OPERATOR $\ell_p \to \ell_q$ NORMS OF GAUSSIAN MATRICES

RAFAŁ LATAŁA AND MARTA STRZELECKA

Abstract. We confirm the conjecture posed by Guédon, Hinrichs, Litvak, and Prochno in 2017 that $\mathbb{E}\|(a_{ij}g_{ij})_{i\le m,j\le n}\colon \ell_p^n \to \ell_q^m\|$ is comparable, up to constants depending only on p and q, to

$$\max_i \|(a_{ij})_j\|_{p^*} + \max_j \|(a_{ij})_i\|_q + \mathbb{E}\max_{i,j}|a_{ij}g_{ij}|$$

provided that $1 \le p < 2 < q$ or $4/3 < p \le 2 \le q < 4$. In the remaining cases, when $p = 2$ and $q \ge 4$ or $1 \le p \le 4/3$ and $q = 2$, we prove it up to a factor of order $\log\log(mn)$.

1. Introduction

Let $A = (a_{ij})_{i\le m,j\le n}$ be a deterministic $m \times n$ matrix and let $p, q \in [1,\infty]$. In this paper we study $\ell_p^n \to \ell_q^m$ norms of centered structured Gaussian random matrices $G_A = (a_{ij}g_{ij})_{i\le m,j\le n}$ with a variance profile $A \circ A = (a_{ij}^2)_{i\le m,j\le n}$, i.e., quantities of the form

$$\|G_A\|_{p\to q} = \|G_A\|_{\ell_p^n\to\ell_q^m} = \sup\Big\{\sum_{i=1}^m\sum_{j=1}^n a_{ij}g_{ij}s_it_j \colon s \in B_{q^*}^m, t \in B_p^n\Big\},$$

where random variables g_{ij} are iid standard Gaussians, q^* denotes the Hölder conjugate of q, i.e., the unique number from $[1,\infty]$ satisfying $\frac{1}{q} + \frac{1}{q^*} = 1$, and B_p^n is the unit ball in the ℓ_p-norm in $\mathbb{R}^n$.

Although the behaviour of random matrices with iid entries is quite well understood, it is not the case for the random matrices with non-trivial variance profile, whose $\ell_p^n \to \ell_q^m$ norms appear naturally in many problems in applied mathematics. However, much effort was made recently to understand $\ell_p^n \to \ell_q^m$ norms of structured random matrices (cf. [2, 17, 12, 6, 18, 11, 13, 1, 4, 15]).

In this paper we focus on two-sided estimates (i.e., lower and upper bounds matching up to a multiplicative constant) for the expectation of $\|G_A\|_{p\to q}$. Such bounds encode much more information than only an order of $\mathbb{E}\|G_A\|_{p\to q}$. They imply two-sided estimates on higher moments and tail bounds for $\|G_A\|_{p\to q}$ (see Corollary 12 below). Moreover, they yield a condition for an infinite Gaussian matrix to be a bounded operator from ℓ_p to ℓ_q (see Corollary 7 below). We also discuss how to generalize the estimates for $\|G_A\|_{p\to q}$ to more general classes of random matrices with independent, but not necessarily Gaussian entries.

Before we move further, let us introduce some more notation. For two nonnegative functions f and g we write $f \lesssim g$ (or $g \gtrsim f$) if there exists an absolute constant C such that $f \le Cg$; the notation $f \sim g$ means that $f \lesssim g \lesssim f$. We write $\lesssim_\alpha$, $\sim_{K,\gamma}$, etc. if the underlying constant depends on the parameters given in the subscripts. Whenever we write $p \ge p_1$ or $p \le p_2$ we mean that $p \in [p_1,\infty]$ or $p \in [1,p_2]$, respectively. By $[m]$ we denote the set $\{1,\dots,m\}$ of the first m positive integers. Let us also denote

$$\operatorname{Log} x = 1 \vee \ln x \quad \text{for } x > 0, \quad \text{and } \operatorname{Log} 0 = 1.$$

If $p = 2 = q$, the $\ell_p^n \to \ell_q^m$ norm coincides with the spectral norm and it is known by [12] that

$$\mathbb{E}\|G_A\|_{2\to 2} \sim \max_i \|(a_{ij})_j\|_2 + \max_j \|(a_{ij})_i\|_2 + \mathbb{E}\max_{i,j} |a_{ij}g_{ij}| \\ \sim \max_i \|(a_{ij}g_{ij})_j\|_2 + \max_i \|(a_{ij}g_{ij})_j\|_2.$$

Moreover, two-sided bounds are also known for extremal values of (p, q), i.e., when $p \in \{1, \infty\}$ or $q \in \{1, \infty\}$ (see [6, Remark 1.4] and [1, Propositions 1.8 and 1.10]). The question whether similar two-sided inequalities hold for other ranges of p and q with arbitrary A was, up to now, entirely open; all known bounds match only up to a logarithmic constant (see [6, 1]) or are valid only in some very special cases (for the trivial structure, i.e., when $a_{ij} = 1$ for all i, j, or, more generally, for tensor structures – this follows by the Chevet inequality). We refer to the introductions to [1] and [9] for more details and an overview of the history of the problem.

From now on, we will consider only the case $1 \le p \le 2 \le q \le \infty$. The following conjecture was formulated in [6] (see [1] for a discussion of other ranges of p and q).

Conjecture 1. *For every $p \le 2 \le q$ and every deterministic $m \times n$ matrix $A = (a_{ij})_{i\le m, j\le n}$,*

$$\mathbb{E}\|G_A\|_{p\to q} \sim_{p,q} \max_i \|(a_{ij})_j\|_{p^*} + \max_j \|(a_{ij})_i\|_q + \mathbb{E}\max_{i,j} |a_{ij}g_{ij}|.$$

The main difficulty in obtaining Conjecture 1 is to prove the upper estimate, since the lower bound is easy. It was shown in [6] that the upper bound holds up to multiplicative constants depending logarithmically on the dimensions. Our main result states that one may skip these logarithmic factors in the range $4/3 < p \le 2 \le q < 4$ and in the range $1 \le p < 2 < q \le \infty$.

Theorem 2. *If $p^*, q \in [2, 4)$ or $p^*, q \in (2, \infty)$, then for every deterministic matrix $A = (a_{ij})_{i\le m, j\le n}$ we have*

$$\begin{aligned}\mathbb{E}\|G_A\|_{p\to q} &\sim_{p,q} \max_i \|(a_{ij})_j\|_{p^*} + \max_j \|(a_{ij})_i\|_q + \mathbb{E}\max_{i,j} |a_{ij}g_{ij}| \\ &\sim \max_i \|(a_{ij})_j\|_{p^*} + \max_j \|(a_{ij})_i\|_q + \max_{k\ge 0} \inf_{|I|=k} \sqrt{\operatorname{Log} k} \max_{(i,j)\notin I} |a_{ij}| \\ &\sim_{p,q} \max_i \|(a_{ij})_j\|_{p^*} + \max_j \|(a_{ij})_i\|_q + \max_{k\ge 0} \inf_{|I|=|J|=k} \sqrt{\operatorname{Log} k} \max_{i\notin I, j\notin J} |a_{ij}| \\ &\sim_{p,q} \mathbb{E}\max_i \|(a_{ij}g_{ij})_j\|_{p^*} + \mathbb{E}\max_j \|(a_{ij}g_{ij})_i\|_q.\end{aligned}$$

Remark 3. The constant in the first lower bound of Theorem 2 does not depend on p and q. Moreover, if $p^* \vee q > 2$, the constant in the first upper bound is at most of order

$$\begin{cases} \dfrac{(p^*\vee q)^{13/2} \operatorname{Log}\big(\frac{p^*\wedge q}{p^*\wedge q - 2}\big)}{(p^*\vee q - 2)^{5/2}} & \text{if } p^*, q > 2 \text{ and } \dfrac{1}{p} + \dfrac{1}{q^*} < 3/2, \\ (p^*\vee q)^{5/2} \operatorname{Log}\Big(\dfrac{p^*\wedge q}{p^*\wedge q - 2}\Big) & \text{if } p^*, q > 2 \text{ and } \dfrac{1}{p} + \dfrac{1}{q^*} \ge 3/2, \\ \dfrac{(p^*\vee q)^{13/2} \operatorname{Log}\big(\frac{2}{4 - p^*\vee q}\big)}{(p^*\vee q - 2)^{5/2}} & \text{if } p^*, q \in [2, 4) \text{ and } p^*\vee q > 2. \end{cases}$$

The latter quantity blows up when p and q approach 2. However, we show in the appendix that the upper bound from Theorem 2 holds with the constant bounded in the range $2 \le p^*, q \le 4 - \delta$ (but blowing up when p^* or q approaches 4).

Remark 4. Recall that the two-sided bounds for $\mathbb{E}\|G_A\|_{p\to q}$ were known before in the cases when $p = 2 = q$ and when p^* or q is infinite; [6, Remark 1.4] implies that for every $p^*, q \ge 2$

$$\mathbb{E}\|G_A\|_{1\to q} = \mathbb{E}\max_j \|(a_{ij}g_{ij})_i\|_q \lesssim \sqrt{q}\max_j \|(a_{ij})_i\|_q + \mathbb{E}\max_{i,j}|a_{ij}g_{ij}|$$

and

$$\mathbb{E}\|G_A\|_{p\to\infty} = \mathbb{E}\max_i \|(a_{ij}g_{ij})_j\|_{p^*} \lesssim \sqrt{p^*}\max_i \|(a_{ij})_j\|_{p^*} + \mathbb{E}\max_{i,j}|a_{ij}g_{ij}|.$$

Theorem 2 and Remark 4 provide an affirmative answer to Conjecture 1, excluding the cases when $p^* = 2$ and $q \in [4,\infty)$ or when $q = 2$ and $p^* \in [4,\infty)$. From our proof of Theorem 2 in the case $p^*, q > 2$ one may deduce the following proposition. It says that in order to prove Conjecture 1 in the remaining ranges, it suffices to derive a weaker dimension dependent bound.

Proposition 5. *Let $\gamma > 0$, $\alpha \ge 1$, and $p \le 2 \le q$. Assume that $p^* \vee q > 2$ and that for every integers m and n, and every deterministic matrix $A = (a_{ij})_{i\le m, j\le n}$,*

$$\mathbb{E}\|G_A\|_{p\to q} \le \alpha\Big(\max_i \|(a_{ij})_j\|_{p^*} + \max_j \|(a_{ij})_i\|_q + \mathrm{Log}^\gamma(mn)\max_{i,j}|a_{ij}|\Big). \tag{1}$$

Then for every integers m and n, and every deterministic matrix $A = (a_{ij})_{i\le m, j\le n}$,

$$\mathbb{E}\|G_A\|_{p\to q} \lesssim \frac{(p^*\vee q)^{13/2}\alpha\,\mathrm{Log}\,\gamma}{(p^*\vee q - 2)^{5/2}}\Big(\max_i \|(a_{ij})_j\|_{p^*} + \max_j \|(a_{ij})_i\|_q$$
$$+ \max_{k\ge 0}\inf_{|I|=k}\sqrt{\mathrm{Log}\,k}\max_{(i,j)\notin I}|a_{ij}|\Big).$$

Unfortunately, we were not able to provide (1) in the remaining range $p^* \wedge q = 2$ and $p^* \vee q \ge 4$. However, our methods yield the following estimate, which is optimal up to a polylog factor.

Theorem 6. *If $p^* \wedge q = 2$ and $p^* \vee q \ge 4$, then for every deterministic matrix $A = (a_{ij})_{i\le m, j\le n}$ we have*

$$\mathbb{E}\|G_A\|_{p\to q} \lesssim_{p,q} \mathrm{Log}\,\mathrm{Log}(mn)\Big(\max_i \|(a_{ij})_j\|_{p^*} + \max_j \|(a_{ij})_i\|_q + \mathbb{E}\max_{i,j}|a_{ij}g_{ij}|\Big).$$

Although we were able to confirm Conjecture 1 in almost whole range $p^*, q \ge 2$, our methods do not allow us to retrieve the exact dependence of $\mathbb{E}\|G_A\|_{p,q}$ on p^* and q in Theorem 2. For example, if $a_{i,j} = 1$ for all $i \le m$ and $j \le n$, then

$$\mathbb{E}\|G_A\|_{p,q} \sim \sqrt{p^*\wedge \mathrm{Log}\,n}\max_i \|(a_{ij})_j\|_{p^*} + \sqrt{q\wedge \mathrm{Log}\,m}\max_j \|(a_{ij})_i\|_q$$

(the third term disappears since, in this case, it is upper bounded by the sum of the first two terms), whereas the constant in the first upper bound in Theorem 2 (for $p^* \wedge q$ separated from 2) grows like $(p^*\vee q)^\gamma$ with $\gamma > 1$. However, we conjecture, that the correct dependence of parameters p and q in the range $p \le 2 \le q$ is the following

$$\mathbb{E}\|G_A\|_{p,q} \overset{?}{\lesssim} \sqrt{p^*\wedge \mathrm{Log}\,n}\max_i \|(a_{ij})_j\|_{p^*} + \sqrt{q\wedge \mathrm{Log}\,m}\max_j \|(a_{ij})_i\|_q$$
$$+ \mathbb{E}\max_{i,j}|a_{ij}g_{ij}|.$$

1.1. **Consequences of the main result.** Let us now present a couple of consequences of Theorem 2. Some of them are immediate and the rest is proven in Section 2.

Theorem 2 easily implies its non-centered counterpart:

$$\mathbb{E}\|(a_{ij}g_{ij}+m_{ij})_{i,j}\|_{p\to q} \sim_{p,q} \mathbb{E}\max_i \|(a_{ij}g_{ij})_j\|_{p^*} + \mathbb{E}\max_j \|(a_{ij}g_{ij})_i\|_q + \|(m_{ij})_{i,j}\|_{p\to q}$$

for every $m_{ij}, a_{ij} \in \mathbb{R}$, $i \leq m$, $j \leq n$, and every p and q satisfying $2 \leq p^*, q < 4$ or $p^*, q > 2$.

Moreover, Theorem 2, Remark 4, and [1, Proposition 1.2] yield the following characterisation of the boundedness of Gaussian linear operators from ℓ_p to ℓ_q whenever $2 \leq p^*, q < 4$ or $p^*, q > 2$. We say that a matrix $B = (b_{ij})_{i,j\in\mathbb{N}}$ defines a bounded operator from ℓ_p to ℓ_q if for all $x \in \ell_p$ the product Bx is well defined, belongs to ℓ_q, and the corresponding linear operator is bounded.

Corollary 7. *Let p and q be such that $2 \leq p^*, q < 4$ or $p^*, q > 2$, and let $(a_{ij})_{i,j\in\mathbb{N}}$ be an infinite deterministic real matrix. The matrix $(a_{ij}g_{ij})_{i,j\in\mathbb{N}}$ defines a bounded linear operator between ℓ_p and ℓ_q almost surely if and only if $\sup_i \|(a_{ij})_j\|_{p^*} < \infty$, $\sup_j \|(a_{ij})_i\|_q < \infty$, and $\mathbb{E}\sup_{i,j\in\mathbb{N}} |a_{ij}g_{ij}| < \infty$.*

Remark 8. The condition $\mathbb{E}\sup_{i,j\in\mathbb{N}} |a_{ij}g_{ij}| < \infty$ in Corollary 7 is equivalent to the deterministic bound

$$\sup_{k\geq 0} \inf_{|I|=k} \sqrt{\operatorname{Log} k} \sup_{(i,j)\in(\mathbb{N}\times\mathbb{N})\setminus I} |a_{ij}| < \infty$$

(see estimate (12) below).

Theorem 2 easily implies two-sided bounds for norms of Gaussian mixtures. We say that a random variable X is a Gaussian mixture if there exists a nonnegative random variable R such that such that X has the same distribution as Rg, where g is a standard Gaussian random variable, independent of R (cf. [5]). The next corollary is an immediate consequence of Theorem 2.

Corollary 9. *Assume that $2 \leq p^*, q < 4$ or $p^*, q > 2$ and let X_{ij}, $i \leq m, j \leq n$, be independent Gaussian mixtures. Then*

$$\mathbb{E}\|(X_{ij})_{i\leq m, j\leq n}\|_{p\to q} \sim_{p,q} \mathbb{E}\max_i \|(X_{ij})_j\|_{p^*} + \mathbb{E}\max_j \|(X_{ij})_i\|_q.$$

We say that X is a symmetric Weibull random variable with (shape) parameter $r \in (0, \infty]$ if X is symmetric and for every $t \geq 0$,

$$\mathbb{P}(|X| \geq t) = e^{-t^r}.$$

Corollary 10. *Let X_{ij}, $i \leq m$, $j \leq n$, be independent symmetric Weibull variables with parameter $r \in (0, 2]$. Then for every p and q satisfying $2 \leq p^*, q < 4$ or $p^*, q \in (2, \infty)$, and every deterministic matrix $A = (a_{ij})_{i\leq m, j\leq n}$ we have*

$$\begin{aligned}
&\mathbb{E}\|(a_{ij}X_{ij})_{i\leq m, j\leq n}\|_{p\to q} \\
&\qquad \sim_{p,q,r} \max_i \|(a_{ij})_j\|_{p^*} + \max_j \|(a_{ij})_i\|_q + \mathbb{E}\max_{i,j} |a_{ij}X_{ij}| \\
&\qquad \sim_r \max_i \|(a_{ij})_j\|_{p^*} + \max_j \|(a_{ij})_i\|_q + \max_{k\geq 0} \inf_{|I|=k} \operatorname{Log}^{1/r} k \max_{(i,j)\notin I} |a_{ij}| \\
&\qquad \sim_{p,q,r} \max_i \|(a_{ij})_j\|_{p^*} + \max_j \|(a_{ij})_i\|_q + \max_{k\geq 0} \inf_{|I|=|J|=k} \operatorname{Log}^{1/r} k \max_{i\notin I, j\notin J} |a_{ij}| \\
&\qquad \sim_{p,q,r} \mathbb{E}\max_i \|(a_{ij}X_{ij})_j\|_{p^*} + \mathbb{E}\max_j \|(a_{ij}X_{ij})_i\|_q.
\end{aligned}$$

We postpone the proof of Corollary 10 to Section 2.

Remark 11. One cannot omit the assumption $r \leq 2$ in Corollary 10. Indeed, in the limit case $r = \infty$ the entries $X_{ij} = \varepsilon_{ij}$ are independent symmetric Bernoulli random variables and it is known that the behaviour of the expected operator norm is different than in the case $r \leq 2$ (see [16]). Moreover, it was conjectured in [11] and proven in [13] up to a factor of order $\log\log\log(mn)$ that

$$\begin{aligned}\mathbb{E}\|(a_{ij}\varepsilon_{ij})_{i,j}\|_{2\to 2} &\overset{?}{\sim} \max_i \|(a_{ij})_j\|_2 + \max_j \|(a_{ij})_i\|_2 \\ &\qquad + \max_{k\geq 0} \inf_{|I|=k} \sup_{\|s\|_2,\|t\|_2\leq 1} \Big\| \sum_{i,j\notin I} a_{ij}\varepsilon_{ij}s_it_j \Big\|_{\operatorname{Log} k} \\ &= \mathbb{E}\max_i \|(a_{ij}\varepsilon_{ij})_j\|_2 + \mathbb{E}\max_j \|(a_{ij}\varepsilon_{ij})_i\|_2 \\ &\qquad + \max_{k\geq 0} \inf_{|I|=k} \sup_{\|s\|_2,\|t\|_2\leq 1} \Big\| \sum_{i,j\notin I} a_{ij}\varepsilon_{ij}s_it_j \Big\|_{\operatorname{Log} k}.\end{aligned}$$

We conjecture that for $p \leq 2 \leq q$,

$$\begin{aligned}\mathbb{E}\|(a_{ij}\varepsilon_{ij})_{i,j}\|_{p\to q} &\overset{?}{\sim}_{p,q} \max_i \|(a_{ij})_j\|_{p^*} + \max_j \|(a_{ij})_i\|_q \\ &\qquad + \max_{k\geq 0} \inf_{|I|=k} \sup_{\|s\|_{q^*},\|t\|_p\leq 1} \Big\| \sum_{i,j\notin I} a_{ij}\varepsilon_{ij}s_it_j \Big\|_{\operatorname{Log} k} \\ &\sim_{p,q} \mathbb{E}\max_i \|(a_{ij}\varepsilon_{ij})_j\|_{p^*} + \mathbb{E}\max_j \|(a_{ij}\varepsilon_{ij})_i\|_q \\ &\qquad + \max_{k\geq 0} \inf_{|I|=k} \sup_{\|s\|_{q^*},\|t\|_p\leq 1} \Big\| \sum_{i,j\notin I} a_{ij}\varepsilon_{ij}s_it_j \Big\|_{\operatorname{Log} k}.\end{aligned} \tag{2}$$

We also believe that the methods of [13] could be adapted to the case $p^*, q \geq 2$ and that – together with Theorem 2 – they would imply (2) up to a polylog factor in the ranges $p^*, q > 2$ and $2 \leq p^*, q < 4$.

The bounds for the expectation of the norm of a random matrix with independent entries satisfying some mild regularity assumptions automatically imply bounds for higher moments as well as for the tails of this norm. Let us state explicitly two such estimates for the structured Gaussian and Weibull random matrices.

Theorem 2 and the Gaussian concentration yield the following moment and tail bounds.

Corollary 12. *If $2 \leq p^*, q < 4$ or $p^*, q > 2$, then for every deterministic matrix $A = (a_{ij})_{i\leq m, j\leq n}$, $\rho \geq 1$, and $t > 0$ we have*

$$\begin{aligned}(\mathbb{E}\|G_A\|_{p\to q}^\rho)^{1/\rho} \sim_{p,q} &\max_i \|(a_{ij})_j\|_{p^*} + \max_j \|(a_{ij})_i\|_q + \max_{k\geq 0}\inf_{|I|=k} \sqrt{\operatorname{Log} k} \max_{(i,j)\notin I} |a_{ij}| \\ &+ \sqrt{\rho}\max_{i,j}|a_{ij}|,\end{aligned}$$

and

$$\begin{aligned}\mathbb{P}\Big(\|G_A\|_{p\to q} \geq C(p,q)\Big(&\max_i \|(a_{ij})_j\|_{p^*} + \max_j \|(a_{ij})_i\|_q \\ &+ \max_{k\geq 0}\inf_{|I|=k} \sqrt{\operatorname{Log} k} \max_{(i,j)\notin I} |a_{ij}|\Big) + t\Big) \leq e^{-t^2/(2\max_{i,j} a_{ij}^2)}.\end{aligned}$$

In the Weibull case, Corollary 10, [8, Theorem 1.1 and Corollary 1.3], and [1, Lemma 2.19] imply the following. (One may also deduce the moreover part, with a constant C_2 depending on p, q and r, from (3) via Markov's inequality.)

Corollary 13. *Let X_{ij}, $i \leq m$, $j \leq n$, be independent symmetric Weibull variables with parameter $r \in (0, 2]$. If $2 \leq p^*, q < 4$ or $p^*, q > 2$, then for every deterministic matrix $A = (a_{ij})_{i\leq m, j\leq n}$, and every $\rho \geq 1$ we have*

$$\begin{aligned}(\mathbb{E}\|(a_{ij}X_{ij})_{i,j}\|_{p\to q}^\rho)^{1/\rho} \sim_{p,q,r} &\max_i \|(a_{ij})_j\|_{p^*} + \max_j \|(a_{ij})_i\|_q \\ &+ \max_{k\geq 0}\inf_{|I|=k} \operatorname{Log}^{1/r} k \max_{(i,j)\notin I}|a_{ij}| + \rho^{1/r}\max_{i,j}|a_{ij}|.\end{aligned} \tag{3}$$

Moreover, for every $t > 0$,

$$\mathbb{P}\Big(\|(a_{ij}X_{ij})_{i,j}\|_{p\to q} \geq C_1(p,q,r)\Big(\max_i \|(a_{ij})_j\|_{p^*} + \max_j \|(a_{ij})_i\|_q + \max_{k\geq 0} \inf_{|I|=k} \operatorname{Log}^{1/r} k \max_{(i,j)\notin I} |a_{ij}|\Big) + t\Big) \leq e^{-t^r/(C_2(r)\max_{i,j}|a_{ij}|^r)}.$$

Remark 14. The upper bound in (3) and, as a consequence, the tail bound from Corollary 13 hold under the weaker assumption that the variables X_{ij} are independent, centered, and have uniformly bounded ψ_r-norm. This follows by a standard argument (see, e.g., the proof of [9, Lemma 2.1]).

The next result is a generalization of [7, Theorem 2]. Its advantage is that we do not need to assume much about the distribution of the entries; however, two additional summands appear in the upper bound.

