The Distribution of Non-Commutative Rademacher Series

S.J. Montgomery-Smith*

Department of Mathematics, University of Missouri, Columbia, MO 65211.

Cursed is the ground because of you;
through painful toil you will eat of it
all the days of your life.
Genesis 3:17b (N.I.V.)

ABSTRACT: We give a formula for the tail of the distribution of the non-commutative Rademacher series, which generalizes the result that is already available in the commutative case. As a result, we are able to calculate the norm of these series in many rearrangement invariant spaces, generalizing work of Pisier and Rodin and Semyonov.

1. Introduction

The Rademacher functions are a sequence of independent random variables r_n such that $\Pr(r_n = \pm 1) = \frac{1}{2}$. These functions have played a very important role in mathematics, finding applications in many parts of analysis, as well as other subjects like electronic engineering.

One of the key inequalities concerning the Rademacher functions is due to Khintchine in 1923 **[Kh]**: if a_n is a sequence of scalars, then for $0 < p < \infty$

$$c_p \left(\sum_{n=1}^\infty |a_n|^2 \right)^{1/2} \le \left(\mathbb{E} \left| \sum_{n=1}^\infty a_n r_n \right|^p \right)^{1/p} \le C_p \left(\sum_{n=1}^\infty |a_n|^2 \right)^{1/2} ,$$

where C_p and c_p are constants that depend upon p only. In particular, $C_p \le c\sqrt{p}$ for $p \ge 1$. (Throughout this paper we will not be rigorous with infinite random sums — an expression such as the one above means that the random variable in the middle converges in L_p if the right hand side is finite.)

* Research supported in part by N.S.F. Grant DMS 9001796.

A.M.S. (1991) subject classification: 60-02, 60G50, 43A46, 46E30.

Clearly, one would desire to find generalizations of such an important inequality. For example, one might like to calculate the norm of the Rademacher series $\sum_{n=1}^\infty a_n r_n$ in Orlicz or Lorentz spaces. An obvious result (at least for real scalars) is the following:

$$\left\| \sum_{n=1}^\infty a_n r_n \right\|_\infty = \sum_{n=1}^\infty |a_n| \,.$$

However, another such generalization follows immediately from Khintchine's inequality. For a random variable f and $0 < p < \infty$, let us denote by $\|f\|_{\exp(t^p)}$ the Orlicz norm calculated using the Orlicz function $e^{t^p} - 1$, i.e.

$$\|f\|_{\exp(t^p)} = \inf\bigl\{\, \lambda : \mathbb{E}\bigl(\exp(|f/\lambda|^p)\bigr) \le 2 \,\bigr\}.$$

Then

$$\left\| \sum_{n=1}^\infty a_n r_n \right\|_{\exp(t^p)} \approx \left(\sum_{n=1}^\infty |a_n|^2 \right)^{1/2} ,$$

whenever $p \le 2$. (Here, as in the rest of the paper, the expression $A \approx B$ means that $c^{-1} A \le B \le cA$ for some constant c.)

In 1975, Rodin and Semyonov **[R–S]** considered the value of the Rademacher series $\sum_{n=1}^\infty a_n r_n$ in other rearrangement invariant spaces. In particular, they showed that

$$\left\| \sum_{n=1}^\infty a_n r_n \right\|_{\exp(t^p)} \approx \|(a_n)\|_{q,\infty} ,$$

(see also **[P1]**), and that

$$\left\| \sum_{n=1}^\infty a_n r_n \right\|_{\exp(t^p),r} \approx \|(a_n)\|_{q,r} ,$$

whenever $p > 2$, $\frac{1}{p} + \frac{1}{q} = 1$ and $0 < r < \infty$. Here

$$\|f\|_{\exp(t^p),r} = \left(\int_0^1 \bigl(\log(1/t)\bigr)^{(r/p)-1} \bigl(f^*(t)\bigr)^r \,\frac{dt}{t} \right)^{1/r} ,$$

$$\|(a_n)\|_{q,r} = \begin{cases} \bigl(\sum_{n=1}^\infty n^{(r/q)-1} a_n^*\bigr)^{1/r} & \text{if } 0 < r < \infty \\ \sup_{n\ge 1} n^{1/p} a_n^* & \text{if } r = \infty, \end{cases}$$

where f^* and a_n^* are the non-increasing rearrangements of $|f|$ and $|a_n|$ respectively.

In fact they were able to show that if X is any symmetric sequence space with Boyd indices strictly between 1 and 2, then there exists a rearrangement invariant space Y on probability space such that

$$\left\| \sum_{n=1}^{\infty} a_n r_n \right\|_Y \approx \|(a_n)\|_X .$$

There still remained the question of finding tail distributions of Rademacher series, that is, to find

$$\Pr\left(\left| \sum_{n=1}^{\infty} a_n r_n \right| > t \right)$$

for every $t > 0$. This was answered in **[Mo]** as follows. Given a sequence $a = (a_n)$, we will define its K-functional with respect to ℓ_1 and ℓ_2 to be

$$K_{1,2}(a,t) = \inf \{ \|a'\|_1 + t \|a''\|_2 : a' + a'' = a \} .$$

These quantities play an important role in the theory of interpolation of spaces (see **[B–S]** or **[B–L]**). They are not so hard to calculate, since there is the following formula due to Holmstedt **[Ho]**:

$$K_{1,2}(a,t) \approx \sum_{n=1}^{[t^2]} a_n^* + t \left(\sum_{n=[t^2]+1}^{\infty} (a_n^*)^2 \right)^{1/2} .$$

Then we have the following results.

$$\Pr\left(\left| \sum_{n=1}^{\infty} a_n r_n \right| > c\, K_{1,2}(t,a) \right) \le c\, e^{-c^{-1} t^2},$$

$$\Pr\left(\left| \sum_{n=1}^{\infty} a_n r_n \right| > c^{-1}\, K_{1,2}(t,a) \right) \ge c^{-1} e^{-ct^2}.$$

(Here, as in the rest of the paper, the letter c denotes a positive constant that changes with each occurrence.) We remark that the hard part of this result, the lower bound, can be deduced from a more general result contained in the book by Ledoux and Talagrand **[L–T]**, namely Theorem 4.15. They obtain a Rademacher version of Sudakov's Theorem.

From this formula, and using known facts about the Hardy operators, it is possible to reproduce all of the results of Rodin and Semyonov. It is interesting to note that in order to obtain the lower bounds of Rodin and Semyonov, one only requires the following estimate to be true:

$$\Pr\left(\left| \sum_{n=1}^{\infty} a_n r_n \right| > c^{-1} \sum_{n=1}^{[t^2]} a_n^* \right) \ge c^{-1} e^{-ct^2}.$$

This bound has an extremely simple proof: simply consider the random event $\{r_{n_k} = \operatorname{sign}(a_{n_k})\}$ for an appropriate sequence $n_1, n_2, \ldots, n_{[t^2]}$.

We might also add that a consequence of the above result is the following. If $t \le c \|a\|_2 / \|a\|_\infty$, then

$$\Pr\left(\left| \sum_{n=1}^\infty a_n r_n \right| > c^{-1} t \|a\|_2 \right) \ge c^{-1} e^{-ct^2}.$$

This result can also be deduced from a result of Kolmogorov [**Ko**] (see also [**L–T**] Chapter 4).

Recently, Hitczenko [**Hi**] used the distribution formula to obtain an asymptotically more accurate version of Khintchine's original inequalities. He showed that

$$\left(\mathbb{E} \left| \sum_{n=1}^\infty a_n r_n \right|^p \right)^{1/p} \approx K_{1,2}(a, \sqrt{p}),$$

for $p \ge 1$, where the constants of approximation do not depend upon p.

It has also been discovered that many of these results have vector valued analogues (see [**D–M**]).

2. The Non-Commutative Rademacher Series

Now we get to the main subject of this paper. Non-commutative Rademacher series arise in a natural way when one considers Fourier series on non-commutative compact groups. For example, a Sidon series on a non-commutative compact group has a distribution equivalent to a non-commutative Rademacher series (see [**F–R**], [**H–R**] and [**A–M**]). They are also the natural things to consider if one wishes to work with random Fourier series on non-commutative compact groups (see [**M–P**]).

