} \|(a_{ij}g_{ij})_i\|_q = 2 \max_{i \leq n} |g_{ii}| \asymp \sqrt{\ln n},$$

whereas Proposition 5.1 and our assumption $p/q > 1$ imply

$$\mathbb{E}\|G_A : \ell_p^n \to \ell_q^n\| \gtrsim \|\mathrm{Id}_n : \ell_{p/2}^n \to \ell_{q/2}^n\|^{1/2} = \sup_{x \in B_{p/2}^n} \left(\sum_{i=1}^n |x_i|^{q/2}\right)^{1/q}$$
$$= \left(\sup_{y \in B_{p/q}^n} \sum_{i=1}^n |y_i|\right)^{1/q} = \left(n^{1/(p/q)^*}\right)^{1/q} \gg \sqrt{\ln n}.$$

Since cases $2 < p \leq q$ and $p \leq q < 2$ are dual (see (1.12)), we give an example for which (5.14) holds only in the first case.

Example 5.7 (for (5.14) in the case $2 < p \leq q$.) Fix p and q satisfying $2 < p \leq q$. Let $m, n \to \infty$ be such that $m^{1/q} \gg n^{1/p^*}$, and let A be an $m \times n$ matrix with all entries equal to 1. For $p > 2$ we have $2(p/2)^* = 2p/(p-2)$. This together with (5.12) implies

$$\mathbb{E} \max_{i \leq m} \|(a_{ij} g_{ij})_j\|_{p^*} + \mathbb{E} \max_{j \leq n} \|(a_{ij} g_{ij})_i\|_q$$

$$\lesssim_{p,q} \max_{i \leq m} \|(a_{ij})_j\|_{p^*} + \max_{j \leq n} \|(a_{ij})_i\|_q + \max_{i \leq m} \sqrt{\ln(i+1)} d_i^{\downarrow}$$

$$= n^{1/p^*} + m^{1/q} + \sqrt{\ln(m+1)} n^{(p-2)/2p} \lesssim m^{1/q} + \sqrt{\ln m}\, n^{\frac{1}{2(p/2)^*}}.$$

On the other hand, Proposition 5.1 and our assumption $p/2 > 1$ imply

$$\mathbb{E}\|G_A : \ell_p^n \to \ell_q^n\| \gtrsim \|A : \ell_{p/2}^n \to \ell_{q/2}^n\|^{1/2} = \sup_{x \in B_{p/2}^n} \left(\sum_{i=1}^m \left| \sum_{j=1}^n x_j \right|^{q/2} \right)^{1/q}$$

$$= m^{1/q} \sup_{x \in B_{p/2}^n} \left(\left| \sum_{j=1}^n x_j \right| \right)^{1/2} = m^{1/q} n^{\frac{1}{2(p/2)^*}}$$

$$\gg m^{1/q} + \sqrt{\ln m}\, n^{\frac{1}{2(p/2)^*}}.$$

5.5 Infinite dimensional Gaussian operators

In this subsection we prove Proposition 1.2 concerning infinite dimensional Gaussian operators. It allows us to see that Conjecture 1 implies Conjecture 2.

Proof of Proposition 1.2 We adapt the proof of [40, Corollary 1.2] to prove Proposition 1.2 in the case $p \leq 2 \leq q$—remaining cases may be proven similarly. Fix $1 \leq p \leq 2 \leq q \leq \infty$ for which (1.14) holds and a deterministic infinite matrix $A = (a_{ij})_{i,j \in \mathbb{N}}$. Using the monotone convergence theorem one can show that a matrix $B = (b_{ij})_{i,j \in \mathbb{N}}$ defines a bounded operator between $\ell_p(\mathbb{N})$ and $\ell_q(\mathbb{N})$ if an only if $\sup_{n \in \mathbb{N}} \|(b_{ij})_{i,j \leq n} : \ell_p^n \to \ell_q^n\| < \infty$. Interpreting $\|B : \ell_p(\mathbb{N}) \to \ell_q(\mathbb{N})\|$ as infinity for matrices which do not define a bounded operator, we have

