University of Warsaw
Faculty of Mathematics, Informatics and Mechanics

Rafał Meller

Two-sided bounds of moments of random chaoses - real and vector case

PhD dissertation

Supervisor
prof. dr hab. Rafał Latała
Institute of Mathematics
University of Warsaw

January 2019

Author's declaration:
I hereby declare that this dissertation is my own work.

January 28, 2019 ....

Rafał Meller

Supervisor's declaration:
The dissertation is ready to be reviewed

January 28, 2019 ....

prof. dr hab. Rafał Latała

Abstract

In the following thesis we are investigating random multilinear forms, which are called random chaoses, defined by

$$S := \sum_{i_1,\ldots,i_d} a_{i_1,\ldots,i_d} X_{i_1} \cdots X_{i_d},$$

where $d \in \mathbb{N}$, $X_1, \ldots, X_n$ are independent random variables, and $a_{i_1,\ldots,i_d} \in F$, where $(F, \|\cdot\|)$ is a Banach space. We want to derive two-sided bounds of $\|S\|_p := (\mathbb{E}\|S\|^p)^{1/p}$ under some conditions about the structure of $(a_{i_1,\ldots,i_d})_{i_1,\ldots,i_d}$ and the distribution of $(X_1, X_2, \ldots)$.

The first part of the thesis concerns the case when $F = \mathbb{R}$ (the real case). In the first chapter we assume that the random variables $X_1, X_2, \ldots$ are nonnegative and their moments do not grow "too fast". In the second chapter we consider the symmetric random variables satisfying the same moment condition.

The second part is dedicated to vector-valued chaoses. We derive certain upper bounds which turn out to be two-sided in a special class of Banach spaces (which includes L_q spaces). In the third and fourth chapters we analyze the case when the random variables $X_1, X_2, \ldots$ have Gaussian distribution and $d = 2$ or $d > 2$ respectively. In the fifth chapter we study variables with log-concave tails and $d = 2$.

AMS 2000 subject classifications: Primary: 60E15. Secondary: 60B11.

Keywords: random chaoses, random quadratic forms, two-sided estimates for tails and moments, logarithmically concave tails, suprema of Gaussian processes, metric entropy.

Rafał Meller
Warsaw, January 2019

Streszczenie

Niniejsza rozprawa poświęcona jest losowym formom wieloliniowym, zwanym chaosami losowymi, które są zdefiniowane jako

$$S := \sum_{i_1,\ldots,i_d} a_{i_1,\ldots,i_d} X_{i_1} \cdots X_{i_d},$$

gdzie $d \in \mathbb{N}$, $X_1, \ldots, X_n$ są niezależnymi zmiennymi losowymi natomiast $a_{i_1,\ldots,i_d}$ są współczynnikami z przestrzeni Banacha $(F, \|\cdot\|)$. Naszym celem jest znalezienie dwustronnych oszacowań momentów całkowych zmiennej S określonych jako $\|S\|_p := (\mathbb{E}\|S\|^p)^{1/p}$. Będziemy przy tym zakładać pewne warunki o strukturze współczynników $(a_{i_1,\ldots,i_d})_{i_1,\ldots,i_d}$ oraz o rozkładzie zmiennych losowych $(X_1, X_2, \ldots)$.

W pierwszej części pracy rozważamy przypadek gdy $F = \mathbb{R}$ (przypadek rzeczywisty). W pierwszym rozdziale zakładamy, że zmienne losowe $X_1, X_2, \ldots$ są nieujemne oraz, że momenty tych zmiennych nie rosną "za szybko". W drugim rozdziale rozważamy symetryczne zmienne losowe, które spełniają ten sam warunek na wzrost momentów.

Druga część pracy poświęcona jest przypadkowi chaosów losowych o wartościach wektorowych. Wyprowadzamy górne oszacowania na momenty, które można odwrócić w pewnej klasie przestrzeni Banacha (zawierającej przestrzenie L_q). Trzeci i czwarty rozdział poświęcony jest przypadkowi, gdy zmienne $X_1, X_2, \ldots$ mają rozkład normalny oraz, gdy odpowiednio $d = 2$ i $d > 2$. W ostatnim rozdziale omówiony jest przypadek zmiennych losowych o logarytmicznie wklęsłych ogonach oraz $d = 2$.

Klasyfikacja tematyczna. 60E15; 60B11.

Słowa kluczowe. Chaosy losowe, losowe formy kwadratowe, dwustronne oszacowania momentów i ogonów, logarytmicznie wklęsłe ogony, suprema procesów gaussowskich, entropia metryczna.

Rafał Meller
Warszawa, Styczeń 2019

Contents

Chapter 0
Introduction

0.1 Formulation of the problem.

Assume that $X_1, X_2, \ldots$ are independent random variables and $a_{i_1,\ldots,i_d} \in F$, where $(F, \|\cdot\|)$ is a Banach space. Then we will call

$$S = \sum_{i_1,\ldots,i_d} a_{i_1,\ldots,i_d} X_{i_1} \cdots X_{i_d}$$

a random (undecoupled) chaos of order $d \in \mathbb{N}$. We say that S is symmetric if r.v's $X_1, X_2, \ldots$ are symmetric. We define p-th moment of S by $\|S\|_p = (\mathbb{E}\|S\|^p)^{1/p}$. We aim at finding two-sided bounds for $\|S\|_p$, so to find a "simpler" expression H which depends on p, $(a_{i_1,\ldots,i_d})_{i_1,\ldots,i_d}$ and the distribution of $X_1, X_2, \ldots$ such that

$$C(d)^{-1} H \leq \|S\|_p \leq C(d) H, \tag{0.1}$$

where $C(d)$ (resp. C) is a constant which depends only on d (resp. is a numerical constant) and differs at each occurrence (if (0.1) holds we will write $S \sim^d H$ for short).

For $d = 1$ the problem is well understood. For $d > 1$ it may be guessed the form of H by the decoupling argument (see below) and iteration. The induction type arguments usually enable to show in an easy way the lower bound in (0.1). The upper bound is typically much harder to establish.

In order to derive (0.1) some conditions about the distribution of $(X_1, X_2, \ldots)$ must be imposed and usually we add some conditions about the structure of $(a_{i_1,\ldots,i_d})_{i_1,\ldots,i_d}$. For example we will consider $X_i \sim \mathcal{N}(0,1)$ (even in this case there are still open problems). In this instance we write $g_1, \ldots, g_n$ instead of $X_1, \ldots, X_n$. Another possibility is to assume that $X_1, \ldots, X_n$ are symmetric random variables with log-concave tails (LCT for short), i.e for any i the function $t \to -\ln \mathbb{P}(|X_i| \geq t) \in [0, \infty]$ is convex on $[0, \infty)$. This class is interesting, since it contains many important distributions such as Rademacher, normal and exponential.