Corollary 15. *If $2 \leq p^*, q < 4$ or $2 < p^*, q < \infty$, then for every matrix $(X_{ij})_{i\leq m, j\leq n}$ with independent centered entries,*

$$\mathbb{E}\|(X_{ij})_{i,j}\|_{p\to q} \lesssim_{p,q} \max_i\Big(\sum_j \mathbb{E}|X_{ij}|^{p^*}\Big)^{1/p^*} + \max_j\Big(\sum_i \mathbb{E}|X_{ij}|^q\Big)^{1/q} + \Big(\sum_{i,j}\mathbb{E}|X_{ij}|^{2p^*}\Big)^{1/(2p^*)} + \Big(\sum_{i,j}\mathbb{E}|X_{ij}|^{2q}\Big)^{1/(2q)} \tag{4}$$

and

$$\mathbb{E}\|(X_{ij})_{i,j}\|_{p\to q} \lesssim_{p,q} \max_i\Big(\sum_j \mathbb{E}|X_{ij}|^{p^*}\Big)^{1/p^*} + \max_j\Big(\sum_i \mathbb{E}|X_{ij}|^q\Big)^{1/q} + \Big(\sum_i\Big(\sum_j \mathbb{E}|X_{ij}|^{2p^*}\Big)^{\frac{p^*\vee q}{p^*}}\Big)^{\frac{1}{2(p^*\vee q)}} + \Big(\sum_j\Big(\sum_i \mathbb{E}|X_{ij}|^{2q}\Big)^{\frac{p^*\vee q}{q}}\Big)^{\frac{1}{2(p^*\vee q)}}. \tag{5}$$

Remark 16. If $p^*, q \in [2,4)$ or $p^*, q \in (2,\infty)$, $\alpha \geq 1$, and independent centered random variables X_{ij} satisfy

$$(\mathbb{E}|X_{ij}|^{2\rho})^{1/(2\rho)} \leq \alpha(\mathbb{E}|X_{ij}|^\rho)^{1/\rho} \qquad \text{for } \rho \in \{p^*, q\}, \tag{6}$$

then estimate (5) yields

$$\begin{aligned}\mathbb{E}\|(X_{ij})_{i,j}\|_{p\to q} &\lesssim_{p,q} n^{1/p^*}\|X_{11}\|_{p*} + m^{1/q}\|X_{11}\|_q \\ &\quad + m^{1/(2(p^*\vee q))}n^{1/(2p^*)}\|X_{11}\|_{2p*} + m^{1/(2q)}n^{1/(2(p^*\vee q))}\|X_{11}\|_{2q} \\ &\lesssim_{p,q,\alpha} n^{1/p^*}\|X_{11}\|_{p^*} + m^{1/q}\|X_{11}\|_q,\end{aligned}$$

where in the last inequality we used the AM-GM inequality and the estimate $\|X_{11}\|_{2(p*\vee q)} \lesssim_{p,q,\alpha} \|X_{11}\|_{p*\wedge q}$, which follows from the assumption (6) by Hölder's inequality. One can repeat the argument from the proof of the lower bound in [10, Proposition 21] to show that under the above assumptions

$$\mathbb{E}\|(X_{ij})_{i,j}\|_{p\to q} \gtrsim_{p,q,\alpha} n^{1/p^*}\|X_{11}\|_{p^*} + m^{1/q}\|X_{11}\|_q,$$

so in fact

$$\mathbb{E}\|(X_{ij})_{i,j}\|_{p\to q} \sim_{p,q,\alpha} n^{1/p^*}\|X_{11}\|_{p^*} + m^{1/q}\|X_{11}\|_q.$$

We refer to [10] for more precise two-sided bounds (with constants not depending on p and q) under the stronger assumption that the entries X_{ij} are iid centered α-regular random variables.

1.2. **Strategy of the proof of the main result.** Throughout the paper we denote

$$D_1 := \max_i \|(a_{ij})_j\|_{p^*}, \quad D_2 := \max_j \|(a_{ij})_i\|_q$$

to avoid long formulas for the first two terms on the right-hand side of our main estimates.

Similarly as in [12], Theorem 2 is a consequence of two weaker estimates given in the following three propositions. The first one is based on the Slepian-Fernique lemma and generalizes van Handel's bound [17]. A similar result for $p^* = q \in [2, 4)$ was obtained in [14].

Proposition 17. *For every* $2 \le p^*, q < 4$ *and every deterministic matrix* $A = (a_{ij})_{i\le m, j\le n}$*,*

$$\mathbb{E}\|G_A\|_{p\to q} \lesssim D_1 + D_2 + \max_{i,j} \operatorname{Log}^{2/(4-p^*\vee q)}(i+j)|a_{ij}|.$$

Although van Handel's method fails in the range $p^* \vee q \ge 4$, we are able to prove, using different ideas, the following counterpart of Proposition 17 in the range $p^*, q > 2$.

Proposition 18. *For every* $2 < p^*, q < \infty$ *and every deterministic matrix* $A = (a_{ij})_{i\le m, j\le n}$*,*

$$\mathbb{E}\|G_A\|_{p\to q} \lesssim \beta(p,q)\big(\sqrt{p^*}D_1 + \sqrt{q}D_2\big) + \beta'(p,q)\max_{i,j}(i+j)^{(8(p^*\vee q))^{-1}}|a_{ij}|,$$

where $\beta'(p,q) = (p^*\vee q)\beta(p,q)$ *and*

$$\beta(p,q) = \begin{cases} (p^*\vee q)^{3/2}\operatorname{Log}\Big(\dfrac{p^*\wedge q}{p^*\wedge q - 2}\Big) & \text{if } \dfrac{1}{p} + \dfrac{1}{q^*} \ge 3/2, \\ \dfrac{(p^*\vee q)^{11/2}\operatorname{Log}\big(\frac{p^*\wedge q}{p^*\wedge q-2}\big)}{(p^*\vee q - 2)^{5/2}} & \text{if } \dfrac{1}{p} + \dfrac{1}{q^*} < 3/2. \end{cases}$$

Note that the third term in the bound from Proposition 18 is often of greater order than the one from Proposition 17. However, it is sufficiently small to exploit the method from [12] and deduce our main result in the range $p^*, q > 2$ from Proposition 18 and the following dimension dependent bound.

Proposition 19. *Assume that* $2 < p^*, q < \infty$ *or* $2 \le p^*, q < 4$*. Then for every deterministic matrix* $A = (a_{ij})_{i\le m, j\le n}$*,*

$$\mathbb{E}\|G_A\|_{p\to q} \lesssim \alpha(p,q)\Big(\sqrt{p^*}D_1 + \sqrt{q}D_2 + \sqrt{\operatorname{Log}(mn)}\max_{i,j}|a_{ij}|\Big), \tag{7}$$

where

$$\alpha(p,q) \le \begin{cases} 1 & \text{if } p^* = q = 2, \\ \dfrac{(p^*\vee q)^{11/2}\operatorname{Log}\big(\frac{2}{4-p^*\vee q}\big)}{(p^*\vee q-2)^{5/2}} & \text{if } p^*, q \in [2,4) \text{ and } p^*\vee q > 2, \\ (p^*\vee q)^{3/2}\operatorname{Log}\Big(\dfrac{p^*\wedge q}{p^*\wedge q-2}\Big) & \text{if } p^*, q \in (2,\infty) \text{ and } \dfrac{1}{p} + \dfrac{1}{q^*} \ge 3/2, \\ \dfrac{(p^*\vee q)^{11/2}\operatorname{Log}\big(\frac{p^*\wedge q}{p^*\wedge q-2}\big)}{(p^*\vee q-2)^{5/2}} & \text{if } p^*, q \in (2,\infty) \text{ and } \dfrac{1}{p} + \dfrac{1}{q^*} < 3/2. \end{cases}$$

In the case $p = q = 2$ estimate (7) was proven by Bandeira and van Handel in [2]. We cannot use their combinatorial approach based on the trace method since for $(p,q) \ne (2,2)$ we no longer deal with the spectral norm. Proving Proposition 19 is one of the main difficulties and novelties of our paper.

In the case $2 \le p^*, q < 4$ estimate (7) follows by Proposition 17 and the exponent reduction procedure described in Subsection 4.2. In this step we use some ideas from [13]. Surprisingly, in the case $p^*, q > 2$ we first derive Proposition 19, and then use it to show Proposition 18.

To prove Proposition 19 in the case $p^*, q > 2$ we exploit the ideas from the proof of the main result of [1] (to estimate the suprema over vectors with small coordinates) together with a net argument (which allows us to estimate the suprema over vectors with small supports). Finally, to deduce Proposition 17 we decompose the matrix A into block diagonal matrices A_k with blocks of a smaller size, and matrices B_l whose norms are easier to control (due to Proposition 36 below), and use Proposition 19 for each A_k separately.

1.3. **Organization of the paper.** In Section 2 we show how Propositions 17–19 imply Theorem 2 and then we prove Remark 3 and Corollaries 10 and 15. Section 3 contains the proof of Proposition 17 and some other estimates derived from Slepian's lemma, necessary for proving Proposition 19 in the range $p^* \wedge q > 2$. In Section 4 we prove Proposition 19. Finally, Section 5 is devoted to the proof of Proposition 18 and Theorem 6; it also contains a sketch of the proof of Proposition 5.

2. Proof of Theorem 2 and its corollaries

In this section we first show how to deduce the most challenging part of Theorem 2 from Propositions 17–19. Then we give the proofs of Theorem 2 and Corollaries 10 and 15.

Proposition 20. *Let $p^*, q \geq 2$, m, n be positive integers and $\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3 \geq 1$. Assume that for every $1 \leq m' \leq m$, $1 \leq n' \leq n$ and every deterministic matrix $A = (a_{ij})_{i\leq m', j\leq n'}$,*

$$\mathbb{E}\|G_A\|_{p\to q} \leq \alpha_1 D_1 + \alpha_2 D_2 + \alpha_3 \sqrt{\operatorname{Log}(m'n')} \max_{i,j} |a_{ij}|$$

and

$$\mathbb{E}\|G_A\|_{p\to q} \leq \beta_1 D_1 + \beta_2 D_2 + \beta_3 \max_{i,j} (i+j)^{(8(p^*\vee q))^{-1}} |a_{ij}|\Big).$$

Then for every deterministic matrix $A = (a_{ij})_{i\leq m, j\leq n}$,

$$\begin{aligned}\mathbb{E}\|G_A\|_{p\to q} \lesssim (\alpha_1 + \beta_1 + \beta_3) D_1 + (\alpha_2 + \beta_2 + \beta_3) D_2 \\ + \alpha_3 \max_{k\geq 0} \inf_{|I|=k} \sqrt{\operatorname{Log} k} \max_{(i,j)\notin I} |a_{ij}|\Big).\end{aligned}$$

We will need the following deterministic lemma about norms of block diagonal matrices.

Lemma 21. *Let $(c_{ij})_{i\leq m, j\leq n}$ be a block diagonal matrix with blocks C_l, and $1 \leq p \leq q$. Then*

$$\|(c_{ij})_{i\leq m, j\leq n}\|_{p\to q} = \max_l \|C_l\|_{p\to q}.$$

Proof. Assume that the blocks C_l, $l \leq l_0$, consist of entries a_{ij} such that $i \in I_l$ and $j \in J_l$. Then

$$\begin{aligned}\|(c_{ij})_{i\leq m, j\leq n}\|_{p\to q} &= \sup_{s\in B_{q^*}^m, t\in B_p^n} \sum_{i,j} c_{ij} s_i t_j \\ &= \sup_{x\in B_{q^*}^{l_0}, y\in B_p^{l_0}} \sum_{l=1}^{l_0} \sup_{s\in x_l B_{q^*}^{I_l}, t\in y_l B_p^{J_l}} \sum_{i\in I_l, j\in J_l} c_{ij} s_i t_j \\ &= \sup_{x\in B_{q^*}^{k_0}, y\in B_p^{l_0}} \sum_{l=1}^{l_0} x_l y_l \|C_l\|_{p\to q} = \sup_{y\in B_p^{l_0}} \Big(\sum_l |y_l|^q \|C_l\|_{p\to q}^q\Big)^{1/q}.\end{aligned}$$

Since $p \le q$, the latter supremum is attained at $y = e_l$ for some $l \le l_0$, so

$$\sup_{y \in B_p^{l_0}} \Big(\sum_l |y_l|^q \|C_l\|_{p\to q}^q \Big)^{1/q} = \max_l \|C_l\|_{p\to q}. \qquad \square$$

Proof of Proposition 20. Let

$$D_\infty := \max_{k\ge 0} \inf_{|I|=k} \max_{(i,j)\notin I} \sqrt{\operatorname{Log} k}|a_{ij}|,$$

$N_0 = 1$, and $N_k = 2^{2^k}$ for $k \ge 1$. Without loss of generality we may assume that $n = m = N_{k_0}$ for some k_0; if necessary, we simply add zero rows and columns.

We follow the ideas of the proof of [12, Theorem 3.9] and [11, Remark 4.5], starting with constructing a suitable permutations $(i_1, \dots, i_{N_{k_0}})$ and $(j_1, \dots, j_{N_{k_0}})$ of $\{1, \dots, N_{k_0}\}$. (Note that the change of order of rows and columns does not change $\mathbb{E}\|G_A\|_{p\to q}$.) Then we decompose the matrix G_A and bound each piece of this decomposition separately, each time using one of the assumptions of the proposition.

In the first step we choose $I_1 = \{i_1, \dots, i_{N_1}\}$ and $J_1 = \{j_1, \dots, j_{N_1}\}$ in such a way that

$$\max_{i\notin I_1} \max_j |a_{ij}| \le \frac{D_\infty}{\sqrt{\operatorname{Log} N_1}} \quad \text{and} \quad \max_{j\notin J_1} \max_i |a_{ij}| \le \frac{D_\infty}{\sqrt{\operatorname{Log} N_1}}.$$

Suppose now that we have selected $I_k = \{i_1, \dots, i_{N_k}\}$ and $J_k = \{i_1, \dots, j_{N_k}\}$ for $k < k_0$. To construct $(i_{N_k+1}, \dots, i_{N_{k+1}})$ we choose first $N_k N_{k-1}$ indices i from $[m] \setminus I_k$ that contain the N_{k-1} largest moduli of entries $|a_{ij}|$ from each column $j \in J_k$. Next, among remaining indices we choose $N_{k+1} - N_k - N_k N_{k-1} \ge N_k$ indices i in such a way that $I_{k+1} = \{i_1, \dots, i_{N_{k+1}}\}$ satisfies

$$\max_{i\notin I_{k+1}} \max_j |a_{ij}| \le \frac{D_\infty}{\sqrt{\operatorname{Log} N_k}}. \tag{8}$$

Similarly, to construct $(j_{N_k+1}, \dots, j_{N_{k+1}})$ we choose first $N_k N_{k-1}$ indices j from $[n] \setminus J_k$ that contain the N_{k-1} largest moduli of entries $|a_{ij}|$ from each row $i \in I_k$. Next, among remaining indices we choose $N_{k+1} - N_k - N_k N_{k-1} \ge N_k$ indices j in such a way that $J_{k+1} = \{j_1, \dots, j_{N_{k+1}}\}$ satisfies

$$\max_{j\notin J_{k+1}} \max_i |a_{ij}| \le \frac{D_\infty}{\sqrt{\operatorname{Log} N_k}}. \tag{9}$$

The above construction implies in particular that for $k \ge 1$,

$$|a_{ij}| \le D_2 N_{k-1}^{-1/q} \le 2D_2 2^{-2^{k-1}/q} \qquad \text{if } j \le N_k \text{ and } i \ge M_k := N_k + N_k N_{k-1}, \tag{10}$$

$$|a_{ij}| \le D_1 N_{k-1}^{-1/p^*} \le 2D_1 2^{-2^{k-1}/p^*} \qquad \text{if } i \le N_k \text{ and } j \ge M_k. \tag{11}$$

We set (see [12, Fig. 1])

$$E_1 := [1, M_1]^2 \cup \bigcup_{k\ge 1} [N_{2k} + 1, M_{2k+1} \wedge N_{k_0}]^2,$$

$$E_2 := \bigcup_{k\ge 1} [N_{2k-1} + 1, M_{2k} \wedge N_{k_0}]^2 \setminus E_1, \qquad E_3 := [1, N_{k_0}]^2 \setminus (E_1 \cup E_2)$$

and write $G_A = U + V + W$, where

$$U_{ij} := a_{ij} g_{ij} \mathbb{1}_{\{(i,j)\in E_1\}}, \quad V_{ij} := a_{ij} g_{ij} \mathbb{1}_{\{(i,j)\in E_2\}}, \quad W_{ij} := a_{ij} g_{ij} \mathbb{1}_{\{(i,j)\in E_3\}}.$$

The matrix U is block diagonal with the first block $U_1 = (X_{ij})_{i,j\in E_{1,1}}$ of dimension $|E_{1,1}| = M_1$ and blocks $U_k = (X_{ij})_{i,j\in E_{1,k}}$ for $k \geq 2$ of dimension $|E_{1,k}| \leq M_{2k-1} - N_{2k-2}$. We have

$$(\mathbb{E}\|U_1\|_{p\to q}^2)^{1/2} \leq (\mathbb{E}\|U_1\|_{2\to 2}^2)^{1/2} \leq \Big(\sum_{i,j\in[1,M_1]} \mathbb{E}g_{ij}^2 a_{ij}^2\Big)^{1/2} \leq M_1 \max_{i,j} |a_{ij}| \lesssim D_\infty.$$

For $k = 2, 3, \ldots$ the Gaussian concentration (see, e.g., [3, Theorem 5.6]), the first assumption of the proposition, and property (8) imply

$$\begin{aligned}
\Big(\mathbb{E}\|U_k\|_{p\to q}^{2^k}\Big)^{2^{-k}} &\lesssim \mathbb{E}\|U_k\|_{p\to q} + 2^{k/2} \max_{i,j\in E_{1,k}} |a_{ij}| \\
&\leq \alpha_1 D_1 + \alpha_2 D_2 + \Big(\alpha_3\sqrt{2\log|E_{1,k}|} + 2^{k/2}\Big) \max_{i,j\in E_{1,k}} |a_{ij}| \\
&\lesssim \alpha_1 D_1 + \alpha_2 D_2 + \alpha_3 D_\infty.
\end{aligned}$$

Thus,

$$M := \sup_k \Big(\mathbb{E}\|U_k\|_{p\to q}^{2^k}\Big)^{2^{-k}} \lesssim \alpha_1 D_1 + \alpha_2 D_2 + \alpha_3 D_\infty,$$

and Lemma 21 yields

$$\begin{aligned}
\mathbb{E}\|U\|_{p\to q} &= \mathbb{E}\sup_{k\geq 1} \|U_k\|_{p\to q} \leq 2M + M\mathbb{E}\sum_{k\geq 1} \frac{\|U_k\|_{p\to q}}{M} \mathbb{1}_{\{\|U_k\|_{p\to q}\geq 2M\}} \\
&\leq 2M + M\mathbb{E}\sum_{k\geq 1} 2^{1-2^k}\Big(\frac{\|U_k\|_{p\to q}}{M}\Big)^{2^k} \leq M\Big(2 + \sum_{k\geq 1} 2^{1-2^k}\Big) \leq 3M \\
&\lesssim \alpha_1 D_1 + \alpha_2 D_2 + \alpha_3 D_\infty.
\end{aligned}$$

In a similar way we show that

$$\mathbb{E}\|V\|_{p\to q} \lesssim \alpha_1 D_1 + \alpha_2 D_2 + \alpha_3 D_\infty.$$

Finally, fix $(i,j) \in E_3$ and take $k \in \{0, 1, \ldots\}$ such that $M_k \leq i < M_{k+1}$, where $M_0 = 1$. Observe that if $M_k \leq i \leq N_{k+1}$, then either $j \leq N_k$ or $j > M_{k+1}$ and if $N_{k+1} + 1 \leq i < M_{k+1}$, then either $j \leq N_k$ or $j > M_{k+2}$. Therefore, either $j \leq N_k$ or $j > M_{k+1}$. If $j \leq N_k$, then by (10)

$$(i+j)^{(8(p^*\vee q))^{-1}} |a_{ij}| \leq 2N_{k+2}^{(8(p^*\vee q))^{-1}} D_2 2^{-2^{k-1}/q} \leq 2D_2.$$

If on the other hand $M_{k+l} < j \leq M_{k+l+1}$ for some $l \geq 1$, then $i < N_{k+l}$ and (11) imply

$$(i+j)^{(8(p^*\vee q))^{-1}} |a_{ij}| \leq 2N_{k+l+2}^{(8(p^*\vee q))^{-1}} D_1 2^{-2^{k+l-1}/p^*} \leq 2D_1.$$

Thus, the second assumption of the proposition yields

$$\begin{aligned}
\mathbb{E}\|W\|_{p\to q} &\leq \beta_1 D_1 + \beta_2 D_2 + \beta_3 \max_{(i,j)\in E_3} (i+j)^{(8(p^*\vee q))^{-1}} |a_{ij}| \\
&\leq (\beta_1 + 2\beta_3) D_1 + (\beta_2 + 2\beta_3) D_2.
\end{aligned}$$ □

Proof of Theorem 2. The two-sided estimate between the expressions on the right-hand side of the first and the last line was proven in [1, Section 5.4]. This also yields the lower bound in the first asserted two-sided estimate.

The second two-sided estimate follows by

$$\mathbb{E}\max_{i,j} |a_{ij} g_{ij}| \sim \max_{k\geq 0} \inf_{|I|=k} \sqrt{\operatorname{Log} k} \max_{(i,j)\notin I} |a_{ij}| \tag{12}$$

(cf. [17, Lemmas 2.3 and 2.4]).

The upper bound from the first two-sided estimate follows by Propositions 17–20 and (12).

Now we move to the proof of the third two-sided estimate from Theorem 2. We will show a more precised bound

$$\sqrt{p^*}D_1 + \sqrt{q}D_2 + D_\infty \sim \sqrt{p^*}D_1 + \sqrt{q}D_2 + D'_\infty, \tag{13}$$

where

$$D'_\infty := \max_{k\geq 0} \inf_{|I|=|J|=k} \max_{i\notin I, j\notin J} \sqrt{\operatorname{Log} k}|a_{ij}|.$$

The lower bound in (13) is trivial, since $I \subset P_1(I) \times P_2(I)$, where P_1 and P_2 are coordinate projections. To establish the upper bound in (13) we need to show that for any $k \geq 1$,

$$\sqrt{\operatorname{Log} k} \inf_{|V|=k} \max_{(i,j)\notin V} |a_{ij}| \lesssim \sqrt{p^*}D_1 + \sqrt{q}D_2 + D'_\infty. \tag{14}$$

If $k \leq 3$, then

$$\sqrt{\operatorname{Log} k} \inf_{|V|=k} \max_{(i,j)\notin V} |a_{ij}| \lesssim \max_{i,j} |a_{ij}| \leq D_1.$$

If $k \geq 3$, then put $k' := \lfloor \sqrt{k/3} \rfloor$ and choose $|I_0| = |J_0| = k'$ such that

$$\inf_{|I|=|J|=k'} \max_{i\notin I, j\notin J} |a_{ij}| = \max_{i\notin I_0, j\notin J_0} |a_{ij}|.$$

Let $V_0 := (I_0 \times J_0) \cup V_1 \cup V_2$, where

$$V_1 := \{(i,j) \colon i \in I_0, |a_{ij}| \geq (k')^{-1/p*} D_1\},$$
$$V_2 := \{(i,j) \colon j \in J_0, |a_{ij}| \geq (k')^{-1/q} D_2\}.$$

Then $|V_0| \leq |I_0||J_0| + |I_0|k' + |J_0|k' = 3(k')^2 \leq k$, so that

$$\begin{aligned}
\inf_{|V|=k} \max_{(i,j)\notin V} |a_{ij}| &\leq \max_{(i,j)\notin V_0} |a_{ij}| \\
&\leq \max_{i\notin I_0, j\notin J_0} |a_{ij}| + \max_{(i,j)\in(I_0\times[n])\setminus V_1} |a_{ij}| + \max_{(i,j)\in([m]\times J_0)\setminus V_2} |a_{ij}| \\
&\leq (\operatorname{Log} k')^{-1/2} D'_\infty + (k')^{-1/p*} D_1 + (k')^{-1/q} D_2.
\end{aligned}$$

We have $\sqrt{\operatorname{Log} k} \lesssim \min\{\sqrt{\operatorname{Log} k'}, \sqrt{p^*}(k')^{1/p^*}, \sqrt{q}(k')^{1/q}\}$, so (14) follows. □

Proof of Remark 3. The order of the constants in the case when $p^*, q > 2$ follows by Propositions 18, 19, and 20.