$$\mathbb{E}\|G_A : \ell_p(\mathbb{N}) \to \ell_q(\mathbb{N})\| = \mathbb{E} \sup_{x \in B_p^\infty} \left(\sum_{i=1}^\infty \left| \sum_{j=1}^\infty a_{ij} g_{ij} x_j \right|^q \right)^{1/q}$$

$$= \mathbb{E} \lim_{n \to \infty} \sup_{x \in B_p^n} \left(\sum_{i=1}^n \left| \sum_{j=1}^n a_{ij} g_{ij} x_j \right|^q \right)^{1/q} = \lim_{n \to \infty} \mathbb{E} \sup_{x \in B_p^n} \left(\sum_{i=1}^n \left| \sum_{j=1}^n a_{ij} g_{ij} x_j \right|^q \right)^{1/q}$$

$$= \lim_{n \to \infty} \mathbb{E} \|(g_{ij} a_{ij})_{i,j \leq n} : \ell_p^n \to \ell_q^n\|$$

and similarly

$$\|A \circ A : \ell_{p/2}(\mathbb{N}) \to \ell_{q/2}(\mathbb{N})\| = \lim_{n \to \infty} \|(a_{ij}^2)_{i,j \leq n} : \ell_{p/2}^n \to \ell_{q/2}^n\|,$$

$$\|(A \circ A)^T : \ell_{q^*/2}(\mathbb{N}) \to \ell_{p^*/2}(\mathbb{N})\| = \lim_{n \to \infty} \|(a_{ji}^2)_{i,j \leq n} : \ell_{q^*/2}^n \to \ell_{p^*/2}^n\|,$$

and

$$\mathbb{E} \sup_{i,j \in \mathbb{N}} |a_{ij} g_{ij}| = \lim_{n \to \infty} \mathbb{E} \sup_{i,j \leq n} |a_{ij} g_{ij}|.$$

Therefore, (1.14) implies the following: $\mathbb{E}\|G_A : \ell_p(\mathbb{N}) \to \ell_q(\mathbb{N})\| < \infty$ if and only if $\|A \circ A : \ell_{p/2}(\mathbb{N}) \to \ell_{q/2}(\mathbb{N})\| < \infty$, $\|(A \circ A)^T : \ell_{q^*/2}(\mathbb{N}) \to \ell_{p^*/2}(\mathbb{N})\| < \infty$, and $\mathbb{E} \sup_{i,j \in \mathbb{N}} |a_{ij} g_{ij}| < \infty$. It thus suffices to prove the following claim: $\|G_A : \ell_p(\mathbb{N}) \to \ell_q(\mathbb{N})\| < \infty$ almost surely if and only if $\mathbb{E}\|G_A : \ell_p(\mathbb{N}) \to \ell_q(\mathbb{N})\| < \infty$.

If $\mathbb{P}(\|G_A : \ell_p(\mathbb{N}) \to \ell_q(\mathbb{N})\| < \infty) < 1$, then $\mathbb{P}(\|G_A : \ell_p(\mathbb{N}) \to \ell_q(\mathbb{N})\| = \infty) > 0$, so $\mathbb{E}\|G_A : \ell_p(\mathbb{N}) \to \ell_q(\mathbb{N})\| = \infty$.

Assume now that $\mathbb{P}(\|G_A : \ell_p(\mathbb{N}) \to \ell_q(\mathbb{N})\| < \infty) = 1$. By (4.23) and (4.24) we know that for every $n \in \mathbb{N}$ there exist finite sets S_n and T_n such that

$$\|G_A : \ell_p(\mathbb{N}) \to \ell_q(\mathbb{N})\| = \sup_{n \in \mathbb{N}} \sup_{x \in B_p^n, y \in B_{q^*}^n} \sum_{i=1}^{n} \sum_{j=1}^{n} y_i a_{ij} g_{ij} x_j$$

$$\asymp \sup_{n} \sup_{x \in S_n, y \in T_n} \sum_{i=1}^{n} \sum_{j=1}^{n} y_i a_{ij} g_{ij} x_j \qquad \text{a.s.}$$

In particular, there exist Gaussian random variables $(\Gamma_k)_{k \in \mathbb{N}}$ such that

$$\|G_A : \ell_p(\mathbb{N}) \to \ell_q(\mathbb{N})\| \asymp \sup_{k \in \mathbb{N}} \Gamma_k \qquad \text{a.s.}$$

