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 $1 \le p \le 2 \le q \le \infty$. This was known before only in the case $p = 1$ or $q = \infty$, and in the spectral case $p = 2 = q$. We also reprove the conjecture in the case $p = 2 = q$ without using spectral theory (which was employed in the previously known proof).

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

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

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 random matrices with a 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, 18, 13, 7, 19, 12, 14, 1, 4, 16]).

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 the 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 10 below). Moreover, they yield a condition for an infinite Gaussian matrix to be a bounded operator from ℓ_p to ℓ_q (see Corollary 5 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 non-negative 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

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

If $p = 2 = q$, the $\ell_p^n \to \ell_q^m$ norm coincides with the spectral norm and it is known by [13] that

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

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 [7, 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 [7, 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 [10] 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 [7] (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 known before that the conjecture holds if $p = 1$ or $q = \infty$. Moreover, for other ranges of p and q, it was shown in [7] 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, so it confirms Conjecture 1.

Theorem 2. *If $p^*, q \in [2, \infty)$, 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, it follows by Propositions 15, 16, and 18 below, that the constant in the first upper bound of Theorem 2 is at most of order

$$\begin{cases}(p^* \vee q)^{13/2} & \text{if } 1/p + 1/q^* < 3/2, \\ (p^* \vee q)^{5/2} & \text{if } 1/p + 1/q^* \ge 3/2.\end{cases}$$

Remark 4. Recall that the two-sided bounds for $\mathbb{E}\|G_A\|_{p\to q}$ were known before if $p = q = 2$ – in this case we provide a new proof, which does not use the spectral methods as the previously known proof did. Moreover, [7, Remark 1.4] implies that for every $p^*, q \ge 2$

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

and

$$\mathbb{E}\|G_A\|_{p\to\infty} = \mathbb{E}\max_i \|(a_{ij}g_{ij})_j\|_{p^*} \lesssim \sqrt{p^*}\max_i \|(a_{ij})_j\|_{p^*} + \mathbb{E}\max_{i,j}|a_{ij}g_{ij}|. \tag{2}$$

This is why in order to prove Conjecture 1 it suffices to restrict ourselves to the case $p^*, q \in [2, \infty)$ in Theorem 2.

Although we were able to confirm Conjecture 1, 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

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

(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 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

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

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 and inequalities (1) and (2) easily imply their 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 $m_{ij}, a_{ij} \in \mathbb{R}$, $i \le m$, $j \le n$, and every $p^*, q \ge 2$.

Moreover, Theorem 2, inequalities (1) and (2), and [1, Proposition 1.2] yield the following characterisation of the boundedness of Gaussian linear operators from ℓ_p to ℓ_q whenever $p^*, q \ge 2$. 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 5. *Let $p^*, q \in [2, \infty]$, 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 6. The condition $\mathbb{E}\sup_{i,j\in\mathbb{N}} |a_{ij}g_{ij}| < \infty$ in Corollary 5 is equivalent to the deterministic bound

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

(see estimate (13) 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 X has the same distribution as Rg, where g is a standard Gaussian random variable, independent of R (cf. [6]). The next corollary is an immediate consequence of Theorem 2 and inequalities (1) and (2). In all the corollaries of this subsection the constants depending on p and q are at most of order q_0^7, provided that $p^*, q \in [2, q_0]$.

Corollary 7. *Assume that $p^*, q \in [2, \infty]$ and let X_{ij}, $i \le m, j \le n$, be independent Gaussian mixtures. Then*

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

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

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

Moreover, if $p^*, q \in [2,\infty]$, then

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

We postpone the proof of Corollary 8 to Section 2.

Remark 9. One cannot omit the assumption $r \leq 2$ in Corollary 8. 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 [17]). Moreover, it was conjectured in [12] and proven in [14] 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 $p \leq 2 \leq q$,

$$\begin{aligned}
\mathbb{E}\|(a_{ij}\varepsilon_{ij})_{i,j}\|_{p\to q} &\overset{?}{\sim}_{p,q} \max_i \|(a_{ij})_j\|_{p^*} + \max_j \|(a_{ij})_i\|_q \\
&\qquad + \max_{k\geq 0}\inf_{|I|=k} \sup_{\|s\|_{q^*},\|t\|_p\leq 1} \Big\|\sum_{i,j\notin I} a_{ij}\varepsilon_{ij}s_it_j\Big\|_{\operatorname{Log} k} \\
&= \mathbb{E}\max_i \|(a_{ij}\varepsilon_{ij})_j\|_{p^*} + \mathbb{E}\max_j \|(a_{ij}\varepsilon_{ij})_i\|_q \\
&\qquad + \max_{k\geq 0}\inf_{|I|=k} \sup_{\|s\|_{q^*},\|t\|_p\leq 1} \Big\|\sum_{i,j\notin I} a_{ij}\varepsilon_{ij}s_it_j\Big\|_{\operatorname{Log} k}.
\end{aligned} \tag{3}$$

We also believe that the methods of [14] could be adapted to the case $p^*, q \geq 2$ and that – together with Theorem 2 – they would imply (3) up to a polylog factor.

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, inequalities (1) and (2), and the Gaussian concentration yield the following moment and tail bounds.

Corollary 10. *If $p^*, q \in [2,\infty]$, then for every deterministic matrix $A = (a_{ij})_{i\le m, j\le n}$, $\rho \ge 1$, and $t > 0$ we have*

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

and

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

In the Weibull case, Corollary 8, [9, 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 (4) via Markov's inequality.)

Corollary 11. *Let X_{ij}, $i \le m$, $j \le n$, be independent symmetric Weibull variables with parameter $r \in (0,2]$. If $p^*, q \in [2,\infty]$, then for every deterministic matrix $A = (a_{ij})_{i\le m, j\le n}$, and every $\rho \ge 1$ we have*

$$(\mathbb{E}\|(a_{ij}X_{ij})_{i,j}\|_{p\to q}^{\rho})^{1/\rho} \sim_{p,q,r} \max_i \|(a_{ij})_j\|_{p^*} + \max_j \|(a_{ij})_i\|_q + \max_{k\ge 0} \inf_{|I|=k} \operatorname{Log}^{1/r} k \max_{(i,j)\notin I} |a_{ij}| + \rho^{1/r} \max_{i,j} |a_{ij}|. \tag{4}$$

Moreover, for every $t > 0$,

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

Remark 12. The upper bound in (4) and, as a consequence, the tail bound from Corollary 11 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 [10, Lemma 2.1]).

The next result is a generalization of [8, 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 13. *If $p^*, q \in [2,\infty)$, then for every matrix $(X_{ij})_{i\le m, j\le n}$ with independent centered entries,*

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

and, as a consequence,

$$\mathbb{E}\|(X_{ij})_{i,j}\|_{p\to q} \lesssim_{p,q} \max_i\Big(\sum_j \mathbb{E}|X_{ij}|^{p^*}\Big)^{1/p^*} + \max_j\Big(\sum_i \mathbb{E}|X_{ij}|^{q}\Big)^{1/q} + \Big(\sum_{i,j} \mathbb{E}|X_{ij}|^{2p^*}\Big)^{1/(2p^*)} + \Big(\sum_{i,j} \mathbb{E}|X_{ij}|^{2q}\Big)^{1/(2q)}. \tag{6}$$

Remark 14. If $p^*, q \in [2,\infty)$, $\alpha \geq 1$, and independent centered random variables X_{ij} satisfy

$$(\mathbb{E}|X_{ij}|^{2\rho})^{1/(2\rho)} \leq \alpha(\mathbb{E}|X_{ij}|^{\rho})^{1/\rho} \qquad \text{for } \rho \in \{p^*, q\}, \tag{7}$$

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 (7) by Hölder's inequality. One can repeat the argument from the proof of the lower bound in [11, 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 [11] 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

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

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

Similarly as in [13], Theorem 2 is a consequence of two weaker estimates given in Propositions 15 and 16.

Proposition 15. *Assume that $p^*, q \in [2,\infty)$. Then for every deterministic matrix $A = (a_{ij})_{i\leq m, j\leq n}$,*

$$\mathbb{E}\|G_A\|_{p\to q} \lesssim \alpha(p,q)\Big(\sqrt{p^*}D_1 + \sqrt{q}D_2 + \sqrt{\mathrm{Log}(mn)}\max_{i,j}|a_{ij}|\Big), \tag{8}$$

where

$$\alpha(p,q) = \begin{cases}(p^*\vee q)^{3/2} & \text{if } 1/p + 1/q^* \geq 3/2,\\ (p^*\vee q)^{11/2} & \text{if } 1/p+1/q^* < 3/2.\end{cases}$$

Proposition 16. *For every $p^*, q \in [2,\infty)$ and every deterministic matrix $A = (a_{ij})_{i\leq m, j\leq n}$,*

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

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

$$\beta(p,q) = \begin{cases}(p^*\vee q)^{3/2} & \text{if } 1/p + 1/q^* \geq 3/2,\\ (p^*\vee q)^{11/2} & \text{if } 1/p+1/q^* < 3/2.\end{cases}$$

In the case $p = q = 2$ estimate (8) was proven by Bandeira and van Handel in [2]. We cannot use their combinatorial approach based on the trace method since for $(p, q) \neq (2, 2)$ we no longer deal with the spectral norm. Proving Proposition 15 is one of the main difficulties and novelties of our paper. Using our new methods we also reprove Proposition 15 in the case $p = q = 2$ without relying on spectral tools.

To prove Proposition 15 we first obtain the following weaker result with a bigger exponent of the logarithm. Then we perform the exponent reduction described in Section 3.2.

Proposition 17. *If $p^*, q \in [2, \infty)$, then*

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

where $\gamma \leq 5$ is a universal constant.

To prove Proposition 17 in the case $(p^*, q) \notin [2, 3]^2$ it suffices – by duality – to consider the case $q \geq p^*$, so that $q > 3$. We split every vector in $B^m_{q^*}$ into a part with a small support and a part with a small supremum norm. To deal with vectors with small supports we use a standard net argument. In order to bound the supremum over vectors in $B^m_{q^*}$ with small coordinates we first introduce a certain interpolation norm of ℓ^m_2- and ℓ^m_q-norms satisfying a crucial hypercontractive property. To estimate the norm of G_A acting between ℓ^n_p and the interpolation space we discretise the B^n_p ball and use an inductive argument; in every induction step we use the Gaussian concentration and the hypercontraction.