In the case $p^*, q \in [2,4)$, the proof of Proposition 18, given in Section 5 below, shows that Proposition 19 implies

$$\mathbb{E}\|G_A\|_{p\to q} \lesssim \beta(p,q)\big(\sqrt{p^*}D_1 + \sqrt{q}D_2\big) + \beta'(p,q) \max_{i,j} (i+j)^{(8(p^*\vee q))^{-1}} |a_{ij}|,$$

with $\beta'(p,q) = (p^* \vee q)\beta(p,q)$ and

$$\beta(p,q) = \frac{(p^*\vee q)^{11/2} \operatorname{Log}\big(\frac{2}{4-p^*\wedge q}\big)}{(p^*\vee q - 2)^{5/2}}.$$

This, together with Propositions 19 and 20, yields the asserted order of constant in the range $p^*, q \in [2,4)$. □

Proof of Corollary 10. Let $s > 0$ be such that $\frac{1}{r} = \frac{1}{2} + \frac{1}{s}$ (if $r = 2$, then $s = \infty$) and let $Y_{ij} = g_{ij}|\widetilde{g}_{ij}|^{2/s}$ where $(\widetilde{g}_{ij})_{i,j}$ is an independent copy of $(g_{ij})_{i,j}$. Then for every $\rho \geq 1$,

$$(\mathbb{E}|X_{ij}|^\rho)^{1/\rho} \sim_r \rho^{1/r} \sim_r (\mathbb{E}|Y_{ij}|^\rho)^{1/\rho}, \tag{15}$$

so Theorem 2 and [12, Lemma 4.7] (applied twice) imply

$$\begin{aligned}\mathbb{E}\|(a_{ij}X_{ij})_{i,j}\|_{p\to q} &\sim_{p,q,r} \mathbb{E}\max_i \|(a_{ij}|g_{ij}|^{2/s})_j\|_{p^*} + \mathbb{E}\max_j \|(a_{ij}|g_{ij}|^{2/s})_i\|_q \\ &\qquad + \mathbb{E}\max_{k\geq 0}\inf_{|I|=k} \sqrt{\operatorname{Log} k} \max_{(i,j)\notin I} |a_{ij}||g_{ij}|^{2/s} \\ &\sim_{p,q,r} \mathbb{E}\max_i \|(a_{ij}X_{ij})_j\|_{p^*} + \mathbb{E}\max_j \|(a_{ij}X_{ij})_i\|_q.\end{aligned}$$

Moreover,

$$\begin{aligned}\mathbb{E}\max_i \|(a_{ij}|g_{ij}|^{2/s})_j\|_{p^*} &= \mathbb{E}\max_i \|(|a_{ij}|^{s/2}g_{ij})_j\|_{2p^*/s}^{2/s} \\ &\sim_r \Big(\mathbb{E}\max_i \|(|a_{ij}|^{s/2}g_{ij})_j\|_{2p^*/s}\Big)^{2/s}\end{aligned}$$

and similarly

$$\mathbb{E}\max_j \|(a_{ij}|g_{ij}|^{2/s})_i\|_q \sim_r \Big(\mathbb{E}\max_j \|(|a_{ij}|^{s/2}g_{ij})_i\|_{2q/s}\Big)^{2/s},$$

so [1, equation (5.11)] yields

$$\begin{aligned}&\mathbb{E}\max_i \|(a_{ij}|g_{ij}|^{2/s})_j\|_{p^*} + \mathbb{E}\max_j \|(a_{ij}|g_{ij}|^{2/s})_i\|_q \\ &\quad\sim_{p,q,r} \max_i \|(|a_{ij}|^{s/2})_j\|_{2p^*/s}^{2/s} + \max_j \|(|a_{ij}|^{s/2})_i\|_{2q/s}^{2/s} + \Big(\mathbb{E}\max_{i,j} |a_{ij}|^{s/2}|g_{ij}|\Big)^{2/s} \\ &\quad\sim_r \max_i \|(a_{ij})_j\|_{p^*} + \max_j \|(a_{ij})_i\|_q + \mathbb{E}\max_{i,j} |a_{ij}||g_{ij}|^{2/s}.\end{aligned}$$

Note also that

$$\begin{aligned}\mathbb{E}\max_{i,j} |a_{ij}||g_{ij}|^{2/s} &\leq \mathbb{E}\max_{k\geq 0}\inf_{|I|=k} \sqrt{\operatorname{Log} k} \max_{(i,j)\notin I} |a_{ij}||g_{ij}|^{2/s} \\ &\sim \mathbb{E}\max_{i,j} |a_{ij}||g_{ij}|^{2/s}|\tilde{g}_{i,j}| = \mathbb{E}\max_{i,j}|a_{ij}Y_{ij}| \sim_r \mathbb{E}\max_{i,j}|a_{ij}X_{ij}|,\end{aligned}$$

where the second estimate follows by the conditional application of (12) and the last one by (15) and [12, Lemma 4.7]. Thus,

$$\mathbb{E}\|(a_{ij}X_{ij})_{i,j}\|_{p\to q} \sim_{p,q,r} \max_i \|(a_{ij})_j\|_{p^*} + \max_j \|(a_{ij})_i\|_q + \mathbb{E}\max_{i,j}|a_{ij}X_{ij}|.$$

Let $Z_{ij} = \varepsilon_{ij}|g_{ij}|^{2/r}$, where ε_{ij}, $i \leq m$, $j \leq n$, are iid symmetric Bernoulli variables, independent of $(g_{ij})_{i,j}$. Then

$$(\mathbb{E}|X_{ij}|^{\rho})^{1/\rho} \sim_r \rho^{1/r} \sim_r (\mathbb{E}|Z_{ij}|^{\rho})^{1/\rho},$$

so [12, Lemma 4.7] and (12) yield

$$\begin{aligned}\mathbb{E}\max_{i,j}|a_{ij}X_{ij}| &\sim_r \mathbb{E}\max_{i,j}|a_{ij}Z_{ij}| = \mathbb{E}\max_{i,j}|a_{ij}||g_{ij}|^{2/r} \\ &\sim_r \big(\mathbb{E}\max_{i,j}|a_{ij}|^{r/2}|g_{ij}|\big)^{2/r} \sim \big(\max_{k\geq 0}\inf_{|I|=k}\sqrt{\operatorname{Log} k}\max_{(i,j)\notin I}|a_{ij}|^{r/2}\big)^{2/r} \\ &= \max_{k\geq 0}\inf_{|I|=k}\operatorname{Log}^{1/r} k \max_{(i,j)\notin I}|a_{ij}|.\end{aligned}$$

Finally, the third two-sided estimate in Corollary 10 may be established in a similar way as in the Gaussian case (see the last part of the proof of Theorem 2). □

Proof of Corollary 15. Let $(\varepsilon_{ij})_{i,j}$ and $(g_{ij})_{i,j}$ be independent matrices with independent symmetric ± 1 and $\mathcal{N}(0,1)$ entries, respectively, and assume that they are independent of $(X_{ij})_{i,j}$. Let $Y_{ij} = g_{ij}X_{ij}$ and $Z_{ij}^{(\rho)} = |Y_{ij}|^\rho - \mathbb{E}|Y_{ij}|^\rho$ for $\rho \in \{p^*, q\}$. Since X_{ij} are centered,

$$\begin{aligned}\mathbb{E}\|(X_{ij})\|_{p\to q} &\le 2\mathbb{E}\|(\varepsilon_{ij}X_{ij})\|_{p\to q} \le \frac{2}{\mathbb{E}|g_{ij}|}\mathbb{E}\|(Y_{ij})\|_{p\to q} \\ &\sim_{p,q} \mathbb{E}\max_i \|(Y_{ij})_j\|_{p^*} + \mathbb{E}\max_j \|(Y_{ij})_i\|_q,\end{aligned}$$

where the last bound follows by Corollary 9.

Since for every $a, b \ge 0$, $(a+b)^{1/q} \le a^{1/q} + b^{1/q}$, we have

$$\mathbb{E}\max_j \|(Y_{ij})_i\|_q \le \max_j\Big(\sum_i \mathbb{E}|Y_{ij}|^q\Big)^{1/q} + \mathbb{E}\max_j\Big|\sum_i (|Y_{ij}|^q - \mathbb{E}|Y_{ij}|^q)\Big|^{1/q}.$$

By the independence of g_{ij} and X_{ij} we have

$$\begin{aligned}\max_j\Big(\sum_i \mathbb{E}|Y_{ij}|^q\Big)^{1/q} &= (\mathbb{E}|g_{1,1}|^q)^{1/q}\max_j\Big(\sum_i \mathbb{E}|X_{ij}|^q\Big)^{1/q} \\ &\sim \sqrt{q}\max_j\Big(\sum_i \mathbb{E}|X_{ij}|^q\Big)^{1/q}.\end{aligned}$$

Moreover, by the Rosenthal inequality we have for every $r \ge 2$,

$$\begin{aligned}\mathbb{E}\max_j\Big|\sum_i Z_{ij}^{(q)}\Big|^{1/q} &\le \mathbb{E}\Big(\sum_j\Big|\sum_i Z_{ij}^{(q)}\Big|^r\Big)^{1/(rq)} \le \Big(\sum_j \mathbb{E}\Big|\sum_i Z_{ij}^{(q)}\Big|^r\Big)^{1/(rq)} \\ &\lesssim_r \Big(\sum_j\Big(\sum_i \mathbb{E}|Z_{ij}^{(q)}|^2\Big)^{r/2} + \sum_{i,j}\mathbb{E}|Z_{ij}^{(q)}|^r\Big)^{1/(rq)}.\end{aligned}$$

Observe that for $u \ge 1$ we have

$$\mathbb{E}|Z_{ij}^{(q)}|^u \le 2^u\mathbb{E}|Y_{ij}|^{qu} = 2^u\mathbb{E}|g_{ij}|^{qu}\mathbb{E}|X_{ij}|^{qu} \le 2^u(qu)^{(qu)/2}\mathbb{E}|X_{ij}|^{qu}.$$

Therefore for any $r \ge 2$,

$$\begin{aligned}\mathbb{E}\max_j \|(Y_{ij})_i\|_q \lesssim_{q,r} &\max_j\Big(\sum_i \mathbb{E}|X_{ij}|^q\Big)^{1/q} \\ &+ \Big(\sum_j\Big(\sum_i \mathbb{E}|X_{ij}|^{2q}\Big)^{r/2} + \sum_{i,j}\mathbb{E}|X_{ij}|^{rq}\Big)^{1/(rq)}.\end{aligned}$$

In a similar way we show that for every $s \ge 2$,

$$\begin{aligned}\mathbb{E}\max_i \|(Y_{ij})_j\|_{p^*} \lesssim_{p,s} &\max_i\Big(\sum_j \mathbb{E}|X_{ij}|^{p^*}\Big)^{1/p^*} \\ &+ \Big(\sum_i\Big(\sum_j \mathbb{E}|X_{ij}|^{2p^*}\Big)^{s/2} + \sum_{i,j}\mathbb{E}|X_{ij}|^{sp^*}\Big)^{1/(sp^*)}.\end{aligned}$$

Estimate (4) follows if we choose $s = r = 2$ and estimate (5) if we take $r = 2(p^*\vee q)/q$ and $s = 2(p^* \vee q)/p^*$. □

3. Bounds following from the Slepian-Fernique lemma

3.1. **Range** $p^*, q \in [2, 4)$**.** Let us begin with a modification of van Handel's argument from [17]. This allows us to prove Proposition 17, i.e., a weaker version of the main result in the range $p^*, q \in [2, 4)$. Let $A \circ A = (a_{ij}^2)_{i\le m, j\le n}$ be the variance profile of $G_A = (a_{ij}g_{ij})_{i\le m, j\le n}$ and

$$B = (b_{ij})_{i,j\le m+n} = \begin{pmatrix} 0 & A\circ A \\ (A\circ A)^T & 0\end{pmatrix}.$$

Let Y be a $(m+n)$-dimensional Gaussian vector with mean 0 and covariance matrix B^- being the negative part of B, i.e., $B^- = -\sum_{i=1}^{m+n}(\lambda_i \wedge 0)u_iu_i^T$, where $B = \sum_{i=1}^{m+n}\lambda_iu_iu_i^T$ is the spectral decomposition of B. The proof of [17, Corollary 4.2] yields that Y_k are Gaussian random variables with variance $\mathbb{E}Y_k^2 \le (\sum_l b_{k,l}^2)^{1/2}$. Hence,

$$
\begin{aligned}
\mathrm{Var}(Y_i) &\le \Big(\sum_j a_{ij}^4\Big)^{1/2}, \quad 1 \le i \le m,\\
\mathrm{Var}(Y_{j+m}) &\le \Big(\sum_i a_{ij}^4\Big)^{1/2}, \quad 1 \le j \le n.
\end{aligned}
\tag{16}
$$

Lemma 22. *For any bounded, nonempty sets $K \subset \mathbb{R}^m$ and $L \subset \mathbb{R}^n$,*

$$
\begin{aligned}
\mathbb{E}\sup_{s\in K, t\in L} a_{ij}g_{ij}s_it_j \le{}& \mathbb{E}\sup_{s\in K,t\in L}\sum_{i=1}^m s_ig_i\sqrt{\sum_{j=1}^n t_j^2a_{ij}^2}\\
&+\mathbb{E}\sup_{s\in K,t\in L}\sum_{j=1}^n t_jg_{m+j}\sqrt{\sum_{i=1}^m s_i^2a_{ij}^2}\\
&+\frac12\mathbb{E}\sup_{s\in K}\sum_{i=1}^m s_i^2Y_i + \frac12\mathbb{E}\sup_{t\in L}\sum_{j=1}^n t_j^2Y_{m+j}.
\end{aligned}
$$

Proof. Let us define the symmetric Gaussian matrix

$$
X = (X_{ij})_{i,j\le m+n} = \begin{pmatrix} 0 & G_A \\ (G_A)^T & 0\end{pmatrix}.
$$

Then B is the variance profile of X. Consider two centered Gaussian processes indexed by $v \in \mathbb{R}^{n+m}$:

$$
G_v = \sum_{k,l\le m+n} X_{kl}v_kv_l \quad \text{and} \quad Z_v = 2\sum_{k=1}^{m+n} v_kg_k\sqrt{\sum_{l=1}^{m+n} v_l^2b_{k,l}^2} + \sum_{k=1}^{m+n} v_k^2Y_k,
$$

It was shown in [17, proof of Theorem 4.1] that for any $v, v' \in \mathbb{R}^{n+m}$, $\mathbb{E}|G_v - G_{v'}|^2 \le \mathbb{E}|Z_v - Z_{v'}|^2$ and, as a consequence, the Slepian-Fernique lemma (see, e.g., [3, Theorem 13.3]) yields

$$
\mathbb{E}\sup_{v\in K\times L} G_v \le \mathbb{E}\sup_{v\in K\times L} Z_v.
$$

Thus, to finish the proof it is enough to observe that

$$
\begin{aligned}
\sup_{v\in K\times L} G_v &= \sup_{v\in K\times L}\sum_{k,l\le m+n} X_{kl}v_kv_l\\
&= \sup_{s\in K,t\in L}\Big(\sum_{i\le m, j\le n}(G_A)_{ij}s_it_j + \sum_{i\le m,j\le n}((G_A)^T)_{ji}t_js_i\Big)\\
&= 2\sup_{s\in K,t\in L}\sum_{i\le m,j\le n} a_{ij}g_{ij}s_it_j
\end{aligned}
$$

and

$$
\begin{aligned}
\sup_{v\in K\times L} Z_v = \sup_{s\in K,t\in L}\Big(&2\sum_{i=1}^m s_ig_i\sqrt{\sum_{j=1}^n t_j^2a_{ij}^2} + 2\sum_{j=1}^n t_jg_{m+j}\sqrt{\sum_{i=1}^m s_i^2a_{ij}^2}\\
&+\sum_{i=1}^m s_i^2Y_i + \sum_{j=1}^n t_j^2Y_{m+j}\Big).
\end{aligned}
$$

□

Corollary 23. *If $p \le 2 \le q$, then*

$$\mathbb{E}\|G_A\|_{\ell_p^n \to \ell_q^m} \le \mathbb{E}\max_{j\le n} \|(a_{ij}g_i)_i\|_q + \mathbb{E}\max_{i\le m}\|(a_{ij}g_j)_j\|_{p^*} + \mathbb{E}\max_{k\le m+n}|Y_k|.$$

Proof. We apply Lemma 22 with $K = B_{q^*}^m$, $L = B_p^n$. Note that

$$\sup_{s\in K}\sum_{i=1}^m s_i^2 Y_i \le \sup_{s\in B_2^m}\sum_{i=1}^m s_i^2 Y_i \le \max_{i\le m}|Y_i|$$

and

$$\sup_{t\in L}\sum_{j=1}^n t_j^2 Y_{m+j} \le \sup_{t\in B_2^n}\sum_{j=1}^n t_j^2 Y_{m+j} \le \max_{j\le n}|Y_{m+j}|.$$

The convexity of the function $u \to \sum_{i=1}^m |g_i|^q\Big(\sum_{j=1}^n |u_j|a_{ij}^2\Big)^{q/2}$ and the fact that $B_1^n = \mathrm{conv}\{\pm e_j\colon j\le n\}$ yield

$$\begin{aligned}\sup_{s\in K, t\in L}\sum_{i=1}^m s_i g_i\sqrt{\sum_{j=1}^n t_j^2 a_{ij}^2} &= \sup_{t\in B_p^n}\Big(\sum_{i=1}^m |g_i|^q\Big(\sum_{j=1}^n t_j^2 a_{ij}^2\Big)^{q/2}\Big)^{1/q}\\ &\le \sup_{u\in B_1^n}\Big(\sum_{i=1}^m |g_i|^q\Big(\sum_{j=1}^n |u_j| a_{ij}^2\Big)^{q/2}\Big)^{1/q}\\ &= \max_{j\le n}\Big(\sum_{i=1}^m |g_i|^q|a_{ij}|^q\Big)^{1/q} = \max_{j\le n}\|(a_{ij}g_i)_i\|_q.\end{aligned}$$

In a similar way we show that

$$\sup_{s\in K,t\in L}\sum_{j=1}^n t_j g_{m+j}\sqrt{\sum_{i=1}^m s_i^2 a_{ij}^2} = \max_i \|(a_{ij}g_{m+j})_j\|_{p^*}. \qquad \square$$

We shall also use the following lemma, which follows by the Gaussian concentration (see, e.g., [3, Theorem 5.6]) and [17, Lemma 2.3], applied with $X_i \coloneqq \sup_{t\in T_i} X_t - \mathbb{E}\sup_{t\in T_i} X_t$ and $\sigma_i \coloneqq \sup_{t\in T_i}(\mathrm{Var}X_t)^{1/2}$.

Lemma 24. *Let $X = (X_t)_{t\in T}$ be a Gaussian process and let $T_1,\dots,T_k$ be nonempty subsets of T such that $T = \bigcup_{i=1}^k T_i$. Then*

$$\mathbb{E}\sup_{t\in T} X_t \le \max_i \mathbb{E}\sup_{t\in T_i} X_t + C\max_i\sqrt{\mathrm{Log}\, i}\,\sup_{t\in T_i}(\mathrm{Var}X_t)^{1/2}.$$

In particular, if X is centered, then

$$\mathbb{E}\sup_{t\in T} X_t \le \max_i \mathbb{E}\sup_{t\in T_i} X_t + C\sqrt{\mathrm{Log}\, k}\,\sup_{t\in T}(\mathbb{E}X_t^2)^{1/2}. \tag{17}$$

Proof of Proposition 17. Since

$$\sup_{x\in B_{q^*}^m}\sqrt{\sum_{i=1}^m x_i^2 a_{ij}^2} = \max_i |a_{ij}|,$$

Lemma 24 yields for every $q\in[2,\infty)$,

$$\begin{aligned}\mathbb{E}\max_{j\le n}\|(a_{ij}g_i)_i\|_q &\lesssim \max_{j\le n}\mathbb{E}\|(a_{ij}g_i)_i\|_q + \max_j\sqrt{\mathrm{Log}\, j}\max_i|a_{ij}|\\ &\le \max_{j\le n}\big(\mathbb{E}\|(a_{ij}g_i)_i\|_q^q\big)^{1/q} + \max_j\sqrt{\mathrm{Log}\, j}\max_i|a_{ij}|\\ &\lesssim \sqrt{q}D_2 + \max_{i,j}\sqrt{\mathrm{Log}\, j}|a_{ij}|.\end{aligned} \tag{18}$$

In a similar way we show that for every $p \in (1,2]$,

$$\mathbb{E}\max_{i\le m}\|(a_{ij}g_j)_j\|_{p^*} \lesssim \sqrt{p^*}D_1 + \max_{i,j}\sqrt{\operatorname{Log} i}|a_{ij}|.$$

By (16) and Lemma 24 we get

$$\mathbb{E}\max_{i\le m}|Y_i| \lesssim \max_i \sqrt{\operatorname{Log} i}\Big(\sum_j a_{ij}^4\Big)^{1/4}.$$

Moreover, for $p^* < 4$,

$$\begin{aligned}\sqrt{\operatorname{Log} i}\Big(\sum_{j=1}^n a_{ij}^4\Big)^{1/4} &\le \sqrt{\operatorname{Log} i}\Big(\sum_{j=1}^n |a_{ij}|^{p^*}\Big)^{1/4}\max_{j\le n}|a_{ij}|^{(4-p^*)/4} \\ &\le \frac{p^*}{4}\Big(\sum_{j=1}^n |a_{ij}|^{p^*}\Big)^{1/p^*} + \Big(1-\frac{p^*}{4}\Big)\max_{i,j}\operatorname{Log}^{2/(4-p^*)}(i)|a_{ij}|.\end{aligned}$$

Similarly, for $q < 4$,

$$\mathbb{E}\max_{j\le n}|Y_{j+m}| \lesssim \frac{q}{4}\Big(\sum_{i=1}^m |a_{ij}|^q\Big)^{1/q} + \Big(1-\frac{q}{4}\Big)\max_{i,j}\operatorname{Log}^{2/(4-q)}(j)|a_{ij}|,$$

so the assertion follows by Corollary 23. □

3.2. **Range** $p^*, q > 2$**.** The first step in the proof of Proposition 19 in the case $p^*, q > 2$ is Proposition 25 below. In this subsection we follow the ideas from [1] – which also use the Slepian-Fernique lemma – to provide some tools to be used in Section 4.1 to prove Proposition 25.

Proposition 25. *If* $p^*, q \in (2,\infty)$*, then*

$$\mathbb{E}\|G_A\|_{p\to q} \lesssim \sqrt{p^*}D_1 + \sqrt{q}D_2 + \operatorname{Log}^{\gamma(p,q)}(mn)\max_{i,j}|a_{ij}|,$$

where $\gamma(p,q) = \frac12 + \max\{\frac{p}{2-p}, \frac{q^*}{2-q^*}\} = \frac12 + \max\{\frac{p^*}{p^*-2}, \frac{q}{q-2}\} \sim \frac{p^*\wedge q}{p^*\wedge q-2}$.

In order to prove Proposition 25 we split each $s \in B_{q^*}^m$ and $t \in B_p^n$ into two parts: one consisting of vectors with coordinates which do not exceed a certain level and the second one consisting of vectors with non-vanishing coordinates having absolute value exceeding the same level. In order to control the supremum over the points with small coordinates, we first replace the sets $B_p^n \cap bB_\infty^n$ and $B_{q^*}^m \cap aB_\infty^m$ by sets whose extremal points have a very 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. More precisely, we substitute $B_p^n \cap bB_\infty^n$ and $B_{q^*}^m \cap aB_\infty^m$ by the sets $K_{p,n,b}$ and $K_{q^*,m,a}$, respectively, where, for a given $b \in [n^{-1/p}, 1]$,

$$K_{p,n,b} = \operatorname{conv}\big\{(|J|^{-1/p}\eta_j \mathbb{1}_{j\in J})_{j\le n} \colon n \ge |J| \ge b^{-p}, \eta_j \in \{-1,1\}\big\}.$$

The next lemma shows that by doing so we loose only a logarithmic factor.