We usually assume about the structure of $(a_{i_1,\ldots,i_d})_{i_1,\ldots,i_d}$ that (observe the first condition does not reduce the generality)

- $(a_{i_1,\ldots,i_d})$ is symmetric, i.e $a_{i_1,\ldots,i_d} = a_{i_{\pi(1)},\ldots,i_{\pi(d)}}$ for all permutations π of $[d] := \{1, \ldots, d\}$,
- $(a_{i_1,\ldots,i_d})$ is tetrahedral, i.e $i_k = i_l$ for $k \neq l$, $k, l \leq d$ implies $a_{i_1,\ldots,i_d} = 0$.

The reason for these conditions is that they enable the decoupling method [7, 8, 14], a basic tool in this area.

Theorem 0.1.1 (Kwapień decoupling theorem). *Let S be a chaos of order d and $(a_{i_1,\ldots,i_d})_{i_1,\ldots,i_d}$ be symmetric and tetrahedral. Take*

$$S' := \sum_{i_1,\ldots,i_d} a_{i_1,\ldots,i_d} X^1_{i_1} \cdots X^d_{i_d},$$

where $(X^k_i)_{i\le n}$, $k=1,\ldots d$ *are independent copies of* $(X_i)_{i\le n}$. *Then for any* $p\ge 1$

$$\|S\|_p \sim^d \|S'\|_p .$$

So instead of bounding $\|S\|_p$ it is enough to bound $\|S'\|_p$. The latter object has a reacher structure which allows us for inductive-type reasoning. Indeed, assume that

$$\left\| \sum_{i_1,\ldots,i_d} a_{i_1,\ldots,i_d} X^1_{i_1} \cdots X^d_{i_d} \right\|_p \sim^d H(d,(a_{i_1,\ldots,i_d}),p).$$

By conditioning we obtain

$$\left\| \sum_{i_1,\ldots,i_{d+1}} a_{i_1,\ldots,i_{d+1}} X^1_{i_1} \cdots X^d_{i_d} X^{d+1}_{i_{d+1}} \right\|_p \sim^d \left\| H\left(d, \left(\sum_{i_{d+1}} a_{i_1,\ldots,i_{d+1}} X^{d+1}_{i_{d+1}}\right), p\right)\right\|_p .$$

By concentration of measure-type arguments the problem of estimating the latter expression can be reduced (often) to the problem of bounding $\left\|H(d,(\sum_{i_{d+1}} a_{i_1,\ldots,i_{d+1}} X^{d+1}_{i_{d+1}}),p)\right\|_1$ which can be expressed as an expectation of suprema of stochastic processes. For this reason the problem of estimating $\|S\|_p$ is deeply connected with the problem of estimating the expectation of suprema of some stochastic processes. The modern approach of dealing with the latter is based on chaining methods which are ubiquitous in the second part of the thesis.

Tail estimates from moments bounds can be deduced in an easy way. Assume that $\|S\|_p \sim h(p)$ for any $p\ge 1$. Then Chebyshev's inequality yields

$$\mathbb{P}\left(|S| \ge Ch(p)\right) \le \left(\frac{\|S\|_p}{Ch(p)}\right)^p \le e^{-p}. \tag{0.2}$$

Surprisingly inequality (0.2) can be reversed if $h(2p)\le Ch(p)$ or equivalently if $\|S\|_{2p} \le C\|S\|_p$ (this condition is satisfied if the chaos S is based on random variables with moments growing at most polynomially, which holds in all of our considerations). By Paley-Zygmund's inequality we have

$$\mathbb{P}\left(|S|\ge C^{-1}h(p)\right) \ge \mathbb{P}\left(|S| \ge \frac{1}{2}\|S\|_p\right) \ge \left(1-\frac{1}{2^p}\right)^2 \left(\frac{\|S\|_p}{\|S\|_{2p}}\right)^p \ge e^{-Cp}.$$

Observe that for $d=1$, $S=\sum a_i X_i$ is just a linear combination of random variables, and thus many classical bounds on moments and tails can be applied (such as Khintchine, Rosenthal, Bernstein, Hoeffding, Prokhorov and Bennett inequalities to name a few). The situation is more sophisticated if two-sided estimates are considered. Moreover, for $d\ge 2$ (observe that for $d=2$ we have $S=\sum_{ij} a_{ij} X_i X_j$) the classical inequalities cannot be applied, since the summands are no longer independent (even in the decoupled case).

We would like to end this section by the following comment. The field of studying chaoses is quite technical, and obtained formulas are not easy to grasp. After spending some time on the problems connected with random chaoses the notation becomes quite natural, but for someone who sees it for the first time it can be quite overwhelming.

0.2 Significance and earlier results

Chaoses appear in many branches of modern probability, e.g. as approximations of multiple stochastic integrals, elements of Fourier expansions in harmonic analysis on the discrete cube (when the underlying variables X_i's are independent Rademachers), in subgraph counting problems for random graphs (in this case X_i's are zero-one random variables), in statistical physics or in statistics. For instance, recently the Hanson-Wright type inequality (i.e estimates for quadratic forms in subgaussian random variables) attracted attention of many statisticians (the relatively recent paper [28] has more than 250 citations in the google scholar database).

The case of $d=1$ and real coefficients can be considered closed due to results of Latała [15]. Assuming that the r.v's $X_1,\ldots,X_n$ are symmetric (and nothing more) Latała [15] proved that

$$\left\|\sum_i X_i\right\|_p \sim \inf\left\{t>0:\ \sum_i \ln\left(\mathbb{E}\left|1+\frac{X_i}{t}\right|^p\right)\le p\right\}, \tag{0.3}$$

(the result holds also under the assumption that $\mathbb{P}(X_1,\ldots,X_n\ge 0)=1$). The point of the above formula is that in many cases the function $t\to\ln\mathbb{E}|1+tX_i|^p$ can be estimated in a simple way.

The case of coefficients from a Banach space $(F,\|\cdot\|)$ is solved under the assumption that moments of $X_1,\ldots,X_n$ grow at most polynomially [21], so in quite satisfying generality. Namely, it is shown in [21] that in this case

$$\left\|\sum a_iX_i\right\|_p \sim^{\beta} \left\|\sum a_iX_i\right\|_1 + \sup_{\substack{\varphi\in F^*\\ \|\varphi\|_*=1}} \left\|\sum \varphi(a_i)X_i\right\|_p,$$

where β is such that $X_1,X_2,\ldots$ satisfy (0.5). The term $\|\sum\varphi(a_i)X_i\|_p$ can be bounded by [15] or Theorem 2.3.1.

The case $d\ge 2$ was much less understood. In particular all results mentioned below (except the last one in this subsection) involve only real-valued chaoses.