The proof of Proposition 17 in the case $(p^*, q) \in [2, 3]^2$ is based on the ideas from the proof of the main result of [18] (a similar result for $p^* = q \in [2, 4)$ was obtained in [15]). However, in this case performing the exponent reduction with constants which do not blow up when p^* and q are close to 2 is much more involved than in the case $(p^*, q) \notin [2, 3]^2$. The exponent reduction in this case is described in Section 5. This is the most challenging part of our new proof of two-sided bounds in the case $p = q = 2$.

Finally, to deduce Proposition 16 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 32 below), and use Proposition 15 for each A_k separately. In the proof of the aforementioned Proposition 32 we use similar tools as in the proof of Proposition 17 in the case $(p^*, q) \notin [2, 3]^2$ for vectors with small coordinates.

1.3. **Organization of the paper.** In Section 2 we show how Propositions 15 and 16 imply Theorem 2 and then we prove Corollaries 8 and 13. In Section 3 we prove Proposition 15 in the case $(p^*, q) \notin [2, 3]^2$. In Section 4 we show how to deduce Proposition 16 from Proposition 15. Finally, in Section 5 we prove Proposition 15 in the case $p^*, q \in [2, 3]$.

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 15 and 16. Then we give the proofs of Theorem 2 and Corollaries 8 and 13. The proofs of Propositions 15 and 16 may be found in the next sections.

Proposition 18. *Let $p^*, q \geq 2$, and $\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3 \geq 1$. Assume that for every integers m and n, and every deterministic matrix $A = (a_{ij})_{i\leq m, j\leq n}$,*

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

and

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

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

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

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

Lemma 19. *Let $(c_{ij})_{i\le m, j\le n}$ be a block diagonal matrix with blocks C_l, and $1\le p\le q$. Then*

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

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

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

Since $p\le q$and $q\ge 1$, the latter supremum is attained at $y=e_l$ for some $l\le l_0$, so

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

Proof of Proposition 18. 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 [13, Theorem 3.9] and [12, 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_kN_{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 $N_{k+1}-N_k-N_kN_{k-1}\ge N_k$ indices i in such a way that $I_{k+1}=\{i_1,\dots,i_{N_{k+1}}\}$ satisfies

$$\max_{i\notin I_{k+1}} \max_j |a_{ij}| \le \frac{D_\infty}{\sqrt{\operatorname{Log} N_k}}. \tag{9}$$

Similarly, to construct $(j_{N_k+1},\dots,j_{N_{k+1}})$ we choose first N_kN_{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 $N_{k+1}-N_k-N_kN_{k-1}\ge N_k$ indices j in such a way that $J_{k+1}=\{j_1,\dots,j_{N_{k+1}}\}$ satisfies

$$\max_{j\notin J_{k+1}}\max_i |a_{ij}| \le \frac{D_\infty}{\sqrt{\operatorname{Log} N_k}}. \tag{10}$$

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

$$|a_{ij}| \le D_2 N_{k-1}^{-1/q} \le 2D_2 2^{-2^{k-1}/q} \qquad \text{if } j\le N_k \text{ and } i\ge M_k := N_k+N_kN_{k-1}, \tag{11}$$

$$|a_{ij}| \le D_1 N_{k-1}^{-1/p^*} \le 2D_1 2^{-2^{k-1}/p^*} \qquad \text{if } i\le N_k \text{ and } j\ge M_k. \tag{12}$$

We set (see [13, 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 $k\ge 2$ of dimension $|E_{1,k}|\le M_{2k-1}-N_{2k-2}$. We have

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

For $k=2,3,\dots$ the Gaussian concentration (see, e.g., [3, Theorem 5.6]), the first assumption of the proposition, and property (9) imply

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

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 19 yields

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

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 $k\in\{0,1,\dots\}$ such that $M_k\le i<M_{k+1}$, where $M_0=1$. Observe that if $M_k\le i\le N_{k+1}$, then either $j\le N_k$ or $j>M_{k+1}$ and if

$N_{k+1}+1 \le i < M_{k+1}$, then either $j \le N_k$ or $j > M_{k+2}$. Therefore, either $j \le N_k$ or $j > M_{k+1}$. If $j \le N_k$, then by (11)

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

If on the other hand $M_{k+l} < j \le M_{k+l+1}$ for some $l \ge 1$, then $i < N_{k+l}$ and (12) imply

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

Thus, the second assumption of the proposition yields

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

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

$$\mathbb{E}\max_{i,j}|a_{ij}g_{ij}| \sim \max_{k\ge 0}\inf_{|I|=k}\sqrt{\operatorname{Log} k}\max_{(i,j)\notin I}|a_{ij}| \tag{13}$$

(cf. [18, Lemmas 2.3 and 2.4]).

The upper bound from the first two-sided estimate follows by Propositions 15, 16 and 18 and estimate (13).

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

$$\sqrt{p^*}D_1 + \sqrt{q}D_2 + D_\infty \sim \sqrt{p^*}D_1 + \sqrt{q}D_2 + D'_\infty, \tag{14}$$

where

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

The lower bound in (14) 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 (14) we need to show that for any $k \ge 1$,

$$\sqrt{\operatorname{Log} k}\inf_{|V|=k}\max_{(i,j)\notin V}|a_{ij}| \lesssim \sqrt{p^*}D_1 + \sqrt{q}D_2 + D'_\infty. \tag{15}$$

If $k \le 3$, then

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

If $k \ge 3$, 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

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

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

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

We have $\sqrt{\operatorname{Log} k} \lesssim \min\{\sqrt{\operatorname{Log} k'}, \sqrt{p^*}(k')^{1/p^*}, \sqrt{q}(k')^{1/q}\}$, so (15) follows. $\square$

Proof of Corollary 8. Assume first that $p^*, q < \infty$. 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 $\rho \geq 1$,

$$(\mathbb{E}|X_{ij}|^\rho)^{1/\rho} \sim_r \rho^{1/r} \sim_r (\mathbb{E}|Y_{ij}|^\rho)^{1/\rho}, \tag{16}$$

so Theorem 2 and [13, 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 (13) and the last one by (16) and [13, 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 [13, Lemma 4.7] and (13) yield

$$\begin{aligned}\mathbb{E}\max_{i,j}|a_{ij}X_{ij}| &\sim_r \mathbb{E}\max_{i,j}|a_{ij}Z_{ij}| = \mathbb{E}\max_{i,j}|a_{ij}||g_{ij}|^{2/r} \\ &\sim_r \big(\mathbb{E}\max_{i,j}|a_{ij}|^{r/2}|g_{ij}|\big)^{2/r} \sim \big(\max_{k\geq 0}\inf_{|I|=k}\sqrt{\operatorname{Log} k}\max_{(i,j)\notin I}|a_{ij}|^{r/2}\big)^{2/r} \\ &= \max_{k\geq 0}\inf_{|I|=k}\operatorname{Log}^{1/r} k\max_{(i,j)\notin I}|a_{ij}|. \end{aligned}\tag{17}$$

Finally, the third two-sided estimate in Corollary 8 may be established in a similar way as in the Gaussian case (see the last part of the proof of Theorem 2).

If $p^* = \infty$ (i.e., $p = 1$), then we proceed similarly, using inequality (1) instead of Theorem 2, to prove that for every $q \in [2, \infty)$,

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

Moreover, if $p^* = q = \infty$, then (17) yields that

$$\mathbb{E}\|(a_{ij}X_{ij})_{i\le m, j\le n}\|_{1\to\infty} \sim_r \mathbb{E}\max_{i,j} |a_{i,j}X_{i,j}| \sim_r \max_{k\ge 0} \inf_{|I|=k} \operatorname{Log}^{1/r} k \max_{(i,j)\notin I} |a_{ij}|.$$

The last case, $q = \infty > p^* \ge 2$, follows by duality.

$\square$

Proof of Corollary 13. 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 7.

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

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

Moreover, by the Rosenthal inequality [5, Theorem 1.5.11] we have for every $r \ge 2$,

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

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 $r \ge 2$,

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

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

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

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

3. Dimension dependent bounds

In this section our aim is to obtain the dimension dependent bound from Proposition 15 in the case $(p^*, q) \notin [2,3]^2$. We begin by proving a weaker estimate, asserted in Proposition 17. Then, in Subsections 3.2 and 3.3, we show how to reduce the exponent of the logarithm appearing in this proposition in the case $p^* \vee q > 2$. The proof of Proposition 15 in the case $(p^*, q) \notin [2,3]^2$ is provided at the end of this section.

The methods provided in this section work in the whole range $p^* \vee q > 2$, but they yield the constant which blows up when $p^* \vee q$ approaches 2. In order to obtain the bound with a universal constant in the case when p^* and q are close to 2 , we need more involved combinatorial arguments and some additional estimates obtained by the Slepian-Fernique lemma to run the dimension reduction. We provide these additional tools in Section 5, together with the proofs of Propositions 15 and 17 in the case $p^*, q \in [2,3]$.

3.1. Proof of Propostion 17 in the case $(p^*, q) \notin [2,3]^2$.

In order to prove Proposition 17 we split each $s \in B^m_{q^*}$ 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. The next propostion shows how to control the suprema over points with large non-zero coordinates.

Proposition 20. *If $\gamma \geq 1$ and $p^*, q \geq 2$, then*

$$\mathbb{E}\sup_{s\in B^m_{q^*}}\sup_{t\in B^n_p}\sum_{i,j} a_{ij}g_{ij}s_it_j\mathbb{1}_{\{|s_i|\geq\gamma^{-1}\}} \lesssim \sqrt{p^*}D_1 + \sqrt{\gamma^{q^*}\operatorname{Log} m}\max_{i,j}|a_{ij}| \tag{18}$$

and

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

In the proof of Proposition 20 we shall use the following standard lemma; we formulate and prove it for the sake of completeness.