Lemma 26. *Assume that* $p \in [1,\infty]$*, and* $b \in (n^{-1/p}, 1]$*. Then*

$$B_p^n \cap bB_\infty^n \subset \big(3 + \operatorname{Log}^{1/p^*}(nb^p)\big)K_{p,n,b}.$$

Proof. Fix a vector $x = (x_1,\dots,x_n) \in B_p^n \cap bB_\infty^n$. We only need to prove that

$$\|x\|_{K_{p,n,b}} \le 3 + \operatorname{Log}^{1/p^*}(nb^p), \tag{19}$$

where

$$\|x\|_{K_{p,n,b}} = \inf\{\lambda > 0 \colon x \in \lambda K_{p,n,b}\}.$$

Since both $K_{p,n,b}$ and $B_p^n \cap bB_\infty^n$ are permutationally invariant and unconditional, we may and do assume that $b \ge x_1 \ge \cdots \ge x_n \ge 0$. Set $j_b \coloneqq \lfloor b^{-p} \rfloor + 1 \le n$ and $x_{n+1} \coloneqq 0$. Then

$$x = \sum_{j=1}^n x_j e_j = \sum_{j=1}^{j_b - 1} (x_j - x_{j_b}) e_j + \sum_{j=j_b}^n (x_j - x_{j+1})(e_1 + \cdots + e_j). \tag{20}$$

Since $j_b \ge b^{-p}$, $j^{-1/p}(e_1 + \cdots + e_j) \in K_{p,n,b}$ for every integer $j \in [j_b, n]$, so $\|e_1 + \cdots + e_j\|_{K_{p,n,b}} \le j^{1/p}$ whenever $j \in [j_b, n]$. This, together with the triangle and Hölder inequalities, yields

$$\begin{aligned}
\Big\| \sum_{j=j_b}^n (x_j - x_{j+1})(e_1 + \cdots + e_j) \Big\|_{K_{p,n,b}} &\le \sum_{j=j_b}^n (x_j - x_{j+1}) j^{1/p} \\
&= j_b^{1/p} x_{j_b} + \sum_{j=j_b+1}^n x_j (j^{1/p} - (j-1)^{1/p}) \\
&\le 1 + \sum_{j=j_b+1}^n x_j j^{1/p - 1} \le 1 + \|x\|_p \Big(\sum_{j=j_b+1}^n \frac{1}{j} \Big)^{1/p^*} \\
&\le 1 + \ln^{1/p^*}(nb^p), \qquad (21)
\end{aligned}$$

where we also used the elementary estimates $j^{1/p} - (j-1)^{1/p} \le j^{\frac{1}{p}-1}$ and $\sum_{j=j_b+1}^n \frac{1}{j} \le \int_{j_b}^n \frac{1}{t} dt = \ln(n/j_b)$.

Since $K_{p,n,b}$ is unconditional,

$$\Big\| \sum_{j=1}^{j_b - 1} (x_j - x_{j_b}) e_j \Big\|_{K_{p,n,b}} \le \Big\| \sum_{j=1}^{j_b} b e_j \Big\|_{K_{p,n,b}} \le b j_b^{1/p} \le 2. \tag{22}$$

Estimates (20)–(22) yield the desired inequality (19). $\square$

The next lemma shows how to bound the suprema over the sets $K_{q^*,m,a}$ and $K_{p,n,b}$.

Lemma 27. *Assume that $a \in (m^{1/q^*}, 1]$, $b \in (n^{-1/p}, 1]$, and $p^*, q \ge 2$. Then*

$$\begin{aligned}
\mathbb{E} \sup_{s \in K_{q^*,m,a}, t \in K_{p,n,b}} &\sum_{i,j} a_{ij} g_{ij} s_i t_j \\
&\lesssim a^{1 - q^*/2} \sqrt{p^*} D_1 + b^{1-p/2} \sqrt{q} D_2 + (a^{1-q^*/2} + b^{1-p/2}) \sqrt{\operatorname{Log}(mn)} \max_{i,j} |a_{ij}|.
\end{aligned}$$

Proof. The proof of [1, Proposition 3.1] (see estimate (3.6) and the last formula on page 3492 therein, based onthe Slepian-Fernique lemma) shows that for every $1 \le k \le m$ and $1 \le l \le n$,

$$\begin{aligned}
\mathbb{E} \max_{I,J} \sup_{\eta \in B_\infty^m} \sup_{\eta' \in B_\infty^n} \sum_{i \in I, j \in J} a_{ij} g_{ij} \eta_i \eta'_j &\lesssim \max_{I,J} \sum_{i \in I} \sqrt{\sum_{j \in J} a_{ij}^2} + \max_{I,J} \sum_{j \in J} \sqrt{\sum_{i \in I} a_{ij}^2} \\
&\quad + \mathbb{E} \max_{I,J} \sum_{i \in I} g_i \sqrt{\sum_{j \in J} a_{ij}^2} + \mathbb{E} \max_{I,J} \sum_{j \in J} \widetilde{g}_j \sqrt{\sum_{i \in I} a_{ij}^2},
\end{aligned}$$

where the maxima and the suprema are taken over all sets $I \subset \{1, \ldots, m\}$, $J \subset \{1, \ldots, n\}$ such that $|I| = k$, $|J| = l$, and $g_1, \ldots, g_m, \widetilde{g}_1, \ldots, \widetilde{g}_n$ are independent standard Gaussian variables. Moreover, for every $l \geq b^{-p}$,

$$\begin{aligned}
\frac{1}{k^{1/q^*} l^{1/p}} \max_{I,J} \sum_{i\in I} \sqrt{\sum_{j\in J} a_{ij}^2} &\leq \sup_{s\in B_{q^*}^m} \sup_{t\in B_p^n \cap bB_\infty^n} \sum_{i=1}^m s_i \sqrt{\sum_{j=1}^n t_j^2 a_{ij}^2} \\
&= \sup_{t\in B_p^n\cap bB_\infty^n} \Big(\sum_{i=1}^m \Big(\sum_{j=1}^n t_j^2 a_{ij}^2\Big)^{q/2}\Big)^{1/q} \\
&\leq b^{1-p/2} \sup_{t\in B_p^n} \Big(\sum_{i=1}^m \Big(\sum_{j=1}^n |t_j|^p a_{ij}^2\Big)^{q/2}\Big)^{1/q} \\
&= b^{1-p/2} \max_j \|(a_{ij})_i\|_q = b^{1-p/2} D_2,
\end{aligned}$$

where in the second inequality we used the convexity of $u \mapsto \sum_{i=1}^m \Big|\sum_{j=1}^n u_j a_{ij}^2\Big|^{q/2}$. Likewise, for every $k \geq a^{-q^*}$,

$$\frac{1}{k^{1/q^*} l^{1/p}} \max_{I,J} \sum_{j\in J} \sqrt{\sum_{i\in I} a_{ij}^2} \leq a^{1-q^*/2} D_1.$$

A similar reasoning, together with (18), shows that for every $l \geq b^{-p}$,

$$\begin{aligned}
\frac{1}{k^{1/q^*} l^{1/p}} \mathbb{E} \max_{I,J} \sum_{i\in I} g_i \sqrt{\sum_{j\in J} a_{ij}^2} &\leq b^{1-p/2} \mathbb{E} \max_j \|(a_{ij} g_i)_i\|_q \\
&\lesssim b^{1-p/2} \big(\sqrt{q} D_2 + \max_{i,j} \sqrt{\operatorname{Log} j}\, |a_{ij}|\big).
\end{aligned}$$

Similarly, for every $k \geq a^{-q^*}$,

$$\frac{1}{k^{1/q^*} l^{1/p}} \mathbb{E} \max_{I,J} \sum_{i\in I} g_i \sqrt{\sum_{j\in J} a_{ij}^2} \lesssim a^{1-q^*/2} \big(\sqrt{p^*} D_1 + \max_{i,j} \sqrt{\operatorname{Log} i}\, |a_{ij}|\big).$$

This implies that for every $a^{-q^*} \leq k \leq m$ and $b^{-p} \leq l \leq n$,

$$\begin{aligned}
\text{(23)} \quad \mathbb{E} W_{k,l} \lesssim a^{1-q^*/2} \sqrt{p^*} D_1 + b^{1-p/2} \sqrt{q} D_2 \\
+ (a^{1-q^*/2} + b^{1-p/2}) \sqrt{\operatorname{Log}(mn)} \max_{i,j} |a_{ij}|,
\end{aligned}$$

where

$$W_{k,l} := \sup_{s\in K_{q^*,m}^{(k)}, t\in K_{p,n}^{(l)}} \sum_{i,j} a_{ij} g_{ij} s_i t_j$$

and

$$K_{p,n}^{(l)} := \{ l^{-1/p} (\eta_j \mathbb{1}_{j\in J}) \colon \eta \in \{-1,1\}^n, J \subset [n], |J| = l\}.$$

Hence, by inequality (17) we obtain

$$\begin{aligned}
\mathbb{E} &\sup_{s\in K_{q^*,m,a}, t\in K_{p,n,b}} \sum_{i,j} a_{ij} g_{ij} s_i t_j = \mathbb{E} \max_{a^{-q^*} \leq k \leq m, b^{-p} \leq l \leq n} W_{k,l} \\
&\lesssim \max_{\substack{a^{-q^*} \leq k \leq m \\ b^{-p} \leq l \leq n}} \mathbb{E} W_{k,l} + \sqrt{\operatorname{Log}(mn)} \sup_{\substack{s\in B_{q^*}^m \cap aB_\infty^m \\ t\in B_p^n \cap bB_\infty^n}} \sqrt{\sum_{i,j} a_{ij}^2 s_i^2 t_j^2} \\
&\leq \max_{\substack{a^{-q^*} \leq k \leq m \\ b^{-p} \leq l \leq n}} \mathbb{E} W_{k,l} + a^{1-q^*/2} b^{1-p/2} \sqrt{\operatorname{Log}(mn)} \sup_{\substack{s\in B_{q^*}^m \cap aB_\infty^m \\ t\in B_p^n \cap bB_\infty^n}} \sqrt{\sum_{i,j} a_{ij}^2 |s_i|^{q^*} |t_j|^p}
\end{aligned}$$

$$= \max_{\substack{a^{-q^*}\le k\le m\\ b^{-p}\le l\le n}} \mathbb{E}W_{k,l} + a^{1-q^*/2}b^{1-p/2}\sqrt{\operatorname{Log}(mn)}\max_{i,j}|a_{ij}|.$$

Thus, (23) and the assumption that $a, b \le 1$ yield the assertion. □

4. Dimension dependent bounds

In this section our aim is to obtain dimension dependent bounds from Proposition 19. We begin by proving an even weaker estimate in the range $p, q > 2$, i.e., Proposition 25. Then, in Subsections 4.2 and 4.3, we show how to reduce the exponent of the logarithm appearing in this proposition (and in Proposition 17). The proof of Proposition 19 is provided at the end of this section.

4.1. **Proof of Propostion 25.** Let us recall that in order to prove Proposition 25 we split each $s \in B_{q^*}^m$ and $t \in B_p^n$ into two parts: one consisting of vectors with coordinates which do not exceed a certain level, and the second one consisting of vectors with non-vanishing coordinates having absolute value exceeding the same level. Results of Section 3.2 allow us to bound the supremum over the points from the first part. The next lemma shows how to control the suprema over points with large non-zero coordinates.

Lemma 28. *If $\gamma \ge 1$ and $p^*, q \ge 2$, then*

$$\mathbb{E}\sup_{s\in B_{q^*}^m}\sup_{t\in B_p^n}\sum_{i,j} a_{ij}g_{ij}s_it_j\mathbb{1}_{\{|s_i|\ge\gamma^{-1}\}} \lesssim \sqrt{p^*}D_1 + \sqrt{\gamma^{q^*}\operatorname{Log} m}\max_{i,j}|a_{ij}| \tag{24}$$

and

$$\mathbb{E}\sup_{s\in B_{q^*}^m}\sup_{t\in B_p^n}\sum_{i,j} a_{ij}g_{ij}s_it_j\mathbb{1}_{\{|t_j|\ge\gamma^{-1}\}} \lesssim \sqrt{q}D_2 + \sqrt{\gamma^{p}\operatorname{Log} n}\max_{i,j}|a_{ij}|.$$

In the proof of Lemma 28 we shall use the following standard lemma; we formulate and prove it for the sake of completeness.

Lemma 29. *For every fixed $s \in B_{q^*}^m$, $q \in [2,\infty]$, and $p \in (1, 2]$,*

$$\mathbb{E}\sup_{t\in B_p^n}\sum_{i,j} a_{ij}g_{ij}s_it_j = \mathbb{E}\Big\|\Big(\sum_i a_{ij}g_{ij}s_i\Big)_j\Big\|_{p^*} \le \sqrt{p^*}D_1. \tag{25}$$

Similarly, for every fixed $t \in B_p^n$, $p \in [1, 2]$, and $q \in [2,\infty)$,

$$\mathbb{E}\sup_{s\in B_{q^*}^m}\sum_{i,j} a_{ij}g_{ij}s_it_j = \Big\|\Big(\sum_j a_{ij}g_{ij}t_j\Big)_i\Big\|_q \le \sqrt{q}D_2. \tag{26}$$

Proof. We have

$$\mathbb{E}\sup_{t\in B_p^n}\sum_{i,j} a_{ij}g_{ij}s_it_j = \mathbb{E}\Big(\sum_j\Big(\sum_i a_{ij}g_{ij}s_i\Big)^{p^*}\Big)^{1/p^*} \le \Big(\mathbb{E}\sum_j\Big(\sum_i a_{ij}g_{ij}s_i\Big)^{p^*}\Big)^{1/p^*}$$
$$= \|g_{1,1}\|_{p^*}\Big(\sum_j\Big(\sum_i a_{ij}^2s_i^2\Big)^{p^*/2}\Big)^{1/p^*}.$$

We have $\|g_{1,1}\|_{p^*} \le \sqrt{p^*}$ and, since $\|s\|_2 \le \|s\|_q \le 1$,

$$\Big(\sum_j\Big(\sum_i a_{ij}^2s_i^2\Big)^{p^*/2}\Big)^{2/p^*} = \Big\|\Big(\sum_i s_i^2a_{ij}^2\Big)_j\Big\|_{p^*/2} \le \sum_i s_i^2\|(a_{ij}^2)_j\|_{p^*/2}$$
$$\le \max_i\|(a_{ij}^2)_j\|_{p^*/2} = \max_i\|(a_{ij})_j\|_{p^*}^2 = D_1^2.$$

Hence, (26) follows. Estimate (25) holds by an analogous argument. □

Proof of Lemma 28. By the symmetry it is enough to show (24). For a fixed nonempty set I let

$$X_I := \|(a_{ij}g_{ij})_{i\in I,j\leq n}\|_{p\to q}.$$

Let T be a 1/2-net in $B^I_{q^*}$ (with respect to the ℓ_{q^*} norm) of cardinality at most $5^{|I|}$. Then

$$X_I \leq 2\sup_{s\in T}\sup_{t\in B_p^n}\sum_{i,j} a_{ij}g_{ij}s_it_j.$$

Thus, inequality (17) yields

$$\mathbb{E}X_I \lesssim \sup_{s\in T}\mathbb{E}\sup_{t\in B_p^n}\sum_{i,j} a_{ij}g_{ij}s_it_j + \sqrt{\log|T|}\sup_{s\in T}\sup_{t\in B_p^n}\Big(\sum_{i,j} a_{ij}^2 s_i^2 t_j^2\Big)^{1/2}.$$

Since $T\subset B^m_{q^*}\subset B^m_2$ and $B^n_p\subset B^n_2$ we have

$$\sup_{s\in T}\sup_{t\in B_p^n}\Big(\sum_{i,j} a_{ij}^2 s_i^2 t_j^2\Big)^{1/2} \leq \max_{i,j}|a_{ij}|. \tag{27}$$

Estimate (25) implies that for any fixed $s\in T$,

$$\mathbb{E}\sup_{t\in B_p^n}\sum_{i,j} a_{ij}g_{ij}s_it_j \leq \sqrt{p^*}D_1,$$

hence,

$$\mathbb{E}X_I \lesssim \sqrt{p^*}D_1 + \sqrt{|I|}\max_{i,j}|a_{ij}|. \tag{28}$$

There are at most $Cm^{\gamma^{q^*}}$ subsets of $[m]$ of cardinality at most γ^{q^*}. Thus, estimates (17), (27), and (28) yield

$$\begin{aligned}\mathbb{E}\max_{I\in\mathcal{I}_2(k),k\leq\gamma^{q^*}} X_I &\lesssim \max_{|I|\leq\gamma^{q^*}}\mathbb{E}X_I + \sqrt{\gamma^{q^*}\operatorname{Log} m}\sup_{s\in B^m_{q^*},t\in B^n_p}\Big(\sum_{i,j} a_{ij}^2 s_i^2 t_j^2\Big)^{1/2}\\ &\lesssim \sqrt{p^*}D_1 + \sqrt{\gamma^{q^*}\operatorname{Log} m}\max_{i,j}|a_{ij}|.\end{aligned}$$

To derive (24) it suffices to note that for $s\in B^m_{q^*}$, $|\{i\colon |s_i|\geq\gamma^{-1}\}|\leq\gamma^{q^*}$, so

$$\mathbb{E}\sup_{s\in B^m_{q^*}}\sup_{t\in B^n_p}\sum_{i,j} a_{ij}g_{ij}s_it_j\mathbb{1}_{\{|s_i|\geq\gamma^{-1}\}} \leq \mathbb{E}\sup_{s\in B^m_{q^*},|\mathrm{supp}(s)|\leq\gamma^{q^*}}\sup_{t\in B^n_p}\sum_{i,j} a_{ij}g_{ij}s_it_j. \qquad \square$$

Proof of Proposition 25. We begin with a simple observation that [1, Proposition 1.8] implies

$$\begin{aligned}\mathbb{E}\sup_{s\in B^m_{q^*}\cap m^{-1/q^*}B^m_\infty}\sup_{t\in B^n_p}\sum_{i,j} a_{ij}g_{ij}s_it_j &\leq m^{-1/q^*}\|G_A\|_{p\to 1}\\ &\lesssim m^{-1/q^*}\sup_{t\in B^n_p}\sum_{i=1}^m\Big(\sum_{j=1}^n a_{ij}^2t_j^2\Big)^{1/2} + \sqrt{p^*}m^{-1/q^*}\Big(\sum_{j=1}^n\Big(\sum_{i=1}^m a_{ij}^2\Big)^{p^*/2}\Big)^{1/p^*}\\ &\leq \sup_{t\in B^n_p}\Big(\sum_{i=1}^m\Big(\sum_{j=1}^n a_{ij}^2t_j^2\Big)^{q/2}\Big)^{1/q} + \sqrt{p^*}\sup_{s\in B_{q^*}}\Big(\sum_{j=1}^n\Big(\sum_{i=1}^m a_{ij}^2s_i^2\Big)^{p^*/2}\Big)^{1/p^*}\\ &= D_2 + D_1,\end{aligned} \tag{29}$$

and, similarly,

$$\mathbb{E}\sup_{s\in B^m_{q^*}}\sup_{t\in B^n_p\cap n^{-1/p}B^m_\infty}\sum_{i,j} a_{ij}g_{ij}s_it_j \lesssim D_1 + \sqrt{q}D_2. \tag{30}$$

If $a := \mathrm{Log}^{-2/(2-q^*)}(mn) > m^{-1/q^*}$ and $b := \mathrm{Log}^{-2/(2-p)}(mn) > n^{-1/p}$, then Lemmas 26 and 27, imply

$$\begin{aligned}\mathbb{E} \sup_{s\in B_{q^*}^m\cap aB_\infty^m} \sup_{t\in B_p^n\cap bB_\infty^n} \sum_{i,j} a_{ij}g_{ij}s_it_j &\lesssim \mathrm{Log}(mn)\mathbb{E} \sup_{\substack{s\in K_{q^*,m,a}\\ t\in K_{p,n,b}}} \sum_{i,j} a_{ij}g_{ij}s_it_j \\ &\lesssim \sqrt{p^*}D_1 + \sqrt{q}D_2 + \sqrt{\mathrm{Log}(mn)}\max_{i,j}|a_{ij}|.\end{aligned}$$

If $a \le m^{-1/q^*}$ or $b \le n^{-1/p}$, then (29) and (30) yield

$$\mathbb{E} \sup_{s\in B_{q^*}^m\cap aB_\infty^m} \sup_{t\in B_p^n\cap bB_\infty^n} \sum_{i,j} a_{ij}g_{ij}s_it_j \lesssim \sqrt{p^*}D_1 + \sqrt{q}D_2.$$

Moreover, Lemma 28 (applied with $\gamma = \mathrm{Log}^{2/(2-q^*)}(mn)$) implies

$$\begin{aligned}&\mathbb{E} \sup_{s\in B_{q^*}^m} \sup_{t\in B_p^n} \sum_{i,j} a_{ij}g_{ij}s_it_j\mathbb{1}_{\{|s_i|\ge \mathrm{Log}^{-2/(2-q^*)}(mn)\}} \\ &\qquad\lesssim \sqrt{p^*}D_1 + \sqrt{q}D_2 + \mathrm{Log}^{1/2+q^*/(2-q^*)}(mn)\max_{i,j}|a_{ij}|\end{aligned}$$

and, similarly,

$$\begin{aligned}&\mathbb{E} \sup_{s\in B_{q^*}^m} \sup_{t\in B_p^n} \sum_{i,j} a_{ij}g_{ij}s_it_j\mathbb{1}_{\{|t_j|\ge \mathrm{Log}^{-2/(2-p)}(mn)\}} \\ &\qquad\lesssim \sqrt{p^*}D_1 + \sqrt{q}D_2 + \mathrm{Log}^{1/2+p/(2-p)}(mn)\max_{i,j}|a_{ij}|.\end{aligned}$$

The last four displayed inequalities yield the assertion. □

4.2. **Exponent reduction.** Proposition 17 and Proposition 25 show that

$$\mathbb{E}\|G_A\|_{p\to q} \lesssim \sqrt{p^*}D_1 + \sqrt{q}D_2 + \mathrm{Log}^{\gamma(p,q)}(mn)\max_{i,j}|a_{ij}|$$

whenever $p^*, q \in [2,4)$ or $p^*, q > 2$, with $\gamma(p,q) =: \gamma$ being a constant depending only on p and q. In this and the next subsection we show how to reduce the exponent γ and get Proposition 19. We do it in a similar way as in [13]. The argument is based on the analysis of the graph associated to the matrix A. To run the exponent reduction procedure we also need to obtain some weaker estimates with constants depending on the degree of this graph (we do this in Subsection 4.3). Since we work in the Gaussian setting and not with bounded Bernoulli entries as in [13], we face some new difficulties. It is possible to deal with them making the advantage of the fact that $p^* \wedge q > 2$ (recall that the case $p = 2 = q$ was solved in [12]).

With an $m \times n$ matrix $A = (a_{ij})_{i\le m, j\le n}$ we associate the set

$$E_A := \{(i,j)\colon a_{ij} \ne 0\} \subset [m]\times[n].$$

We set

$$\begin{aligned}d_{1,A} &:= \max_{i\le m}|\{j\le n\colon (i,j)\in E_A\}|,\\ d_{2,A} &:= \max_{j\le n}|\{i\le m\colon (i,j)\in E_A\}|,\\ d_A &:= d_{1,A}\vee d_{2,A}.\end{aligned}$$

We do not assume that A is symmetric, but we may treat $([m],[n],E_A)$ as a bipartite graph. Then d_A is its degree. We write $i \sim_A j$ if $(i,j)\in E_A$, and $I \sim_A j$ (for $I\subset[m]$) if there exists $i\in I$ such that $i\sim_A j$.

By $\rho = \rho_A$ we denote the distance on $[m]\cup[n]$ induced by E_A. A subset $I\subset[m]$ (resp. $J\subset[n]$) is called r-connected if $\rho_A(i, I\setminus\{i\}) \le r$ for every $i\in I$ (resp. $\rho_A(j, J\setminus\{j\}) \le r$ for every $j \in J$). Equivalently a set is r-connected if it is a connected subset of the graph $G_A(r) := ([m],[n],E_A(r))$, where $(i,j)\in E_A(r)$ if and only if $\rho_A(i,j)\le r$.

We denote by $\mathcal{I}_r(k) = \mathcal{I}_r(k, A)$ (resp. $\mathcal{J}_r(k) = \mathcal{J}_r(k, A)$) the family of all r-connected subsets of $[m]$ (resp. $[n]$) of cardinality k. Note that the maximal degree of $G_A(r)$ is at most $d_A + d_A(d_A - 1) + \cdots + d_A(d_A - 1)^{r-1} \leq d_A^r$. Thus, [13, Lemma 11] implies

$$|\mathcal{I}_r(k)| \leq m4^k d_A^{rk} \quad \text{and} \quad |\mathcal{J}_r(k)| \leq n4^k d_A^{rk}. \tag{31}$$

For $I \subset [m]$ we define

$$I' = \{j \in [n] \colon \ \exists_{i \in I}\, (i, j) \in E_A\}$$

In a similar way we define $J' \subset [m]$ for $J \subset [n]$.