For arbitrary d, two-sided estimates were known in the following cases:

- Gaussian chaoses [16],
- Chaoses based on nonnegative r.v's with LCT and nonnegative $a_{i_1,\ldots,i_d}$'s [19],
- Chaoses based on symmetric r.v's with logarithmically convex tails [13].

Łochowski [26] derived (generalizing the earlier result [3]) also the two-sided bound for arbitrary d in the real case for chaoses based on symmetric random variables with LCT. However, his bounds involve suprema of some stochastic processes indexed by sets which depend on p. Such expressions are in general very hard to estimate.

In the symmetric real-valued case for small d, two-sided estimates were known in the following cases:

- $d=2$, and chaoses based on symmetric r.v's with LCT [17],
- $d=3$, and chaoses based on symmetric r.v's with LCT [1]

In the case of $d>1$ and vector-valued chaoses only result concerning the Gaussian chaoses was known (cf. Subsection 0.3.3). However, this bound involves quantity (the expectation of suprema of norms evaluated on a Gaussian vector) which is hard to deal with.

0.3 Overview of the thesis

The following dissertation consists of five chapters containing results from the research conducted at the Mathematical Institute of University of Warsaw from February 2016 to December 2018. The chapters are based mainly (except for the fifth chapter) on published or submitted articles as follows:

- Chapter 1 R. Meller, *Two-sided moment estimates for a class of nonnegative chaoses*, Statistics & Probability Letters 119 (2016), 213–219;
- Chapter 2 R. Meller, *Tail and moment estimates for a class of random chaoses of order two*, to appear in Studia Mathematica (2018);
- Chapter 3 R. Adamczak, R. Latala, R. Meller, *Hanson-Wright inequality in Banach spaces*, submitted;
- Chapter 4 R. Adamczak, R. Latala, R. Meller, *Moments of Gaussian chaoses in Banach spaces*, in preparation.

The fifth chapter contains observations of the author of this thesis concerning chaoses based on symmetric r.v's with LCT and with values in L_q spaces. This doctoral dissertation uses many results contained in various publications. For the convenience of the reader most of the quatoed results are gathered in the Appendix (Chapter 6).

We will now briefly discuss the main results presented in the thesis. Since in all of them we assume conditions which enable the decoupling method (with a small exception in the third and fourth chapter) we will limit ourself mostly to the decoupled case.

0.3.1 Two-sided moment estimates for a class of nonnegative chaoses

Assume that for any $i_1,\ldots,i_d$ we have $a_{i_1,\ldots,i_d} \geq 0$, $\mathbb{P}(\forall_{i,j} X_i^j \geq 0) = 1$, and for any i,j, $\mathbb{E}X_i^j = 1$ and X_i^j has LCT. We define $N_i^j(t) = -\ln \mathbb{P}(X_i^j > t) \in [0,\infty]$. Łochowski and Latała [19] showed that

$$\left\|S'\right\|_p \sim^d \sup \sum_{i_1,\ldots,i_d} a_{i_1,\ldots,i_d} \prod_{j=1}^{d} (1+v_{i_j}^j), \tag{0.4}$$

where the supremum is taken over vectors $v^1,\ldots,v^d \in \mathbb{R}_+^n$ such that for any $j = 1,2,\ldots,d$

$$\sum_i N_i^j(v_i^j) \leq p.$$

In particular for $d = 1$ they obtained that

$$\left\|\sum_i a_i X_i^1\right\|_p \sim \sup\left\{\sum_i a_i(1+v_i^1) \mid \sum_i N_i^1(v_i^1) \leq p\right\}.$$

Łochowski showed in his PHD thesis that chaoses based on i.i.d zero-one random variables satisfy (0.4) with a small correction. Namely he showed that

$$C(d)^{-1}\ln^{-d}(\frac{1}{\alpha})\sup\sum_{i_1,\ldots,i_d} a_{i_1,\ldots,i_d}\prod_{k=1}^{d}(\alpha+v^k_{i_k})\le \|S'\|_p$$
$$\le C(d)\ln^{d}(\frac{1}{\alpha})\sup\sum_{i_1,\ldots,i_d} a_{i_1,\ldots,i_d}\prod_{k=1}^{d}(\alpha+v^k_{i_k}),$$

where $\alpha=\mathbb{P}(X^j_i=1)$ (in the above formula we have $(\alpha+v^k_{i_k})$ instead of $(1+v^k_{i_k})$ since the r.v's $(X^j_i)_{ij}$ are not normalized). He also showed that the constant in the upper bound must depend on α. If $\mathbb{P}(X=1)=1-\mathbb{P}(X=0)=\alpha$ then for any $p\ge 1$, we have $\|X\|_{2p}\le \alpha^{-1/2}\|X\|_p$. This justified the question whether (0.4) holds (with a constant depending on β) for nonnegative variables satisfying

$$\left\|X^j_i\right\|_{2p}\le \beta\left\|X^j_i\right\|_p. \tag{0.5}$$

We give a positive answer which is the main result of the first chapter.

Theorem 0.3.1. *Let $(X^j_i)_{i\le n,j\le d}$ be independent nonnegative random variables satisfying* (0.5) *and $\mathbb{E}X^j_i=1$. Then for any nonnegative coefficients $(a_{i_1,\ldots,i_d})_{i_1,\ldots,i_d\le n}$ we have*

$$\|S'\|_p\sim^{d,\beta}\sup\sum_{i_1,\ldots,i_d} a_{i_1,\ldots,i_d}\prod_{j=1}^{d}(1+v^j_{i_j}),$$

where the supremum is taken over $v^1,\ldots,v^d\in\mathbb{R}^n_+$ such that for any $j=1,2,\ldots,d$

$$\sum_i N^j_i(v^j_i)\le p.$$

It is a simple fact that moments of a symmetric r.v. X with LCT satisfy (0.5) with $\beta=2$, so Theorem 0.3.1 generalizes [19]. Moreover, (0.5) arises naturally in the paper of Latała and Strzelecka [21] as a sufficient condition (and even necessary in the i.i.d case) for comparison of weak and strong moments of the random variable $\sup_{t\in T\subset\mathbb{R}^n}\sum t_iX_i$. Lastly it is shown in Chapter 1 (see Remark 1.2.2 therein) that if $\ln\mathbb{P}(|X|\ge Ktx)\le t^\alpha\mathbb{P}(|X|\ge x)$ for any $t,x\ge 1$ and some constants K,α, then the random variable X satisfies (0.5) with $\beta=\beta(K,\alpha)$. Thus this condition can be verified in many examples by an easy computation.