They were considered by Figà-Talamanca and Rider in [**F–R**], where they showed the non-commutative analogue of the Khintchine inequalities (see also [**H–R**]). Many results about them are also given in [**M–P**].

We let M_d denote the vector space of d-dimensional matrices (i.e. $d \times d$ matrices), and we let O_d denote the multiplicative subgroup of orthogonal matrices. Let d_n be a sequence of positive integers, let A_n be a d_n-dimensional matrix, and let ϵ_n be a sequence of independent random variables such that ϵ_n takes values in O_{d_n} uniformly distributed with respect to the Haar measure. If A is a d-dimensional matrix, we denote by $\operatorname{tr}(A)$ the trace of A, that is, the sum of the diagonal entries of A.

Then a non-commutative Rademacher series is a random variable of the following form:

$$S_\epsilon = \sum_{n=1}^\infty d_n \operatorname{tr}(\epsilon_n A_n).$$

If A is a d-dimensional matrix, we define the singular values to be the eigenvalues of $\sqrt{A^*A}$, where A^* is the transpose of A. We define the Schatten norms on M_d as follows: if $A \in M_d$, set $\|A\|_p$ equal to the usual ℓ_p sequence norm of the singular values of A. Thus $\|A\|_\infty$ is the usual operator norm of A on d-dimensional Hilbert space, $\|A\|_2$ is the Hilbert–Schmidt norm of A, and $\|A\|_1$ is the trace class norm of A.

The result of Figà-Talamanca and Rider **[F–R]** is the following:

$$c_p \left(\sum_{n=1}^\infty d_n \, \|A_n\|_2^2 \right)^{1/2} \le \left\| \sum_{n=1}^\infty d_n \ \mathrm{tr}(\epsilon_n A_n) \right\|_p \le C_p \left(\sum_{n=1}^\infty d_n \, \|A_n\|_2^2 \right)^{1/2} ,$$

for $0 < p < \infty$. Here $C_p \le c\sqrt{p}$ for $p \ge 1$. From this one can obtain the result

$$\left\| \sum_{n=1}^\infty d_n \ \mathrm{tr}(\epsilon_n A_n) \right\|_{\exp(t^2)} \approx \left(\sum_{n=1}^\infty d_n \, \|A_n\|_2^2 \right)^2 .$$

It is also true that

$$\left\| \sum_{n=1}^\infty d_n \ \mathrm{tr}(\epsilon_n A_n) \right\|_\infty = \sum_{n=1}^\infty d_n \, \|A_n\|_1 .$$

Let s denote the vector formed in the following manner. First list the singular values of A_n, repeating each singular value d_n times. Combine these into one long list, rearranging them into decreasing order. Then the above results can be written in the more suggestive form:

$$\|S_\epsilon\|_p \approx \|S_\epsilon\|_{\exp(t^2)} \approx \|s\|_2 \quad (0 < p < \infty) \qquad \text{and} \qquad \|S_\epsilon\|_\infty \approx \|s\|_1 .$$

Pisier **[P1]** was able to obtain partial non-commutative versions of the results of Rodin and Semyonov. He showed that

$$\|S_\epsilon\|_{\exp(t^p)} \le c \, \|s\|_{q,\infty} ,$$

where $p > 2$ and $\frac{1}{p} + \frac{1}{q} = 1$. He was not able to obtain the lower bound.

The purpose of this paper is to show that all of these results for the commutative Rademacher series also apply to the non-commutative case. The main result is the following formulae for the distribution of the non-commutative Rademacher series.

THEOREM 2.1. *The distribution of S_ϵ is given by the following formulae.*

i) $$\Pr(S_\epsilon > c\, K_{1,2}(t,s)) \le c\, e^{-c^{-1} t^2} .$$

ii) $$\Pr(S_\epsilon > c^{-1}\, K_{1,2}(t,s)) \ge c^{-1} e^{-ct^2} .$$

COROLLARY 2.2. *We have the following for all $t > 0$:*

$$c^{-1}\,\mathrm{Pr}\left(\sum_{n=1}^{\infty} s_n r_n > ct\right) \le \mathrm{Pr}\left(\sum_{n=1}^{\infty} d_n\ \mathrm{tr}(\epsilon_n A_n) > t\right) \le c\ \mathrm{Pr}\left(\sum_{n=1}^{\infty} s_n r_n > c^{-1} t\right).$$

Now we are able to obtain the following results immediately from the commutative case.

COROLLARY 2.3. *We have the following inequalities.*

i) $\|S_\epsilon\|_{\exp(t^p)} \approx \|s\|_{q,\infty}$ *for* $p > 2$ *and* $\frac{1}{p} + \frac{1}{q} = 1$;

ii) $\|S_\epsilon\|_{\exp(t^p),r} \approx \|s\|_{q,r}$ *for* $p > 2$, $\frac{1}{p} + \frac{1}{q} = 1$ *and* $0 < r < \infty$;

iii) $\|S_\epsilon\|_p \approx K_{1,2}(s, \sqrt{p})$ *for* $1 \le p < \infty$ *with with constants of approximation independent of* p.

The proof of Theorem 2.1 is split into two halves. In the first half we make a number of changes to the problem, by showing that the problem is equivalent to a similar result involving Gaussian matrices.

The second half contains the meat of the argument. Part (i) of Theorem 2.1 is essentially the result of Figà-Talamanca and Rider combined with a fairly straightforward interpolation argument. It is part (ii) that provides the difficulties. The argument proceeds by considering four cases according to the nature of the sequence s.

The arguments used in this paper assume that the matrices A_n are real valued, but it is very easy to extend the results to the complex case as well. Instead of using the non-commutative Rademacher functions, one should use the non-commutative Steinhaus random variables, that is, ξ_n, where ξ_n is uniformly distributed over the d_n-dimensional unitary matrices with respect to Haar measure. Then comparison results from **[M–P]** combined with Lemma 3.9 below will give the results.

This is probably a hard paper to read. It certainly was a hard paper to write. As it says in Genesis 3:17, we eat of the ground through painful toil.

3. The Proof of Theorem 2.1 — Part I

The first observation is that we may assume that all the matrices are diagonal with entries from the non-negative reals. This follows because A_n may be factored $A_n = U_n D_n V_n$, where U_n and V_n are elements of O_{d_n}, and D_n is diagonal with entries from the non-negative reals. But $\operatorname{tr}(\epsilon_n U_n D_n V_n) = \operatorname{tr}(V_n \epsilon_n U_n D_n)$, and $V_n \epsilon_n U_n$ has the same law as ϵ_n.

Thus we will assume that

$$A_n = \begin{pmatrix} a_1^n & 0 & 0 & \cdots & 0 \\ 0 & a_2^n & 0 & \cdots & 0 \\ 0 & 0 & a_3^n & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & a_{d_n}^n \end{pmatrix},$$

where a_1^n, $a_2^n, \dots, a_{d_n}^n$ are the singular values of A_n.

Let G_n be the matrix

$$G_n = \frac{1}{\sqrt{d_n}} \begin{pmatrix} g_{1,1}^n & g_{1,2}^n & g_{1,3}^n & \cdots & g_{1,d_n}^n \\ g_{2,1}^n & g_{2,2}^n & g_{2,3}^n & \cdots & g_{2,d_n}^n \\ g_{3,1}^n & g_{3,2}^n & g_{3,3}^n & \cdots & g_{3,d_n}^n \\ \vdots & \vdots & \vdots & & \vdots \\ g_{d_n,1}^n & g_{d_n,2}^n & g_{d_n,3}^n & \cdots & g_{d_n,d_n}^n \end{pmatrix},$$

where $(g_{i,j}^n)$ is a sequence of independent Gaussian random variables of mean 0 and variance 1. We would like to compare S_ϵ with the random variable

$$S_G = \sum_{n=1}^\infty d_n \operatorname{tr}(G_n A_n).$$

This random variable is particularly easy to understand — it is simply a Gaussian random variable:

$$S_G = \sum_{n=1}^\infty \sqrt{d_n} \sum_{i=1}^{d_n} g_{i,i}^n a_i^n .$$

Unfortunately, this random variable is too large to give us the lower bounds required for Theorem 2.1 part (ii). To get around this problem, we split G_n as follows:

$$G_n = G_n' + G_n'',$$

where

$$G_n' = G_n \chi_{\|G_n\|_\infty \le \lambda} \quad \text{and} \quad G_n'' = G_n \chi_{\|G_n\|_\infty > \lambda}.$$

Here λ is a universal constant. For the proof to work, λ needs to be sufficiently large. As we proceed, we will make it clear where the restrictions on λ are required.