Therefore, we may apply [35, (1.2)] to see that there exists $\varepsilon > 0$ such that $\mathbb{E} \exp(\varepsilon \|G_A : \ell_p(\mathbb{N}) \to \ell_q(\mathbb{N})\|^2) < \infty$, so $\mathbb{E}\|G_A : \ell_p(\mathbb{N}) \to \ell_q(\mathbb{N})\| < \infty$, which completes the proof of the claim. $\qquad \square$

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

1. Achlioptas, D., Mcsherry, F.: Fast computation of low-rank matrix approximations, J. ACM **54**(2), 9-es, (2007)
2. Adamczak, R., Latała, R., Litvak, A.E., Pajor, A., Tomczak-Jaegermann, N.: Chevet type inequality and norms of submatrices. Studia Math. **210**(1), 35–56 2012)
3. Adamczak, R., Latała, R., Puchała, Z., Życzkowski, K.: Asymptotic entropic uncertainty relations, J. Math. Phys. **57**(3), 032204, 24, (2016)
4. Ailon, N., Chazelle, B.: The fast Johnson-Lindenstrauss transform and approximate nearest neighbors. SIAM J. Comput. **39**(1), 302–322 (2009)
5. Akemann, G., Baik, J., Di Francesco, P. (eds.): The Oxford handbook of random matrix theory. Oxford University Press, Oxford (2015)
6. Anderson, G.W., Guionnet, A., Zeitouni, O.: An introduction to random matrices, Cambridge Studies in Advanced Mathematics, vol. 118. Cambridge University Press, Cambridge (2010)
7. Bandeira, A.S., van Handel, R.: Sharp nonasymptotic bounds on the norm of random matrices with independent entries. Ann. Probab. **44**(4), 2479–2506 (2016)
8. Bennett, G.: Schur multipliers. Duke Math. J. **44**(3), 603–639 (1977)
9. Bennett, G., Goodman, V., Newman, C.M.: Norms of random matrices. Pac. J. Math. **59**(2), 359–365 (1975)
10. Benyamini, Y., Gordon, Y.: Random factorization of operators between Banach spaces. J. Anal. Math. **39**, 45–74 (1981)
11. Boucheron, S., Lugosi, G., Massart, P.: Concentration inequalities, Oxford University Press, Oxford. A nonasymptotic theory of independence, With a foreword by Michel Ledoux (2013)
12. Bourgain, J., Dirksen, S., Nelson, J.: Toward a unified theory of sparse dimensionality reduction in Euclidean space. Geom. Funct. Anal. **25**(4), 1009–1088 (2015)
13. Carl, B., Maurey, B., Puhl, J.: Grenzordnungen von absolut-(r, p)-summierenden Operatoren. Math. Nachr. **82**, 205–218 (1978)
14. Chafaï, D., Guédon, O., Lecué, G., Pajor, A.: Interactions between compressed sensing random matrices and high dimensional geometry, Panoramas et Synthèses [Panoramas and Syntheses], vol. 37. Société Mathématique de France, Paris (2012)
15. Davidson, K.R., Szarek, S.J.: Local Operator Theory, Random Matrices and Banach Spaces, Handbook of the Geometry of Banach Spaces, vol. I, pp. 317–366. North-Holland, Amsterdam (2001)
16. Foucart, S., Rauhut, H.: A Mathematical Introduction to Compressive Sensing. Applied and Numerical Harmonic Analysis, Birkhäuser/Springer, New York (2013)
17. Friedland, O., Youssef, P.: Approximating matrices and convex bodies, Int. Math. Res. Not. IMRN (8), 2519–2537, (2019)
18. Gluskin, E. D.: Norms of random matrices and diameters of finite-dimensional sets, Mat. Sb. (N.S.) **120(162)**(2), 180–189, 286, (1983)
19. Gluskin, E.D., Kwapień, S.: Tail and moment estimates for sums of independent random variables with logarithmically concave tails. Studia Math. **114**(3), 303–309 (1995)
20. Goldstine, H.H., von Neumann, J.: Numerical inverting of matrices of high order. II. Proc. Am. Math. Soc. **2**, 188–202 (1951)
21. Gordon, Y.: Some inequalities for Gaussian processes and applications. Israel J. Math. **50**(4), 265–289 (1985)
22. Gordon, Y., Litvak, A.E., Schütt, C., Werner, E.M.: Geometry of spaces between polytopes and related zonotopes. Bull. Sci. Math. **126**(9), 733–762 (2002)
23. Guédon, O., Hinrichs, A., Litvak, A. E., Prochno, J.: On the expectation of operator norms of random matrices, Geometric aspects of functional analysis, Lecture Notes in Math., vol. 2169, Springer, Cham, (2017), pp. 151–162
24. Guédon, O., Mendelson, S., Pajor, A., Tomczak-Jaegermann, N.: Majorizing measures and proportional subsets of bounded orthonormal systems. Rev. Mat. Iberoam. **24**(3), 1075–1095 (2008)
25. Guédon, O., Rudelson, M.: L_p-moments of random vectors via majorizing measures. Adv. Math. **208**(2), 798–823 (2007)
26. Haagerup, U.: The best constants in the Khintchine inequality, Studia Math. **70**(3) (1981), 231–283 (1982)