The next proposition reveals how one may reduce the exponent at the logarithmic term to deduce the desired bound (7) from a weaker estimate depending on $d_{1,A}$ and $d_{2,A}$.

Proposition 30. *Let $p^*, q \in [2, \infty)$, $0 \leq \gamma_1 < \frac{1}{p^*}$, $0 \leq \gamma_2 < \frac{1}{q}$, and $\alpha_1, \alpha_2, \alpha_3 \geq 1$ be such that for every $m \times n$ matrix A,*

$$\mathbb{E}\|G_A\|_{p\to q} \leq \alpha_1 D_1 + \alpha_2 D_2 + \alpha_3 (d_{1,A}^{\gamma_1} + d_{2,A}^{\gamma_2} + \sqrt{\operatorname{Log}(mn)}) \max_{i,j} |a_{ij}|. \tag{32}$$

Then for every $m \times n$ matrix A

$$\mathbb{E}\|G_A\|_{p\to q} \lesssim \frac{\operatorname{Log}\operatorname{Log}(mn)}{-\ln((\gamma_1 p^*) \vee (\gamma_2 q)))} \cdot \Big((\alpha_1 + \alpha_3) D_1 + (\alpha_2 + \alpha_3) D_2 + \alpha_3 \sqrt{\operatorname{Log}(mn)} \max_{i,j} |a_{ij}|\Big).$$

Moreover, if we assume additionally that for every $m \times n$ matrix A,

$$\mathbb{E}\|G_A\|_{p\to q} \leq \alpha_1 D_1 + \alpha_2 D_2 + \alpha_3 \operatorname{Log}^{\gamma_3}(mn) \max_{i,j} |a_{ij}| \tag{33}$$

for some $\gamma_3 > 0$, then for every $m \times n$ matrix A,

$$\mathbb{E}\|G_A\|_{p\to q} \lesssim \frac{\operatorname{Log} \gamma_3}{-\ln((\gamma_1 p^*) \vee (\gamma_2 q)))} \cdot \Big((\alpha_1 + \alpha_3) D_1 + (\alpha_2 + \alpha_3) D_2 + \alpha_3 \sqrt{\operatorname{Log}(mn)} \max_{i,j} |a_{ij}|\Big).$$

Proof. Let $\delta = ((p^*\gamma_1) \vee (q\gamma_2))^{-1} - 1$. Then $p^*\gamma_1, q\gamma_2 \leq (1+\delta)^{-1}$. Let $k_0 = k_0(p, q)$ be the smallest positive integer satisfying

$$\frac{1}{2}(1+\delta)^{k_0+1} \geq \operatorname{Log}(mn) \geq \max\{\operatorname{Log}(n^{\gamma_1}), \operatorname{Log}(m^{\gamma_2})\},$$

so that $k_0 \lesssim -\operatorname{Log}\operatorname{Log}(mn)/\ln((\gamma_1 p^*) \vee (\gamma_2 q)))$.

Define

$$u_k := (e \operatorname{Log}(mn))^{-\frac{1}{2}(1+\delta)^{k+1}}, \quad k = 0, 1, \ldots, k_0.$$

Let

$$M := D_1 + D_2, \quad \widetilde{M} := \alpha_1 D_1 + \alpha_2 D_2 + \alpha_3 \sqrt{\operatorname{Log}(mn)} \max_{i,j} |a_{ij}|.$$

Define matrices $A_k = (a_{ij}(k))_{i\leq m, j\leq n}$, $k = 0, 1, \ldots, k_0 + 1$ by

$$a_{ij}(0) = a_{ij} \mathbb{1}_{\{|a_{ij}| \geq u_0 M\}}, \qquad a_{ij}(k_0 + 1) = a_{ij} \mathbb{1}_{\{|a_{ij}| < u_{k_0} M\}},$$
$$\text{and} \quad a_{ij}(k) = a_{ij} \mathbb{1}_{\{u_k M \leq |a_{ij}| < u_{k-1} M\}} \quad \text{for } k = 1, \ldots, k_0.$$

Then $A = \sum_{k=0}^{k_0+1} A_k$, so

$$\|G_A\|_{p\to q} \leq \sum_{k=0}^{k_0+1} \|G_{A_k}\|_{p\to q}.$$

Observe that for any $u \ge 0$,

$$uM \max_i |\{j\colon |a_{ij}| \ge uM\}|^{1/p^*} \le \max_i \|(a_{ij})_j\|_{p^*} \le M,$$
$$uM \max_j |\{i\colon |a_{ij}| \ge uM\}|^{1/q} \le \max_j \|(a_{ij})_i\|_q \le M.$$

Thus,

$$d_{1,A_k} \le u_k^{-p^*},\ d_{2,A_k} \le u_k^{-q} \quad \text{for } k = 0,1,\dots,k_0.$$

Since

$$d_{1,A_0}^{\gamma_1} + d_{2,A_0}^{\gamma_2} \le u_0^{-p^*\gamma_1} + u_0^{-q\gamma_2} \le 2u_0^{-1/(1+\delta)} = 2\sqrt{e\operatorname{Log}(mn)},$$

assumption (32) (applied to the matrix A_0) yields

$$\|G_{A_0}\|_{p\to q} \lesssim \widetilde{M}.$$

Moreover, assumption (32), applied to the matrix A_k, yields for $1 \le k \le k_0$,

$$\begin{aligned}\|G_{A_k}\|_{p\to q} &\le \widetilde{M} + \alpha_3 u_{k-1} M(d_{1,A_k}^{\gamma_1} + d_{2,A_k}^{\gamma_2}) \le \widetilde{M} + \alpha_3 M u_{k-1}(u_k^{-p^*\gamma_1} + u_k^{-q\gamma_2}) \\ &\le \widetilde{M} + 2\alpha_3 M u_{k-1} u_k^{-1/(1+\delta)} = \widetilde{M} + 2\alpha_3 M.\end{aligned}$$

Finally, applying (32) to the matrix A_{k_0+1} and using the trivial bounds $d_{1,A_{k_0+1}} \le n$, $d_{1,A_{k_0+1}} \le m$, we get

$$\|G_{A_{k_0+1}}\|_{p\to q} \le \widetilde{M} + \alpha_3(n^{\gamma_1} + m^{\gamma_2}) u_{k_0} M \le \widetilde{M} + 2\alpha_3 M.$$

To conclude the moreover part of the proposition, we define k_0 differently by

$$k_0 := \inf\{k \ge 1\colon (1+\delta)^{k+1} \ge 2\gamma_3\}$$

and apply the assumption (33) to estimate $\|G_{A_{k_0+1}}\|_{p\to q}$ from above by

$$\alpha_1 D_1 + \alpha_2 D_2 + 2\alpha_3 M.$$

We finish the proof by noting that $k_0 \lesssim \frac{\operatorname{Log}\gamma_3}{-\ln((\gamma_1 p^*)\vee(\gamma_2 q))}$. □

4.3. **Degree dependent bounds.** In order to provide a weaker degree dependent estimate (32) we are going to use the following proposition. For $\gamma > 1$ we define the set of γ-flat vectors from B_u^r with support I by

$$K(u,r,I,\gamma) = \Big\{s \in B_u^r\colon \operatorname{supp}(s) = I,\ \sup_{i\in I}|s_i| \le \gamma \inf_{i\in I}|s_i|\Big\}.$$

Note that if $s \in K(u,r,I,\gamma)$, then $\max_{i\in I}|s_i| \le \gamma |I|^{-1/r}$. We also put

$$Y_{k,l}(\gamma) = Y_{k,l}(\gamma, A, p, q) = \max_{I\in\mathcal{I}_4(k)} \max_{J\in\mathcal{J}_4(l)} \sup_{s\in K(q^*,m,I,\gamma)} \sup_{t\in K(p,n,J,\gamma)} \sum_{i\in I, j\in J} a_{ij} g_{ij} s_i t_j.$$

Proposition 31. *For every $\varepsilon \in (0,1]$ and $1 \le p \le 2 \le q \le \infty$,*

$$\mathbb{E}\|G_A\|_{p\to q} \lesssim \frac{1}{\varepsilon}\mathbb{E} \max_{1\le k\le m}\max_{1\le l\le n} Y_{k,l}(d_A^\varepsilon) + \mathbb{E}\max_{i,j}|a_{ij}g_{ij}| \tag{34}$$

$$\lesssim \frac{1}{\varepsilon}\Big(\max_{1\le k\le m}\max_{1\le l\le n} \mathbb{E} Y_{k,l}(d_A^\varepsilon) + \sqrt{\operatorname{Log}(mn)}\|(a_{ij})\|_\infty\Big). \tag{35}$$

Proof. Fix $\varepsilon \in (0,1]$. For $s \in B_{q^*}^m$, $t \in B_p^n$ and $k,l = 1,2,\dots$ define

$$I_k(s) := \{i \in [m]\colon\ d_A^{-k\varepsilon} < |s_i| \le d_A^{(1-k)\varepsilon}\},$$
$$J_l(t) := \{j \in [n]\colon\ d_A^{-l\varepsilon} < |t_j| \le d_A^{(1-l)\varepsilon}\}.$$

Then

$$\sum_{k\ge1} d_A^{-q^*k\varepsilon}|I_k(s)| \le \|s\|_{q^*}^{q^*}, \quad \sum_{l\ge1} d_A^{-pl\varepsilon}|J_l(t)| \le \|t\|_p^p$$

and

$$\mathbb{E}\|G_A\|_{p\to q} = \mathbb{E}\sup_{\|s\|_{q^*}\le 1}\sup_{\|t\|_p\le 1}\sum_{k,l\ge 1}\sum_{i\in I_k(s)}\sum_{j\in J_l(t)} a_{ij}g_{ij}s_it_j.$$

Define

$$M_A = \max_{i,j}|a_{ij}g_{ij}|.$$

Observe that for any $s\in B^m_{q^*}$ and $t\in B^n_p$,

$$\begin{aligned}\Big|\sum_{k\ge 1}\sum_{l\ge k+1/\varepsilon+1}\sum_{i\in I_k(s)}\sum_{j\in J_l(t)} a_{ij}g_{ij}s_it_j\Big| &\le \sum_{k\ge 1}\sum_{i\in I_k(s)}|s_i|\sum_{l\ge k+1/\varepsilon+1}\sum_{j\in J_l(t)}|a_{ij}g_{ij}||t_j| \\ &\le M_A\sum_{k\ge 1}\sum_{i\in I_k(s)}|s_i|d_A^{-k\varepsilon-1}\sum_j \mathbb{1}_{E_A}(i,j) \\ &\le M_A\sum_{k\ge 1}\sum_{i\in I_k(s)}s_i^2 = M_A\|s\|_2^2\le M_A\|s\|_{q^*}^2\le M_A.\end{aligned}$$

Similarily,

$$\Big|\sum_{l\ge 1}\sum_{k\ge l+1/\varepsilon+1}\sum_{i\in I_k(s)}\sum_{j\in J_l(t)} a_{ij}g_{ij}s_it_j\Big|\le M_A\|t\|_2^2\le M_A.$$

Therefore, it is enough to estimate

$$\begin{aligned}&\sum_{|r|\le 1/\varepsilon+1}\mathbb{E}\sup_{\|s\|_{q^*}\le 1}\sup_{\|t\|_p\le 1}\sum_{\substack{k,l\ge 1\\ l-k=r}}\sum_{i\in I_k(s)}\sum_{j\in J_l(t)} a_{ij}g_{ij}s_it_j \\ &\qquad\le (2/\varepsilon+3)\max_{|r|\le 1/\varepsilon+1}\mathbb{E}\sup_{\|s\|_{q^*}\le 1}\sup_{\|t\|_p\le 1}\sum_{\substack{k,l\ge 1\\ l-k=r}}\sum_{i\in I_k(s)}\sum_{j\in J_l(t)} a_{ij}g_{ij}s_it_j.\end{aligned}$$

Let us fix $|r|\le 1/\varepsilon+1$, $s\in B^m_{q^*}$, and $t\in B^n_p$. For $k,l\ge 1$ with $l-k=r$ let $I_{k,1},\ldots,I_{k,u_k}$ be 2-connected components of $I_k(s)\cap J_l(t)'$ and $J_{k,u} := \{j\in J_l(t)\colon I_{k,u}\sim_A j\}$, $1\le u\le u_k$. Then the sets $(I_{k,u})_{k\ge 1,1\le u\le u_k}$ are nonempty pairwise disjoint 2-connected subsets of $[m]$ and the sets $(J_{k,u})_{k\ge 1,1\le u\le u_k}$ are nonempty pairwise disjoint 4-connected subsets of $[n]$. Define vectors $s_{k,u},\bar{s}_{k,u}\in\mathbb{R}^m$ and $t_{k,u},\bar{t}_{k,u}\in\mathbb{R}^n$ by

$$\begin{aligned}s_{k,u} &:= (s_i\mathbb{1}_{\{i\in I_{k,u}\}})_{i\le m},\ \bar{s}_{k,u} := \frac{s_{k,u}}{\|s_{k,u}\|_{q^*}},\\ t_{k,u} &:= (t_j\mathbb{1}_{\{j\in J_{k,u}\}})_{j\le n},\ \bar{t}_{k,u} := \frac{t_{k,u}}{\|t_{k,u}\|_p}.\end{aligned}$$

Since $I_{k,u}\subset I_k(s)$ and $J_{k,u}\subset J_l(t)$, s_i and t_j do not vanish if $i\in I_{k,u}$ and $j\in J_{k,u}$. Thus, $\bar{s}_{k,u}\in K(q^*,m,I_{k,u},d_A^\varepsilon)$ and $\bar{t}_{k,u}\in K(p,n,J_{k,u},d_A^\varepsilon)$, so

$$\begin{aligned}\sum_{\substack{k,l\ge 1\\ l-k=r}}\sum_{i\in I_k(s)}\sum_{j\in J_l(t)} a_{ij}g_{ij}s_it_j &= \sum_k\sum_{1\le u\le u_k}\sum_{i\in I_{k,u}}\sum_{j\in J_{k,u}} a_{ij}g_{ij}s_it_j \\ &\le \sum_k\sum_{1\le u\le u_k} Y_{|I_{k,u}|,|J_{k,u}|}(d_A^\varepsilon)\|s_{k,u}\|_{q^*}\|t_{k,u}\|_p \\ &\le \max_{1\le\widetilde{k}\le m,1\le\widetilde{l}\le n} Y_{\widetilde{k},\widetilde{l}}(d_A^\varepsilon)\sum_k\sum_{1\le u\le u_k}\|s_{k,u}\|_{q^*}\|t_{k,u}\|_p.\end{aligned}$$

Moreover,

$$\sum_k\sum_{1\le u\le u_k}\|s_{k,u}\|_{q^*}\|t_{k,u}\|_p\le\Big(\sum_k\sum_{1\le u\le u_k}\|s_{k,u}\|_{q^*}^2\Big)^{1/2}\Big(\sum_k\sum_{1\le u\le u_k}\|t_{k,u}\|_p^2\Big)^{1/2}$$

$$\le \Big(\sum_k \sum_{1\le u\le u_k} \|s_{k,u}\|_{q^*}^{q^*}\Big)^{1/2} \Big(\sum_k \sum_{1\le u\le u_k} \|t_{k,u}\|_p^p\Big)^{1/2}$$
$$= \|s\|_{q^*}^{q^*/2} \|t\|_p^{p/2} \le 1.$$

The above calculations show that

$$\|G_A\|_{p\to q} \lesssim \frac{1}{\varepsilon} \max_{1\le k\le m, 1\le l\le n} Y_{k,l}(d_A^\varepsilon) + M_A,$$

which yield bound (34).

Moreover, $\mathbb{E} M_A \lesssim \mathrm{Log}^{1/2}(mn)\|(a_{ij})\|_\infty$ and, by estimate (17),

$$\begin{aligned}\mathbb{E} \max_{1\le k\le m, 1\le l\le n} Y_{k,l}(d_A^\varepsilon) &\lesssim \max_{1\le k\le m, 1\le l\le n} \mathbb{E} Y_{k,l}(d_A^\varepsilon) \\ &\quad + \sqrt{\mathrm{Log}(mn)} \sup_{s\in B_{q^*}^m} \sup_{t\in B_p^n} \Big(\sum_{i,j} a_{ij}^2 s_i^2 t_j^2\Big)^{1/2} \\ &= \max_{1\le k\le m, 1\le l\le n} \mathbb{E} Y_{k,l}(d_A^\varepsilon) + \sqrt{\mathrm{Log}(mn)}\|(a_{ij})\|_\infty,\end{aligned}$$

so estimate (35) follows. □

Now we need a bound for the expectation of $Y_{k,l}(\gamma)$. It is derived in the following proposition.

Proposition 32. *For every $p^*, q \in [2,\infty)$ and $\gamma \ge 1$,*

$$\begin{aligned}\mathbb{E} Y_{k,l}(\gamma) &\lesssim \sqrt{p^*} D_1 + \sqrt{q} D_2 + \sqrt{\mathrm{Log}(mn)} \max_{i,j} |a_{ij}| \\ &\quad + \gamma^{3/2} k^{-1/q^*} l^{-1/p} \min\{k l^{1/2}, k^{1/2} l, k d_{1,A}^{1/2}, l d_{2,A}^{1/2}\} \sqrt{\mathrm{Log}(d_A)} \max_{i,j} |a_{ij}|. \end{aligned} \tag{36}$$

Proof. Observe that

$$Y_{k,l}(\gamma) \le \max_{I\in \mathcal{I}_4(k)} X_I,$$

where

$$\begin{aligned} X_I &= X_I(p,q,k,l,\gamma) \\ &:= \sup\Big\{\sum_{i\in I, j} a_{ij} g_{ij} s_i t_j \colon s \in B_{q*}^I \cap \gamma k^{-1/q^*} B_\infty^I,\ t \in B_p^n \cap \gamma l^{-1/p} B_\infty^n\Big\}.\end{aligned}$$

Define

$$\begin{aligned}\sigma_{k,l} &= \sigma_{k,l}(p,q,\gamma) \\ &:= \sup\Big\{\sqrt{\sum_{i,j} a_{ij}^2 s_i^2 t_j^2} \colon s \in B_{q*}^m \cap \gamma k^{-1/q^*} B_\infty^m,\ t \in B_p^n \cap \gamma l^{-1/p} B_\infty^n\Big\}.\end{aligned}$$

Let us fix $I \in \mathcal{I}_4(k)$ and choose a 1/2-net T in $B_{q^*}^I \cap \gamma k^{-1/q^*} B_\infty^I$ (with respect to the norm determined by this set) of cardinality 5^k. Then

$$\mathbb{E} X_I \le 2\mathbb{E} \sup_{s\in T} \sup_{t\in B_p^n \cap \gamma l^{-1/p} B_\infty^m} \sum_{i\in I, j} a_{ij} g_{ij} s_i t_j.$$

By inequality (17) we have

$$\mathbb{E} X_I \lesssim \sup_{s\in T} \mathbb{E} \sup_{t\in B_p^n \cap \gamma l^{-1/p} B_\infty^m} \sum_{i\in I, j} a_{ij} g_{ij} s_i t_j + \sqrt{\mathrm{Log}\,|T|}\sigma_{k,l}.$$

By (25)

$$\sup_{s\in T} \mathbb{E} \sup_{t\in B_p^n \cap \gamma l^{-1/p} B_\infty^m} \sum_{i\in I, j} a_{ij} g_{ij} s_i t_j \le \sqrt{p^*} D_1.$$

Thus,

$$\mathbb{E}X_I \lesssim \sqrt{p^*}D_1 + \sqrt{k}\sigma_{k,l}.$$

Applying estimate (17) again and using (31) we get

$$\begin{aligned}\mathbb{E}Y_{k,l} &\lesssim \max_{I\in\mathcal{I}_4(k)} \mathbb{E}X_I + \sqrt{\operatorname{Log}|\mathcal{I}_4(k)|}\sigma_{k,l}\\ &\lesssim \sqrt{p^*}D_1 + (\sqrt{k\operatorname{Log}(d_A)} + \sqrt{\operatorname{Log} m})\sigma_{k,l}. \qquad (37)\end{aligned}$$

To estimate $\sigma_{k,l}$ observe first that $B^I_{q^*} \subset B^I_2$ and $B^n_p \subset B^n_2$, thus

$$\sup_{t\in B^I_{q*}} \sup_{s\in B^n_p} \sum_{i\in I,j} a_{ij}^2 s_i^2 t_j^2 \le \max_{i,j} |a_{ij}|^2.$$

Moreover, for any $s \in B^m_{q*} \cap \gamma k^{-1/q^*} B^m_\infty$ and $t \in B^n_p \cap \gamma l^{-1/p} B^n_\infty$ we have

$$\begin{aligned}\sum_{i,j} a_{ij}^2 s_i^2 t_j^2 &\le \|s\|_\infty^{2-q^*}\|t\|_\infty^{2-p} \max_{i,j}|a_{ij}|^2 \sum_{i,j}|s_i|^{q^*}|t_j|^p\\ &\le \gamma^{4-q^*-p}k^{1-2/q^*}l^{1-2/p}\max_{i,j}|a_{ij}|^2\\ &\le \gamma^2 k^{1-2/q^*}l^{1-2/p}\max_{i,j}|a_{ij}|^2\end{aligned}$$

and

$$\begin{aligned}\sum_{i,j} a_{ij}^2 s_i^2 t_j^2 &\le \|s\|_\infty^{2-q^*}\|t\|_\infty^2 \sum_{i,j} a_{ij}^2|s_i|^{q*} \le \|s\|_\infty^{2-q^*}\|t\|_\infty^2 \max_i \sum_j a_{ij}^2\\ &\le \|s\|_\infty^{2-q^*}\|t\|_\infty^2 d_{1,A}\max_{i,j}|a_{ij}|^2 \le \gamma^{4-q^*}k^{1-2/q^*}l^{-2/p}d_{1,A}\max_{i,j}|a_{ij}|^2\\ &\le \gamma^3 k^{1-2/q^*}l^{-2/p}d_{1,A}\max_{i,j}|a_{ij}|^2.\end{aligned}$$

Therefore

$$\sigma_{k,l} \le \min\{1, \gamma k^{1/2-1/q^*}l^{1/2-1/p}, \gamma^{3/2}k^{1/2-1/q^*}l^{-1/p}d_{1,A}^{1/2}\}\max_{i,j}|a_{ij}|. \qquad (38)$$

Estimates (37) and (38) yield

$$\begin{aligned}\mathbb{E}Y_{k,l}(\gamma) &\lesssim \sqrt{p^*}D_1 + \sqrt{\operatorname{Log} m}\max_{i,j}|a_{ij}|\\ &\quad + \gamma^{3/2}k^{1-1/q^*}l^{-1/p}\min\{l^{1/2}, d_{1,A}^{1/2}\}\sqrt{\operatorname{Log}(d_A)}\max_{i,j}|a_{ij}|.\end{aligned}$$

In a similar way we show that

$$\begin{aligned}\mathbb{E}Y_{k,l}(\gamma) &\lesssim \sqrt{q}D_2 + \sqrt{\operatorname{Log} n}\max_{i,j}|a_{ij}|\\ &\quad + \gamma^{3/2}k^{-1/q^*}l^{1-1/p}\min\{k^{1/2}, d_{2,A}^{1/2}\}\sqrt{\operatorname{Log}(d_A)}\max_{i,j}|a_{ij}|. \qquad \square\end{aligned}$$

To make use of the two previous propositions we consider the cases $\frac{1}{p} + \frac{1}{q^*} \ge 3/2$ and $\frac{1}{p} + \frac{1}{q^*} \le 3/2$ separately.