The crucial idea in the proof of Theorem 0.3.1 was that any random variable X^j_i which satisfies (0.5) is "almost" a product of $\lceil\ln\beta\rceil$ independent variables with log-concave tails. Informally

$$X^j_i\approx\prod_{k=1}^{\lceil\ln\beta\rceil}Y^j_i(k), \tag{0.6}$$

where $(Y^j_i(k))_{i,j,k}$ are independent nonnegative r.v's with LCT. Thus

$$\|S'\|_p=\left\|\sum_{i_1,\ldots,i_d}a_{i_1,\ldots,i_d}X^1_{i_1}\cdots X^d_{i_d}\right\|_p\approx\left\|\sum_{i_1,\ldots,i_d}a_{i_1,\ldots,i_d}\prod_{l=1}^{d}\prod_{k=1}^{\lceil\ln\beta\rceil}Y^l_{i_l}(k)\right\|_p,$$

where the latter is chaos based on r.v's with LCT and the result from [19] can be applied.

0.3.2 *Tail and moment estimates for a class of random chaoses of order two*

We restrict our attention to the case $d=2$ so that $S'=\sum_{ij} a_{ij} X_i^1 X_j^2$. We assume that for any i,j, $\mathbb{E}(X_i^1)^2 = \mathbb{E}(X_j^2)^2 = 1$, the bound (0.5) holds and X_i^1, X_j^2 are symmetric (in particular a_{ij} can be negative). We define

$$\hat{N}_i^1(t) = \begin{cases} t^2 & \text{for } |t| < 1 \\ -\ln \mathbb{P}(|X_i^1| \geq |t|) & \text{for } |t| \geq 1 \end{cases},$$

$\hat{N}_j^2(t)$ is defined analogously. We denote

$$\|(a_{ij})\|_{X^1,X^2,p} = \sup\left\{ \sum_{ij} a_{ij} x_i y_j \;\Big|\; \sum_i \hat{N}_i^1(x_i) \leq p,\ \sum_j \hat{N}_j^2(y_j) \leq p \right\}, \tag{0.7}$$

$$\|(a_i)\|_{X^1,p} = \sup\left\{ \sum_i a_i x_i \;\Big|\; \sum_i \hat{N}_i^1(x_i) \leq p \right\},\quad \|(a_j)\|_{X^2,p} = \sup\left\{ \sum_j a_j y_j \;\Big|\; \sum_j \hat{N}_j^2(y_j) \leq p \right\}. \tag{0.8}$$

The main result of the chapter is the following theorem.

Theorem 0.3.2. *Under the above assumptions for any $p \geq 1$ we have*

$$\left\| \sum_{ij} a_{ij} X_i^1 X_j^2 \right\|_p \sim^{\beta} \|(a_{ij})\|_{X^1,X^2,p} + \left\| \left(\sqrt{\sum_j a_{ij}^2} \right) \right\|_{X^1,p} + \left\| \left(\sqrt{\sum_i a_{ij}^2} \right) \right\|_{X^2,p}. \tag{0.9}$$

The above theorem generalizes Latała's result [17]. It was a surprise that in both cases (chaoses based on r.v's with LCT and chaoses based on r.v's satisfying (0.5)) the moments can be estimated in the same way (cf. [17, Theorem 1]). The proof of Theorem 0.3.2 is based on three ideas (listed in chronological order as they were invented).

1. Use the decomposition (0.6) (in the symmetric case we decompose r.v's into a product of symmetric r.v's with LCT).
2. Introduce a family of technical norms and establish the moment estimates in the case $d=1$, namely show that $\left\|\sum_i a_i X_i^1\right\|_p \sim^{\beta} \|(a_i)\|_{X^1,p}$.
3. Reduce (0.9) to the case where $(X_j^2)_j$ have LCT.

Contrary to the nonnegative case the first idea is not enough. After using our decomposition we obtain

$$\left\| \sum_{ij} a_{ij} X_i^1 X_j^2 \right\|_p \approx \left\| \sum_{ij} a_{ij} \prod_{k=1}^{\lceil \ln \beta \rceil} X_i^1(k) \cdot \prod_{k=1}^{\lceil \ln \beta \rceil} X_i^2(k) \right\|_p =: \left\| S'_{dec} \right\|_p.$$

The main problem is that there is no counterpart of [19] in the symmetric case. The bounds in the general symmetric LCT case are known only for chaoses of order $d \leq 3$, while if $\beta \geq e^2$ then S'_{dec} is a chaos of order at least 4. However, S'_{dec} has a simple coefficient structure (support is concentrated on the generalized diagonal) and this property plays a fundamental role in the proof of (0.9).

The chapter naturally falls into two parts. In the first part we study the case of $d=1$. It would be expected that Latała's result (0.3) can be applied directly, but this is not this case (at least we did not manage to do it). Instead a direct tedious proof is presented. In the second part of the chapter, firstly we develop some decomposition lemmas concerning processes based on the symmetric r.v's satisfying (0.5) (which in our opinion are an independent interest). Then by applying few ideas and invoking results obtained in [1] we are able to show Theorem 0.3.2.

0.3.3 Hanson-Wright inequality in Banach spaces

This chapter describe an attempt to find two-sided estimates for moments of vector-valued Gaussian chaoses of order 2. In this instance $S' = \sum_{ij} a_{ij} g_i g'_j$ and $a_{ij} \in F$, where $(F, \|\cdot\|)$ is a Banach space. Such bounds were known in the literature in various contexts, even for arbitrary degree of the Gaussian chaos cf. [3, 5, 22, 24]. However, they involve quantities which are hard to work with (expectations of the suprema of norms evaluated on Gaussian vectors). In particular case $d = 2$ it was known that for any $p \geq 1$,

$$\left\|\sum_{ij} a_{ij} g_i g'_j\right\|_p \sim \mathbb{E}\left\|\sum_{ij} a_{ij} g_i g'_j\right\| + \sqrt{p}\,\mathbb{E}\sup_{\|x\|_2\leq 1}\left\|\sum_{ij} a_{ij} x_i g_j\right\| + p\sup_{\|x\|_2\leq 1, \|y\|_2\leq 1}\left\|\sum_{ij} a_{ij} x_i y_j\right\|. \tag{0.10}$$

The estimation of the problematic term $\mathbb{E}\sup_{\|x\|_2\leq 1}\left\|\sum_{ij} a_{ij} x_i g_j\right\|$ by quantities, which in many situations can be handled more easily, is the crucial result from the third chapter.