27. Hinrichs, A., Krieg, D., Novak, E., Prochno, J., Ullrich, M.: On the power of random information, Multivariate Algorithms and Information-Based Complexity (F. J. Hickernell and P. Kritzer, eds.), De Gruyter, Berlin/Boston, (1994), pp. 43–64

28. Hinrichs, A., Krieg, D., Novak, E., Prochno, J., Ullrich, M.: Random sections of ellipsoids and the power of random information. Trans. Am. Math. Soc. **374**(12), 8691–8713 (2021)

29. Hinrichs, A., Prochno, J., Sonnleitner, M.: Random sections of ℓ_p-ellipsoids, optimal recovery and Gelfand numbers of diagonal operators, (2021)

30. Hinrichs, A., Prochno, J., Vybíral, J.: Gelfand numbers of embeddings of Schatten classes. Math. Ann. **380**(3–4), 1563–1593 (2021)

31. Hitczenko, P., Montgomery-Smith, S.J., Oleszkiewicz, K.: Moment inequalities for sums of certain independent symmetric random variables. Studia Math. **123**(1), 15–42 (1997)

32. Hoeffding, W.: Probability inequalities for sums of bounded random variables. J. Am. Statist. Assoc. **58**, 13–30 (1963)

33. Krieg, D., Ullrich, M.: Function values are enough for L_2-approximation. Found. Comput. Math. **21**(4), 1141–1151 (2021)

34. Kwapień, S.: Decoupling inequalities for polynomial chaos. Ann. Probab. **15**(3), 1062–1071 (1987)

35. Landau, H.J., Shepp, L.A.: On the supremum of a Gaussian process. Sankhyā Ser. A **32**, 369–378 (1970)

36. Latała, R.: Tail and moment estimates for sums of independent random vectors with logarithmically concave tails. Studia Math. **118**(3), 301–304 (1996)

37. Latała, R.: Some estimates of norms of random matrices. Proc. Am. Math. Soc. **133**(5), 1273–1282 (2005)

38. Latała, R., Strzelecka, M.: Comparison of weak and strong moments for vectors with independent coordinates. Mathematika **64**(1), 211–229 (2018)

39. Latała, R., Świątkowski, W.: Norms of randomized circulant matrices, Electron. J. Probab. **27**, Paper No. 80, 23, (2022)

40. Latała, R., van Handel, R., Youssef, P.: The dimension-free structure of nonhomogeneous random matrices. Invent. Math. **214**(3), 1031–1080 (2018)

41. Ledoux, M.: The Concentration of Measure Phenomenon, Mathematical Surveys and Monographs, vol. 89. American Mathematical Society, Providence, RI (2001)

42. Ledoux, M.: Deviation inequalities on largest eigenvalues, Geometric aspects of functional analysis, Lecture Notes in Math., Springer. Berlin **2007**, 167–219 (1910)

43. Ledoux, M., Talagrand, M.: Probability in Banach spaces, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], vol. 23, Springer-Verlag, Berlin, (1991), Isoperimetry and processes

44. Linial, N., London, E., Rabinovich, Y.: The geometry of graphs and some of its algorithmic applications. Combinatorica **15**(2), 215–245 (1995)