Corollary 33. *Suppose that $p^*, q \in [2,\infty)$ are such that $\frac{1}{p} + \frac{1}{q^*} \ge 3/2$. Then for every $\varepsilon \in (0,1]$,*

$$\mathbb{E}\|G_A\|_{p\to q} \lesssim \varepsilon^{-1}\Big(\sqrt{p^*}D_1 + \sqrt{q}D_2 + (\varepsilon^{-1/2}d_A^\varepsilon + \sqrt{\operatorname{Log}(mn)})\max_{i,j}|a_{ij}|\Big).$$

Proof. By Proposition 31 we have

$$\mathbb{E}\|G_A\|_{p\to q} \lesssim \frac{1}{\varepsilon}\Big(\max_{1\le k\le m}\max_{1\le l\le n}\mathbb{E}Y_{k,l}(d_A^{\varepsilon/3}) + \sqrt{\operatorname{Log}(mn)}\|(a_{ij})\|_\infty\Big).$$

Observe that in the case $p^*, q \in [2,\infty)$, $\frac{1}{p}+\frac{1}{q^*} \geq 3/2$ we have for any k, l,

$$\begin{aligned} k^{-1/q^*} l^{-1/p} \min\{kl^{1/2}, k^{1/2}l, kd_{1,A}^{1/2}, ld_{2,A}^{1/2}\} &\leq \min\{k^{1-1/q^*} l^{1/2-1/p}, k^{1/2-1/q^*} l^{1-1/p}\} \\ &\leq (k \wedge l)^{3/2-1/q^*-1/p} \leq 1. \end{aligned}$$

Hence estimate (36) yields that

$$\mathbb{E} Y_{k,l}(d_A^{\varepsilon/3}) \lesssim \sqrt{p^*} D_1 + \sqrt{q} D_2 + (d_A^{\varepsilon/2}\sqrt{\operatorname{Log} d_A} + \sqrt{\operatorname{Log}(mn)}) \max_{i,j} |a_{ij}|.$$

Finally we observe that $\sup_{x\geq 1} x^{-\varepsilon} \operatorname{Log} x \lesssim \varepsilon^{-1}$, so $d_A^{\varepsilon/2}\sqrt{\operatorname{Log} d_A} \lesssim d_A^{\varepsilon}\varepsilon^{-1/2}$. □

Corollary 34. *If $p^*, q \in [2,\infty)$ are such that $\frac{1}{p}+\frac{1}{q^*} \leq \frac{3}{2}$ and $q \vee p^* > 2$, then*

$$\begin{aligned} \mathbb{E}\|G_A\|_{p\to q} \lesssim \frac{(q\vee p^*)^3}{q\vee p^*-2}\Big[&\sqrt{p^*}D_1+\sqrt{q}D_2 \\ &+ \Big(\frac{(q\vee p^*)^{3/2}}{(q\vee p^*-2)^{1/2}}\big(d_{1,A}^{\frac{3}{2p^*}-\frac{q}{2(p^*+q)}} + d_{2,A}^{\frac{3}{2q}-\frac{p^*}{2(p^*+q)}}\big) + \sqrt{\operatorname{Log}(mn)}\Big)\max_{i,j}|a_{ij}|\Big]. \end{aligned}$$

Proof. We proceed as in the proof of the previous corollary. The only difference is a more delicate estimate of

$$\pi_{k,l} := k^{-1/q^*} l^{-1/p} \min\{kl^{1/2}, k^{1/2}l, kd_{1,A}^{1/2}, ld_{2,A}^{1/2}\}.$$

Suppose $2 \leq p^* \leq q$ and $q > 2$ (the case when $2 \leq q \leq p^*$ and $p^* > 2$ follows by duality). Let $\rho > 0$ be a constant to be chosen later. We consider two cases.

Case 1. $k \leq \rho l$. Assumption $\frac{1}{p}+\frac{1}{q^*} \leq \frac{3}{2}$ implies that $\frac{1}{2}+\frac{1}{p^*}-\frac{1}{q^*} \geq 0$, and so

$$\begin{aligned} \pi_{k,l} &\leq k^{1/p^*-1/q^*} \min\{k^{1/2}(k/l)^{1/2-1/p^*}, d_{1,A}^{1/2}(k/l)^{1/p}\} \\ &\leq k^{1/p^*-1/q^*}\rho^{1/2-1/p^*}\big(\min\{k,\rho d_{1,A}\}\big)^{1/2} \\ &\leq (\rho d_{1,A})^{1/p^*-1/q^*+1/2}\rho^{1/2-1/p^*} = \rho^{1/q} d_{1,A}^{1/p^*+1/q-1/2}. \end{aligned}$$

Case 2. $k \geq \rho l$. Then

$$\begin{aligned} \pi_{k,l} &\leq k^{1/p^*-1/q^*}\big(\min\{k,d_{2,A}\}\big)^{1/2}(l/k)^{1/p^*} \leq \rho^{-1/p^*}k^{1/p^*-1/q^*}\big(\min\{k,d_{2,A}\}\big)^{1/2} \\ &\leq \rho^{-1/p^*} d_{2,A}^{1/p^*-1/q^*+1/2} = \rho^{-1/p^*} d_{2,A}^{1/p^*+1/q-1/2}. \end{aligned}$$

Choosing $\rho := (d_{2,A}/d_{1,A})^{1-p^*q/(2(p^*+q))}$ we obtain in both cases

$$\pi_{k,l} \leq d_{1,A}^{1/p^*-q/(2(p^*+q))} d_{2,A}^{1/q-p^*/(2(p^*+q))} \leq \frac{1}{2}\Big(d_{1,A}^{2/p^*-q/(p^*+q)} + d_{2,A}^{2/q-p^*/(p^*+q)}\Big).$$

Therefore, Propositions 31 and 32 (applied to $\varepsilon := \varepsilon/3$ and $\gamma := d_A^{\varepsilon/3}$) yield for every $\varepsilon \in (0,1)$,

$$\begin{aligned} \mathbb{E}\|G_A\|_{p\to q} \lesssim \frac{1}{\varepsilon}\Big(&\sqrt{p^*}D_1+\sqrt{q}D_2 \\ &+ \Big(d_A^{\frac{\varepsilon}{2}}\sqrt{\operatorname{Log} d_A}\Big(d_{1,A}^{\frac{2}{p^*}-\frac{q}{p^*+q}} + d_{2,A}^{\frac{2}{q}-\frac{p^*}{p^*+q}}\Big) + \sqrt{\operatorname{Log}(mn)}\Big)\max_{i,j}|a_{ij}|\Big). \end{aligned}$$

Recall that

$$d_A^{\varepsilon/2}\sqrt{\log d_A} \lesssim \varepsilon^{-1/2} d_A^{\varepsilon},$$

and $d_A = d_{1,A} \vee d_{2,A}$. Note that $\frac{2}{p^*}-\frac{q}{p^*+q} < \frac{1}{p^*}$ and $\frac{2}{q}-\frac{p^*}{p^*+q} < \frac{1}{q}$, so we may take

$$\varepsilon_0 := \frac{1}{2}\min\Big\{\frac{q}{p^*+q}-\frac{1}{p^*}, \frac{p^*}{p^*+q}-\frac{1}{q}\Big\} \quad \text{and} \quad \varepsilon := \frac{p^*}{q}\varepsilon_0.$$

For such a choice of ε we have $\frac{2}{p^*}-\frac{q}{p^*+q}+\varepsilon \leq \frac{2}{p^*}-\frac{q}{p^*+q}+\varepsilon_0 \leq \frac{3}{2p^*}-\frac{q}{2(p^*+q)} < \frac{1}{p^*}$ and $\frac{2}{q}-\frac{p^*}{p^*+q}+\varepsilon \leq \frac{2}{q}-\frac{p^*}{p^*+q}+\varepsilon_0 \leq \frac{3}{2q}-\frac{p^*}{2(p^*+q)} < \frac{1}{q}$. Moreover, the assumption

$1/p+1/q^*\le 3/2$ yields that $1/p^*+1/q\ge 1/2$ and thus $2/p^*-q/(p^*+q)\ge 0$ and $2/q-p^*/(p^*+q)\ge 0$. Hence, the AM-GM inequality implies

$$\begin{aligned} d_{1,A}^{\frac{2}{p^*}-\frac{q}{p^*+q}}d_{2,A}^{\varepsilon} &\le \frac{\frac{2}{p^*}-\frac{q}{p^*+q}}{\frac{2}{p^*}-\frac{q}{p^*+q}+\varepsilon_0}d_{1,A}^{\frac{2}{p^*}-\frac{q}{p^*+q}+\varepsilon_0}+\frac{\varepsilon_0}{\frac{2}{p^*}-\frac{q}{p^*+q}+\varepsilon_0}d_{2,A}^{\frac{\varepsilon}{\varepsilon_0}(\frac{2}{p^*}-\frac{q}{p^*+q}+\varepsilon_0)} \\ &\le d_{1,A}^{\frac{2}{p^*}-\frac{q}{p^*+q}+\varepsilon_0}+d_{2,A}^{\frac{2}{q}-\frac{p^*}{p^*+q}+\varepsilon_0} \end{aligned}$$

and

$$\begin{aligned} d_{1,A}^{\varepsilon}d_{2,A}^{\frac{2}{q}-\frac{p^*}{p^*+q}} &\le \frac{\varepsilon_0}{\frac{2}{q}-\frac{p^*}{p^*+q}+\varepsilon_0}d_{1,A}^{\frac{\varepsilon}{\varepsilon_0}(\frac{2}{q}-\frac{p^*}{p^*+q}+\varepsilon_0)}+\frac{\frac{2}{q}-\frac{p^*}{p^*+q}}{\frac{2}{q}-\frac{p^*}{p^*+q}+\varepsilon_0}d_{2,A}^{\frac{2}{q}-\frac{p^*}{p^*+q}+\varepsilon_0} \\ &\le d_{1,A}^{\frac{\varepsilon}{\varepsilon_0}(\frac{2}{q}-\frac{p^*}{p^*+q})+\varepsilon}+d_{2,A}^{\frac{2}{q}-\frac{p^*}{p^*+q}+\varepsilon_0}\le d_{1,A}^{\frac{\varepsilon_0}{\varepsilon}(\frac{2}{q}-\frac{p^*}{p^*+q})+\varepsilon}+d_{2,A}^{\frac{2}{q}-\frac{p^*}{p^*+q}+\varepsilon_0} \\ &\le d_{1,A}^{\frac{2}{p^*}-\frac{q}{p^*+q}+\varepsilon_0}+d_{2,A}^{\frac{2}{q}-\frac{p^*}{p^*+q}+\varepsilon_0}. \end{aligned}$$

To finish the proof it suffices to note that $\varepsilon^{-1}\lesssim\frac{q^3}{(p^*-1)(q-1)-1}\le\frac{q^3}{q-2}$. □

Proof of Proposition 19. The assertion follows by

- Remark 4 if $p=q=2$,
- Corollary 34 and Propositions 17 and 30 if $p^*,q\in[2,4)$ and $p^*\vee q>2$ (note that in this case we have $1/p+1/q^*<3/2$),
- Corollary 33, applied with $\varepsilon=(2(p^*\vee q))^{-1}$, and Propositions 25 and 30 if $p^*,q\in(2,\infty)$ and $1/p+1/q^*\ge 3/2$.
- Corollary 34 and Propositions 25 and 30 if $p^*,q\in(2,\infty)$ and $1/p+1/q^*<3/2$.

To get the claimed bounds on $\alpha(p,q)$ we observe that in the case $1/p+1/q^*<3/2$ we have $1>\frac{3}{2}-\frac{qp^*}{2(p^*+q)}\ge 1/2$, so that

$$-\ln\Big(\frac{3}{2}-\frac{qp^*}{2(p^*+q)}\Big)\sim\frac{qp^*}{2(p^*+q)}-\frac{1}{2}\gtrsim\frac{p^*\vee q-2}{p^*\vee q}.$$ □

Moreover, Corollaries 33 and 34, and Proposition 30 imply the following.

Proposition 35. *Suppose that $p^*=2$ and $q\in[4,\infty)$ or that $q=2$ and $p^*\in[4,\infty)$. Then*

$$\mathbb{E}\|G_A\|_{p\to q}\gtrsim_{p,q}\operatorname{Log}\operatorname{Log}(mn)\Big(D_1+D_2+\sqrt{\operatorname{Log}(mn)}\max_{i,j}|a_{ij}|\Big).$$

5. Proof of Proposition 18

To prove Proposition 18 we decompose the underlying matrix A into block diagonal matrices A_k (with blocks of appropriately small size) and matrices B_l whose norm may be controlled by the following proposition providing a crude, but dimension-independent bound.

Proposition 36. *Let $p^*,q\in[2,\infty)$. Then for every partition $J_1,J_2,\ldots,J_{k_0}$ of $[n]$,*

$$\mathbb{E}\|G_A\|_{p\to q}\le\mathbb{E}\|G_A\|_{2\to q}\lesssim\sqrt{q}D_2+\max_{1\le k\le k_0}\max_{1\le l\le k}k^3|J_l|^{1/2}\max_{i\le m,j\in J_k}|a_{ij}|. \tag{39}$$

Similarily, for every partition $I_1,\ldots,I_{k_0}$ of $[m]$,

$$\mathbb{E}\|G_A\|_{p\to q}\le\mathbb{E}\|G_A\|_{p\to 2}\lesssim\sqrt{p^*}D_1+\max_{1\le k\le k_0}\max_{1\le l\le k}k^3|I_l|^{1/2}\max_{i\in I_k,j\le n}|a_{ij}|. \tag{40}$$

In order to show Proposition 36 we need the following two lemmas; they allow us to perform the induction step.

Lemma 37. *For every $a,b\in\mathbb{R}^m$, $c\in\mathbb{R}$ and $q\geq 2$ we have*

$$\mathbb{E}\Big(\Big(\sum_{i\leq m}|a_i+b_ig_i|^q\Big)^{2/q}+c^2\Big)^{1/2}\leq\Big(\Big(\sum_{i\leq m}|a_i|^q\Big)^{2/q}+q\Big(\sum_{i\leq m}|b_i|^q\Big)^{2/q}+c^2\Big)^{1/2}.$$

Proof. Jensen's inequality implies that

$$\mathbb{E}\Big(\Big(\sum_{i\leq m}|a_i+b_ig_i|^q\Big)^{2/q}+c^2\Big)^{1/2}\leq\Big(\Big(\sum_{i\leq m}\mathbb{E}|a_i+b_ig_i|^q\Big)^{2/q}+c^2\Big)^{1/2}.$$

The hypercontractivity of Gaussian variables and the triangle inequality in $L_{q/2}$ yield

$$\begin{aligned}\Big(\sum_{i\leq m}\mathbb{E}|a_i+b_ig_i|^q\Big)^{2/q}&\leq\Big(\sum_{i\leq m}(\mathbb{E}|a_i+\sqrt{q}b_ig_i|^2)^{q/2}\Big)^{2/q}=\|(a_i^2+qb_i^2)_i\|_{q/2}\\&\leq\|(a_i^2)_i\|_{q/2}+q\|(b_i^2)_i\|_{q/2}\\&=\Big(\sum_{i\leq m}|a_i|^q\Big)^{2/q}+q\Big(\sum_{i\leq m}|b_i|^q\Big)^{2/q}.\end{aligned}$$ □

Lemma 38. *For every $q\geq 2$, $J\subset J'\subset[n]$, a finite set $T\subset B_2^{J'}$ and functions $b=(b_1,\ldots,b_m)\colon T\to\mathbb{R}^m$ and $c\colon T\to\mathbb{R}$ we have*

$$\begin{aligned}&\mathbb{E}\sup_{t\in T}\Big(\Big(\sum_{i\leq m}\Big|b_i(t)+\sum_{j\in J}a_{ij}g_{ij}t_j\Big|^q\Big)^{2/q}+c(t)^2\Big)^{1/2}\\&\leq\sup_{t\in T}\mathbb{E}\Big(\Big(\sum_{i\leq m}\Big|b_i(t)+\sum_{j\in J}a_{ij}g_{ij}t_j\Big|^q\Big)^{2/q}+c(t)^2\Big)^{1/2}+C\sqrt{\operatorname{Log}|T|}\max_{i\leq m,j\in J}|a_{ij}|.\end{aligned}$$

Proof. Observe that

$$\Big(\Big(\sum_{i\leq m}\Big|b_i(t)+\sum_{j\in J}a_{ij}g_{ij}t_j\Big|^q\Big)^{2/q}+c(t)^2\Big)^{1/2}=\sup_{s\in B_{q^*}^m,v\in B_2^2}X_{s,v,t},$$

where

$$X_{s,v,t}:=v_1\sum_{i\leq m}s_i\Big(b_i(t)+\sum_{j\in J}a_{ij}g_{ij}t_j\Big)+v_2c(t).$$

We have

$$\begin{aligned}\sup_{t\in T}\sup_{s\in B_{q^*}^m,v\in B_2^2}\operatorname{Var}(X_{s,v,t})&=\sup_{t\in T}\sup_{s\in B_{q^*}^m,v\in B_2^2}v_1^2\sum_{i\leq m,j\in J}a_{ij}^2s_i^2t_j^2\\&\leq\sup_{t\in B_2^J}\sup_{s\in B_{q^*}^m}\sum_{i\leq m,j\in J}a_{ij}^2s_i^2t_j^2=\max_{i\leq m,j\in J}|a_{ij}|^2.\end{aligned}$$

Hence the assertion follows by Lemma 24. □

Proof of Proposition 36. We will prove (39); the second estimate (40) follows by duality.

For $1\leq k\leq k_0$ let $T_{1,k}$ be a 2^{-2k}-net in $B_2^{J_k}$ of cardinality at most $2^{4k|J_k|}$. Set

$$J_{\leq k}:=\bigcup_{l\leq k}J_l,\quad J_{>k}:=\bigcup_{l>k}J_l$$

and for $t\in\mathbb{R}^n$,

$$\pi_k(t):=(t_j)_{j\in J_k}\in\mathbb{R}^{J_k},\quad\pi_{\leq k}(t):=(t_j)_{j\in J_{\leq k}}\in\mathbb{R}^{J_{\leq k}}.$$

Define

$$\begin{aligned}T_{1,\leq k}&:=\{t\in B_2^{J_{\leq k}}\colon\ \pi_l(t)\in T_{1,l},1\leq l\leq k\},\\T_1&:=T_{1,\leq k_0}=\{t\in B_2^n\colon\ \pi_k(t)\in T_{1,k},1\leq k\leq k_0\}.\end{aligned}$$

Then T_1 is a $\frac{1}{2}$-net in B_2^n, and hence,

$$\mathbb{E}\|G_A\|_{p\to q} \le \mathbb{E}\|G_A\|_{2\to q} \le 2\mathbb{E}\sup_{t\in T_1}\Big(\sum_i\Big|\sum_j a_{ij}g_{ij}t_j\Big|^q\Big)^{1/q}.$$

We will show by the reverse induction on k that for $0\le k\le k_0$,

$$\begin{aligned}\mathbb{E}\sup_{t\in T_1}\Big(\sum_i\Big|\sum_j a_{ij}g_{ij}t_j\Big|^q\Big)^{1/q} &\le \mathbb{E}\sup_{t\in T_1}\Big(\Big(\sum_i\Big|\sum_{j\in J_{\le k}} a_{ij}g_{ij}t_j\Big|^q\Big)^{2/q} + qD_2^2\sum_{j\in J_{>k}} t_j^2\Big)^{1/2} \\ &\quad + C\sum_{l>k}\sqrt{\operatorname{Log}|T_{1,\le l}|}\max_i\max_{j\in J_l}|a_{ij}|. \end{aligned}\tag{41}$$

For $k=k_0$ inequality (41) holds with equality. Observe that

$$\begin{aligned}&\sup_{t\in T_1}\Big(\Big(\sum_i\Big|\sum_{j\in J_{\le k}} a_{ij}g_{ij}t_j\Big|^q\Big)^{2/q} + qD_2^2\sum_{j\in J_{>k}} t_j^2\Big)^{1/2} \\ &\qquad = \sup_{\tilde{t}\in T_{1,\le k}}\Big(\Big(\sum_i\Big|\sum_{j\in J_{\le k}} a_{ij}g_{ij}\tilde{t}_j\Big|^q\Big)^{2/q} + c(\tilde{t})^2\Big)^{1/2},\end{aligned}$$

where

$$c(\tilde{t}) := \sqrt{q}D_2\sup_{t\in T_1\colon \pi_{\le k}(t)=\tilde{t}}\Big(\sum_{j\in J_{>k}} t_j^2\Big)^{1/2}.$$

Let $\mathbb{E}_{J_k}$ denote the integration with respect to random variables $(g_{ij})_{i\le m, j\in J_k}$. Conditional application of Lemma 38 yields

$$\begin{aligned}&\mathbb{E}_{J_k}\sup_{\tilde{t}\in T_{1,\le k}}\Big(\Big(\sum_i\Big|\sum_{j\in J_{\le k}} a_{ij}g_{ij}\tilde{t}_j\Big|^q\Big)^{2/q} + c(\tilde{t})^2\Big)^{1/2} \\ &\le \sup_{\tilde{t}\in T_{1,\le k}}\mathbb{E}_{J_k}\Big(\Big(\sum_i\Big|\sum_{j\in J_{\le k}} a_{ij}g_{ij}\tilde{t}_j\Big|^q\Big)^{2/q} + c(\tilde{t})^2\Big)^{1/2} + C\sqrt{\operatorname{Log}|T_{1,\le k}|}\max_i\max_{j\in J_k}|a_{ij}|.\end{aligned}$$

Lemma 37, applied conditionally, implies that for any $\tilde{t}\in T_{1,\le k}$,

$$\begin{aligned}&\mathbb{E}_{J_k}\Big(\Big(\sum_i\Big|\sum_{j\in J_{\le k}} a_{ij}g_{ij}\tilde{t}_j\Big|^q\Big)^{2/q} + c(\tilde{t})^2\Big)^{1/2} \\ &\qquad \le \Big(\Big(\sum_i\Big|\sum_{j\in J_{\le k-1}} a_{ij}g_{ij}\tilde{t}_j\Big|^q\Big)^{2/q} + \tilde{c}(\tilde{t})^2\Big)^{1/2},\end{aligned}$$

where

$$\tilde{c}(\tilde{t})^2 = q\Big(\sum_i\Big|\sum_{j\in J_k} a_{ij}^2\tilde{t}_j^2\Big|^{q/2}\Big)^{2/q} + c(\tilde{t})^2.$$

Note that

$$\begin{aligned}\Big(\sum_i\Big|\sum_{j\in J_k} a_{ij}^2\tilde{t}_j^2\Big|^{q/2}\Big)^{2/q} &\le \sum_{j\in J_k}\tilde{t}_j^2\sup_{x\in B_1^{J_k}}\Big(\sum_i\Big|\sum_{j\in J_k} a_{ij}^2x_j\Big|^{q/2}\Big)^{2/q} \\ &= \sum_{j\in J_k}\tilde{t}_j^2\max_{j\in J_k}\Big(\sum_i|a_{ij}|^q\Big)^{2/q} \le D_2^2\sum_{j\in J_k}\tilde{t}_j^2,\end{aligned}$$

so

$$\tilde{c}(\tilde{t})^2 \le qD_2^2\sup_{t\in T_1\colon \pi_{\le k}(t)=\tilde{t}}\sum_{j\in J_{>k-1}} t_j^2.$$

Thus,

$$
\begin{aligned}
&\mathbb{E}\sup_{t\in T_1}\Big(\Big(\sum_i\Big|\sum_{j\in J_{\le k}} a_{ij}g_{ij}t_j\Big|^q\Big)^{2/q}+qD_2^2\sum_{j\in J_{>k}}t_j^2\Big)^{1/2}\\
&\le \mathbb{E}\sup_{\tilde t\in T_{1,\le k}}\Big(\Big(\sum_i\Big|\sum_{j\in J_{\le k-1}} a_{ij}g_{ij}\tilde t_j\Big|^q\Big)^{2/q}+\tilde c(\tilde t)^2\Big)^{1/2}+C\sqrt{\operatorname{Log}|T_{1,\le k}|}\max_i\max_{j\in J_k}|a_{ij}|\\
&\le \mathbb{E}\sup_{t\in T_1}\Big(\Big(\sum_i\Big|\sum_{j\in J_{\le k-1}} a_{ij}g_{ij}t_j\Big|^q\Big)^{2/q}+qD_2^2\sum_{j\in J_{>k-1}}t_j^2\Big)^{1/2}\\
&\qquad + C\sqrt{\operatorname{Log}|T_{1,\le k}|}\max_i\max_{j\in J_k}|a_{ij}|,
\end{aligned}
$$

so the reverse induction step immediately follows.

Estimate (41) for $k=0$ gives

$$
\mathbb{E}\sup_{t\in T_1}\Big(\sum_i\Big(\sum_j a_{ij}g_{ij}t_j\Big)^q\Big)^{1/q}\le\sqrt{q}D_2+C\sum_{k=1}^{k_0}\sqrt{\operatorname{Log}|T_{1,\le k}|}\max_i\max_{j\in J_k}|a_{ij}|.
$$

We have

$$
\operatorname{Log}|T_{1,\le k}|\le\sum_{l=1}^{k}\operatorname{Log}|T_{1,l}|\lesssim\sum_{l=1}^{k}l|J_l|,
$$

so that

$$
\begin{aligned}
\sum_{k=1}^{k_0}\sqrt{\operatorname{Log}|T_{1,\le k}|}\max_i\max_{j\in J_k}|a_{ij}| &\lesssim \sum_{1\le l\le k\le k_0}\sqrt{l}|J_l|^{1/2}\max_i\max_{j\in J_k}|a_{ij}|\\
&\le \max_{1\le l\le k\le k_0}k^3|J_l|^{1/2}\max_i\max_{j\in J_k}|a_{ij}|\cdot\sum_{1\le l\le k\le k_0}\frac{\sqrt{l}}{k^3}\\
&\lesssim \max_{1\le l\le k\le k_0}k^3|J_l|^{1/2}\max_i\max_{j\in J_k}|a_{ij}|. \qquad \square
\end{aligned}
$$

Now we are ready to prove Proposition 18.