Lemma 21. *For every fixed $s \in B^m_{q^*}$, $q \in [2,\infty]$, and $p \in (1,2]$,*

$$\mathbb{E}\sup_{t\in B^n_p}\sum_{i,j} a_{ij}g_{ij}s_it_j = \mathbb{E}\Big\|\Big(\sum_i a_{ij}g_{ij}s_i\Big)_j\Big\|_{p^*} \leq \sqrt{p^*}D_1. \tag{19}$$

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

$$\mathbb{E}\sup_{s\in B^m_{q^*}}\sum_{i,j} a_{ij}g_{ij}s_it_j = \Big\|\Big(\sum_j a_{ij}g_{ij}t_j\Big)_i\Big\|_q \leq \sqrt{q}D_2. \tag{20}$$

Proof. We have

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

Moreover, $\|g_{1,1}\|_{p^*}\le\sqrt{p^*}$ 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, (20) follows. Estimate (19) holds by an analogous argument. □

We shall also use the following lemma, which follows by the Gaussian concentration (see, e.g., [3, Theorem 5.6]) and [18, Lemma 2.3], applied with $X_i:=\sup_{t\in T_i}X_t-\mathbb{E}\sup_{t\in T_i}X_t$ and $\sigma_i:=\sup_{t\in T_i}(\mathrm{Var}X_t)^{1/2}$.

Lemma 22. *Let $X=(X_t)_{t\in T}$ be a Gaussian process and let $T_1,\ldots,T_k$ be nonempty subsets of T such that $T=\bigcup_{i=1}^kT_i$. 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

$$\mathbb{E}\sup_{t\in T}X_t\le\max_i\mathbb{E}\sup_{t\in T_i}X_t+C\sqrt{\mathrm{Log}\,k}\sup_{t\in T}(\mathbb{E}X_t^2)^{1/2}.\tag{21}$$

Proof of Proposition 20. By the symmetry it is enough to show (18). For a fixed nonempty set I let

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

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

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

Thus, inequality (21) yields

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

Since $S\subset B_{q^*}^m\subset B_2^m$ and $B_p^n\subset B_2^n$ we have

$$\sup_{s\in S}\sup_{t\in B_p^n}\Big(\sum_{i,j}a_{ij}^2s_i^2t_j^2\Big)^{1/2}\le\max_{i,j}|a_{ij}|.\tag{22}$$

Estimate (19) implies that for any fixed $s\in S$,

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

hence,

$$\mathbb{E}X_I\lesssim\sqrt{p^*}D_1+\sqrt{|I|}\max_{i,j}|a_{ij}|.\tag{23}$$

There are at most $Cm^{\gamma^{q^*}}$ subsets of $[m]$ of cardinality at most γ^{q^*}. Thus, estimates (21), (22), and (23) yield

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

To derive (18) it suffices to note that for $s\in B^m_{q^*}$, $|\{i\colon |s_i|\ge\gamma^{-1}\}|\le\gamma^{q^*}$, so

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

The next proposition shows how to control the suprema over points with small coordinates. It is a crucial step in the proof of Proposition 17.

Proposition 23. *If $q>2$, $p^*\ge 2$, and $a\in(0,1]$, then*

$$\mathbb{E}\sup_{s\in B^m_{q^*}\cap aB^m_\infty}\sup_{t\in B^n_p}\sum_{i\le m,j\le n} a_{ij}g_{ij}s_it_j \lesssim \sqrt{q}D_2 + a^{(2-q^*)/2}D_1\operatorname{Log}^{3/2} n. \tag{24}$$

If $p^*>2$, $q\ge 2$, and $b\in(0,1]$, then

$$\mathbb{E}\sup_{s\in B^m_{q^*}}\sup_{t\in B^n_p\cap bB^n_\infty}\sum_{i\le m,j\le n} a_{ij}g_{ij}s_it_j \lesssim \sqrt{p^*}D_1 + b^{(2-p)/2}D_2\operatorname{Log}^{3/2} m. \tag{25}$$

To prove Proposition 23 we need the following two lemmas. For a given $\lambda\in(0,1]$, let us first define the following interpolation norm on $\mathbb{R}^m$,

$$\|x\|_{2,q,\lambda} := \inf\big\{\big(\lambda^2\|y\|_2^2+\|z\|_q^2\big)^{1/2}\colon\ x=y+z\big\}.$$

Observe that $B^m_{q^*}\cap aB^m_\infty\subset a^{(2-q^*)/2}B^m_2$, thus

$$\sup_{s\in B^m_{q^*}\cap aB^m_\infty}\sum_{i\le m}s_ix_i \le \sqrt{2}\|x\|_{2,q,a^{(2-q^*)/2}}. \tag{26}$$

The following lemma presents a crucial property of this norm. This will allow us to perform the induction step in the proof of Proposition 23. We will use its second assertion later, in Section 4.

Lemma 24. *For every $q\ge 2$, $b,c\in\mathbb{R}^m$, $d\ge 0$, and $\lambda\in(0,1]$,*

$$\mathbb{E}\big(d+\|(b_i+c_ig_i)_{i\le m}\|^2_{2,q,\lambda}\big)^{1/2} \le \big(d+\|b\|^2_{2,q,\lambda}+q\|c\|_q^2\big)^{1/2} \tag{27}$$

and

$$\mathbb{E}\Big(d+\|(b_i+c_ig_i)_{i\le m}\|_q^2\Big)^{1/2} \le \big(d+\|b\|_q^2+q\|c\|_q^2\big)^{1/2}. \tag{28}$$

Proof. Let $b',b''\in\mathbb{R}^m$. Then Jensen's inequality implies

$$\begin{aligned}&\mathbb{E}\big(d+\lambda^2\|b'\|_2^2+\|(b_i''+c_ig_i)_{i\le m}\|_q^2\big)^{1/2}\\ &\qquad\qquad \le \Big(d+\lambda^2\|b'\|_2^2+\Big(\sum_{i\le m}\mathbb{E}|b_i''+c_ig_i|^q\Big)^{2/q}\Big)^{1/2}.\end{aligned}$$

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

$$\begin{aligned}\Big(\sum_{i\le m}\mathbb{E}|b_i''+c_ig_i|^q\Big)^{2/q} &\le \Big(\sum_{i\le m}(\mathbb{E}|b_i''+\sqrt{q}c_ig_i|^2)^{q/2}\Big)^{2/q} = \|(b_i''^2+qc_i^2)_i\|_{q/2}\\ &\le \|(b_i''^2)_i\|_{q/2}+q\|(c_i^2)_i\|_{q/2} = \|b''\|_q^2+q\|c\|_q^2.\end{aligned}$$

Thus,

$$\mathbb{E}\big(d+\lambda^2\|b'\|_2^2+\|(b_i''+c_ig_i)_{i\le m}\|_q^2\big)^{1/2}\le\big(d+\lambda^2\|b'\|_2^2+\|b''\|_q^2+q\|c\|_q^2\big)^{1/2}.$$

Inequality (27) follows by taking b' and b'' such that

$$\|b\|_{2,q,\lambda}^2=\lambda^2\|b'\|_2^2+\|b''\|_q^2,$$

and observing that

$$\|(b_i+c_ig_i)_{i\le m}\|_{2,q,\lambda}^2\le\lambda^2\|b'\|_2^2+\|(b_i''+c_ig_i)_{i\le m}\|_q^2,$$

while inequality (28) follows by taking $b'=0$ and $b''=b$. □

For a given $p\ge 1$ and $k=0,1,\dots$ let

$$T_{\le k}=\big\{t\in\mathbb{R}^n:\ \|t\|_p\le 1,|t_j|\in\{0,1,1/2,\dots,1/2^k\},\ 1\le j\le n\big\}.$$

Choose $k_0=k_0(n):=\lceil\log_2 n\rceil+2\sim\operatorname{Log} n$. It is standard to see that for any norm on $\mathbb{R}^n$

$$\sup_{\|t\|_p\le 1}\|t\|\le 4\max_{t\in T_{\le k_0}}\|t\|. \tag{29}$$

Indeed, we decompose $t\in B_p^n$ as $t=t'+t''$, where

$$t_i'=\begin{cases}2^{-k}\operatorname{sgn}(t_i) & \text{if } 2^{-k}\le|t_i|<2^{-k+1} \text{ for } k=0,1,\dots,k_0,\\ 0 & \text{if } |t_i|<2^{-k_0}.\end{cases}$$

and $t''=t-t'$. Then $t'\in T_{\le k_0}$. Moreover, $\|t''\|_p\le 3/4$, since $|t_i''|\le\frac12|t_i|$ if $|t_i|\ge 2^{-k_0}$ and $\sum_{i:\ |t_i|<2^{-k_0}}|t_i|^p\le 2^{-k_0p}n\le 2^{-2p}$. Hence,

$$\sup_{\|t\|_p\le1}\|t\|\le\max_{t'\in T_{\le k_0}}\|t'\|+\sup_{\|t''\|_p\le 3/4}\|t''\|_p=\max_{t\in T_{\le k_0}}\|t\|+\frac34\sup_{\|t\|_p\le1}\|t\|_p.$$

The next lemma provides the crucial bound we need in the proof of Proposition 23.

Lemma 25. *For any $p\le 2$ and $k=0,1,\dots$ we have*

$$\mathbb{E}\max_{t\in T_{\le k}}\Big\|\Big(\sum_{j\le n}t_ja_{ij}g_{ij}\Big)_{i\le m}\Big\|_{2,q,\lambda}\le\sqrt{q}D_2+C\lambda D_1\sum_{l=0}^k 2^{-lp/2}\sqrt{\operatorname{Log}|T_{\le l}|}.$$

Proof. We will establish by induction on k the following stronger bound:

$$\mathbb{E}\max_{t\in T_{\le k}}\Big(q(1-\|t\|_2^2)D_2^2+\big\|\big(\sum_{j\le n}t_ja_{ij}g_{ij}\big)_{i\le m}\big\|_{2,q,\lambda}^2\Big)^{1/2} \le\sqrt{q}D_2+C\lambda D_1\sum_{l=0}^k 2^{-lp/2}\sqrt{\operatorname{Log}|T_{\le l}|}. \tag{30}$$

We will first show the induction step, i.e., that for $k=1,2,\dots$,

$$\begin{aligned}\mathbb{E}\max_{t\in T_{\le k}}\Big(q(1-\|t\|_2^2)D_2^2+\big\|\big(\sum_{j\le n}t_ja_{ij}g_{ij}\big)_{i\le m}\big\|_{2,q,\lambda}^2\Big)^{1/2}\\ \le\mathbb{E}\max_{t\in T_{\le k-1}}\Big(q(1-\|t\|_2^2)D_2^2+\big\|\big(\sum_{j\le n}t_ja_{ij}g_{ij}\big)_{i\le m}\big\|_{2,q,\lambda}^2\Big)^{1/2}\\ +C\lambda D_1 2^{-kp/2}\sqrt{\operatorname{Log}|T_{\le k}|}.\end{aligned} \tag{31}$$