45. Matlak, D.: Oszacowania norm macierzy losowych, Master's thesis, Uniwersytet Warszawski, (2017)

46. Naor, A., Regev, O., Vidick, T.: Efficient rounding for the noncommutative Grothendieck inequality. Theory Comput. **10**, 257–295 (2014)

47. Puchała, Z., Rudnicki, Ł., Życzkowski, K.: Majorization entropic uncertainty relations, J. Phys. A **46**(27), 272002, 12, (2013)

48. Rauhut, H.: Compressive sensing and structured random matrices, Theoretical foundations and numerical methods for sparse recovery, Radon Ser. Comput. Appl. Math., vol. 9, Walter de Gruyter, Berlin, (2010), pp. 1–92

49. Riemer, S., Schütt, C.: On the expectation of the norm of random matrices with non-identically distributed entries. Electron. J. Probab. **18**(29), 13 (2013)

50. Rudelson, M., Vershynin, R.: Sampling from large matrices: an approach through geometric functional analysis, J. ACM **54**(4), Art. 21, 19, (2007)

51. Seginer, Y.: The expected norm of random matrices. Combin. Probab. Comput. **9**(2), 149–166 (2000)

52. Slepian, D.: The one-sided barrier problem for Gaussian noise. Bell Syst. Tech. J. **41**, 463–501 (1962)

53. So, A. M.-C.: Moment inequalities for sums of random matrices and their applications in optimization, Math. Program. **130**(1), Ser. A, 125–151, (2011)

54. Spielman, D. A., Teng, S.-H.: Nearly-linear time algorithms for graph partitioning, graph sparsification, and solving linear systems, Proceedings of the Thirty-Sixth Annual ACM Symposium on Theory of Computing (New York, NY, USA), STOC '04, Association for Computing Machinery, (2004), p. 81-90

55. Strzelecka, M.: Estimates of norms of log-concave random matrices with dependent entries, Electron. J. Probab. **24**, Paper No. 107, 15, (2019)
56. Talagrand, M.: A new look at independence. Ann. Probab. **24**(1), 1–34 (1996)
57. Tropp, J.A.: Norms of random submatrices and sparse approximation. C. R. Math. Acad. Sci. Paris **346**(23–24), 1271–1274 (2008)
58. Tropp, J.A.: On the conditioning of random subdictionaries. Appl. Comput. Harmon. Anal. **25**(1), 1–24 (2008)
59. Tropp, J.A.: User-friendly tail bounds for sums of random matrices. Foundations of Computational Mathematics **12**(4), 389–434 (2012)
60. Tropp, J. A.: An introduction to matrix concentration inequalities, Foundations and Trends® in Machine Learning **8**(1-2), 1–230, (2015)
61. van Handel, R.: On the spectral norm of Gaussian random matrices. Trans. Am. Math. Soc. **369**(11), 8161–8178 (2017)
62. van Handel, R.: Structured random matrices, Convexity and concentration, IMA Vol. Math. Appl., vol. 161, Springer, New York, (2017), pp. 107–156
63. Vershynin, R.: Introduction to the Non-asymptotic Analysis of Random Matrices, Compressed Sensing, pp. 210–268. Cambridge Univ. Press, Cambridge (2012)
64. Vershynin, R: High-dimensional probability, Cambridge Series in Statistical and Probabilistic Mathematics, vol. 47, Cambridge University Press, Cambridge, (2018), An introduction with applications in data science, With a foreword by Sara van de Geer
65. von Neumann, J., Goldstine, H.H.: Numerical inverting of matrices of high order. Bull. Am. Math. Soc. **53**(11), 1021–1099 (1947)
66. Wigner, E.P.: Characteristic vectors of bordered matrices with infinite dimensions. Ann. Math. (2) **62**, 548–564 (1955)
67. Wigner, E.P.: Characteristic vectors of bordered matrices with infinite dimensions. II. Ann. Math. **2**(65), 203–207 (1957)
68. Wigner, E.P.: On the distribution of the roots of certain symmetric matrices. Ann. Math. **2**(67), 325–327 (1958)
69. Wishart, J.: The generalised product moment distribution in samples from a normal multivariate population. Biometrika **20A**(1/2), 32–52 (1928)