Proposition 0.3.3. *In the above setting for any $p \geq 1$, we have*

$$\mathbb{E}\sup_{\|x\|_2\leq 1}\left\|\sum_{ij} a_{ij} x_i g_j\right\| \leq p^{-1/2}\left(\mathbb{E}\left\|\sum_{ij} a_{ij} g_i g'_j\right\| + \mathbb{E}\left\|\sum_{ij} a_{ij} g_{ij}\right\|\right) + \sup_{\|x\|_2\leq 1}\mathbb{E}\left\|\sum_{ij} a_{ij} x_i g_j\right\| + \sup_{\|x\|_2\leq 1}\left\|\sum_{ij} a_{ij} x_{ij}\right\| + p^{1/2}\sup_{\|x\|_2\leq 1, \|y\|_2\leq 1}\left\|\sum_{ij} a_{ij} x_i y_j\right\|. \tag{0.11}$$

Unfortunately, (0.11) is to weak to provide a complete description of moments of a vector-valued Gaussian chaoses of order 2, because it is not true that in any Banach space $(F, \|\cdot\|)$

$$\left\|\sum_{ij} a_{ij} g_i g'_j\right\|_p \geq \frac{1}{C}\mathbb{E}\left\|\sum_{ij} a_{ij} g_{ij}\right\|.$$

The reason for appearance of the term $\mathbb{E}\left\|\sum_{ij} a_{ij} g_{ij}\right\|$ in (0.11) is the lackness of proper entropy estimates of certain sets in the Hilbert Schmidt norm. Moreover, it can be shown that in general such estimates cannot exist. But if

$$\mathbb{E}\left\|\sum_{ij} a_{ij} g_{ij}\right\| \leq \alpha\mathbb{E}\left\|\sum_{ij} a_{ij} g_i g'_j\right\| \quad \text{for any matrix } (a_{ij})_{ij} \text{ with values in } F, \tag{0.12}$$

where α is a constant which depends only on the Banach space F, (0.11) is enough to obtain two-sided bounds which are the main result of the third chapter.

Theorem 0.3.4. *Let $(a_{ij})_{ij}$ be a matrix with values in the Banach space $(F, \|\cdot\|)$. Then for any $p \geq 1$*

$$\left\|\sum_{ij} a_{ij} g_i g'_j\right\|_p \le C\Bigg(\mathbb{E}\left\|\sum_{ij} a_{ij} g_i g'_j\right\| + \mathbb{E}\left\|\sum_{ij} a_{ij} g_{ij}\right\| + \sqrt{p} \sup_{\|x\|_2\le 1} \mathbb{E}\left\|\sum_{ij} a_{ij} x_i g_j\right\|$$

$$+ \sqrt{p} \sup_{\|x\|_2\le 1} \|\sum_{ij} a_{ij} x_{ij}\| + p \sup_{\|x\|_2\le 1, \|y\|_2\le 1} \left\|\sum_{ij} a_{ij} x_i y_j\right\|\Bigg). \tag{0.13}$$

Additionally if (0.12) *holds then*

$$\left\|\sum_{ij} a_{ij} g_i g'_j\right\|_p \sim^{\alpha} \quad \mathbb{E}\left\|\sum_{ij} a_{ij} g_i g'_j\right\| + \sqrt{p} \sup_{\|x\|_2\le 1} \mathbb{E}\left\|\sum_{ij} a_{ij} x_i g_j\right\|$$

$$+ \sqrt{p} \sup_{\|x\|_2\le 1} \|\sum_{ij} a_{ij} x_{ij}\| + p \sup_{\|x\|_2\le 1, \|y\|_2\le 1} \left\|\sum_{ij} a_{ij} x_i y_j\right\|. \tag{0.14}$$

It is an easy exercise to show that (0.12) holds (with $\alpha \sim q$) in the L_q spaces. Moreover, in the case of L_q spaces the non deterministic expressions can be replaced by deterministic ones. We also expect that (0.14) is true in an arbitrary Banach space (with the constant independent on the geometry of the space) but we are unable to show this. However, we managed to show (0.14) up to a $\ln p$ factor (see Theorem 3.1.3 in the third chapter).

By standard arguments (0.13) yields the following Hanson-Wright type inequality in the Banach space (observe that in the assertion an undecoupled chaos occurs). In the below statement δ_{ij} stands for the Kronecker delta.

Theorem 0.3.5. *Let $X_1, X_2, \ldots$ be independent zero-mean $\alpha-$subgaussian random variables and $(a_{ij})_{ij}$ be a symmetric matrix with values in a normed space $(F, \|\cdot\|)$. Then for*

$$t > C\alpha^2 \left(\mathbb{E}\left\|\sum_{ij} a_{ij}(g_i g_j - \delta_{ij})\right\| + \mathbb{E}\left\|\sum_{i\neq j} a_{ij} g_{ij}\right\|\right)$$

we have

$$\mathbb{P}\left(\left\|\sum_{ij} a_{ij}(X_i X_j - \mathbb{E}(X_i X_j))\right\| \ge t\right) \le 2\exp\left(-\frac{1}{C}\min\left\{\frac{t^2}{\alpha^4 U^2}, \frac{t}{\alpha^2 V}\right\}\right), \tag{0.15}$$

where

$$U = \sup_{\|x\|_2\le 1} \mathbb{E}\left\|\sum_{i\neq j} a_{ij} x_i g_j\right\| + \sup_{\|(x_{ij})\|_2\le 1} \left\|\sum_{ij} a_{ij} x_{ij}\right\|,$$

$$V = \sup_{\|x\|_2\le 1, \|y\|_2\le 1} \left\|\sum_{ij} a_{ij} x_i y_j\right\|.$$

It is easy to check that in the case $F = \mathbb{R}$ Theorem 0.3.5 implies the classical Hanson-Wright bound (see Remark 3.1.8 in the third chapter).

0.3.4 Moments of Gaussian chaoses in Banach spaces

In this chapter we discuss moments of vector valued Gaussian chaoses of order bigger than 2 and L_q-valued exponential chaoses of an arbitrary order. The latter is deduced from the Gaussian case. The exponential random variable is "almost" a product of two standard Gaussian variables thus any exponential chaos of order d can be approximated by a Gaussian one of order $2d$. There are two reasons why we decide to divide our results concerning vector valued Gaussian chaoses. First, we wanted to make the case (probably the most interesting one) $d=2$ transparent. Secondly, the proofs of upper bounds for chaoses of orders $d=2$ and $d>2$ differ significantly.

The main idea (which arises from (0.11)) is to introduce Gaussian variables indexed by a group of indexes instead of singletons. For example, our main result Theorem 4.2.1 takes for $d=3$ the following form (we decided to quote the particular case of Theorem 4.2.1 in order to avoid introducing a new notation in the introduction).