We are also going to introduce the following random variables. We let G_n^{d} denote the d_n-dimensional matrix consisting only of the diagonal entries of G_n, and we let $G_n^{\mathrm{ad}} = G_n - G_n^{\mathrm{d}}$ be the d_n-dimensional matrix consisting of the off-diagonal entries of G_n. We can also split G_n in the following manner:

$$G_n = G_n^* + G_n^{**},$$

where

$$G_n^* = G_n(\chi_{\|G_n^{\mathrm{d}}\|_\infty \le \lambda})(\chi_{\|G_n^{\mathrm{ad}}\|_\infty \le \lambda}) \quad \text{and} \quad G_n^{**} = G_n \chi_{\|G_n^{\mathrm{d}}\|_\infty \vee \|G_n^{\mathrm{ad}}\|_\infty > \lambda}.$$

The strategy will be to compare S_ϵ with the random variables

$$S_{G'} = \sum_{n=1}^\infty d_n \ \mathrm{tr}(G'_n A_n) = \sum_{n=1}^\infty d_n \ \mathrm{tr}(G_n A_n)\chi_{\|G_n\|_\infty \le \lambda},$$
$$S_{G^*} = \sum_{n=1}^\infty d_n \ \mathrm{tr}(G^*_n A_n) = \sum_{n=1}^\infty d_n \ \mathrm{tr}(G_n A_n)\chi_{\|G_n^{\mathrm{d}}\|_\infty \le \lambda}\chi_{\|G_n^{\mathrm{ad}}\|_\infty \le \lambda}.$$

Now let us present the results that we will be requiring. Note that we denote the commutative Rademacher functions by r_n, so that they will not be confused with the non-commutative Rademacher functions ϵ_n.

The first pair of results we present are comparison principles. Let us suppose that V_n is a sequence of random variables taking values in M_{d_n} for which the sequence $(a_n V_n)$ has the same law as (V_n) for any $a_n \in O_{d_n}$. We note that the random variables ϵ_n, G_n, G'_n and G''_n all have this property. We also suppose that x_n is a sequence of d_n-dimensional matrices taking values in a Banach space B. The first result is selected parts from **[M–P]**, Proposition V.2.1.

Lemma 3.1. *Suppose that T_n is any element of M_{d_n}. Then for $1 \le p < \infty$ we have that*

$$\left(\mathbb{E}\left\|\sum_{n=1}^\infty \mathrm{tr}(T_n V_n x_n)\right\|^p\right)^{1/p} \le \sup_{n\ge 1}\|T_n\|_\infty \left(\mathbb{E}\left\|\sum_{n=1}^\infty \mathrm{tr}(V_n x_n)\right\|^p\right)^{1/p}.$$

We also have the following.

$$\left(\mathbb{E}\left\|\sum_{n=1}^\infty \mathrm{tr}(\epsilon_n x_n)\right\|^p\right)^{1/p} \le \sup_{n\ge 1}\left\|(\mathbb{E}\,|V_n|)^{-1}\right\|_\infty \left(\mathbb{E}\left\|\sum_{n=1}^\infty \mathrm{tr}(V_n x_n)\right\|^p\right)^{1/p}.$$

The next lemma is simply the commutative version of the same result, and may be found in **[M–P]**, Theorem 4.9.

Lemma 3.2. *Suppose that v_n is a sequence of independent, real valued, symmetric random variables, and that x_n is a sequence of values from a Banach space B. Then for all $1 \le p < \infty$ we have*

$$\left(\mathbb{E} \left\| \sum_{n=1}^{\infty} x_n r_n \right\|^p \right)^{1/p} \le \sup_{n \ge 1} (\mathbb{E}\,|v_n|)^{-1} \left(\mathbb{E} \left\| \sum_{n=1}^{\infty} x_n v_n \right\|^p \right)^{1/p} .$$

Next we present a couple of reflection principles. The first will enable us to remove some of the elements of s. Let us suppose that V_n is random variable taking its values in M_{d_n} such that the sequence $(D_n V_n)$ has the same law as (V_n) for any sequence of diagonal matrices D_n whose diagonal entries are ± 1. Notice that all the random variables we have introduced have this property: ϵ_n, G_n, G'_n, G''_n, G^*_n and G^{**}_n. Recall that we have supposed that the matrices A_n are diagonal.

Lemma 3.3. *Suppose that A'_n is a sequence of diagonal matrices such that for each $n \ge 1$, each entry of A'_n is either the same as the corresponding entry as A_n or it is 0. Then for all $t > 0$ we have that*

$$\Pr\left(\sum_{n=1}^{\infty} d_n \ \mathrm{tr}(V_n A'_n) > t \right) \le 2 \ \Pr\left(\sum_{n=1}^{\infty} d_n \ \mathrm{tr}(V_n A_n) > t \right) .$$

Proof: Notice that the random variables

$$\sum_{n=1}^{\infty} d_n \ \mathrm{tr}(V_n A_n) \quad \text{and} \quad \sum_{n=1}^{\infty} d_n \ \mathrm{tr}(V_n (2A'_n - A_n))$$

have the same law. Thus

$$\begin{aligned} \Pr\left(\sum_{n=1}^{\infty} d_n \ \mathrm{tr}(V_n A'_n) > t \right) &\le \Pr\left(\sum_{n=1}^{\infty} d_n \ \mathrm{tr}(V_n A_n) > t \right) + \Pr\left(\sum_{n=1}^{\infty} d_n \ \mathrm{tr}(V_n (2A'_n - A_n)) > t \right) \\ &= 2 \ \Pr\left(\sum_{n=1}^{\infty} d_n \ \mathrm{tr}(V_n A_n) > t \right) , \end{aligned}$$

as required. ∎

The next lemma is simply the commutative version of the above result, and is essentially the same as **[Ka]** Chapter 2 Theorem 5.

LEMMA 3.4. *Let x_n be any sequence of elements from a Banach space B, and let α_n be a sequence of values taking only the values 0 or 1. Then for all $t > 0$ we have*

$$\Pr\left(\left\|\sum_{n=1}^{\infty} \alpha_n x_n r_n\right\| > t\right) \le 2 \Pr\left(\left\|\sum_{n=1}^{\infty} x_n r_n\right\| > t\right).$$

Now we present results concerning the behavior of the non-commutative Gaussian random variables.

LEMMA 3.5. *For sufficiently large λ, the following is true.*

$$\mathbb{E}\left\|G_n\right\|_\infty \approx \mathbb{E}\left\|G_n'\right\|_\infty \approx 1.$$

Proof: The statement $\mathbb{E}\left\|G_n\right\|_\infty \approx 1$ is given in **[M–P]**, Proposition 1.5. That $\mathbb{E}\left\|G_n'\right\|_\infty \approx 1$ for sufficiently large λ then follows by the monotone convergence theorem. □

Let Id_n denote the d_n-dimensional identity matrix.

LEMMA 3.6. *There exists constants c_n and c_n' that are uniformly bounded above and below such that for sufficiently large λ we have that*

$$\mathbb{E}\left|G_n\right| = c_n \mathrm{Id}_n \quad \textit{and} \quad \mathbb{E}\left|G_n'\right| = c_n' \mathrm{Id}_n.$$

Proof: The first statement is from **[M–P]**, Corollary 1.8. The second statement has entirely the same proof. □

The next lemma uses a result of C. Borell **[Bo]** (see also **[P2]** or **[L–T]**).