For $t \in T_{\leq k}$ define

$$\begin{aligned}
J_k(t) &:= \{j \leq n \colon |t_j| = 2^{-k}\}, \quad \pi_k(t) := (t_j \mathbb{1}_{\{j \in J_k\}})_j, \\
J_{<k}(t) &:= \{j \leq n \colon |t_j| > 2^{-k}\}, \quad \pi_{<k}(t) := (t_j \mathbb{1}_{\{j \in J_{<k}\}})_j, \\
X_i(k,t) &:= \sum_{j \in J_k(t)} t_j a_{ij} g_{ij}, \quad Y_i(k,t) := \sum_{j \in J_{<k}(t)} t_j a_{ij} g_{ij}, \\
Z_k(t) &:= \Big(q(1 - \|t\|_2^2) D_2^2 + \big\| \big(X_i(k,t) + Y_i(k,t) \big)_{i \leq m} \big\|_{2,q,\lambda}^2 \Big)^{1/2}, \\
Z_{<k}(t) &:= \Big(q(1 - \|t\|_2^2 + \|\pi_k(t)\|_2^2) D_2^2 + \big\| \big(Y_i(k,t) \big)_{i \leq m} \big\|_{2,q,\lambda}^2 \Big)^{1/2}.
\end{aligned}$$

Observe that

$$\max_{t \in T_{\leq k}} \Big(q(1 - \|t\|_2^2) D_2^2 + \big\| \big(\sum_{j \leq n} t_j a_{ij} g_{ij} \big)_{i \leq m} \big\|_{2,q,\lambda}^2 \Big)^{1/2} = \max_{t \in T_{\leq k}} Z_k(t) \tag{32}$$

and

$$\max_{t \in T_{\leq k-1}} \Big(q(1 - \|t\|_2^2) D_2^2 + \big\| \big(\sum_{j \leq n} t_j a_{ij} g_{ij} \big)_{i \leq m} \big\|_{2,q,\lambda}^2 \Big)^{1/2} = \max_{t \in T_{\leq k}} Z_{<k}(t). \tag{33}$$

Let $\mathbb{E}_{k,t}$ detones the integration with respect to variables $(g_{i,j})_{i \leq m, j \in J_k(t)}$. Conditional application of (27) yields for any $t \in T_{\leq k}$,

$$\mathbb{E}_{k,l} Z_k(t) \leq \Big(q(1 - \|t\|_2^2) D_2^2 + q \|((\mathbb{E} X_i^2(k,t))^{1/2})_{i \leq m}\|_q^2 + \big\| \big(Y_i(k,t) \big)_{i \leq m} \big\|_{2,q,\lambda}^2 \Big)^{1/2}.$$

Convexity of the function $x \to |x|^{q/2}$ yields

$$\begin{aligned}
\|((\mathbb{E} X_i^2(k,t))^{1/2})_{i \leq m}\|_q^q &= \sum_{i \leq m} \Big| \sum_{j \in J_k(t)} a_{ij}^2 t_j^2 \Big|^{q/2} \leq \|\pi_k(t)\|_2^q \sum_{j \in J_k(t)} \frac{t_j^2}{\|\pi_k(t)\|_2^2} \sum_{i \leq m} |a_{ij}|^q \\
&\leq \|\pi_k(t)\|_2^q \max_{j \in J_k(t)} \sum_{i \leq m} |a_{ij}|^q \leq \|\pi_k(t)\|_2^q D_2^q.
\end{aligned}$$

Hence

$$\mathbb{E}_{k,t} Z_k(t) \leq Z_{<k}(t).$$

For a fixed $d \geq 0$ and $y \in \mathbb{R}^m$ define functions $f \colon \mathbb{R}^m \to \mathbb{R}$ and $f_1, \ldots, f_m \colon \mathbb{R}^{|J_k(t)|} \to \mathbb{R}$ via

$$f(z) := \Big(d + \big\| \big(z_i + y_i \big)_{i \leq m} \big\|_{2,q,\lambda}^2 \Big)^{1/2}, \quad f_i(x_i) := \sum_{j \in J_k(t)} a_{ij} t_j x_{i,j}.$$

Then f is Lipschitz with a constant $\sup_{\|z\|_2 \leq 1} \|z\|_{2,q,\lambda} \leq \lambda$ and, since $(\frac{p}{2-p})^* = \frac{p^*}{2}$, f_i is Lipschitz with a constant

$$\begin{aligned}
\Big(\sum_{j \in J_k(t)} a_{ij}^2 t_j^2 \Big)^{1/2} &\leq 2^{-kp/2} \Big(\sum_{j \leq n} a_{ij}^2 |t_j|^{2-p} \Big)^{1/2} \leq 2^{-kp/2} \sup_{y \in B^n_{p/(2-p)}} \Big(\sum_{j \leq n} a_{ij}^2 y_j \Big)^{1/2} \\
&\leq 2^{-kp/2} D_1.
\end{aligned}$$

Hence, the function $x \to f(f_1(x_1), \ldots, f_m(x_m))$ is $\lambda 2^{-kp/2} D_1$-Lipschitz on $\mathbb{R}^{|J_k| m}$. Therefore, the Gaussian concentration yields

$$\mathbb{P}_{k,t}\big(Z_k(t) \geq Z_{<k}(t) + u \lambda 2^{-kp/2} D_1 \big) \leq e^{-u^2/2} \quad \text{for } u \geq 0.$$

and

$$\mathbb{P}\big(\max_{t\in T_{\le k}} Z_k(t) \ge \max_{t\in T_{\le k}} Z_{<k}(t) + u\lambda 2^{-kp/2}D_1\big)$$
$$\le \sum_{t\in T_{\le k}} \mathbb{E}\mathbb{P}_{k,t}(Z_k(t) \ge Z_{<k}(t) + u\lambda 2^{-kp/2}D_1) \le |T_{\le k}|e^{-u^2/2}.$$

Integration by parts yields

$$\mathbb{E}\Big(\max_{t\in T_{\le k}} Z_k(t) - \max_{t\in T_{\le k}} Z_{<k}(t)\Big)_+ \le C\lambda 2^{-kp/2}D_1\sqrt{\log|T_{\le k}|},$$

so

$$\mathbb{E}\max_{t\in T_{\le k}} Z_k(t) \le \mathbb{E}\max_{t\in T_{\le k}} Z_{<k}(t) + C\lambda 2^{-kp/2}D_1\sqrt{\log|T_{\le k}|},$$

which, in view of (32) and (33), implies (31).

In a similar (in fact a bit simpler) way we show that

$$\mathbb{E}\max_{t\in T_{\le 0}}\Big(q(1-\|t\|_2^2)D_2^2 + \big\|\big(\sum_{j\le n} t_j a_{ij}g_{ij}\big)_{i\le m}\big\|_{2,q,\lambda}^2\Big)^{1/2} \le \sqrt{q}D_2 + CD_1\sqrt{\operatorname{Log}|T_{\le 0}|},$$

which together with (31) completes the inductive proof of (30). □

Lemma 26. *We have* $\log|T_{\le k}| \lesssim 2^{pk}\operatorname{Log}(n(k+1))$ *for* $p<\infty$ *and* $k=0,1,\ldots$.

Proof. Each $t\in T_{\le k}$ has support of cardinality at most 2^{pk}. Every nonzero coordinate of such t may be chosen in at most $2k+1$ ways, so

$$|T_{\le k}| \le \sum_{0\le l\le 2^{pk}} \binom{n}{l}(2k+1)^l \le \sum_{0\le l\le 2^{pk}} (2n(k+1))^l \le (2n(k+1))^{2^{pk}+1}. \qquad \square$$

Proof of Proposition 23. It suffices to prove (24), since (25) follows by duality.

Choose $k_0 = k_0(n) := \log_2 n + 1 \sim \operatorname{Log} n$. Then by (26), (29), and Lemmas 25 and 26 we get

$$\begin{aligned}\mathbb{E}\sup_{s\in B_{q^*}^m\cap aB_\infty^n}\sup_{t\in B_2^n}\sum_{i\le m,j\le n} a_{ij}g_{ij}s_it_j &\le \sqrt{2}\mathbb{E}\sup_{t\in B_2^n}\Big\|\Big(\sum_{j\le n} a_{ij}g_{ij}t_j\Big)_{i\le m}\Big\|_{2,q,a^{(2-q^*)/2}}\\ &\lesssim \mathbb{E}\sup_{t\in T_{\le k_0}}\Big\|\Big(\sum_{j\le n} a_{ij}g_{ij}t_j\Big)_{i\le m}\Big\|_{2,q,a^{(2-q^*)/2}}\\ &\lesssim \sqrt{q}D_2 + a^{(2-q^*)/2}D_1\sum_{l=0}^{k_0} 2^{-pk/2}\sqrt{\operatorname{Log}|T_{\le k}|}\\ &\lesssim \sqrt{q}D_2 + a^{(2-q^*)/2}D_1(k_0+1)\sqrt{\operatorname{Log} n}\\ &\lesssim \sqrt{q}D_2 + a^{(2-q^*)/2}D_1\operatorname{Log}^{3/2} n. \qquad \square\end{aligned}$$

Now we are ready to prove Proposition 17 in the case $(p^*,q)\notin[2,3]^2$. We postpone the proof in the case $(p^*,q)\in[2,3]^2$ to Section 5.

Proof of Proposition 17 in the case $(p^*,q)\notin[2,3]^2$. By duality it suffices to consider the case $p^*\le q$. Then $q>3$.