Proof of Proposition 18. Let $r=8(p^*\vee q)$ and

$$
D'_\infty:=\max_{i,j}(i+j)^{1/r}|a_{ij}|.
$$

Define the sequence $(n_k)_{k\ge0}$ by

$$
n_0:=0,\ n_k:=\lceil e^{r^k}\rceil\quad\text{for } k\ge1.
$$

Then

$$
\sqrt{\operatorname{Log}(n_kn_l)}\le\sqrt{\operatorname{Log}n_k}+\sqrt{\operatorname{Log}n_l}\lesssim r(n_{k-1}+n_{l-1}+2)^{1/r}\text{ for } k,l\ge1. \tag{42}
$$

For $k=1,2,\ldots$ define

$$
\begin{aligned}
I_k&:=\{i\in[m]\colon n_{k-1}<i\le n_k\},\\
J_k&:=\{j\in[n]\colon n_{k-1}<j\le n_k\}.
\end{aligned}
$$

Then

$$
\max_{i\in I_k,j\in J_l}(n_{k-1}+n_{l-1}+2)^{1/r}|a_{ij}|\le\max_{i\in I_k,j\in J_l}(i+j)^{1/r}|a_{ij}|\le D'_\infty. \tag{43}
$$

Let

$$
A_{k,l}:=(a_{ij})_{i\in I_k,j\in J_l}.
$$

Then Proposition 19 and estimates (42) and (43) yield

$$\begin{aligned}\mathbb{E}\|G_{A_{k,l}}\|_{p\to q} &\lesssim \alpha(p,q)\Big(\sqrt{p^*}D_1+\sqrt{q}D_2+\sqrt{\mathrm{Log}(|I_k||J_l|)}\max_{i\in I_k,j\in J_l}|a_{ij}|\Big)\\ &\lesssim \alpha(p,q)(\sqrt{p^*}D_1+\sqrt{q}D_2+rD'_\infty). \qquad (44)\end{aligned}$$

For $r\in\mathbb{Z}$ set

$$A_r=\sum_{k\geq \max\{0,-r\}}(a_{ij}\mathbb{1}_{\{i\in I_k,j\in J_{k+r}\}})_{i\leq m,j\leq n}.$$

For every r (after deleting some zero rows and columns) A_r is block diagonal with blocks $A_{k,k+r}$. Hence, Lemmas 21 and 24, together with estimates (43) and (44), imply

$$\begin{aligned}\mathbb{E}\|G_{A_r}\|_{p\to q}&=\mathbb{E}\max_k\|G_{A_{k,k+r}}\|_{p\to q}\\ &\leq \max_k\mathbb{E}\|G_{A_{k,k+r}}\|_{p\to q}+\max_k\sqrt{\mathrm{Log}\,k}\max_{i\in I_k,j\in J_{k+r}}|a_{ij}|\\ &\lesssim \alpha(p,q)(\sqrt{p^*}D_1+\sqrt{q}D_2+rD'_\infty).\end{aligned}$$

We have

$$A=\sum_{r=-1}^{1}A_r+B_1+B_2,$$

where

$$B_1=\sum_l\sum_{k\geq l+2}(a_{ij}\mathbb{1}_{\{i\in I_k,j\in J_l\}}),\qquad B_2=\sum_k\sum_{l\geq k+2}(a_{ij}\mathbb{1}_{\{i\in I_k,j\in J_l\}}).$$

Estimate (39) applied to the matrix B_1 yields

$$\begin{aligned}\mathbb{E}\|G_{B_1}\|_{p\to q}&\lesssim \sqrt{q}D_2+\max_l\max_{l'\leq l}l^3\sqrt{|J_{l'}|}\max_{k\geq l+2}\max_{i\in I_k}\max_{j\in J_l}|a_{ij}|\\ &\leq \sqrt{q}D_2+\max_l l^3\sqrt{n_l}\max_{k\geq l+2}\max_{i\in I_k}\max_{j\in J_l}|a_{ij}|.\end{aligned}$$

Observe that if $i\in I_k$, $j\in J_l$ and $k\geq l+2$ then

$$l^3\sqrt{n_l}\lesssim n_{l+1}^{1/r}\leq n_{k-1}^{1/r}\leq i^{1/r}\leq (i+j)^{1/r},$$

Hence,

$$\mathbb{E}\|G_{B_1}\|_{p\to q}\lesssim\sqrt{q}D_2+D'_\infty.$$

Similarily, estimate (40) implies

$$\mathbb{E}\|G_{B_2}\|_{p\to q}\lesssim\sqrt{p}D_1+D'_\infty. \qquad \square$$

Proceeding similarly as in the proof of Proposition 18 we may show that Proposition 35 implies

$$\mathbb{E}\|G_A\|_{p\to q}\lesssim_{p,q}\mathrm{Log}\,\mathrm{Log}(mn)\Big(D_1+D_2+\max_{i,j}(i+j)^{(8(p^*\vee q))^{-1}}|a_{ij}|\Big)$$

when $p^*=2$ and $q\geq 4$ or $q=2$ and $p^*\geq 4$. This estimate, together with Proposition 35, Proposition 20 and estimate (12), yields Theorem 6.

Sketch of the proof of Proposition 5. In a similar way as in the proof of Proposition 19 (using assumption (1) instead of Proposition 25) we show that

$$\mathbb{E}\|G_A\|_{p\to q}\lesssim \alpha'(p,q)\Big(\sqrt{p^*}D_1+\sqrt{q}D_2+\sqrt{\mathrm{Log}(mn)}\max_{i,j}|a_{ij}|\Big),$$

where

$$\alpha'(p,q)=\begin{cases}(p^*\vee q)^{3/2}\alpha\operatorname{Log}\gamma & \text{if } \dfrac{1}{p}+\dfrac{1}{q^*}\ge 3/2,\\ \dfrac{(p^*\vee q)^{11/2}\alpha\operatorname{Log}\gamma}{(p^*\vee q-2)^{5/2}} & \text{if } \dfrac{1}{p}+\dfrac{1}{q^*}< 3/2\end{cases}\le \frac{(p^*\vee q)^{11/2}\alpha\operatorname{Log}\gamma}{(p^*\vee q-2)^{5/2}}.$$

Then, as in the proof of Proposition 18, we obatin the dimension-free estimate

$$\mathbb{E}\|G_A\|_{p\to q}\lesssim (p^*\vee q)\alpha'(p,q)\Big(D_1+D_2+\max_{i,j}(i+j)^{(8(p^*\vee q))^{-1}}|a_{ij}|\Big)$$

and the assertion follows by Proposition 20. □

Appendix. Estimates in the range $p^*,q\in[2,4)$ with the constant bounded in a neighbourhood of the point $(p,q)=(2,2)$

The aim of this appendix is to prove that constants in Theorem 2 are bounded if $p^*,q\in[2,4)$ and $p^*\vee q$ is separated from 4. Namely, the following result holds.

Theorem 39. *If $\delta\in(0,1]$ and $p^*,q\in[2,4-\delta)$, then for every deterministic matrix $A=(a_{ij})_{i\le m,j\le n}$ we have*

$$\begin{aligned}\mathbb{E}\|G_A\|_{p\to q}&\sim_\delta \max_i\|(a_{ij})_j\|_{p^*}+\max_j\|(a_{ij})_i\|_q+\mathbb{E}\max_{i,j}|a_{ij}g_{ij}|\\ &\sim \max_i\|(a_{ij})_j\|_{p^*}+\max_j\|(a_{ij})_i\|_q+\max_{k\ge0}\inf_{|I|=k}\sqrt{\operatorname{Log}k}\max_{(i,j)\notin I}|a_{ij}|\\ &\sim \max_i\|(a_{ij})_j\|_{p^*}+\max_j\|(a_{ij})_i\|_q+\max_{k\ge0}\inf_{|I|=|J|=k}\sqrt{\operatorname{Log}k}\max_{i\notin I,j\notin J}|a_{ij}|\\ &\sim \mathbb{E}\max_i\|(a_{ij}g_j)_j\|_{p^*}+\mathbb{E}\max_j\|(a_{ij}g_i)_i\|_q.\end{aligned}$$

Moreover, the constant in the first lower bound does not depend on δ and the constant in the first upper bound is bounded by $(C/\delta)^{2/\delta}$.

The proof of Theorem 2 shows that to establish Theorem 39 it is enough to prove that for $p^*,q\in[2,4-\delta)$

$$\mathbb{E}\|G_A\|_{p\to q}\le\Big(\frac{C}{\delta}\Big)^{2/\delta}\Big(D_1+D_2+\max_{k\ge0}\inf_{|I|=k}\sqrt{\operatorname{Log}k}\max_{(i,j)\notin I}|a_{ij}|\Big). \tag{45}$$

The crucial new tool we need to establish the above bound is the following result, which allows a similar exponent reduction as Corollaries 33 and 34 in Section 4.3.

Proposition 40. *If $p^*,q\in[2,4-\delta)$ and $\varepsilon\in(0,1/2]$, then*

$$\mathbb{E}\|G_A\|_{p\to q}\lesssim\varepsilon^{-2}\Big[D_1+D_2+\Big(\Big(\varepsilon^{-1/2}+\Big(\frac{C}{\delta}\Big)^{2/\delta}\Big)d_A^{1/4+\varepsilon}+\sqrt{\operatorname{Log}(mn)}\Big)\max_{i,j}|a_{ij}|\Big].$$

We may deduce the desired estimate (45) from Proposition 40 (applied with $\varepsilon=(4-p^*\vee q)/(8(p^*\vee q))\sim 4-p^*\vee q\ge\delta$); to do so we may use the same arguments as in the proof of Theorem 2 in the case $p^*,q\in[2,4)$ with the constant given in Remark 3.

The rest of the appendix is devoted to the proof of Proposition 40. To this end we will need the following modification of Proposition 17.

Proposition 41. *If $p^*,q\in[2,4)$ and $a,b\in(0,1]$, then*

$$\begin{aligned}&\mathbb{E}\sup_{s\in B_{q^*}^m\cap aB_\infty^m}\ \sup_{t\in B_p^n\cap bB_\infty^n}\sum_{i,j}a_{ij}g_{ij}s_it_j\\ &\quad\lesssim D_1+D_2+\Big(a^{(2-q^*)/2}\operatorname{Log}^{\frac{2}{4-p^*}}(ma^{q^*})+b^{(2-p)/2}\operatorname{Log}^{\frac{2}{4-q}}(nb^p)\Big)\max_{i,j}|a_{ij}|.\end{aligned}$$

The proof of Proposition 41 is based on Lemma 22 and the following quite standard lemma (cf. [13, Lemma 14] for cases $\rho = 1, 2$).

Lemma 42. *Suppose that $1 \leq \rho \leq \rho_0$, $c \in (0,1]$, $m_j, \sigma_j \geq 0$ and nonnegative random variables $Z_1, \ldots, Z_n$ satisfy*

$$\mathbb{P}(Z_j \geq m_j + t\sigma_j) \leq e^{-t^2/2} \quad \text{for every } 1 \leq j \leq n.$$

Then

$$\mathbb{E} \sup_{u \in B_2^n \cap cB_\infty^n} \Big(\sum_{j=1}^n u_j^2 Z_j^\rho\Big)^{1/\rho} \lesssim_{\rho_0} \max_j m_j + \sqrt{\operatorname{Log}(nc^2)} \max_j \sigma_j.$$

Proof. Let k be a positive integer such that $\frac{1}{k+1} < c^2 \leq \frac{1}{k}$ and $(Z_1^*, \ldots, Z_n^*)$ be the nonincreasing rearrangement of $(Z_1, \ldots, Z_n)$. Let $m = \max_j m_j$, $\sigma = \max_j \sigma_j$, and $M = (m + \sqrt{2\operatorname{Log}(n/k)}\sigma)^\rho$. Then

$$\begin{aligned}
\Big(\frac{1}{k}\sum_{l=1}^k \mathbb{E}(Z_l^*)^\rho\Big)^{1/\rho} &\leq 2\Big(M + \frac{1}{k}\sum_{l=1}^n \mathbb{E}(Z_l - m - \sqrt{2\operatorname{Log}(n/k)}\sigma)_+^\rho\Big)^{1/\rho} \\
&\lesssim \Big(M + \frac{1}{k}\sum_{l=1}^n \sigma^\rho \int_0^\infty \rho t^{\rho-1}\, \mathbb{P}\big(Z_l \geq m + \sigma(\sqrt{2\operatorname{Log}(n/k)} + t)\big)dt\Big)^{1/\rho} \\
&\leq \Big(M + \frac{n}{k}\sigma^\rho \int_0^\infty \rho t^{\rho-1} e^{-(t+\sqrt{2\operatorname{Log}(n/k)})^2/2} dt\Big)^{1/\rho} \\
&\leq \Big(M + \sigma^\rho \int_0^\infty \rho t^{\rho-1} e^{-t^2/2} dt\Big)^{1/\rho} \lesssim_{\rho_0} m + \sqrt{\operatorname{Log}(n/k)}\sigma,
\end{aligned}$$

so

$$\begin{aligned}
\mathbb{E} \sup_{u \in B_2^n \cap cB_\infty^n} \Big(\sum_{j=1}^n u_j^2 Z_j^\rho\Big)^{1/\rho} &\leq \mathbb{E}\Big(\frac{1}{k}\sum_{l=1}^k (Z_l^*)^\rho\Big)^{1/\rho} \lesssim_{\rho_0} m + \sqrt{\operatorname{Log}(n/k)}\sigma \\
&\sim m + \sqrt{\operatorname{Log}(nc^2)}\sigma.
\end{aligned}$$

□

Proof of Proposition 41. We apply Lemma 22 with $K = B_{q^*}^m \cap aB_\infty^m$ and $L = B_p^n \cap bB_\infty^n$. Observe that $\|s\|_2 \leq \|s\|_\infty^{(2-q^*)/2} \|s\|_{q^*}^{q^*/2}$. This, together with a similar calculation for ℓ_p-norms, yields

$$K \subset a^{(2-q^*)/2}(B_2^m \cap a^{q^*/2}B_\infty^m), \quad L \subset b^{(2-p)/2}(B_2^n \cap b^{p/2}B_\infty^n). \tag{46}$$

Estimate (16) implies that $Y_1, \ldots, Y_m$ are centered Gaussians with $\operatorname{Var}(Y_i) \leq \|(a_{ij})_j\|_4^2$. Thus, (46) and Lemma 42, applied with $Z_j = |Y_j|$, $\rho = 1$, $c = a^{q^*/2}$, $m_j = 0$ and $\sigma_j = \|(a_{ij})_j\|_4$, yield that

$$\begin{aligned}
\mathbb{E} \sup_{s \in K} \sum_{i=1}^m s_i^2 Y_i &\leq a^{2-q^*} \mathbb{E} \sup_{s \in B_2^m \cap a^{q^*/2}B_\infty^m} \sum_{i=1}^m s_i^2 |Y_i| \\
&\lesssim a^{2-q^*} \sqrt{\operatorname{Log}(ma^{q^*})} \max_i \|(a_{ij})_j\|_4 \\
&\leq a^{2-q^*} \sqrt{\operatorname{Log}(ma^{q^*})} \max_i \|(a_{ij})_j\|_{p^*}^{p^*/4} \max_{i,j} |a_{ij}|^{(4-p^*)/4} \\
&\leq \frac{p^*}{4} D_1 + \frac{4-p^*}{4} a^{4(2-q^*)/(4-p^*)} \operatorname{Log}^{2/(4-p^*)}(ma^{q^*}) \max_{i,j} |a_{ij}| \\
&\leq D_1 + a^{(2-q^*)/2} \operatorname{Log}^{2/(4-p^*)}(ma^{q^*}) \max_{i,j} |a_{ij}|.
\end{aligned}$$

In a similar way we show that

$$\sup_{t\in L}\sum_{j=1}^n t_j^2 Y_{m+j} \lesssim b^{2-p}\sqrt{\operatorname{Log}(nb^p)}\max_j \|(a_{ij})_i\|_4$$
$$\le D_2 + b^{(2-p)/2}\operatorname{Log}^{2/(4-q)}(nb^p)\max_{i,j}|a_{ij}|.$$

By (46) and the convexity of the function $x\mapsto |x|^{q/2}$ we get

$$\sup_{s\in K,t\in L}\sum_{i=1}^m s_i g_i\sqrt{\sum_{j=1}^n t_j^2 a_{ij}^2} \le \sup_{t\in L}\Big(\sum_{i=1}^m |g_i|^q\Big(\sum_{j=1}^n t_j^2 a_{ij}^2\Big)^{q/2}\Big)^{1/q}$$
$$\le b^{(2-p)/2}\sup_{t\in B_2^n\cap b^{p/2}B_\infty^n}\Big(\sum_{i=1}^m |g_i|^q\Big(\sum_{j=1}^n t_j^2 a_{ij}^2\Big)^{q/2}\Big)^{1/q}$$
$$\le b^{(2-p)/2}\sup_{t\in B_2^n\cap b^{p/2}B_\infty^n}\Big(\sum_{j=1}^n t_j^2\|(a_{ij}g_i)_i\|_q^q\Big)^{1/q}\Big(\sum_{j=1}^n t_j^2\Big)^{1/2-1/q}$$
$$\le b^{(2-p)/2}\sup_{t\in B_2^n\cap b^{p/2}B_\infty^n}\Big(\sum_{j=1}^n t_j^2\|(a_{ij}g_i)_i\|_q^q\Big)^{1/q}. \tag{47}$$

Note that

$$\mathbb{E}\|(a_{ij}g_i)_i\|_q \le (\mathbb{E}\|(a_{ij}g_i)_i\|_q^q)^{1/q} \le \sqrt{q}\|(a_{ij})_i\|_q,$$

and

$$\sup_{s\in B_2^n}\|(a_{ij}s_i)_i\|_q = \max_i |a_{ij}|.$$

Therefore, the Gaussian concentration (see, e.g., [3, Theorem 5.6]) yields

$$\mathbb{P}\big(\|(a_{ij}g_i)_i\|_q \ge \sqrt{q}\|(a_{ij})_i\|_q + t\max_i|a_{ij}|\big) \le e^{-t^2/2}.$$

Estimate (47) and Lemma 42, applied with $Z_j = \|(a_{ij}g_i)_i\|_q$, $\rho = q$, $c = b^{p/2}$, and $\rho_0 = 4$, imply

$$\sup_{s\in K,t\in L}\sum_{i=1}^m s_i g_i\sqrt{\sum_{j=1}^n t_j^2 a_{ij}^2} \lesssim \max_j \|(a_{ij})_i\|_q + b^{(2-p)/2}\sqrt{\operatorname{Log}(nb^p)}\max_{i,j}|a_{i.j}|$$
$$\le D_2 + b^{(2-p)/2}\operatorname{Log}^{2/(4-q^*)}(nb^p)\max_{i,j}|a_{i.j}|.$$

In a similar way we show that

$$\sup_{s\in K,t\in L}\sum_{j=1}^n t_j g_j\sqrt{\sum_{i=1}^m s_i^2 a_{ij}^2} \lesssim D_1 + a^{(2-q^*)/2}\operatorname{Log}^{2/(4-p)}(ma^{q^*})\max_{i,j}|a_{i.j}|. \qquad \square$$

We also use the following consequence of Lemma 27 and Proposition 31.

Corollary 43. *If $2\le p^*, q<\infty$, then*

$$\mathbb{E}\|G_A\|_{p\to q}$$
$$\lesssim \operatorname{Log}(d_A)\Big(\sqrt{p^*}\max_i \|(a_{ij})_j\|_{p^*} + \sqrt{q}\max_j \|(a_{ij})_i\|_q + \sqrt{\operatorname{Log}(nm)}\max_{i,j}|a_{ij}|\Big).$$

Proof. If $I\subset[m]$, $J\subset[n]$, and $\gamma>1$, then for every $(b_i)_{i=1}^m$ and $(c_j)_{j=1}^n$,

$$\sup_{s\in K(q^*,m,I,\gamma)}\sum_{i=1}^m b_i s_i \le \gamma \sup_{s\in K_{q^*,m,1}}\sum_{i=1}^m b_i s_i$$

and

$$\sup_{t\in K(p,n,J,\gamma)}\sum_{j=1}^{n}c_jt_j\le\gamma\sup_{t\in K_{p,n,1}}\sum_{j=1}^{n}c_jt_j.$$

Hence,

$$Y_{k,l}(d_A^{\varepsilon})\le d_A^{2\varepsilon}\sup_{s\in K_{q^*,m,1}}\sup_{t\in K_{p,n,1}}\sum_{i,j}a_{ij}g_{ij}s_it_j.$$

We choose $\varepsilon=1/\operatorname{Log}(d_A)$ and apply Lemma 27 (with $a=b=1$) and Proposition 31. □

As in Section 4.3, we estimate $\mathbb{E}\|G_A\|_{p\to q}$ using Propostion 31, so we need to upper bound $\mathbb{E}Y_{k,l}(d_A^{\varepsilon})$. For $I\subset[m]$ and $\alpha\ge 0$ define

$$B(p,I,\alpha)=B(p,I,\alpha,A):=\Big\{t\in B_p^n\colon\ \max_{i\in I}\sum_{j=1}^{n}a_{ij}^2t_j^2\le\alpha^2\Big\},$$

and for $I\subset[m]$ and $J\subset[n]$ let

$$I''=\{i\in[m]\colon\ \exists_{j\in I'}\,(i,j)\in E_A\}\quad\text{ and }\quad J''=\{j\in[n]\colon\ \exists_{i\in J'}\,(i,j)\in E_A\}.$$

Without loss of generality we may assume that matrix A has no zero rows and columns, so $I\subset I''$ and $J\subset J''$ for any $I\subset[m]$ and $J\subset[n]$.

Lemma 44. *For every $p\le 2$, $1\le r\le k\le m$ and $1\le l\le n$ we have*

$$\begin{aligned}\mathbb{E}Y_{k,l}(d_A^{\varepsilon})&\\ \le\mathbb{E}&\max_{I\in\mathcal{I}_4(k)}\max_{I_0\subset I,|I_0|=r}\max_{J\in\mathcal{J}_4(l)}\sup_{s\in K(q^*,m,I,d_A^{\varepsilon})}\sup_{t\in K(p,n,J,d_A^{\varepsilon})}\sum_{i\in I_0''\cap I,j\in I_0'\cap J}a_{ij}g_{ij}s_it_j\\ &+\mathbb{E}\max_{I\in\mathcal{I}_4(k)}\sup_{s\in B_{q^*}^m}\sup_{t\in B(p,I,\alpha_r)}\sum_{i\in I,j\in[n]}a_{ij}g_{ij}s_it_j,\end{aligned}\tag{48}$$

where

$$\alpha_r=\min\{1,d_A^{\varepsilon}l^{1/2-1/p}r^{-1/2}\}\max_{i,j}|a_{ij}|.$$

Proof. Let us fix $I\in\mathcal{I}_4(k)$ and $J\in\mathcal{J}_4(l)$ and consider the following greedy algorithm with output being a subset $I_0=\{i_1,\ldots,i_r\}$ of I of size r.

- In the first step we pick a vertex $i_1\in I$ with maximal number of neighbours in J.
- Once we chose $\{i_1,\ldots,i_u\}$ for $u<r$, we pick $i_{u+1}\in I\setminus\{i_1,\ldots,i_u\}$ with maximal number of neighbours in $J\setminus\{i_1,\ldots,i_u\}'$.

If l_u denotes the number of neighbours of i_u in $J\setminus\{i_1,\ldots,i_{u-1}\}'$, then $l_1\ge l_2\ge\ldots\ge l_r$, so $rl_r\le|J|=l$. Hence, using this algorithm we get a subset $I_0\subset I$ with cardinality r such that for every $i\in I\setminus I_0$, $|\{j\in J\setminus I_0'\colon(i,j)\in E_A\}|\le l/r$.