Theorem 0.3.6. *Assume that $(a_{ijk})_{ijk}$ is a symmetric matrix with values in a normed space $(F,\|\cdot\|)$. Then for any $p\geq 1$, we have*

$$\begin{aligned} C^{-1}\Big\|\sum_{ijk} a_{ijk} g_i^1 g_j^2 g_k^3\Big\|_p \leq & \mathbb{E}\Big\|\sum_{ijk} a_{ijk} g_i^1 g_j^2 g_k^3\Big\| + \mathbb{E}\Big\|\sum_{ijk} a_{ijk} g_{ij}^1 g_k^2\Big\| + \mathbb{E}\Big\|\sum_{ijk} a_{ijk} g_{ijk}\Big\| \\ & + p^{1/2}\Big(\sup_{\|x\|_2\leq 1}\mathbb{E}\Big\|\sum_{ijk} a_{ijk} g_i^1 g_j^2 x_k\Big\| + \sup_{\|x\|_2\leq 1}\mathbb{E}\Big\|\sum_{ijk} a_{ijk} g_{ij}^1 x_k\Big\| \\ & + \sup_{\|x\|_2\leq 1}\mathbb{E}\Big\|\sum_{ijk} a_{ijk} g_i x_{jk}\Big\| + \sup_{\|x\|_2\leq 1}\Big\|\sum_{ijk} a_{ijk} x_{ijk}\Big\|\Big) \\ & + p\Big(\sup_{\|x\|_2,\|y\|_2\leq 1}\mathbb{E}\Big\|\sum_{ijk} a_{ijk} g_i x_j y_k\Big\| + \sup_{\|x\|_2,\|y\|_2\leq 1}\Big\|\sum_{ijk} a_{ijk} x_{ij} y_k\Big\|\Big) \\ & + p^{3/2}\sup_{\|x\|_2,\|y\|_2,\|z\|_2\leq 1}\Big\|\sum_{ijk} a_{ijk} x_i y_j z_k\Big\|. \end{aligned} \tag{0.16}$$

It is worth emphasizing that the bound obtained in Theorem 4.2.1 (so also (0.16)) holds in an arbitrary Banach space with a constant dependent only on the order of chaos (in particular independent of the geometry of the Banach space). Unfortunately, as for the case $d=2$ discussed above there is no chance in reversing it in general. However, it can be done (for any d) if (0.12) holds, so in particular (as we mentioned above) in L_q spaces (in which case all non deterministic expressions can be replaced by deterministic ones). It is a content of Theorem 4.2.1 and Corollary 4.2.8 which for $d=3$ takes the following form.

Theorem 0.3.7. *Assume that* (0.12) *holds in the normed space* $(F,\|\cdot\|)$ *and let* $(a_{ijk})_{ijk}$ *be a symmetric matrix with values in* F*. Then for any* $p\geq 1$*, we have*

$$\left\|\sum_{ijk} a_{ijk} g_i^1 g_j^2 g_k^3\right\|_p \sim^{\alpha} \mathbb{E}\left\|\sum_{ijk} a_{ijk} g_i^1 g_j^2 g_k^3\right\|$$
$$+p^{1/2}\left(\sup_{\|x\|_2\le 1}\mathbb{E}\left\|\sum_{ijk} a_{ijk} g_i^1 g_j^2 x_k\right\| + \sup_{\|x\|_2\le 1}\mathbb{E}\left\|\sum_{ijk} a_{ijk} g_i x_{jk}\right\| + \sup_{\|x\|_2\le 1}\left\|\sum_{ijk} a_{ijk} x_{ijk}\right\|\right)$$
$$+p\left(\sup_{\|x\|_2,\|y\|_2\le 1}\mathbb{E}\left\|\sum_{ijk} a_{ijk} g_i x_j y_k\right\| + \sup_{\|x\|_2,\|y\|_2\le 1}\left\|\sum_{ijk} a_{ijk} x_{ij} y_k\right\|\right)$$
$$+p^{3/2}\sup_{\|x\|_2,\|y\|_2,\|z\|_2\le 1}\left\|\sum_{ijk} a_{ijk} x_i y_j z_k\right\|.$$

Moreover, using methods from [2] we were able to generalize our result concerning decoupled chaoses, to any polynomial in Gaussian variables with coefficients from the normed space $(F,\|\cdot\|)$ (which in particular includes the case of undecoupled chaoses i.e, $S=\sum a_{i_1,\dots,i_d} g_{i_1,\dots,i_d}$). Our proof of the main results uses techniques introduced in [16]. Notably the heart of the proof is an estimation of the expectation of the supremum of a certain Gaussian process, namely the estimation of

$$\mathbb{E}\sup_{(x^2,\dots,x^d)\in U}\left\|\sum_{i_1,\dots,i_d} a_{i_1,\dots,i_d} g_{i_1} x_{i_2}^2\cdots x_{i_d}^d\right\| := F_d(U),$$

where $U\subset (B_2^n)^{d-1}$. Unfortunately, the proof is very technical and there is a simple reason behind it. Observe that for canonical Gaussian processes we have

$$G(T) := \mathbb{E}\sup_{t\in T}\sum_i t_i g_i = \mathbb{E}\sup_{t\in T}\sum_i (t_i+v_i) g_i = G(T+v),$$

where $T\subset\mathbb{R}^n$ and $v\in\mathbb{R}^n$. Thus the functional G is translation invariant. This is not the case concerning the functional F_d resulting in technical problems in the chaining argument. In particular for $d=2$ the functional F_2 takes the form $F_2(U)=\mathbb{E}\sup_{x\in U}\left\|\sum_{ij} a_{ij} g_i x_j\right\|$, and it is not true that $F_2(U+v)=F_2(U)$.

0.3.5 Moments estimates for some types of chaos in Banach spaces

We follow the notation used in Subsection 0.3.2, but we assume this time that $X_1^1, X_2^2, \dots, X_1^2, X_2^2, \dots$ are symmetric, independent r.v's with LCT and $(a_{ij})_{ij}$ is a matrix with values in the Banach space $(F,\|\cdot\|)$. We want to estimate the moments of a vector-valued chaos $S'=\sum_{ij} a_{ij} X_i^1 X_j^2$. While proving the results described in Subsection 0.3.3 it was clear that by modification of the proof of the entropy estimates, the entropy estimates needed in the aforementioned case can be proved. However, it was also clear that the equivalent of the [1, Theorem 7.2] is needed. Without going into details we would like to decompose a "big" set into a sum of "not too many" sets which are "small". By "small" set we mean that a supremum of a certain Gaussian process indexed by this set is "small". The problem is that in the vector case the "big" set has a worse structure than in the real case. Unfortunately, we did not manage to derive an equivalent of [1, Theorem 7.2]. Our main idea is that a simpler decomposition can be used under the assumption that the r.v's $X_1^2, X_2^2, \dots$ are subgaussian. As a result we obtain the following upper bound on moments of S' (with the additional subgaussianity assumption) with values in a general Banach space. Before we state it we recall the norms $\|\cdot\|_{X^1,X^2,p}, \|\cdot\|_{X^1,p}, \|\cdot\|_{X^2,p}$ defined in (0.7) and (0.8).