Let $a := \operatorname{Log}^{-3/(2-q^*)}(mn)$. Then Proposition 23 implies

$$\begin{aligned}\mathbb{E}\sup_{s\in B_{q^*}^m\cap aB_\infty^m}\sup_{t\in B_p^n}\sum_{i,j} a_{ij}g_{ij}s_it_j &\lesssim \sqrt{q}D_2 + a^{(2-q^*)/2}\operatorname{Log}^{3/2}(mn)D_1\\ &= \sqrt{q}D_2 + D_1.\end{aligned}$$

Moreover, Proposition 20 (applied with $\gamma = \mathrm{Log}^{3/(2-q^*)}(mn)$) yields

$$\mathbb{E} \sup_{s\in B_{q^*}^m} \sup_{t\in B_p^n} \sum_{i,j} a_{ij} g_{ij} s_i t_j \mathbb{1}_{\{|s_i|\ge \mathrm{Log}^{-3/(2-q^*)}(mn)\}}$$
$$\lesssim \sqrt{p^*} D_1 + \mathrm{Log}^{1/2+3q^*/(2(2-q^*))}(mn) \max_{i,j} |a_{ij}|.$$

The two displayed inequalities yield the assertion (with $\gamma = 5$, since $q^* \le 3/2$). $\square$

3.2. **Exponent reduction.** Proposition 17 shows that

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

In this and the next subsection we show how to reduce the exponent γ and obtain Proposition 15. The argument is based on the analysis of the graph associated to the matrix A. Such an analysis was performed in the Bernoulli case in [14]. 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 3.3 in the case when $p^* \vee q > 2$, and in Section 5 when $p^*, q \in [2, 4)$). Since we work in the Gaussian setting and not with bounded Bernoulli entries as in [14], we face some new difficulties. It is possible to deal with them making the advantage of the fact that $p^* \wedge q > 2$, whereas in the case $p^*, q \in [2, 4)$ one may use more involved combinatorial arguments and exploit the bounds following from the Slepian-Fernique lemma. The exponent reduction procedure in the latter case is run in Section 5.

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

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

We set

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

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 $[m] \sqcup [n]$ induced by E_A. A subset $I \subset [m]$ or $J \subset [n]$ is called r-connected if it is a connected subset of the graph $G_A(r) := ([m] \sqcup [n], E_A(r))$, where $(u,v) \in E_A(r)$ if and only if $\rho_A(u,v) \le r$ for $u, v \in [m] \sqcup [n]$.

We denote by $\mathcal{I}_r(k) = \mathcal{I}_r(k, A)$ (respectively, $\mathcal{J}_r(k) = \mathcal{J}_r(k, A)$) the family of all r-connected subsets of $[m]$ (respectively, $[n]$) of cardinality k. Note that the maximal degree of $G_A(r)$ is at most $d_A + d_A(d_A - 1) + \cdots + d_A(d_A-1)^{r-1} \le d_A^r$. Thus, [14, Lemma 11] implies

$$|\mathcal{I}_r(k)| \le m 4^k d_A^{rk} \quad \text{and} \quad |\mathcal{J}_r(k)| \le n 4^k d_A^{rk}. \tag{34}$$

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 (8) from a weaker estimate depending on $d_{1,A}$ and $d_{2,A}$.

Proposition 27. *Let $p^*,q\in[2,\infty)$, $0\le\gamma_1<\frac{1}{p^*}$, $0\le\gamma_2<\frac{1}{q}$, $\gamma_3>0$ and $\alpha_1,\alpha_2,\alpha_3\ge 1$ be such that for every $m\times n$ matrix A,*

$$\mathbb{E}\|G_A\|_{p\to q}\le\alpha_1 D_1+\alpha_2 D_2+\alpha_3\bigl(d_{1,A}^{\gamma_1}+d_{2,A}^{\gamma_2}+\sqrt{\operatorname{Log}(mn)}\bigr)\max_{i,j}|a_{ij}| \tag{35}$$

and

$$\mathbb{E}\|G_A\|_{p\to q}\le\alpha_1 D_1+\alpha_2 D_2+\alpha_3\operatorname{Log}^{\gamma_3}(mn)\max_{i,j}|a_{ij}|. \tag{36}$$

Then for every $m\times n$ matrix A,

$$\mathbb{E}\|G_A\|_{p\to q}\lesssim\frac{\operatorname{Log}\gamma_3}{-\ln((\gamma_1p^*)\vee(\gamma_2q)))}\cdot\Bigl((\alpha_1+\alpha_3)D_1+(\alpha_2+\alpha_3)D_2+\alpha_3\sqrt{\operatorname{Log}(mn)}\max_{i,j}|a_{ij}|\Bigr).$$

Proof. Let $\delta=((p^*\gamma_1)\vee(q\gamma_2))^{-1}-1$. Then $p^*\gamma_1,q\gamma_2\le(1+\delta)^{-1}$. Let

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

and define

$$u_k:=(\operatorname{Log}(mn))^{-\frac12(1+\delta)^{k+1}},\quad k=0,1,\ldots,k_0.$$

Let

$$M:=D_1+D_2,\quad \widetilde{M}:=\alpha_1D_1+\alpha_2D_2+\alpha_3\sqrt{\operatorname{Log}(mn)}\max_{i,j}|a_{ij}|.$$

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

$$a_{ij}(0)=a_{ij}\mathbb{1}_{\{|a_{ij}|\ge u_0M\}},\qquad a_{ij}(k_0+1)=a_{ij}\mathbb{1}_{\{|a_{ij}|<u_{k_0}M\}},$$
$$\text{and}\quad a_{ij}(k)=a_{ij}\mathbb{1}_{\{u_kM\le|a_{ij}|<u_{k-1}M\}}\quad\text{for } k=1,\ldots,k_0.$$

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

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

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,

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

Since

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

assumption (35) (applied to the matrix A_0) yields

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

Moreover, assumption (35), applied to the matrix A_k, yields for $1\le k\le k_0$,

$$\begin{aligned}\|G_{A_k}\|_{p\to q}&\le\widetilde{M}+\alpha_3u_{k-1}M(d_{1,A_k}^{\gamma_1}+d_{2,A_k}^{\gamma_2})\le\widetilde{M}+\alpha_3Mu_{k-1}(u_k^{-p^*\gamma_1}+u_k^{-q\gamma_2})\\&\le\widetilde{M}+2\alpha_3Mu_{k-1}u_k^{-1/(1+\delta)}=\widetilde{M}+2\alpha_3M.\end{aligned}$$

Finally, we apply assumption (36) to get

$$\|G_{A_{k_0+1}}\|_{p\to q}\le\alpha_1D_1+\alpha_2D_2+\alpha_3M.$$

We finish the proof by noting that $k_0\lesssim\frac{\operatorname{Log}\gamma_3}{-\ln((\gamma_1p^*)\vee(\gamma_2q))}$. □

3.3. **Degree dependent bounds.** In order to provide a weaker degree dependent estimate (35) 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 28. *For every $\varepsilon \in (0,1]$ and $1 \le p \le 2 \le q \le \infty$,*

$$\mathbb{E}\|G_A\|_{p\to q} \lesssim \frac{1}{\varepsilon} \mathbb{E} \max_{1\le k\le m} \max_{1 \le l \le n} Y_{k,l}(d_A^\varepsilon) + \mathbb{E}\max_{i,j} |a_{ij} g_{ij}| \tag{37}$$

$$\lesssim \frac{1}{\varepsilon}\Big(\max_{1\le k\le m} \max_{1 \le l \le n} \mathbb{E} Y_{k,l}(d_A^\varepsilon) + \sqrt{\operatorname{Log}(mn)} \|(a_{ij})\|_\infty \Big). \tag{38}$$

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

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

Then

$$\sum_{k\ge 1} d_A^{-q^* k\varepsilon} |I_k(s)| \le \|s\|_{q^*}^{q^*}, \quad \sum_{l \ge 1} d_A^{-pl\varepsilon} |J_l(t)| \le \|t\|_p^p$$

and

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

Define

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

Observe that for any $s \in B_{q^*}^m$ and $t \in B_p^n$,

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

Similarily,

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

Therefore, it is enough to estimate

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

Let us fix $|r| \le 1/\varepsilon + 1$, $s \in B_{q^*}^m$, and $t \in B_p^n$. For $k,l \ge 1$ with $l-k=r$ let $I_{k,1},\dots,I_{k,u_k}$ be 2-connected components of $I_k(s) \cap J_l(t)'$ and $J_{k,u} \coloneqq \{j \in$

$J_l(t)\colon I_{k,u} \sim_A j\}$, $1 \le u \le u_k$. 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

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

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

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

Moreover,

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

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 (37).

Moreover, $\mathbb{E} M_A \lesssim \mathrm{Log}^{1/2}(mn) \|(a_{ij})\|_\infty$ and, by estimate (21),

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

so estimate (38) follows. □

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

Proposition 29. *For every $p^*, q \in [2,\infty)$ and $\gamma \ge 1$,*

$$\begin{aligned}\mathbb{E} Y_{k,l}(\gamma) &\lesssim \sqrt{p^*} D_1 + \sqrt{q} D_2 + \sqrt{\mathrm{Log}(mn)} \max_{i,j} |a_{ij}| \\ &\quad + \gamma^{3/2} k^{-1/q^*} l^{-1/p} \min\{k l^{1/2}, k^{1/2} l, k d_{1,A}^{1/2}, l d_{2,A}^{1/2}\} \sqrt{\mathrm{Log}(d_A)} \max_{i,j} |a_{ij}|. \qquad (39)\end{aligned}$$

Proof. Observe that

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

where

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

Define

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

Let us fix $I\in\mathcal{I}_4(k)$ and choose a $1/2$-net S 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

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

By inequality (21) we have

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

By (19)

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

Thus,

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

Applying estimate (21) again and using (34) we get

$$\begin{aligned}\mathbb{E}Y_{k,l} &\lesssim \max_{I\in\mathcal{I}_4(k)}\mathbb{E}X_I + \sqrt{\operatorname{Log}|\mathcal{I}_4(k)|}\sigma_{k,l}\\ &\lesssim \sqrt{p^*}D_1 + \big(\sqrt{k\operatorname{Log}(d_A)} + \sqrt{\operatorname{Log} m}\big)\sigma_{k,l}.\end{aligned}\tag{40}$$

To estimate $\sigma_{k,l}$ observe first that $B_{q^*}^I\subset B_2^I$ and $B_p^n\subset B_2^n$, thus

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

Moreover, for any $s\in B_{q*}^m\cap\gamma k^{-1/q^*}B_\infty^m$ and $t\in B_p^n\cap\gamma l^{-1/p}B_\infty^n$ we have