Observe that if $i\in I$ and $j\in I_0'\cap J$ are such that $a_{ij}\neq 0$, then $i\in I_0''\cap I$, and if $i\in I$ and $j\in J\setminus I_0'$ are such that $a_{ij}\neq0$, then $i\in I\setminus I_0$. Hence, for any $s\in B_{q^*}^m$ and $t\in B_p^n$,

$$\sum_{i\in I,j\in J}a_{ij}g_{ij}s_it_j=\sum_{i\in I_0''\cap I,j\in I_0'\cap J}a_{ij}g_{ij}s_it_j+\sum_{i\in I\setminus I_0,j\in J\setminus I_0'}a_{ij}g_{ij}s_it_j.$$

Moreover, for any $i\in I\setminus I_0$ and $t\in K(p,n,J,d_A^{\varepsilon})$,

$$\sum_{j\in J\setminus I_0'}a_{ij}^2t_j^2\le\max_{i,j}a_{ij}^2\min\Big\{1,\frac{l}{r}\max_{j\in J}t_j^2\Big\}\le\max_{i,j}a_{ij}^2\min\{1,d_A^{2\varepsilon}l^{1-2/p}r^{-1}\}=\alpha_r^2.$$

Therefore,

$$\max_{I\in\mathcal{I}_4(k)}\max_{J\in\mathcal{J}_4(l)}\sup_{s\in K(q^*,m,I,d_A^\varepsilon)}\sup_{t\in K(p,n,J,d_A^\varepsilon)}\sum_{i\in I\setminus I_0,j\in J\setminus I_0'} a_{ij}g_{ij}s_it_j$$
$$\le \max_{I\in\mathcal{I}_4(k)}\sup_{s\in B_{q^*}^m}\sup_{t\in B(p,I,\alpha_r)}\sum_{i\in I,j\in[n]} a_{ij}g_{ij}s_it_j. \qquad \square$$

We begin by estimating the second term on the right-hand side of (48).

Lemma 45. *For $p^*,q\in[2,\infty)$, $1\le k\le m$, and $\alpha\in[0,1]$ we have*

$$\mathbb{E}\max_{I\in\mathcal{I}_4(k)}\sup_{s\in B_{q^*}^m}\sup_{t\in B(p,I,\alpha)}\sum_{i\in I,j\in[n]} a_{ij}g_{ij}s_it_j \lesssim \sqrt{p^*}D_1+\alpha\big(\sqrt{k\operatorname{Log} d_A}+\sqrt{\operatorname{Log} m}\big).$$

Proof. Observe that for any $I\in\mathcal{I}_4(k)$, $s\in B_{q^*}^m\subset B_2^m$, and $t\in B(p,I,\alpha)$ we have

$$\sum_{i\in I,j\in[n]} a_{ij}^2s_i^2t_j^2\le \max_{i\in I}\sum_j a_{ij}^2t_j^2\le\alpha^2.$$

Define for $s\in B_{q^*}^m$,

$$D(p,s,\alpha):=\Big\{t\in B_p^n\colon \sum_{i,j}a_{ij}^2s_i^2t_j^2\le\alpha^2\Big\}.$$

For $I\in\mathcal{I}_4(k)$ let us choose a $1/2$-net T_I in $B_{q^*}^I=\{s\in B_{q^*}^m\colon \operatorname{supp}(s)\subset I\}$ (in ℓ_{q^*}-norm) of cardinality at most 5^k and put $T:=\bigcup_{I\in\mathcal{I}_4(k)}T_I$. Then, by (31), $|T|\le 5^k|\mathcal{I}_4(k)|\le m20^kd_A^{4k}$, so

$$\max_{I\in\mathcal{I}_4(k)}\sup_{s\in B_{q^*}^m}\sup_{t\in B(p,I,\alpha)}\sum_{i\in I,j\in[n]} a_{ij}g_{ij}s_it_j\le 2\max_{I\in\mathcal{I}_4(k)}\sup_{s\in T_I}\sup_{t\in B(p,I,\alpha)}\sum_{i\in I,j\in[n]} a_{ij}g_{ij}s_it_j$$
$$\le 2\max_{s\in T}\sup_{t\in D(p,s,\alpha)}\sum_{i,j}a_{ij}g_{ij}s_it_j.$$

Thus, estimate (17) implies

$$\mathbb{E}\max_{I\in\mathcal{I}_4(k)}\sup_{s\in B_{q^*}^m}\sup_{t\in B(p,I,\alpha)}\sum_{i\in I,j\in[n]} a_{ij}g_{ij}s_it_j\lesssim \max_{s\in T}\mathbb{E}\sup_{t\in D(p,s,\alpha)}\sum_{i,j}a_{ij}g_{ij}s_it_j$$
$$+\sqrt{\operatorname{Log}|T|}\max_{s\in T}\sup_{t\in D(p,s,\alpha)}\Big(\sum_{i,j}a_{ij}^2s_i^2t_j^2\Big)^{1/2}.$$

For any fixed $s\in T\subset B_{q^*}^m$ estimate (25) yields

$$\mathbb{E}\sup_{t\in D(p,s,\alpha)}\sum_{i,j}a_{ij}g_{ij}s_it_j\le\mathbb{E}\sup_{t\in B_p^n}\sum_{i,j}a_{ij}g_{ij}s_it_j\le\sqrt{p^*}D_1.$$

Moreover, $\sqrt{\operatorname{Log}|T|}\lesssim\sqrt{\operatorname{Log} m}+\sqrt{k\operatorname{Log}(d_A)}$ and

$$\max_{s\in T}\sup_{t\in D(p,s,\alpha)}\Big(\sum_{i,j}a_{ij}^2s_i^2t_j^2\Big)^{1/2}\le\alpha. \qquad \square$$

Now we estimate the first term on the right-hand side of (48).

Lemma 46. *If $l \ge r$, $\varepsilon \in (0,1/2]$, and $p^*, q \in [2,4)$, then*

$$\begin{aligned}&\mathbb{E} \max_{I\subset[m]} \max_{I_0\subset I, |I_0|=r} \max_{J\in\mathcal{J}_4(l)} \sup_{s\in K(q^*,m,I,d_A^\varepsilon)} \sup_{t\in K(p,n,J,d_A^\varepsilon)} \sum_{i\in I_0''\cap I, j\in I_0'\cap J} a_{ij}g_{ij}s_it_j \\ &\quad \lesssim \varepsilon^{-1}\Big[D_1+D_2 \\ &\qquad + \Big((5\operatorname{Log}(d_A))^{2/(4-p^*\vee q)} + \sqrt{\operatorname{Log}(mn)} + \sqrt{\operatorname{Log} d_A}\, d_A^{1/2+3\varepsilon}\Big(\frac{r}{l}\Big)^{1/p}\Big)\max_{i,j}|a_{ij}|\Big].\end{aligned}$$

Proof. Let us fix $I\subset[m]$, $I_0\subset I$ with $|I_0|=r$, $J\in\mathcal{J}_4(l)$, $s\in K(q^*,m,I,d_A^\varepsilon)$, and $t\in K(p,n,J,d_A^\varepsilon)$. Let $I_{0,1},\dots,I_{0,V}$ be 4-connected componets of I_0. Then $(I'_{0,1},\dots,I'_{0,V})$ is a partition of I'_0, and $(I''_{0,1},\dots,I''_{0,V})$ is a partition of I''_0. Hence,

$$\sum_{\substack{i\in I_0''\cap I\\ j\in I_0'\cap J}} a_{ij}g_{ij}s_it_j = \sum_{v=1}^V \sum_{\substack{i\in I_{0,v}''\cap I\\ j\in I_{0,v}'\cap J}} a_{ij}g_{ij}s_it_j.$$

Let

$$\mathcal{V} = \Big\{v\le V\colon |I'_{0,v}\cap J| < d_A^{-2\varepsilon p}\frac{l}{r}|I_{0,v}| = d_A^{-2\varepsilon p}\frac{|J|}{|I_0|}|I_{0,v}|\Big\}.$$

Then

$$\begin{aligned}\sum_{v\in\mathcal{V}} \sum_{\substack{i\in I_{0,v}''\cap I\\ j\in I_{0,v}'\cap J}} a_{ij}g_{ij}s_it_j &\le \|G_A\|_{p\to q}\sum_{v\in\mathcal{V}} \|(s_i)_{i\in I''_{0,v}\cap I}\|_{q^*}\|(t_j)_{j\in I'_{0,v}\cap J}\|_p \\ &\le \|G_A\|_{p\to q}\Big(\sum_{v\in\mathcal{V}}\|(s_i)_{i\in I''_{0,v}\cap I}\|_{q^*}^2\Big)^{1/2}\Big(\sum_{v\in\mathcal{V}}\|(t_j)_{j\in I'_{0,v}\cap J}\|_p^2\Big)^{1/2} \\ &\le \|G_A\|_{p\to q}\Big(\sum_{v\in\mathcal{V}}\|(s_i)_{i\in I''_{0,v}\cap I}\|_{q^*}^{q^*}\Big)^{1/q^*}\Big(\sum_{v\in\mathcal{V}}\|(t_j)_{j\in I'_{0,v}\cap J}\|_p^p\Big)^{1/p} \\ &\le \|G_A\|_{p\to q}\|s\|_{q^*}\Big(\sum_{v\in\mathcal{V}}\frac{d_A^{\varepsilon p}}{|J|}|I'_{0,v}\cap J|\Big)^{1/p} \\ &\le d_A^{-\varepsilon}\|G_A\|_{p\to q}\Big(\sum_{v\in\mathcal{V}}\frac{1}{|I_0|}|I_{0,v}|\Big)^{1/p} \le d_A^{-\varepsilon}\|G_A\|_{p\to q}.\end{aligned}$$

For a nonempty $I_0\subset[m]$, $1\le u\le n$, and $\gamma\ge 1$ define

$$\begin{aligned}&Z_{I_0,u}(\gamma) \\ &:= \max_{I_0\subset I\subset[m]} \max_{J\subset[n], |J\cap I_0'|\ge u} \sup_{s\in K(q^*,m,I,\gamma)} \sup_{t\in K(p,n,J,\gamma)} \frac{\sum_{i\in I_0''\cap I, j\in I_0'\cap J} a_{ij}g_{ij}s_it_j}{\|(s_i)_{i\in I_0''\cap I}\|_{q^*}\|(t_j)_{j\in I_0'\cap J}\|_p}.\end{aligned}$$

Moreover, for $1\le r\le m$, $1\le u\le n$, and $\gamma\ge1$, set

$$Z_{r,u}(\gamma) := \max_{I_0\in\mathcal{I}_4(r)} Z_{I_0,u}(\gamma),$$

and for $\alpha\in(0,1]$ and $\gamma\ge1$, let

$$Z_\alpha(\gamma) := \max_{1\le r\le m}\max_{u\ge\alpha r} Z_{r,u}(\gamma).$$

Observe that for $v\notin\mathcal{V}$, $|I'_{0,v}\cap J|\ge\alpha|I_{0,v}|$, where $\alpha = d_A^{-2\varepsilon p}\frac{l}{r}\ge d_A^{-2}$, so

$$\sum_{v\notin\mathcal{V}} \sum_{\substack{i\in I_{0,v}''\cap I\\ j\in I_{0,v}'\cap J}} a_{ij}g_{ij}s_it_j \le Z_\alpha(d_A^\varepsilon)\sum_{v\notin\mathcal{V}}\|(s_i)_{i\in I''_{0,v}\cap I}\|_{q^*}\|(t_j)_{j\in I'_{0,v}\cap J}\|_p$$

$$
\begin{aligned}
&\le Z_\alpha(d_A^\varepsilon)\Big(\sum_v \|(s_i)_{i\in I''_{0,v}\cap I}\|_{q^*}^2\Big)^{1/2}\Big(\sum_v \|(t_j)_{j\in I'_{0,v}\cap J}\|_p^2\Big)^{1/2}\\
&\le Z_\alpha(d_A^\varepsilon)\Big(\sum_v \|(s_i)_{i\in I''_{0,v}\cap I}\|_{q^*}^{q^*}\Big)^{1/q^*}\Big(\sum_v \|(t_j)_{j\in I'_{0,v}\cap J}\|_p^p\Big)^{1/p}\\
&\le Z_\alpha(d_A^\varepsilon)\|s\|_{q^*}\|t\|_p \le Z_\alpha(d_A^\varepsilon).
\end{aligned}
$$

The above argument shows that

$$
\begin{aligned}
\max_{I\in\mathcal{I}_4(k)}\ \max_{I_0\subset I,|I_0|=r}\ \max_{J\in\mathcal{J}_4(l)}\ \sup_{s\in K(q^*,m,I,d_A^\varepsilon)}\ \sup_{t\in K(p,n,J,d_A^\varepsilon)}\ &\sum_{i\in I''_0\cap I,j\in I'_0\cap J} a_{ij}g_{ij}s_it_j\\
&\le d_A^{-\varepsilon}\|G_A\|_{p\to q} + Z_\alpha(d_A^\varepsilon).
\end{aligned}
$$

Corollary 43 yields that

$$
\begin{aligned}
d_A^{-\varepsilon}\mathbb{E}\|G_A\|_{p\to q} &\lesssim \sup_{x\ge 1}(x^{-\varepsilon}\operatorname{Log} x)\Big(D_1+D_2+\sqrt{\operatorname{Log}(mn)}\max_{i,j}|a_{ij}|\Big)\\
&\sim \varepsilon^{-1}\Big(D_1+D_2+\sqrt{\operatorname{Log}(mn)}\max_{i,j}|a_{ij}|\Big),
\end{aligned}
$$

Therefore, the assertion follows by part iii) of the next lemma. □

Lemma 47. *Let $\gamma\ge 1$ and $p^*,q\in[2,4)$.*
i) If $I_0\subset[m]$ and $1\le u\le n$ satisfy $u\ge |I_0|d_A^{-2}$, then

$$
\mathbb{E}Z_{I_0,u}(\gamma)\lesssim D_1+D_2+(3\operatorname{Log}(d_A\gamma))^{2/(4-p^*\vee q)}\max_{i,j}|a_{ij}|.
$$

ii) If $1\le r\le m$ and $1\le u\le n$ satisfy $u\ge rd_A^{-2}$, then

$$
\begin{aligned}
\mathbb{E}Z_{r,u}(\gamma)&\lesssim D_1+D_2\\
&\quad+\Big((3\operatorname{Log}(d_A\gamma))^{2/(4-p^*\vee q)}+\sqrt{\operatorname{Log} m}+\sqrt{r\operatorname{Log} d_A}\gamma d_A^{1/2}u^{-1/p}\Big)\max_{i,j}|a_{ij}|.
\end{aligned}
$$

iii) If $\alpha\ge d_A^{-2}$, then

$$
\begin{aligned}
\mathbb{E}Z_{\alpha}(\gamma)&\lesssim D_1+D_2\\
&\quad+\Big((3\operatorname{Log}(d_A\gamma))^{2/(4-p^*\vee q)}+\sqrt{\operatorname{Log}(mn)}+\sqrt{\operatorname{Log} d_A}\gamma d_A^{1/2}\alpha^{-1/p}\Big)\max_{i,j}|a_{ij}|.
\end{aligned}
$$

Proof. Let us fix $I_0\subset I\subset[m]$, $J\subset[n]$ with $|I_0|=r$, $|J\cap I'_0|\ge u$, $s\in K(q^*,m,I,\gamma)$ and $t\in K(p,n,J,\gamma)$. Define

$$
\bar{s}=\|(s_i)_{i\in I''_0\cap I}\|_{q^*}^{-1}(s_i)_{i\in I''_0\cap I},\quad \bar{t}=\|(t_j)_{j\in I'_0\cap J}\|_p^{-1}(t_j)_{j\in I'_0\cap J}.
$$

Recall that $I_0\subset I''_0$, so $\|\bar{s}\|_{q^*}=1$, $\|\bar{s}\|_\infty\le\gamma|I''_0\cap I|^{-1/q^*}\le\gamma|I_0|^{-1/q^*}\le\gamma r^{-1/q^*}$, $\|\bar{t}\|_p=1$, and $\|\bar{t}\|_\infty\le\gamma|I'_0\cap J|^{-1/p}\le\gamma u^{-1/p}\le\gamma d_A^{2/p}r^{-1/p}$. Hence, part i) of the assertion follows by Proposition 41 applied with the matrix $(a_{ij})_{i\in I''_0,j\in I'_0}$, $m=|I''_0|\le d_A^2r$, $n=|I'_0|\le d_Ar$, $a=(\gamma r^{-1/q^*})\wedge 1$ and $b=(\gamma d_A^{2/p}r^{-1/p})\wedge 1$.

To show part ii) observe that

$$
\begin{aligned}
\sum_{i\in I''_0\cap I,j\in I'_0\cap J} a_{ij}^2\bar{s}_i^2\bar{t}_j^2 &\le \max_{i\in I}\sum_{j\in J}a_{ij}^2\bar{t}_j^2\le\max_{i,j}a_{ij}^2\min\{1,d_A\|\bar{t}\|_\infty^2\}\\
&\le \max_{i,j}a_{ij}^2\min\{1,d_A\gamma^2u^{-2/p}\}.
\end{aligned}
$$

Moreover, estimate (31) yields $\sqrt{\operatorname{Log}|\mathcal{I}_4(r)|}\lesssim\sqrt{\operatorname{Log} m}+\sqrt{r\operatorname{Log} d_A}$, so part ii) follows by part i) and estimate (17).

Part iii) easily follows from ii) and another application of (17). □

Corollary 48. *If* $1 \le k \le m$, $1 \le l \le n$, $p^*, q \in [2,4)$, *and* $\varepsilon \in (0,1/2]$, *then*

$$\mathbb{E}Y_{k,l}(d_A^\varepsilon) \lesssim \varepsilon^{-1}\Big[D_1 + D_2 + \Big(\sqrt{\operatorname{Log}(mn)} + \Big(\varepsilon^{-1/2} + \Big(\frac{10}{4-p^*\vee q}\Big)^{2/(4-p^*\vee q)}\Big)d_A^{1/4+4\varepsilon}\Big)\max_{i,j}|a_{ij}|\Big].$$

Proof. By symmetry we may assume that $l \ge k$.

First assume that $k \ge d_A^{1/2}$. Lemmas 44–46 yield that for every $1 \le r \le k$,

$$\begin{aligned}\mathbb{E}Y_{k,l}(d_A^\varepsilon) \lesssim \varepsilon^{-1}\Big[&D_1 + D_2 \\ &+ \Big((5\operatorname{Log}(d_A))^{2/(4-p^*\vee q)} + \sqrt{\operatorname{Log}(mn)}\Big)\max_{i,j}|a_{ij}| \\ &+ \Big(\sqrt{\operatorname{Log} d_A}\, d_A^{1/2+3\varepsilon}\Big(\frac{r}{l}\Big)^{1/p} + \sqrt{k\operatorname{Log} d_A}\, d_A^\varepsilon l^{1/2-1/p} r^{-1/2}\Big)\max_{i,j}|a_{ij}|\Big].\end{aligned}$$

Moreover,

$$(5\operatorname{Log}(d_A))^{2/(4-p^*\vee q)} \lesssim \Big(\frac{10}{4-p^*\vee q}\Big)^{2/(4-p^*\vee q)} d_A^{1/4},$$

$$\sqrt{\operatorname{Log} d_A} \lesssim \varepsilon^{-1/2} d_A^\varepsilon, \tag{49}$$

and

$$\begin{aligned}\inf_{1\le r\le k}\Big(d_A^{1/2+3\varepsilon}\Big(\frac{r}{l}\Big)^{1/p} + \sqrt{k}d_A^\varepsilon l^{1/2-1/p}r^{-1/2}\Big) &\le \inf_{1\le r\le k}\Big(d_A^{1/2+3\varepsilon}\Big(\frac{r}{k}\Big)^{1/2} + \sqrt{k}d_A^\varepsilon r^{-1/2}\Big) \\ &\lesssim d_A^{1/4+3\varepsilon},\end{aligned}$$

where the last estimate follows by taking $r = \lfloor k d_A^{-1/2}\rfloor \in [1,k]$.

Now assume that $k < d_A^{1/2}$. Recall that for a fixed nonempty set I,

$$X_I = X_I(A,p,q) := \|(a_{ij}g_{ij})_{i\in I, j\in[n]}\|_{p\to q}.$$

Note that

$$Y_{k,l}(d_A^\varepsilon) \le \max_{I\in\mathcal{I}_4(k)} X_I.$$

Therefore, estimates (17), (28), and $|\mathcal{I}_4(k)| \le m4^k d_A^{4k}$ (see (31)) yield

$$\begin{aligned}\mathbb{E}Y_{k,l}(d_A^\varepsilon) &\lesssim \max_{I\in\mathcal{I}_4(k)}\mathbb{E}X_I + \sqrt{\operatorname{Log}|\mathcal{I}_4(k)|}\max_{I\in\mathcal{I}_4(k)}\sup_{\|s\|_{q^*}\le 1}\sup_{\|t\|_p\le 1}\Big(\sum_{i,j}a_{ij}^2 s_i^2 t_j^2\Big)^{1/2}. \\ &\lesssim D_1 + (\sqrt{k} + \sqrt{\operatorname{Log} m} + \sqrt{k\operatorname{Log} d_A})\max_{i,j}|a_{ij}|. \qquad (50)\end{aligned}$$

The assertion easily follows by the assumption that $k < d_A^{1/2}$ and estimate (49). ☐

Proof of Proposition 40. We apply Proposition 31 and Corollary 48 (with $\varepsilon/4$ instead of ε). ☐

Acknowledgements. Part of this work was carried out while the first named author was visiting the Hausdorff Research Institute for Mathematics, Univerity of Bonn. The hospitality of HIM and of the organizers of the program *Boolean Analysis in Computer Science* is gratefully acknowledged.

References

1. R. Adamczak, J. Prochno, M. Strzelecka, and M. Strzelecki, *Norms of structured random matrices*, Math. Ann. **388** (2024), no. 4, 3463–3527. MR 4721775
2. A. S. Bandeira and R. van Handel, *Sharp nonasymptotic bounds on the norm of random matrices with independent entries*, Ann. Probab. **44** (2016), no. 4, 2479–2506. MR 3531673
3. S. Boucheron, G. Lugosi, and P. Massart, *Concentration inequalities*, Oxford University Press, Oxford, 2013, A nonasymptotic theory of independence, With a foreword by Michel Ledoux. MR 3185193
4. T. Brailovskaya and R. van Handel, *Extremal random matrices with independent entries and matrix superconcentration inequalities*, preprint, arXiv:2401.06284, (2024).
5. A. Eskenazis, P. Nayar, and T. Tkocz, *Gaussian mixtures: entropy and geometric inequalities*, Ann. Probab. **46** (2018), no. 5, 2908–2945. MR 3846841
6. O. Guédon, A. Hinrichs, A. E. Litvak, and J. Prochno, *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. MR 3645120
7. R. Latała, *Some estimates of norms of random matrices*, Proc. Amer. Math. Soc. **133** (2005), no. 5, 1273–1282. MR 2111932
8. R. Latała and M. Strzelecka, *Comparison of weak and strong moments for vectors with independent coordinates*, Mathematika **64** (2018), no. 1, 211–229. MR 3778221
9. ———, *Chevet-type inequalities for subexponential Weibull variables and estimates for norms of random matrices*, Electron. J. Probab. **29** (2024), Paper No. 97, 19. MR 4761377
10. ———, *Operator $\ell_p \to \ell_q$ norms of random matrices with iid entries*, J. Funct. Anal. **288** (2025), no. 3, Paper No. 110720, 38. MR 4818302
11. R. Latała and W. Świątkowski, *Norms of randomized circulant matrices*, Electron. J. Probab. **27** (2022), Paper No. 80, 23. MR 4441144
12. R. Latała, R. van Handel, and P. Youssef, *The dimension-free structure of nonhomogeneous random matrices*, Invent. Math. **214** (2018), no. 3, 1031–1080. MR 3878726
13. R. Latała, *On the spectral norm of rademacher matrices*, preprint, arXiv:2405.13656, (2024).
14. D. Matlak, *Oszacowania norm macierzy losowych*, Master's thesis, Uniwersytet Warszawski, 2017.
15. K. Ramanan and X. Xie, *The ℓ_r-Levy-Grothendieck problem and $r \to p$ norms of Levy matrices*, preprint, arXiv:2404.18299, (2024).
16. Y. Seginer, *The expected norm of random matrices*, Combin. Probab. Comput. **9** (2000), no. 2, 149–166. MR 1762786
17. R. van Handel, *On the spectral norm of Gaussian random matrices*, Trans. Amer. Math. Soc. **369** (2017), no. 11, 8161–8178. MR 3695857
18. ———, *Structured random matrices*, Convexity and concentration, IMA Vol. Math. Appl., vol. 161, Springer, New York, 2017, pp. 107–156. MR 3837269

Institute of Mathematics, University of Warsaw, Banacha 2, 02–097 Warsaw, Poland.

Email address: `rlatala@mimuw.edu.pl`

Institute of Mathematics, University of Warsaw, Banacha 2, 02–097 Warsaw, Poland.

Email address: `martast@mimuw.edu.pl`