THEOREM 3.7. *Let X be a mean 0 Gaussian random variable taking values in a Banach space B. Let*

$$\sigma = \sup_{\|\phi\|_{B^*} \le 1} \left(\mathbb{E}\left|\phi(X)\right|^2\right)^{1/2}.$$

Then for all $t > 0$ we have

$$\Pr\left(\left|\|X\| - \mathbb{E}\|X\|\right| \ge t\sigma\right) \le c\, e^{-c^{-1}t^2}.$$

LEMMA 3.8. *For t larger than some universal constant, we have that*

$$\Pr\left(\left\|G_n\right\|_\infty > t\right) \le c\, e^{-c^{-1} d_n t^2}.$$

Proof: Since $\mathbb{E}\|G_n\|_\infty \approx 1$, by Theorem 3.7, it is sufficient to show that if $\|A\|_1 \le 1$, then

$$\left(\mathbb{E}\left|\operatorname{tr}(A^t G_n)\right|^2\right)^{1/2} \le \frac{1}{\sqrt{d_n}}.$$

But

$$\operatorname{tr}(A^t G_n) = \frac{1}{\sqrt{d_n}} \sum_{i=1}^{d_n} \sum_{j=1}^{d_n} g_{i,j}^n a_{i,j}^n,$$

which is a Gaussian variable of variance

$$\frac{1}{d_n} \sum_{i=1}^{d_n} \sum_{j=1}^{d_n} (a_{i,j}^n)^2 = \frac{1}{d_n} \|A\|_2^2 \le \frac{1}{d_n} \|A\|_1^2 .$$

□

Now we present a principle from **[dP–M]** (see also **[A–M]**) that allows us to obtain results about distributions from L_p norm results.

Lemma 3.9. *Let X and Y be two random variables taking values in the positive reals such that the following holds. Whenever X_m and Y_m are independent random variables with the same law as X and Y respectively, for all $M \in \mathbb{N}$ we have that*

i)
$$\mathbb{E} \sup_{1\le m\le M} X_m \le c\, \mathbb{E} \sup_{1\le m\le M} Y_m,$$

ii)
$$\left(\mathbb{E} \sup_{1\le m\le M} Y_m^2\right)^{1/2} \le c\, \mathbb{E} \sup_{1\le m\le M} Y_m.$$

Then it follows that for all $t > 0$ that

$$\Pr(X > t) \le c \Pr(Y > c^{-1} t).$$

We can use this to prove a distributional comparison principle.

Lemma 3.10. *Suppose that v_n is a sequence of real valued symmetric independent random variables, such that for any sequence of vectors x_n from a Banach space B*

$$\left(\mathbb{E}\left\|\sum_{n=1}^\infty x_n v_n\right\|^2\right)^{1/2} \le c\, \mathbb{E}\left\|\sum_{n=1}^\infty x_n v_n\right\|.$$

Suppose also that

$$\mathbb{E}|v_n| \ge c^{-1}.$$

Then for any sequence of scalars a_n and for all $t > 0$ we have

$$\Pr\left(\left|\sum_{n=1}^{\infty} a_n r_n\right| \ge t\right) \le c \Pr\left(\left|\sum_{n=1}^{\infty} a_n v_n\right| \ge c^{-1} t\right).$$

Proof: Let us set $X = |\sum_{n=1}^{\infty} a_n r_n|$ and $Y = |\sum_{n=1}^{\infty} a_n v_n|$. Let $r_{n,m}$ be independent copies of r_n and $v_{n,m}$ be independent copies of v_n for $1 \le m \le M$, and let $x_{n,m} \in \ell_\infty^M$ be defined by

$$x_{n,m} = (0, 0, \dots, a_n, \dots, 0) \qquad \text{(the } a_n \text{ is in the } m\text{th position).}$$

Then notice that

$$\sup_{1 \le m \le M} X_m = \left\| \sum_{m=1}^{M} \sum_{n=1}^{\infty} r_{n,m} x_{n,m} \right\|_{\ell_\infty^M},$$

$$\sup_{1 \le m \le M} Y_m = \left\| \sum_{m=1}^{M} \sum_{n=1}^{\infty} d_n \, v_{n,m} x_{n,m} \right\|_{\ell_\infty^M}.$$

From Lemma 3.2, it then follows that for $p = 1$, 2 that

$$\left(\mathbb{E} \sup_{1 \le m \le M} X_m^p\right)^{1/p} \le c \left(\mathbb{E} \sup_{1 \le m \le M} Y_m^p\right)^{1/p},$$

and by hypothesis we have that

$$\left(\mathbb{E} \sup_{1 \le m \le M} Y_m^2\right)^{1/2} \le c \, \mathbb{E} \sup_{1 \le m \le M} Y_m.$$

Thus we may apply Lemma 3.9 and the result follows. □

The following lemma is an immediate corollary of **[M–P]**, Theorem V.2.7.

LEMMA 3.11. *If x_n is a sequence of d_n-dimensional matrices with entries in a Banach space B, then*

$$\mathbb{E} \left\| \sum_{n=1}^{\infty} d_n \operatorname{tr}(\epsilon_n x_n) \right\| \approx \left(\mathbb{E} \left\| \sum_{n=1}^{\infty} d_n \operatorname{tr}(\epsilon_n x_n) \right\|^2 \right)^{1/2}.$$

Now we are ready to proceed with the main part of this section. We are going to use these results to show the following.

LEMMA 3.12. *For sufficiently large λ, the following holds for all $t > 0$:*

$$c^{-1}\Pr(S_{G'} > ct) \le \Pr(S_\epsilon > t) \le c\,\Pr(S_{G'} > c^{-1}t),$$
$$c^{-1}\Pr(S_{G^*} > ct) \le \Pr(S_\epsilon > t) \le c\,\Pr(S_{G^*} > c^{-1}t).$$

First we will show the L_p-norm version of this result.

LEMMA 3.13. *For sufficiently large λ, and any $1 \le p < \infty$ we have that*

$$\left(\mathbb{E}\left\|\sum_{n=1}^\infty \operatorname{tr}(\epsilon_n x_n)\right\|^p\right)^{1/p} \approx \left(\mathbb{E}\left\|\sum_{n=1}^\infty \operatorname{tr}(G'_n x_n)\right\|^p\right)^{1/p}.$$

Proof: To show the left hand side is bounded by a constant times the right hand side is easy. We apply the second part of Lemma 3.1 with $V_n = G'_n$, using Lemma 3.6.

Next we show that the right hand side is bounded by a constant times the left hand side. Let us suppose that the random variables G_n are independent of the random variables ϵ_n. Without loss of generality, we may suppose that measure space upon which the random variables exist is a product measure of Ω_ϵ and Ω_G, and that the random variables ϵ_n depend only upon the Ω_ϵ co-ordinate, and that the random variables G_n depend only upon the Ω_G co-ordinate. Let us denote integration with respect to the Ω_ϵ co-ordinate by $\mathbb{E}_\epsilon$ and integration with respect to the Ω_G co-ordinate by $\mathbb{E}_G$.

For each $\omega_G \in \Omega_G$, by Lemma 3.1, we have that

$$\left(\mathbb{E}_\epsilon\left\|\sum_{n=1}^\infty \operatorname{tr}(G'_n(\omega_G)\epsilon_n x_n)\right\|^p\right)^{1/p} \le \sup_{n\ge1}\|G'_n(\omega_G)\|_\infty \left(\mathbb{E}_\epsilon\left\|\sum_{n=1}^\infty \operatorname{tr}(\epsilon_n x_n)\right\|^p\right)^{1/p}.$$

Now we take L_p norms of both sides with respect with respect to the Ω_G co-ordinate. We note that the sequence $(G'_n\epsilon_n)$ has the same joint law as (G'_n), and that $\|G'_n\|_\infty \le \lambda$. Hence we obtain that

$$\left(\mathbb{E}\left\|\sum_{n=1}^\infty \operatorname{tr}(G'_n x_n)\right\|^p\right)^{1/p} \le \lambda\left(\mathbb{E}\left\|\sum_{n=1}^\infty \operatorname{tr}(\epsilon_n x_n)\right\|^p\right)^{1/p},$$

as desired. □

Now we need to be able to compare $S_{G'}$ with S_{G^*}.