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

and

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

Therefore

$$\sigma_{k,l} \le \min\{1,\gamma k^{1/2-1/q^*}l^{1/2-1/p},\gamma^{3/2}k^{1/2-1/q^*}l^{-1/p}d_{1,A}^{1/2}\}\max_{i,j}|a_{ij}|.\tag{41}$$

Estimates (40) and (41) yield

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

In a similar way we show that

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

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

Corollary 30. *Suppose that $p^*, q\in[2,\infty)$ are such that $\frac{1}{p}+\frac{1}{q^*}\geq 3/2$. 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 28 we have

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

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,

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

Hence estimate (39) 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 for $\varepsilon\in(0,1)$,

$$\sup_{x\geq 1} x^{-\varepsilon}\operatorname{Log} x = \frac{1}{1\wedge(e\varepsilon)} \leq \frac{1}{\varepsilon}, \tag{42}$$

so $d_A^{\varepsilon/2}\sqrt{\operatorname{Log} d_A}\leq d_A^{\varepsilon}\varepsilon^{-1/2}$. $\square$

Corollary 31. *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*

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

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

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

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. $k \ge \rho l$. Then

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

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

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

Therefore, Propositions 28 and 29 (applied to $\varepsilon \coloneqq \varepsilon/3$ and $\gamma \coloneqq d_A^{\varepsilon/3}$) yield for every $\varepsilon \in (0,1)$,

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

Recall that estimate (42) implies

$$d_A^{\varepsilon/2}\sqrt{\operatorname{Log} d_A} \le \varepsilon^{-1/2}d_A^{\varepsilon},$$

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

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

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

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

and

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

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

Now we are ready to prove Proposition 15 in the range $(p,q) \notin [2,3]^2$. The case $(p,q) \in [2,3]^2$ is considered in Section 5.

Proof of Proposition 15 in the case $(p^*,q) \notin [2,3]^2$. The assertion follows by

- Corollary 30, applied with $\varepsilon = (2(p^*\vee q))^{-1}$, and Propositions 17 and 27 if $1/p+1/q^* \ge 3/2$.
- Corollary 31 and Propositions 17 and 27 if $1/p+1/q^* < 3/2$.

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

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

4. Proposition 15 implies Proposition 16

To prove Proposition 16 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 32. *Let $p^*, q \in [2,\infty)$. Then for every partition $J_1, J_2, \ldots, J_{k_0}$ of $[n]$,*

$$\mathbb{E}\|G_A\|_{p\to q} \le \mathbb{E}\|G_A\|_{2\to q} \lesssim \sqrt{q}D_2 + \max_{1\le k\le k_0}\max_{1\le l\le k} k^3|J_l|^{1/2}\max_{i\le m, j\in J_k}|a_{ij}|. \tag{43}$$

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

$$\mathbb{E}\|G_A\|_{p\to q} \le \mathbb{E}\|G_A\|_{p\to 2} \lesssim \sqrt{p^*}D_1 + \max_{1\le k\le k_0}\max_{1\le l\le k} k^3|I_l|^{1/2}\max_{i\in I_k, j\le n}|a_{ij}|. \tag{44}$$

In order to show Proposition 32 we need estimate (28) and the following lemma; they allow us to perform the induction step.

Lemma 33. *For every $q \ge 2$, $J \subset J' \subset [n]$, a finite set $T \subset B_2^{J'}$ and functions $c = (c_1, \ldots, c_m)\colon T \to \mathbb{R}^m$ and $d\colon T \to [0,\infty)$ we have*

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

Proof. Observe that

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

where

$$X_{s,v,t} := v_1\sum_{i\le m} s_i\Big(c_i(t) + \sum_{j\in J} a_{ij}g_{ij}t_j\Big) + v_2\sqrt{d(t)}.$$

We have

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

Hence the assertion follows by Lemma 22. □

Proof of Proposition 32. We will prove (43); the second estimate (44) follows by duality.

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

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

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

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

Define

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

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 $0\le k\le k_0$,

$$\mathbb{E}\sup_{t\in T_1}\Big(\sum_i\Big|\sum_j a_{ij}g_{ij}t_j\Big|^q\Big)^{1/q} \le \mathbb{E}\sup_{t\in T_1}\Big(\Big(\sum_i\Big|\sum_{j\in J_{\le k}} a_{ij}g_{ij}t_j\Big|^q\Big)^{2/q} + qD_2^2\sum_{j\in J_{>k}} t_j^2\Big)^{1/2}$$
$$+ C\sum_{l=k+1}^{k_0}\sqrt{\operatorname{Log}|T_{1,\le l}|}\max_i\max_{j\in J_l}|a_{ij}|. \tag{45}$$

For $k=k_0$ inequality (45) holds with equality. Observe that

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

where

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

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 33 yields

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

Inequality (28), applied conditionally, implies that for any $\tilde t\in T_{1,\le k}$,

$$\mathbb{E}_{J_k}\Big(\Big(\sum_i\Big|\sum_{j\in J_{\le k}} a_{ij}g_{ij}\tilde t_j\Big|^q\Big)^{2/q} + d(\tilde t)\Big)^{1/2}$$
$$\le \Big(\Big(\sum_i\Big|\sum_{j\in J_{\le k-1}} a_{ij}g_{ij}\tilde t_j\Big|^q\Big)^{2/q} + \tilde d(\tilde t)\Big)^{1/2},$$

where

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

Note that

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

so

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

Thus,

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

so the reverse induction step immediately follows.

Estimate (45) 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 that Proposition 16 follows from Proposition 15.

Proof of Proposition 16. Let $r=8(p^*\vee q)$ and

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

Define the sequence $(n_k)_{k\ge 0}$ by

$$
n_0:=0,\ n_k:=\lceil e^{r^k}\rceil\quad\text{for } k\ge 1.
$$

Then

$$
\sqrt{\operatorname{Log}(n_kn_l)}\le\sqrt{\operatorname{Log}n_k}+\sqrt{\operatorname{Log}n_l}\lesssim r(n_{k-1}+n_{l-1}+2)^{1/r}\ \text{for } k,l\ge 1. \tag{46}
$$

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

$$
\max_{i\in I_k,j\in J_l}(n_{k-1}+n_{l-1}+2)^{1/r}|a_{ij}|\le\max_{i\in I_k,j\in J_l}(i+j)^{1/r}|a_{ij}|\le D'_\infty. \tag{47}
$$

Let

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

Then Proposition 15 and estimates (46) and (47) yield

$$
\begin{aligned}
\mathbb{E}\|G_{A_{k,l}}\|_{p\to q} &\lesssim \alpha(p,q)\Big(\sqrt{p^*}D_1+\sqrt{q}D_2+\sqrt{\mathrm{Log}(|I_k||J_l|)}\max_{i\in I_k, j\in J_l}|a_{ij}|\Big)\\
&\lesssim \alpha(p,q)(\sqrt{p^*}D_1+\sqrt{q}D_2+rD'_\infty). \qquad (48)
\end{aligned}
$$

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

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

For every u (after deleting some zero rows and columns) A_u is block diagonal with blocks $A_{k,k+u}$. Hence, Lemmas 19 and 22, together with estimates (47) and (48), imply

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

We have

$$A=\sum_{u=-1}^{1}A_u+B_1+B_2,$$

where

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

Estimate (43) applied to the matrix B_1 yields

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

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

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

Hence,

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

Similarily, estimate (44) implies

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

5. Estimates in the range $p^*, q\in[2,4)$

The aim of this section is to give proofs of Propositions 15 and 17 in the range $p^*, q\in[2,3]$. The arguments presented in this section work in the whole regime $p^*, q\in[2,4)$, but they yield a constant which blows up when p^* or q approaches 4.

The crucial new tool we need to establish Proposition 15 in the case $p^*, q\in[2,3]$ is the following result, which allows to run the exponent reduction; this is a counterpart of Corollaries 30 and 31 from Section 3.3.

Proposition 34. *If $p^*, q\in[2,4-\delta)$ and $\varepsilon\in(0,1/2]$, then*

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

To prove Proposition 34 (and Proposition 17 for $p^*, q\in[2,3]$) we will need the following result.

Proposition 35. *If $p^*, q \in [2,4)$ and $a, b \in (0,1]$, then*

$$\mathbb{E} \sup_{s \in B^m_{q^*} \cap a B^m_\infty} \sup_{t \in B^n_p \cap b B^n_\infty} \sum_{i,j} a_{ij} g_{ij} s_i t_j$$
$$\lesssim D_1 + D_2 + \Big(a^{(2-q^*)/2} \operatorname{Log}^{\frac{2}{4-p^*}}(m a^{q^*}) + b^{(2-p)/2} \operatorname{Log}^{\frac{2}{4-q}}(n b^p) \Big) \max_{i,j} |a_{ij}|.$$

Proof of Propositions 15 and 17 for $2 \le p^*, q \le 3$. Proposition 35, applied with $a = b = 1$, yields Proposition 17 (with $\gamma = 2$). Propositions 15 follows from Proposition 27 together with Propositions 17 and 34. □

To prove Proposition 35 we will use a modification of van Handel's argument from [18]. 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 an $(m+n)$-dimensional Gaussian vector with mean 0 and covariance matrix B^- being the negative part of B, i.e., $B^- = -\sum_{i=1}^{m+n} (\lambda_i \wedge 0) u_i u_i^T$, where $B = \sum_{i=1}^{m+n} \lambda_i u_i u_i^T$ is the spectral decomposition of B. The proof of [18, 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,

$$\operatorname{Var}(Y_i) \le \Big(\sum_j a_{ij}^4 \Big)^{1/2}, \quad 1 \le i \le m,$$
$$\operatorname{Var}(Y_{j+m}) \le \Big(\sum_i a_{ij}^4 \Big)^{1/2}, \quad 1 \le j \le n. \tag{49}$$

The proof of Proposition 35 uses the following two lemmas. The first one is based on the ideas from [18] and the second one is quite standard (cf. [14, Lemma 14] for cases $\rho = 1, 2$).