Theorem 0.3.8. *Suppose that $X_1^1, X_2^1, \ldots, X_1^2, X_2^2, \ldots$ are symmetric and independent r.v's with LCT such that $\mathbb{E}(X_i^j)^2 = 1$ for $j = 1,2,\ i \geq 1$. Assume also that $X_1^2, X_2^2, \ldots$ are α-subgausian. Let $(a_{ij})_{ij}$ be a matrix with values in the Banach space $(F, \|\cdot\|)$. Then for any $p \geq 1$, we have*

$$C^{-1}\left\|\left\|\sum_{ij} a_{ij} X_i^1 X_j^2\right\|\right\|_p \leq \mathbb{E}\left\|\sum_{ij} a_{ij} X_i^1 X_j^2\right\| + \alpha\left(\mathbb{E}\left\|\sum_{ij} a_{ij} g_{ij}\right\| + \mathbb{E}\left\|\sum_{ij} a_{ij} g_i \mathcal{E}_j\right\|\right)$$
$$+ \sup\left\{\mathbb{E}\left\|\sum_{ij} a_{ij} x_i^1 X_j^2\right\| \mid \sum_i \hat{N}_i^1(x_i^1) \leq p\right\}$$
$$+ \sup\left\{\mathbb{E}\left\|\sum_{ij} a_{ij} X_i^1 x_j^2\right\| \mid \sum_j \hat{N}_j^2(x_j^2) \leq p\right\}$$
$$+ \sup_{\substack{\varphi \in F^* \\ \|\varphi\|_* \leq 1}} \left\|\left(\sqrt{\sum_i \varphi(a_{ij})^2}\right)_j\right\|_{X^2,p} + \sup_{\substack{\varphi \in F^* \\ \|\varphi\|_* \leq 1}} \|(\varphi(a_{ij}))_{ij}\|_{X^1,X^2,p}, \qquad (0.17)$$

where $(F^*, \|\cdot\|_*)$ is the dual space to $(F, \|\cdot\|)$.

As in the Gaussian case, this bound turns out to be two-sided in the case of L_q-values chaoses. The crucial property of L_q space, which is used is that for any chaos S' based on the variables with LCT we have $\mathbb{E}\|S'(t)\|_{L_q(T)} \sim^q \left\|\sqrt{\mathbb{E}|S'(t)|^2}\right\|_{L_q(T)}$ (observe that the latter expression is computable).

Theorem 0.3.9. *Let $X_1^1, X_2^2, \ldots, X_1^2, X_2^2, \ldots$ be as in Theorem 0.3.8 and $(a_{ij})_{ij}$ be a matrix with values in $L_q(T, d\mu)$. Then for any $p \geq 1$, we have*

$$\left\|\left\|\sum_{ij} a_{ij} X_i^1 X_j^2\right\|\right\|_p \sim^{q,\alpha} \mathbb{E}\left\|\sum_{ij} a_{ij} X_i^1 X_j^2\right\|_{L_q} + \sup\left\{\mathbb{E}\left\|\sum_{ij} a_{ij} x_i^1 X_j^2\right\|_{L_q} \mid \sum_i \hat{N}_i^1(x_i^1) \leq p\right\}$$
$$+ \sup\left\{\mathbb{E}\left\|\sum_{ij} a_{ij} X_i^1 x_j^2\right\|_{L_q} \mid \sum_j \hat{N}_j^2(x_j^2) \leq p\right\} \qquad (0.18)$$
$$+ \sup_{\substack{\varphi \in L_q^* \\ \|\varphi\|_* \leq 1}} \left\|\left(\sqrt{\sum_i \varphi(a_{ij})^2}\right)_j\right\|_{X^2,p} + \sup_{\substack{\varphi \in L_q^* \\ \|\varphi\|_* \leq 1}} \|(\varphi(a_{ij}))_{ij}\|_{X^1,X^2,p}, \qquad (0.19)$$

where $(L_q^*, \|\cdot\|_*) = (L_{q^*}, \|\cdot\|_{L_{q^*}})$ is the dual space to $L_q(T, d\mu)$.

Without quoting the result, we will only mention that in the case of the Hilbert space $L_2(T, d\mu)$ the assumption about subgaussanity can be dropped.

In this chapter we have applied most of the techniques used and developed in previous chapters. For this reason it is a good summary of this PHD dissertation.

0.4 Acknowledgment

First and foremost, I would like to thank my supervisor Rafał Latała for all the time devoted to me, for his professionalism and for teaching me the probability theory.

Secondly, I would like to thank my wife for supporting me while writing this dissertation. It was her efforts that allowed me to find the right amount of time to work.

Then I would like to thank my friends: Marta na Michał Strzeleccy, for their help especially in filling out various forms and applications; Katarzyna Kmieć, for the language correction of my works; Grzegorz Bokota, for helping with Latex.

I would also like to thank the whole Section of Probability Theory in our Department for creating a good atmosphere for studying.

Last but not least I would like to thank the National Science Center in Poland for supporting my research via grant NCN 2015/18/A/ST1/00553 and NCN 2018/29/N/ST1/00454. I acknowledge also the financial support of Warsaw Center of Mathematics and Computer Sciences.

Part I
Real case

Chapter 1
Two-sided moment estimates for a class of nonnegative chaoses

In this Chapter we study homogeneous tetrahedral chaoses of order $d \in \mathbb{N}$, i.e. random variables of the form

$$S = \sum_{1 \le i_1,\ldots,i_d \le n} a_{i_1,\ldots,i_d} X_{i_1} \cdots X_{i_d},$$

where $(a_{i_1,\ldots,i_d})$ is a multi-indexed symmetric array of real numbers such that $a_{i_1,\ldots,i_d} = 0$ if $i_l = i_m$ for some $m \neq l$, $m, l \le d$. We derive two-sided bounds for $\|S\|_p$ under the assumption that the coefficients $(a_{i_1,\ldots,i_d})$ are nonnegative and (X_i) are independent, nonnegative and satisfy the following moment condition for some $k \in \mathbb{N}$,

$$\|X_i\|_{2p} \le 2^k \|X_i\|_p \quad \text{for every } p \ge 1. \tag{1.1}$$

As we mentioned in the introduction, the main idea is that if a r.v. X_i satisfy (1.1) then it is comparable with a product of k i.i.d. variables with logarithmically concave tails. In this way the problem reduces to the result of Latała and Łochowski [19] which gives two-sided bounds for moments of nonnegative chaoses generated by r.v's with logarithmically concave tails.

1.1 Notation and main results

We set $\|Y\|_p = (\mathbb{E}|Y|^p)^{1/p}$ for a real r.v. Y and $p \ge 1$, $\log(x) = \log_2(x)$ and ln stands for the natural logarithm. By C, t_0 (sometimes $C(k,d), t_0(k,d)$) we denote constants that may depend on k, d and may vary from line to line. We write $A \sim^{k,d} B$ if $A \cdot C(k,d) \ge B$ and $B \cdot C(k,d) \ge A$.