LEMMA 3.14. *Let A be a d-dimensional matrix, let A^{d} be the matrix taking only the diagonal entries from A, and let A^{ad} be the matrix taking only the non-diagonal entries from A, so that $A = A^{\mathrm{d}} + A^{\mathrm{ad}}$. Then*

$$\frac{1}{2}\max\{\left\|A^{\mathrm{d}}\right\|_\infty, \left\|A^{\mathrm{ad}}\right\|_\infty\} \le \|A\|_\infty \le 2\max\{\left\|A^{\mathrm{d}}\right\|_\infty, \left\|A^{\mathrm{ad}}\right\|_\infty\}.$$

Proof: The right hand inequality follows immediately from the triangle inequality. To show the left hand inequality, note that

$$\left\|A^{\mathrm{d}}\right\|_\infty = \sup_{1\le i\le d} |a_{i,i}| \le \|A\|_\infty .$$

Finally,

$$\left\|A^{\mathrm{ad}}\right\|_\infty \le \|A\|_\infty + \left\|A^{\mathrm{d}}\right\|_\infty \le 2\|A\|_\infty ,$$

as required. □

LEMMA 3.15. *For all $t > 0$ we have*

$$\frac{1}{2}\Pr\left(\left\|\sum_{n=1}^\infty \mathrm{tr}(G_n x_n)\chi_{\|G_n\|_\infty\le\lambda/2} r_n\right\| > t\right) \le \Pr\left(\left\|\sum_{n=1}^\infty \mathrm{tr}(\epsilon_n x_n)\right\| > t\right)$$
$$\le 2\Pr\left(\left\|\sum_{n=1}^\infty \mathrm{tr}(G_n x_n)\chi_{\|G_n\|_\infty\le 2\lambda} r_n\right\| > t\right).$$

Proof: As with the proof of Lemma 3.13, we suppose that the random variables G_n are independent of the random variables r_n. We suppose that measure space upon which the random variables exist is a product measure of Ω_r and Ω_G, and that the random variables r_n depend only upon the Ω_r co-ordinate, and that the random variables G_n depend only upon the Ω_G co-ordinate. Let us denote measure with respect to the Ω_r co-ordinate by $\Pr_r$.

By Lemma 3.14, the numbers

$$\frac{\chi_{\|G_n^{\mathrm{d}}\|_\infty\le\lambda}\chi_{\|G_n^{\mathrm{ad}}\|_\infty\le\lambda}}{\chi_{\|G_n\|_\infty\le 2\lambda}} \quad\text{and}\quad \frac{\chi_{\|G_n\|_\infty\le\lambda/2}}{\chi_{\|G_n^{\mathrm{d}}\|_\infty\le\lambda}\chi_{\|G_n^{\mathrm{ad}}\|_\infty\le\lambda}}$$

take the values 0 or 1. Thus, by Lemma 3.4, it follows that for each $\omega_G \in \Omega_G$ and all $t > 0$ that

$$\begin{aligned}
&\frac{1}{2}\ \mathrm{Pr}_r \left(\left\| \sum_{n=1}^\infty \mathrm{tr}(G_n(\omega_G) x_n)(\chi_{\|G_n(\omega_G)\|_\infty \le \lambda/2}) r_n \right\| > t \right) \\
&\le c\ \mathrm{Pr}_r \left(\left\| \sum_{n=1}^\infty \mathrm{tr}(G_n^*(\omega_G) x_n) r_n \right\| > t \right) \\
&\le 2\ \mathrm{Pr}_r \left(\left\| \sum_{n=1}^\infty \mathrm{tr}(G_n(\omega_G) x_n)(\chi_{\|G_n(\omega_G)\|_\infty \le 2\lambda}) r_n \right\| > t \right).
\end{aligned}$$

Now, taking expectations on both sides with respect to Ω_G, the result follows. ☐

Now we are ready to combine these results.

Proof of Lemma 3.12: We first prove the first inequality. In order to apply Lemma 3.9, let us set $X = |S_{G'}|$ and $Y = |S_\epsilon|$. Let $\epsilon_{n,m}$ be independent copies of ϵ_n and $G_{n,m}$ be independent copies of G_n for $1 \le m \le M$, and let $x_{n,m}$ be diagonal matrices with diagonal entries in ℓ_∞^M:

$$x_i^{n,m} = (0, 0, \dots, a_i^n, \dots, 0) \qquad \text{(the } a_i^n \text{ is in the } m\text{th position)}.$$

Then notice that

$$\begin{aligned}
\sup_{1 \le m \le M} X_m &= \left\| \sum_{m=1}^M \sum_{n=1}^\infty d_n\ \mathrm{tr}(G'_{n,m} x_{n,m}) \right\|_{\ell_\infty^M}, \\
\sup_{1 \le m \le M} Y_m &= \left\| \sum_{m=1}^M \sum_{n=1}^\infty d_n\ \mathrm{tr}(\epsilon_{n,m} x_{n,m}) \right\|_{\ell_\infty^M}.
\end{aligned}$$

From Lemma 3.13, it then follows that for $p = 1$, 2 that

$$\left(\mathbb{E} \sup_{1 \le m \le M} X_m^p \right)^{1/p} \approx \left(\mathbb{E} \sup_{1 \le m \le M} Y_m^p \right)^{1/p},$$

and from Lemma 3.11, we have that

$$\left(\mathbb{E} \sup_{1 \le m \le M} Y_m^2 \right)^{1/2} \approx \mathbb{E} \sup_{1 \le m \le M} Y_m.$$

Thus, we also have that

$$\left(\mathbb{E} \sup_{1 \le m \le M} X_m^2 \right)^{1/2} \approx \mathbb{E} \sup_{1 \le m \le M} X_m.$$

Thus we may apply Lemma 3.9 twice, once with the roles of X_m and Y_m reversed, and the result follows.

The second inequality now follows from Lemma 3.15. ☐

4. The Proof of Theorem 2.1 — Part II

We will first show the first half of Theorem 2.1.

Proposition 4.1. *The following is true for all $t > 0$.*

$$\Pr(S_\epsilon > c\,K_{1,2}(t,s)) \le c\,e^{-c^{-1}t^2}.$$

Proof: Choose sequences s' and s'' such that $s = s' + s''$ and

$$K_{1,2}(t,s) \ge \tfrac{1}{2}(\|s'\|_1 + t\,\|s''\|_2).$$

We may assume that if a certain number occurs several times in the sequence s, then for each occurrence this number is split identically between s' and s''. Thus we may know that there exist sequences of matrices (A'_n) and (A''_n) such that $A_n = A'_n + A''_n$ and such s' comes from repeating d_n times the singular values of A'_n, and s'' comes from repeating d_n times the singular values of A''_n.

From the result of Figà-Talamanca and Rider, we know that

$$\Pr\left(\sum_{n=1}^\infty d_n \operatorname{tr}(\epsilon_n A''_n) > ct\,\|s''\|_2\right) \le \frac{\sqrt{p^p}}{t^p} \le c\,e^{-c^{-1}t^2}.$$

(Here we chose $p = t^2/2$). It is also clearly evident that

$$\sum_{n=1}^\infty d_n \operatorname{tr}(\epsilon_n A'_n) \le \sum_{n=1}^\infty d_n\ \|A'_n\|_1 = \|s'\|_1 .$$

Thus

$$\begin{aligned}
\Pr(S_\epsilon > 2c\,K_{1,2}(t,s)) &\le \Pr\left(\sum_{n=1}^\infty d_n \operatorname{tr}(\epsilon_n A'_n) + \sum_{n=1}^\infty d_n \operatorname{tr}(\epsilon_n A''_n) > c(\|s'\|_1 + t\,\|s''\|_2)\right)\\
&\le \Pr\left(\sum_{n=1}^\infty d_n \operatorname{tr}(\epsilon_n A''_n) > ct\,\|s''\|_2\right)\\
&\le c\,e^{-c^{-1}t^2},
\end{aligned}$$

and the result follows. □

Now we finally come to the hard part of this paper: to show the second part of Theorem 2.1. We will proceed by considering three cases. All of the arguments will make heavy use of the approximation

$$c^{-1}e^{-ct^2} \leq \Pr(g > t) \leq ce^{-c^{-1}t^2} \qquad (t > 0),$$

whenever g is a Gaussian random variable of mean 0 and variance 1.