Lemma 36. *For any bounded, nonempty sets $K \subset \mathbb{R}^m$ and $L \subset \mathbb{R}^n$,*

$$\mathbb{E} \sup_{s \in K, t \in L} a_{ij} g_{ij} s_i t_j \le \mathbb{E} \sup_{s \in K, t \in L} \sum_{i=1}^m s_i g_i \sqrt{\sum_{j=1}^n t_j^2 a_{ij}^2}$$
$$+ \mathbb{E} \sup_{s \in K, t \in L} \sum_{j=1}^n t_j g_{m+j} \sqrt{\sum_{i=1}^m s_i^2 a_{ij}^2}$$
$$+ \frac{1}{2} \mathbb{E} \sup_{s \in K} \sum_{i=1}^m s_i^2 Y_i + \frac{1}{2} \mathbb{E} \sup_{t \in L} \sum_{j=1}^n t_j^2 Y_{m+j}.$$

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 $v \in \mathbb{R}^{n+m}$:

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

It was shown in [18, 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

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

□

Lemma 37. *Suppose that* $1\le \rho\le \rho_0$, $c\in(0,1]$, $m_j,\sigma_j\ge 0$ *and nonnegative random variables* $Z_1,\dots,Z_n$ *satisfy*

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

Then

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

Proof. Let k be a positive integer such that $\frac{1}{k+1}<c^2\le\frac1k$ and $(Z_1^*,\dots,Z_n^*)$ be the nonincreasing rearrangement of $(Z_1,\dots,Z_n)$. Let $m=\max_j m_j$, $\sigma=\max_j\sigma_j$, and $M=(m+\sqrt{2\,\mathrm{Log}(n/k)}\sigma)^\rho$. Then

$$\begin{aligned}\Big(\frac1k\sum_{l=1}^k\mathbb{E}(Z_l^*)^\rho\Big)^{1/\rho}&\le 2\Big(M+\frac1k\sum_{l=1}^n\mathbb{E}(Z_l-m-\sqrt{2\,\mathrm{Log}(n/k)}\sigma)_+^\rho\Big)^{1/\rho}\\ &\lesssim\Big(M+\frac1k\sum_{l=1}^n\sigma^\rho\int_0^\infty\rho t^{\rho-1}\,\mathbb{P}\big(Z_l\ge m+\sigma(\sqrt{2\,\mathrm{Log}(n/k)}+t)\big)dt\Big)^{1/\rho}\\ &\le\Big(M+\frac nk\sigma^\rho\int_0^\infty\rho t^{\rho-1}e^{-(t+\sqrt{2\,\mathrm{Log}(n/k)})^2/2}dt\Big)^{1/\rho}\\ &\le\Big(M+\sigma^\rho\int_0^\infty\rho t^{\rho-1}e^{-t^2/2}dt\Big)^{1/\rho}\lesssim_{\rho_0} m+\sqrt{\mathrm{Log}(n/k)}\sigma,\end{aligned}$$

so

$$\begin{aligned}\mathbb{E}\sup_{u\in B_2^n\cap cB_\infty^n}\Big(\sum_{j=1}^n u_j^2Z_j^\rho\Big)^{1/\rho}&\le\mathbb{E}\Big(\frac1k\sum_{l=1}^k(Z_l^*)^\rho\Big)^{1/\rho}\lesssim_{\rho_0} m+\sqrt{\mathrm{Log}(n/k)}\sigma\\ &\sim m+\sqrt{\mathrm{Log}(nc^2)}\sigma.\end{aligned}$$

□

Proof of Proposition 35. We apply Lemma 36 with $K=B_{q^*}^m\cap aB_\infty^m$ and $L=B_p^n\cap bB_\infty^n$. Observe that $\|s\|_2\le\|s\|_\infty^{(2-q^*)/2}\|s\|_{q^*}^{q^*/2}$. This, together with a similar calculation for ℓ_p-norms, yields

$$K\subset a^{(2-q^*)/2}(B_2^m\cap a^{q^*/2}B_\infty^m),\quad L\subset b^{(2-p)/2}(B_2^n\cap b^{p/2}B_\infty^n). \tag{50}$$

Estimate (49) implies that $Y_1,\dots,Y_m$ are centered Gaussians with $\mathrm{Var}(Y_i)\le\|(a_{ij})_j\|_4^2$. Thus, (50) and Lemma 37, applied with $Z_j=|Y_j|$, $\rho=1$, $c=a^{q^*/2}$, $m_j=0$ and $\sigma_j=\|(a_{ij})_j\|_4$, yield that

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

In a similar way we show that

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

By (50) and the convexity of the function $x\mapsto|x|^{q/2}$ we get

$$\begin{aligned}\sup_{s\in K,t\in L}\sum_{i=1}^m s_ig_i\sqrt{\sum_{j=1}^n t_j^2a_{ij}^2}&\le\sup_{t\in L}\Big(\sum_{i=1}^m|g_i|^q\Big(\sum_{j=1}^n t_j^2a_{ij}^2\Big)^{q/2}\Big)^{1/q}\\ &\le b^{(2-p)/2}\sup_{t\in B_2^n\cap b^{p/2}B_\infty^n}\Big(\sum_{i=1}^m|g_i|^q\Big(\sum_{j=1}^n t_j^2a_{ij}^2\Big)^{q/2}\Big)^{1/q}\\ &\le b^{(2-p)/2}\sup_{t\in B_2^n\cap b^{p/2}B_\infty^n}\Big(\sum_{j=1}^n t_j^2\|(a_{ij}g_i)_i\|_q^q\Big)^{1/q}\Big(\sum_{j=1}^n t_j^2\Big)^{1/2-1/q}\\ &\le b^{(2-p)/2}\sup_{t\in B_2^n\cap b^{p/2}B_\infty^n}\Big(\sum_{j=1}^n t_j^2\|(a_{ij}g_i)_i\|_q^q\Big)^{1/q}.\end{aligned}\tag{51}$$

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 (51) and Lemma 37, applied with $Z_j=\|(a_{ij}g_i)_i\|_q$, $\rho=q$, $c=b^{p/2}$, and $\rho_0=4$, imply

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

In a similar way we show that

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

Now we go back to the proof of Proposition 34. We will need a refinement of the arguments from the proof of the main result of [1]. Let

$$K_{p,n} := \operatorname{conv}\Big\{\frac{1}{|J|^{1/p}}\big(\varepsilon_j \mathbb{1}_{\{j\in J\}}\big)_{j=1}^{n} : J\subset\{1,\dots,n\}, J\neq\emptyset, (\varepsilon_j)_{j=1}^{n}\in\{-1,1\}^n\Big\}.$$

Recall [1, Lemma 2.3].

Lemma 38. *If $p\geq 1$, then $B_p^n\subset \ln(en)^{1/p^*}K_{p,n}$.*

Following the ideas from [1] we obtain the following bound.

Lemma 39. *Assume that $p^*, q\geq 2$. Then*

$$\mathbb{E}\sup_{s\in K_{q^*,m},t\in K_{p,n}}\sum_{i,j} a_{ij}g_{ij}s_i t_j \lesssim \sqrt{p^*}D_1 + \sqrt{q}D_2 + \sqrt{\operatorname{Log}(mn)}\max_{i,j}|a_{ij}|.$$

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

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

where the maxima are taken over all sets $I\subset\{1,\dots,m\}$, $J\subset\{1,\dots,n\}$ such that $|I|=k$, $|J|=l$, and $g_1,\dots,g_m,\widetilde{g}_1,\dots,\widetilde{g}_n$ are independent standard Gaussian variables. Moreover,

$$\frac{1}{k^{1/q^*}l^{1/p}}\max_{I,J}\sum_{i\in I}\sqrt{\sum_{j\in J}a_{ij}^2} \leq \sup_{s\in B_{q^*}^m}\sup_{t\in B_p^n}\sum_{i=1}^{m} s_i\sqrt{\sum_{j=1}^{n} t_j^2 a_{ij}^2} = D_2.$$

Similarly,

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

and

$$\frac{1}{k^{1/q^*}l^{1/p}}\mathbb{E}\max_{I,J}\sum_{i\in I} g_i\sqrt{\sum_{j\in J}a_{ij}^2}\leq \mathbb{E}\max_j\|(a_{ij}g_i)_i\|_q. \tag{52}$$

Estimate (21) yields for every $q\in[2,\infty)$,

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

This and (52) yield

$$\frac{1}{k^{1/q^*}l^{1/p}}\mathbb{E}\max_{I,J}\sum_{i\in I} g_i\sqrt{\sum_{j\in J}a_{ij}^2}\lesssim \sqrt{q}D_2 + \sqrt{\operatorname{Log} n}\max_{i,j}|a_{ij}|.$$

Similarly,

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

This implies that for every $k,l\geq 1$,

$$\mathbb{E}W_{k,l}\lesssim\sqrt{p^*}D_1+\sqrt{q}D_2+\sqrt{\operatorname{Log}(mn)}\max_{i,j}|a_{ij}|, \tag{53}$$

where

$$W_{k,l}\coloneqq\sup_{s\in K^{(k)}_{q^*,m},t\in K^{(l)}_{p,n}}\sum_{i,j}a_{ij}g_{ij}s_it_j$$

and

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

Hence, by inequality (21) we obtain

$$\begin{aligned}\mathbb{E}\sup_{s\in K_{q^*,m},t\in K_{p,n}}\sum_{i,j}a_{ij}g_{ij}s_it_j&=\mathbb{E}\max_{1\leq k\leq m,1\leq l\leq n}W_{k,l}\\&\lesssim\max_{1\leq k\leq m,1\leq l\leq n}\mathbb{E}W_{k,l}+\sqrt{\operatorname{Log}(mn)}\sup_{s\in B^m_{q^*},t\in B^n_p}\sqrt{\sum_{i,j}a_{ij}^2s_i^2t_j^2}\\&=\max_{1\leq k\leq m,1\leq l\leq n}\mathbb{E}W_{k,l}+\sqrt{\operatorname{Log}(mn)}\max_{i,j}|a_{ij}|.\end{aligned}$$

Thus, (53) yields the assertion. □

As in Section 3.3, we estimate $\mathbb{E}\|G_A\|_{p\to q}$ using Propostion 28, so we need to upper bound $\mathbb{E}Y_{k,l}(d_A^\varepsilon)$. The rest of this section is devoted to do so. We begin with an easy consequence of Lemma 39 and Proposition 28.