Our use of the letter c becomes confusing at this point. Thus from now on we will use subscripts on the letter c to denote different values. However, the same subscripted letter c may take different values from result to result and proof to proof.

The first case will be dealt with by the following result.

PROPOSITION 4.2. *For sufficiently large λ, the exist numbers c_1 and c_2 such that for all integers $t \geq 1$.*

$$\Pr(S_{G^*} > c_1^{-1} \sum_{m=1}^{t} s_m) \geq e^{-c_2 t}.$$

Proof: We suppose that $s_1, s_2, \ldots, s_t$ is made up as follows: for each $n \geq 1$ and $1 \leq i \leq d_n$, we pick $0 \leq K_{n,i} \leq d_n$. Then the sequence $(s_1,\ s_2, \ldots,\ s_t)$ consists of the a_i^n, each one repeated $K_{n,i}$ times. Let L be the number of pairs (n,i) such that we have $K_{n,i} \neq 0$. Define the following events.

$$\begin{aligned}
B_{n,i} &= \Big\{ \sqrt{d_n} g_{i,i}^n a_i^n \geq c_1^{-1} K_{n,i} a_i^n \Big\}, \\
C_n &= \Big\{ \sqrt{d_n}\, \big\|G_n^{\mathrm{d}}\big\|_\infty = \sup_{1 \leq i \leq d_n} \big|g_{i,i}^n\big| \leq \sqrt{d_n}\lambda \Big\}, \\
D_n &= \Big\{ \big\|G_n^{\mathrm{ad}}\big\|_\infty \leq \lambda \Big\}.
\end{aligned}$$

By Lemma 3.3, we are really asking for a lower bound for the probability of the event

$$\sum_{n=1}^{\infty} \sum_{i=1}^{d_n} \sqrt{d_n} g_{i,i}^n a_i^n (\chi_{\|G_n^{\mathrm{d}}\|_\infty \leq \lambda})(\chi_{\|G_n^{\mathrm{ad}}\|_\infty \leq \lambda}) \geq c_1^{-1} \sum_{n=1}^{\infty} \sum_{i=1}^{d_n} K_{n,i} a_i^n.$$

However, we notice that this event contains

$$\bigcap_{(n,i):K_{n,i} \neq 0} B_{n,i} \cap C_n \cap D_n.$$

Now,

$$\bigcap_{(n,i):K_{n,i} \neq 0} B_{n,i} \cap C_n = \bigcap_{(n,i):K_{n,i} \neq 0} \Big\{ c_1^{-1} K_{n,i}/\sqrt{d_n} \leq g_{i,i}^n \leq \sqrt{d_n}\lambda \Big\}.$$

Since $K_{n,i} \le d_n$, if λ and c_1 are chosen large enough, then we see that

$$\Pr\left(\bigcap_{(n,i):K_{n,i}\neq 0} B_{n,i} \cap C_n \right) \ge c_3^{-L} \exp\left(-c_4 \sum_{n=1}^{\infty} \sum_{i=1}^{d_n} K_{n,i}^2/d_n \right).$$

Event D_n is independent of $B_{n,i}$ and C_n. By Lemma 3.8, it follows that for each number n that

$$\Pr(\left\|G_n^{\mathrm{ad}}\right\|_\infty \le \lambda) \ge \Pr(\|G_n\|_\infty \le \lambda/2) \ge c_5^{-1}.$$

Thus

$$\Pr\left(\bigcap_{(n,i):K_{n,i}\neq 0} B_{n,i} \cap C_n \cap D_n \right) \ge (c_3 c_5)^{-L} \exp\left(-c_4 \sum_{n=1}^{\infty} \sum_{i=1}^{d_n} K_{n,i}^2/d_n \right).$$

Now $K_{n,i}^2/d_n \le K_{n,i}$, and further, if $u \ge 1$, then $(c_3c_5)^{-1} e^{-c_4 u} \ge e^{-c_2 u}$. Hence the probability that we require is bounded below by

$$\begin{aligned}
(c_3c_5)^{-L} \exp\left(-c_4 \sum_{n=1}^{\infty} \sum_{i=1}^{d_n} K_{n,i} \right) &= \prod_{(n,i):K_{n,i}\neq 0} \left((c_3c_5)^{-1} \exp(-c_4 K_{n,i}) \right) \\
&\ge \prod_{(n,i):K_{n,i}\neq 0} \exp(-c_2 K_{n,i}) \\
&= \exp\left(-c_2 \sum_{n=1}^{\infty} \sum_{i=1}^{d_n} K_{n,i} \right) \\
&= e^{-c_2 t},
\end{aligned}$$

as desired. □

Now we are ready for the second case.

Proposition 4.3. *Fix $t > 0$. Suppose that there is a number c_1 such that for all $n \ge 1$ and $1 \le i \le d_n$ that either $a_i^n = 0$ or*

$$c_1^{-1} \|s\|_2 \le t\sqrt{d_n} a_i^n \le c_1 \sqrt{d_n}\ \|s\|_2 .$$

Then for sufficiently large λ, there are numbers c_2 and c_3, depending only on c_1 and λ, such that

$$\Pr(S_{G^*} \ge c_2^{-1} t \|s\|_2) \ge e^{-c_3 t^2}.$$

Proof: The second case has a very similar proof to the first case. First, without loss of generality, we may suppose that $A_n \neq 0$ for all $n \geq 1$. Also recall that

$$\|s\|_2 = \left(\sum_{n=1}^{\infty} \sum_{i=1}^{d_n} d_n (a_i^n)^2 \right)^{1/2} .$$

Define the events

$$B_{n,i} = \begin{cases} \left\{ c_2^{-1} t a_i^n \sqrt{d_n} / \|s\|_2 \leq g_{i,i}^n \leq \sqrt{d_n} \lambda \right\} & \text{if } a_i^n \neq 0, \\ \left\{ g_{i,i}^n \leq \sqrt{d_n} \lambda \right\} & \text{if } a_i^n = 0, \end{cases}$$

$$C_n = \left\{ \left\| G_n^{\mathrm{ad}} \right\|_\infty \leq \lambda \right\}.$$

By Lemma 3.3, we are looking for a lower bound for the event

$$\sum_{n=1}^{\infty} \sum_{i=1}^{d_n} \sqrt{d_n} g_{i,i}^n a_i^n \chi_{\|G_n^{\mathrm{d}}\|_\infty \leq \lambda} \chi_{\|G_n^{\mathrm{ad}}\|_\infty \leq \lambda} \geq c_2^{-1} t \|s\|_2 ,$$

This event contains

$$\bigcap_{(n,i)} B_{n,i} \cap C_n .$$

Let us first consider $\Pr(B_{n,i})$ in the case when $a_i^n \neq 0$. Since $t\sqrt{d_n} a_i^n \leq c_1 \sqrt{d_n} \|s\|_2$, we see that for sufficiently large λ that $2c_2^{-1} t a_i^n \sqrt{d_n} / \|s\|_2 \leq \sqrt{d_n} \lambda$, and hence

$$\Pr(B_{n,i}) \geq c_4^{-1} \exp\left(-c_5 t^2 \frac{d_n (a_i^n)^2}{\|s\|_2^2} \right) .$$

Now $c_1^{-1} \|s\|_2 \leq t\sqrt{d_n} a_i^n$, and if $u > c_1^{-2}$, then $c_4 e^{-c_5 u} \geq e^{-c_6 u}$, and hence

$$\Pr(B_{n,i}) \geq \exp\left(-c_6 t^2 \frac{d_n (a_i^n)^2}{\|s\|_2^2} \right)$$

Hence for each $n \geq 1$

$$\Pr\left(\bigcap_{i : a_i^n \neq 0} B_{n,i} \right) \geq \exp\left(-c_6 t^2 \sum_{i=1}^{d_n} \frac{d_n (a_i^n)^2}{\|s\|_2^2} \right) .$$

Also

$$\Pr\left(C_n \cap \bigcap_{i : a_i^n = 0} B_{n,i} \right) \geq \Pr(\|G_n\| \leq \lambda/2) \geq c_7^{-1} ,$$

where the last inequality follows from Lemma 3.8 if λ is sufficiently large. Since $c_7 e^{-c_6 u} \geq e^{-c_3 u}$ whenever $u > c_1^{-2}$, it follows that

$$\Pr\left(C_n \cap \bigcap_i B_{n,i}\right) \geq \exp\left(-c_3 t^2 \sum_{i=1}^{d_n} \frac{d_n (a_i^n)^2}{\|s\|_2^2}\right).$$

Hence,

$$\Pr\left(\bigcap_{(n,i)} B_{n,i} \cap C_n\right) \geq \exp\left(-c_3 t^2 \sum_{n=1}^{\infty} \sum_{i=1}^{d_n} \frac{d_n (a_i^n)^2}{\|s\|_2^2}\right) = e^{-c_3 t^2},$$

as desired. □

Now for the third case. The argument that follows was suggested by the proof of Proposition 4.13 in **[L–T]**.