Corollary 40. *If $2\leq p^*,q<\infty$, then*

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

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

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

and

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

Hence,

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

We choose $\varepsilon=1/\operatorname{Log}(d_A)$ and apply Lemma 39 and Proposition 28. □

To derive better bounds for $\mathbb{E}Y_{k,l}(d_A^\varepsilon)$ we first need to introduce some additional notation. For $I\subset[m]$ and $\alpha\geq 0$ define

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

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

$$\begin{aligned}
&\mathbb{E} Y_{k,l}(d_A^\varepsilon) \\
&\quad \le \mathbb{E} \max_{I \in \mathcal{I}_4(k)} \max_{I_0 \subset I, |I_0|=r} \max_{J \in \mathcal{J}_4(l)} \sup_{s \in K(q^*,m,I,d_A^\varepsilon)} \sup_{t \in K(p,n,J,d_A^\varepsilon)} \sum_{i \in I_0'' \cap I, j \in I_0' \cap J} a_{ij} g_{ij} s_i t_j \\
&\qquad + \mathbb{E} \max_{I \in \mathcal{I}_4(k)} \sup_{s \in B_{q^*}^m} \sup_{t \in B(p,I,\alpha_r)} \sum_{i \in I, j \in [n]} a_{ij} g_{ij} s_i t_j,
\end{aligned} \tag{54}$$

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 $r l_r \le |J| = l$. 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 $a_{ij} \neq 0$, then $i \in I \setminus I_0$. Hence, for any $s \in B_{q^*}^m$ and $t \in B_p^n$,

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

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

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

Therefore,

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

We begin by estimating the second term on the right-hand side of (54).

Lemma 42. *For $p^*, q \in [2, \infty)$, $1 \le k \le m$, and $\alpha \in [0,1]$ we have*

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

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

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

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

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

For $I \in \mathcal{I}_4(k)$ let us choose a $1/2$-net S_I in $B_{q^*}^I = \{s \in B_{q^*}^m \colon \mathrm{supp}(s) \subset I\}$ (in ℓ_{q^*}-norm) of cardinality at most 5^k and put $S := \bigcup_{I\in\mathcal{I}_4(k)} S_I$. Then, by (34), $|S| \le 5^k |\mathcal{I}_4(k)| \le m 20^k d_A^{4k}$, so

$$\begin{aligned}\max_{I\in\mathcal{I}_4(k)} \sup_{s\in B_{q^*}^m} \sup_{t\in B(p,I,\alpha)} \sum_{i\in I, j\in[n]} a_{ij} g_{ij} s_i t_j &\le 2 \max_{I\in\mathcal{I}_4(k)} \sup_{s\in S_I} \sup_{t\in B(p,I,\alpha)} \sum_{i\in I, j\in[n]} a_{ij} g_{ij} s_i t_j \\ &\le 2 \max_{s\in S} \sup_{t\in D(p,s,\alpha)} \sum_{i,j} a_{ij} g_{ij} s_i t_j.\end{aligned}$$

Thus, estimate (21) implies

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

For any fixed $s \in S \subset B_{q^*}^m$ estimate (19) yields

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

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

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

Now we estimate the first term on the right-hand side of (54).

Lemma 43. *If $l \ge r$, $\varepsilon \in (0,1/2]$, and $p^*, q \in [2,4)$, then*

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

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 $I_{0,1}, \ldots, I_{0,V}$ be 4-connected componets of I_0. Then $(I_{0,1}', \ldots, I_{0,V}')$ is a partition of I_0', and $(I_{0,1}'', \ldots, I_{0,V}'')$ is a partition of I_0''. Hence,

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

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 $\gamma\ge 1$, 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

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

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 40 and estimate (42) yield that

$$\begin{aligned}
d_A^{-\varepsilon}\mathbb{E}\|G_A\|_{p\to q}&\lesssim\sup_{x\ge1}(x^{-\varepsilon}\operatorname{Log}x)\Big(D_1+D_2+\sqrt{\operatorname{Log}(mn)}\max_{i,j}|a_{ij}|\Big)\\
&\sim\varepsilon^{-1}\Big(D_1+D_2+\sqrt{\operatorname{Log}(mn)}\max_{i,j}|a_{ij}|\Big),
\end{aligned}$$

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

Lemma 44. *Let $\gamma \geq 1$ and $p^*, q \in [2,4)$.*
i) If $I_0 \subset [m]$ and $1 \leq u \leq n$ satisfy $u \geq |I_0| d_A^{-2}$, 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 $1 \leq r \leq m$ and $1 \leq u \leq n$ satisfy $u \geq r d_A^{-2}$, then

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

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

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

Proof. Let us fix $I_0 \subset I \subset [m]$, $J \subset [n]$ with $|I_0| = r$, $|J \cap I_0'| \geq u$, $s \in K(q^*, m, I, \gamma)$ 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 $\|\bar{s}\|_{q^*} = 1$, $\|\bar{s}\|_\infty \leq \gamma |I_0'' \cap I|^{-1/q^*} \leq \gamma |I_0|^{-1/q^*} \leq \gamma r^{-1/q^*}$, $\|\bar{t}\|_p = 1$, and $\|\bar{t}\|_\infty \leq \gamma |I_0' \cap J|^{-1/p} \leq \gamma u^{-1/p} \leq \gamma d_A^{2/p} r^{-1/p}$. Hence, part i) of the assertion follows by Proposition 35 applied with the matrix $(a_{ij})_{i \in I_0'', j \in I_0'}$, $m = |I_0''| \leq d_A^2 r$, $n = |I_0'| \leq d_A r$, $a = (\gamma r^{-1/q^*}) \wedge 1$ and $b = (\gamma d_A^{2/p} r^{-1/p}) \wedge 1$.

To show part ii) observe that

$$\sum_{i \in I_0'' \cap I, j \in I_0' \cap J} a_{ij}^2 \bar{s}_i^2 \bar{t}_j^2 \leq \max_{i \in I} \sum_{j \in J} a_{ij}^2 \bar{t}_j^2 \leq \max_{i,j} a_{ij}^2 \min\{1, d_A \|\bar{t}\|_\infty^2\} \leq \max_{i,j} a_{ij}^2 \min\{1, d_A \gamma^2 u^{-2/p}\}.$$

Moreover, estimate (34) 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 (21).

Part iii) easily follows from ii) and another application of (21). □

Corollary 45. *If $1 \leq k \leq m$, $1 \leq l \leq n$, $p^*, q \in [2,4)$, and $\varepsilon \in (0, 1/2]$, then*

$$\mathbb{E} Y_{k,l}(d_A^\varepsilon) \lesssim \varepsilon^{-1} \Big[D_1 + D_2 + \Big(\sqrt{\operatorname{Log}(mn)} + \Big(\varepsilon^{-1/2} + \Big(\frac{40}{4 - p^* \vee q} \Big)^{2/(4-p^*\vee q)} \Big) d_A^{1/4+4\varepsilon} \Big) \max_{i,j} |a_{ij}| \Big].$$

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

First assume that $k \geq d_A^{1/2}$. Lemmas 41–43 yield that for every $1 \leq r \leq k$,

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

Moreover, estimate (42) implies

$$(5\operatorname{Log}(d_A))^{2/(4-p^*\vee q)} \leq \Big(\frac{40}{4 - p^* \vee q} \Big)^{2/(4-p^*\vee q)} d_A^{1/4},$$

$$\sqrt{\operatorname{Log} d_A} \lesssim \varepsilon^{-1/2} d_A^\varepsilon, \tag{55}$$

and

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

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

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 (21), (23), and $|\mathcal{I}_4(k)|\le m4^kd_A^{4k}$ (see (34)) yield

$$\mathbb{E}Y_{k,l}(d_A^{\varepsilon})\lesssim\max_{I\in\mathcal{I}_4(k)}\mathbb{E}X_I+\sqrt{\operatorname{Log}|\mathcal{I}_4(k)|}\max_{I\in\mathcal{I}_4(k)}\sup_{\|s\|_{q^*}\le1}\sup_{\|t\|_p\le1}\Big(\sum_{i,j}a_{ij}^2s_i^2t_j^2\Big)^{1/2}.$$

$$\lesssim D_1+(\sqrt{k}+\sqrt{\operatorname{Log}m}+\sqrt{k\operatorname{Log}d_A})\max_{i,j}|a_{ij}|. \tag{56}$$

The assertion follows easily by the assumption that $k<d_A^{1/2}$ and estimate (55). □

Proof of Proposition 34. We apply Proposition 28 and Corollary 45 (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. V. H. de la Peña and E. Giné, *Decoupling. from dependence to independence*, Springer-Verlag, New York, 1999.
6. A. Eskenazis, P. Nayar, and T. Tkocz, *Gaussian mixtures: entropy and geometric inequalities*, Ann. Probab. **46** (2018), no. 5, 2908–2945. MR 3846841
7. 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
8. R. Latała, *Some estimates of norms of random matrices*, Proc. Amer. Math. Soc. **133** (2005), no. 5, 1273–1282. MR 2111932
9. 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
10. ———, *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
11. ———, *Operator $\ell_p\to\ell_q$ norms of random matrices with iid entries*, J. Funct. Anal. **288** (2025), no. 3, Paper No. 110720, 38 pp. MR 4818302
12. R. Latała and W. Świątkowski, *Norms of randomized circulant matrices*, Electron. J. Probab. **27** (2022), Paper No. 80, 23. MR 4441144
13. 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

14. R. Latała, *On the spectral norm of rademacher matrices*, preprint, arXiv:2405.13656, (2024).
15. D. Matlak, *Oszacowania norm macierzy losowych*, Master's thesis, Uniwersytet Warszawski, 2017.
16. K. Ramanan and X. Xie, *The ℓ_r-Levy-Grothendieck problem and $r \to p$ norms of Levy matrices*, preprint, arXiv:2404.18299, (2024).
17. Y. Seginer, *The expected norm of random matrices*, Combin. Probab. Comput. **9** (2000), no. 2, 149–166. MR 1762786
18. R. van Handel, *On the spectral norm of Gaussian random matrices*, Trans. Amer. Math. Soc. **369** (2017), no. 11, 8161–8178. MR 3695857
19. ———, *Structured random matrices*, Convexity and concentration, IMA Vol. Math. Appl., vol. 161, Springer, New York, 2017, pp. 107–156. MR 3837269

University of Warsaw, Institute of Mathematics, Banacha 2, 02–097 Warsaw, Poland.
E-mail address: `rlatala@mimuw.edu.pl`

University of Warsaw, Institute of Mathematics, Banacha 2, 02–097 Warsaw, Poland.
E-mail address: `martast@mimuw.edu.pl`