Proposition 4.4. *Fix $t > 0$. Suppose that there is a number c_1 such that for all $n \geq 1$ and $1 \leq i \leq d_n$ that either $a_i^n = 0$ or*

$$c_1^{-1} \|s\|_2 / \sqrt{d_n} \leq t\sqrt{d_n} a_i^n \leq c_1 \|s\|_2 .$$

Then for sufficiently large λ, there are numbers c_2 and c_3, depending only on c_1 and λ, such that

$$\Pr(S_{G'} \geq t \|s\|_2) \geq c_2^{-1} e^{-c_3 t^2}.$$

Proof: First note that for $u > 0$ that

$$\begin{aligned}\Pr\left(d_n \operatorname{tr}(A_n G_n)\chi_{\|G_n\|_\infty > \lambda} > u\right) &\leq \min\left\{\Pr\left(d_n \operatorname{tr}(A_n G_n) > u\right), \Pr\left(\|G_n\|_\infty > \lambda\right)\right\} \\ &\leq c_4 \min\left\{\exp\left(-\frac{c_4^{-1} u^2}{d_n \|A_n\|_2^2}\right), \exp(-c_4^{-1}\lambda^2 d_n)\right\}.\end{aligned}$$

(Here we used Lemma 3.8.) Now let

$$\theta = \frac{t\lambda^{1/2}}{\|s\|_2}.$$

Since $c_1^{-1} \|s\|_2 / \sqrt{d_n} \leq t\sqrt{d_n} a_i^n \leq c_1 \|s\|_2$ whenever $a_i^n \neq 0$, it follows that

$$c_1^{-2} \frac{\|s\|_2^2}{d_n} \leq t^2 d_n \sum_{i : a_i^n \neq 0} (a_i^n)^2 \leq c_1^2 d_n \|s\|_2^2 ,$$

i.e., $c_1^{-1} \|s\|_2 / d_n \le t \, \|A_n\|_2 \le c_1 \|s\|_2$. Hence

$$c_1^{-1} \lambda^{1/2} / d_n \le \theta \|A_n\|_2 \le c_1 \lambda^{1/2}.$$

Now

$$\begin{aligned} \mathbb{E}\left(\exp\left(\theta d_n \operatorname{tr}(A_n G_n) \chi_{\|G_n\|_\infty > \lambda}\right)\right) &\le 1 + \int_0^\infty \theta e^{\theta u} \Pr\left(d_n \operatorname{tr}(A_n G_n) \chi_{\|G_n\|_\infty > \lambda} > u\right) \, du \\ \le 1 + c_4 \int_0^{\lambda d_n \|A_n\|_2} \theta e^{\theta u} \exp(-c_4^{-1} \lambda^2 d_n) \, du &+ c_4 \int_{\lambda d_n \|A_n\|_2}^\infty \theta e^{\theta u} \exp\left(-\frac{c_4^{-1} u^2}{d_n \|A_n\|_2^2}\right) \, du. \end{aligned}$$

Furthermore,

$$\begin{aligned} \int_{\lambda d_n \|A_n\|_2}^\infty \theta e^{\theta u} \exp\left(-\frac{c_4^{-1} u^2}{d_n \|A_n\|_2^2}\right) \, du &\le \int_{\lambda d_n \|A_n\|_2}^\infty \theta e^{\theta u} \exp\left(-\frac{c_4^{-1} \lambda u}{\|A_n\|_2}\right) \, du \\ &= \frac{\theta \|A_n\|_2}{c_4^{-1} \lambda - \theta \|A_n\|_2} \exp\left(\left(\theta \|A_n\|_2 - c_4^{-1} \lambda\right) \lambda d_n\right) \\ &\le c_5 \exp(-c_6^{-1} \lambda^2 d_n), \end{aligned}$$

when λ is sufficiently large, because $\theta \|A_n\|_2 \le c_1 \lambda^{1/2}$. Similarly

$$\begin{aligned} \int_0^{\lambda d_n \|A_n\|_2} \theta e^{\theta u} \exp(-c_4^{-1} \lambda^2 d_n) \, du &= \exp\left(-c_4^{-1} \lambda^2 d_n + \lambda \theta d_n \|A_n\|_2\right) \\ &\le \exp(-c_6^{-1} \lambda^2 d_n), \end{aligned}$$

when λ is sufficiently large.

Now,

$$1 \le c_1^2 \lambda^{-1} \theta^2 d_n^2 \|A_n\|_2^2,$$

and since $e^{-u} \le 1/u$ for $u > 0$,

$$\exp(-c_6^{-1} \lambda^2 d_n) \le c_6 \lambda^{-2} d_n^{-1},$$

and so

$$\exp(-c_6^{-1} \lambda^2 d_n) \le c_1^2 c_6 \lambda^{-3} \theta^2 d_n \|A_n\|_2^2.$$

Hence

$$\begin{aligned} \mathbb{E}\left(\exp\left(\theta d_n \operatorname{tr}(A_n G_n) \chi_{\|G_n\|_\infty > \lambda}\right)\right) &\le 1 + c_4 (1 + c_5) c_1^2 c_6 \lambda^{-3} \theta^2 d_n \|A_n\|_2^2 \\ &\le \exp(c_7 \lambda^{-3} \theta^2 d_n \|A_n\|_2^2). \end{aligned}$$

Now,

$$\begin{aligned}
\mathbb{E}(\exp(\theta S_{G''})) &= \mathbb{E}\left(\prod_{n=1}^{\infty} \exp\bigl(\theta d_n \operatorname{tr}(A_n G''_n)\bigr)\right) \\
&= \prod_{n=1}^{\infty} \mathbb{E}\Bigl(\exp\bigl(\theta d_n \operatorname{tr}(A_n G''_n)\bigr)\Bigr) \\
&\le \prod_{n=1}^{\infty} \exp(c_7 \lambda^{-3} \theta^2 d_n \|A_n\|_2^2) \\
&= \exp(c_7 \lambda^{-3} \theta^2 \|s\|_2^2).
\end{aligned}$$

So,

$$\Pr\Bigl(\exp(\theta S_{G''}) > \exp(c_7 \lambda^{-3} \theta^2 \|s\|_2^2 + \lambda^{1/2} t^2)\Bigr) \le e^{-\lambda^{1/2} t^2},$$

that is,

$$\Pr\Bigl(S_{G''} > (1 + c_7 \lambda^{-5/2}) t \|s\|_2\Bigr) \le e^{-\lambda^{1/2} t^2}.$$

If $\lambda > c_7^{-2/5}$, then

$$\Pr(S_{G''} > 2t \|s\|_2) \le e^{-\lambda^{1/2} t^2}.$$

To finish, we note that

$$\Pr(S_{G'} > t \|s\|_2) \ge \Pr(S_G > 2t \|s\|_2) - \Pr(S_{G''} > t \|s\|_2) \ge c_8^{-1} e^{-c_8 t^2} - e^{-\lambda^{1/2} t^2/4}.$$

Thus if λ is sufficiently large, then

$$\Pr(S_{G'} > t \|s\|_2) \ge c_2^{-1} e^{-c_3 t^2}$$

for $t > \frac{1}{2}$.

If $t \le \frac{1}{2}$, then we can use the following inequality (see **[Ka]**, Chapter 1): if X is a positive random variable, then

$$\Pr(X \ge \|X\|_1 / 2) \ge \frac{\|X\|_1^2}{4 \|X\|_2^2}.$$

Take $X = |S_\epsilon|^2$. By the result of Figà-Talamanca and Rider, and Lemma 3.13 it follows that $\|X\|_2 \le c_9 \|X\|_1$, and the result follows. $\square$

The fourth case follows by comparing the non-commutative case with the commutative case.

PROPOSITION 4.5. *Fix $t > 0$. Suppose that there is a number c_1 such that for all $n \geq 1$ and $1 \leq i \leq d_n$*

$$t\sqrt{d_n} a_i^n \leq c_1 \|s\|_2 / \sqrt{d_n}.$$

Then there is a number c_2, depending only on c_1, such that

$$\Pr(S_\epsilon \geq c_2^{-1} t \|s\|_2) \geq c_2^{-1} e^{-c_2 t^2}.$$

Proof: We apply Lemma 3.10 with

$$v_n = \frac{\sqrt{d_n}\,\mathrm{tr}(A_n \epsilon_n)}{\|A_n\|_2}$$

and $a_n = \sqrt{d_n}\,\|A_n\|_2$ to deduce that

$$\Pr\left(S_\epsilon > t\|s\|_2\right) \geq c_3^{-1} \Pr\left(\sum_{n=1}^\infty \sqrt{d_n}\,\|A_n\|_2\, r_n > c_3 t \|s\|_2\right).$$

From the hypothesis, we have that

$$t\sqrt{d_n}\,\|A_n\|_2 \leq c_1 \|s\|_2\,,$$

and hence by the commutative Rademacher series result, it follows that

$$\Pr\left(\sum_{n=1}^\infty \sqrt{d_n}\,\|A_n\|_2\, r_n > c_4^{-1} t \|s\|_2\right) \geq c_4^{-1} e^{-c_4 t^2},$$

as required. ☐

Now we are finally ready to put the pieces together. Let us restate the theorem we are attempting to prove.

THEOREM 2.1. *The distribution of S_ϵ is given by the following formulae.*

i) $$\Pr(S_\epsilon > c\,K_{1,2}(t,s)) \leq c\,e^{-c^{-1}t^2}.$$

ii) $$\Pr(S_\epsilon > c^{-1}\,K_{1,2}(t,s)) \geq c^{-1} e^{-ct^2}.$$

Proof: Part (i) is simply Proposition 4.1. To prove part (ii), we may suppose that $t \geq 1$. Now, note that we have the following bound:

$$K_{1,2}(t,s) \leq \sum_{m=1}^{[t^2]} s_m + t \left(\sum_{m=[t^2]+1}^{\infty} (s_m)^2 \right)^{1/2} .$$

Then we have two possibilities.

Case 1: The first possibility is that

$$\sum_{m=1}^{[t^2]} s_m \geq \tfrac{1}{2} K_{1,2}(t,s).$$

In that case, the result follows by Lemma 3.12 and Proposition 4.2.

Case 2: Otherwise, we know that

$$t \left(\sum_{m=[t^2]+1}^{\infty} (s_m)^2 \right)^{1/2} \geq \tfrac{1}{2} K_{1,2}(t,s).$$

We also know that

$$t \left(\sum_{m=[t^2]+1}^{\infty} (s_m)^2 \right)^{1/2} \geq \sum_{m=1}^{[t^2]} s_m \geq [t^2] s_{[t^2]},$$

since the sequence (s_m) is in decreasing order. Hence, if $m \geq [t^2]$, we have that

$$2ts_m \leq \left(\sum_{m=[t^2]+1}^{\infty} (s_m)^2 \right)^{1/2} .$$

Let M be the least number m such that $s_m = s_{[t^2]+1}$. Let us replace the matrices A_n with matrices A'_n that drop the entries that correspond to s_m for $m < M$. Thus the new sequence s' formed satisfies the following.

$$\|s'\|_2 \geq \tfrac{1}{2} K_{1,2}(t,s) \quad \text{and} \quad 2ts'_m \leq \|s'\|_2 .$$

Thus, we may replace the matrices A_n with A'_n, and, using Lemma 3.3, we are reduced to showing the following: subject to the restriction that

$$2ta_i^n \leq \|s\|_2 ,$$

we desire to show that

$$\Pr(S_\epsilon \geq 2c^{-1} t \left\| s \right\|_2) \geq 2c^{-1} e^{-ct^2}.$$

To prove this, we will split the entries of the matrices into three parts. Let

$$\begin{aligned} B_1 &= \{\, (n,i) : \left\| s \right\|_2 < t\sqrt{d_n} a_i^n \,\}, \\ B_2 &= \{\, (n,i) : \left\| s \right\|_2 / \sqrt{d_n} < t\sqrt{d_n} a_i^n \leq \left\| s \right\|_2 \,\}, \\ B_3 &= \{\, (n,i) : t\sqrt{d_n} a_i^n \leq \left\| s \right\|_2 / \sqrt{d_n} \,\}. \end{aligned}$$

Then for one of $j = 1$, 2, 3, we have that

$$\sum_{(n,i)\in B_j} d_n (a_i^n)^2 \geq \tfrac{1}{3} \left\| s \right\|_2^2 .$$

In that case, we can replace the matrices A_n with matrices that only take those entries that are in the set B_j. Now the result follows by Lemma 3.12, Lemma 3.3 and Proposition 4.3, 4.4 or 4.5.

∎

Acknowledgements

The author would like to thank Nakhlé Asmar for both useful discussions and warm friendship while this paper was being prepared. He would also like to express gratitude to the referee for useful corrections and comments.

References

A–M N. Asmar and S.J. Montgomery-Smith, On the distribution of Banach valued Sidon spectral functions, *Arkiv Mat.* (*to appear*).

B–S C. Bennett and R. Sharpley, *Interpolation of Operators*, Academic Press.

B–L J. Bergh and J. Löfström, *Interpolation Spaces*, Springer-Verlag, 1976.

Bo C. Borell, The Brunn–Minkowski inequality in Gauss Space, *Invent. Math.* **30** (*1975*) *207–216.*

dP–M V.H. de la Peña and S.J. Montgomery-Smith, Decoupling inequalities for tail probabilities of multilinear forms of symmetric and hypercontractive variables, *submitted.*

D–M. S.J. Dilworth and S.J. Montgomery-Smith, The distribution of vector-valued Rademacher series, *Annals Prob.* (*to appear*).

F–R A. Figà-Talamanca and D. Rider, A theorem of Littlewood and lacunary series for compact groups, *Pacific J. Math.* **16** (*1966*), *505–514.*

H–R E. Hewit and K.A. Ross, *Abstract Harmonic Analysis II*, Springer–Verlag, 1970.

Hi P. Hitczenko, Domination inequality for martingale transforms of a Rademacher sequence, *Israel J. Math.* (*to appear*).

Ho T. Holmstedt, Interpolation of quasi-normed spaces, *Math. Scand.* **26** (*1970*), *177–199.*

Ka J-P. Kahane, *Some Random Series of Functions*, (2nd. Ed.) Cambridge studies in advanced mathematics 5, 1985.

Kh A. Khintchine, Über dyadische Brüche, *Math. Z.* **18** (*1923*), *109–116.*

Ko A.N. Kolmogorov, Über das Gesetz des iterieten Logarithmus, *Math. Ann.* **101** (*1929*), *126–135.*

L–T M. Ledoux and M. Talagrand, *Isoperimetry and Processes in Probability in a Banach Space*, Springer–Verlag, 1991.

M–P M.B. Marcus and G. Pisier, *Random Fourier Series with Applications to Harmonic Analysis*, Princeton University Press, 1981.

Mo S.J. Montgomery-Smith, The distribution of Rademacher sums, *Proc. A.M.S.* **109** (*1990*), *517–522.*

P1 G. Pisier, De nouvelles caractérisations des ensembles de Sidon, Mathematical Analysis and Applications, *Advances in Math. Suppl. Stud.*, *7B* (*1981*), *686–725.*

P2 G. Pisier, *Probabilistic Methods in the Geometry of Banach Spaces*, Springer–Verlag, 1986.

R–S V.A. Rodin and E.M. Semyonov, Rademacher series in symmetric spaces, Analyse Math. **1** (1975), 207–222.