Let $\{X_i^{(1)}\}, \ldots, \{X_i^{(d)}\}$ be independent r.v's. We set

$$N_i^{(r)}(t) = -\ln \mathbb{P}(X_i^{(r)} \ge t).$$

We say that $X_i^{(r)}$ has logarithmically concave tails if the function $N_i^{(r)}$ is convex. We put

$$B_p^{(r)} = \left\{ v \in \mathbb{R}_+^n \mid \sum_{i=1}^n N_i^{(r)}(v_i) \le p \right\}$$

and

$$\left\|(a_{i_1,\ldots,i_d})\right\|_p = \sup \left\{ \sum_{1 \le i_1,\ldots,i_d \le n} a_{i_1,\ldots,i_d} \prod_{r=1}^d \left(1 + v_{i_r}^{(r)}\right) \mid \left(v_i^{(r)}\right) \in B_p^{(r)} \right\}.$$

The main result of this chapter is the following theorem.

Theorem 1.1.1. *Let* $(X_i^{(r)})_{r\le d, i\le n}$ *be independent nonnegative random variables satisfying* (1.1) *and* $\mathbb{E}X_i^{(r)}=1$. *Then for any nonnegative coefficients* $(a_{i_1,\ldots,i_d})_{i_1,\ldots,i_d\le n}$ *we obtain*

$$\frac{1}{C(k,d)}\left\|(a_{i_1,\ldots,i_d})\right\|_p \le \left\|S'\right\|_p \le C(k,d)\left\|(a_{i_1,\ldots,i_d})\right\|_p,$$

where

$$S' = \sum_{1\le i_1,\ldots,i_d\le n} a_{i_1,\ldots,i_d} X_{i_1}^{(1)}\cdots X_{i_d}^{(d)}.$$

Theorem 1.1.1 yields the following two-sided bounds for tails of random chaoses.

Theorem 1.1.2. *Under the assumptions of Theorem 1.1.1 there exist constants* $0<c(k,d), C(k,d)<\infty$ *depending only on* k *and* d *such that for any* $p\ge 1$ *we have*

$$\mathbb{P}\left(S' \ge C(k,d)\left\|(a_{i_1,\ldots,i_d})\right\|_p\right) \le e^{-p}$$

and

$$\mathbb{P}\left(S' \ge c(k,d)\left\|(a_{i_1,\ldots,i_d})\right\|_p\right) \ge \min(c(k,d), e^{-p}).$$

Proof. Theorem 1.1.1 and Chebyshev's inequality yield

$$\mathbb{P}\left(S' \ge C(k,d)\left\|(a_{i_1,\ldots,i_d})\right\|_p\right) \le \mathbb{P}\left((S')^p \ge e^p\mathbb{E}(S')^p\right) \le e^{-p}.$$

Now observe that Theorem 1.1.1 and inequality (1.3) below imply that

$$\left\|S'\right\|_{2p} \le C(k,d)\left\|S'\right\|_p.$$

Theorem 1.1.1, the Paley–Zygmund inequality (see Corollary 6.1) and the above inequality give

$$\mathbb{P}\left(S' \ge c(k,d)\left\|(a_{i_1,\ldots,i_d})\right\|_p\right) \ge \mathbb{P}\left((S')^p \ge c(k,d)\mathbb{E}(S')^p\right) \ge c(k,d)\left(\frac{\|S'\|_p}{\|S'\|_{2p}}\right)^p \ge \min(c(k,d), e^{-p}).$$

□

Now we present two-sided bounds for decoupled chaoses. We define in this case $N_i(t) = -\ln\mathbb{P}(X_i\ge t)$,

$$B_p = \left\{ v\in\mathbb{R}_+^n \mid \sum_{i=1}^n N_i(v_i)\le p \right\}$$

and

$$\left\|(a_{i_1,\ldots,i_d})\right\|_p' = \sup\left\{ \sum_{1\le i_1,\ldots,i_d\le n} a_{i_1,\ldots,i_d}\prod_{r=1}^d\left(1+v_{i_r}^{(r)}\right) \mid \left(v_i^{(r)}\right)\in B_p \right\}.$$

Theorem 1.1.3. *Let* $(X_i)_{i\le n}$ *be nonnegative independent r.v's satisfying* (1.1) *and* $\mathbb{E}X_i=1$. *Then for any symmetric array of nonnegative coefficients* $(a_{i_1,\ldots,i_d})_{i_1,\ldots,i_d\le n}$ *such that*

$$a_{i_1,\ldots,i_d}=0 \text{ if } i_l=i_m \text{ for some } m\ne l,\ m,l\le d \tag{1.2}$$

we get

$$\frac{1}{C(k,d)}\left\|(a_{i_1,\ldots,i_d})\right\|_p' \le \|S\|_p \le C(k,d)\left\|(a_{i_1,\ldots,i_d})\right\|_p',$$

where $S=\sum_{1\le i_1,\ldots,i_d\le n} a_{i_1,\ldots,i_d}X_{i_1}\cdots X_{i_d}$. *Moreover,*

$$\mathbb{P}\left(S \geq C(k,d)\left\|(a_{i_1,\dots,i_d})\right\|_p'\right) \leq e^{-p},$$
$$\mathbb{P}\left(S \geq c(k,d)\left\|(a_{i_1,\dots,i_d})\right\|_p'\right) \geq \min(c(k,d), e^{-p}).$$

Proof. Let $S' = \sum a_{i_1,\dots,i_d} X_{i_1}^{(1)} \cdot \ldots \cdot X_{i_d}^{(d)}$ be the decoupled version of S. By Theorem 6.6 moments and tails of S and S' are comparable up to constants which depend only on d. Hence Theorem 1.1.3 follows by Theorems 1.1.1 and 1.1.2. □

1.2 Preliminaries

In this section let us study properties of nonnegative r.v's satisfying the condition (1.1). We will assume normalization $\mathbb{E}X = 1$ and define $N(t) = -\ln \mathbb{P}(X \geq t)$.

Lemma 1.2.1. *There exists a constant $C = C(k)$ such that for any $x \geq 1$, and $t \geq 1$ we have $N(Ctx) \geq t^{\frac{1}{k}} N(x)$. $C = 8^{k+1}$ may be taken.*

Observe that Lemma 1.2.1 implies that

$$\left\|(a_{i_1,\dots,i_d})\right\|_{2p} \leq C(k,d)\left\|(a_{i_1,\dots,i_d})\right\|_p. \tag{1.3}$$

Proof of Lemma 1.2.1. Our purpose is to show that

$$\mathbb{P}(X \geq Ctx) \leq \mathbb{P}(X \geq x)^{t^{\frac{1}{k}}} \text{ for } t \geq 1,\ x \geq 1. \tag{1.4}$$

It is enough to prove the assertion for $x < \frac{\|X\|_\infty}{2}$ because for $x \geq \frac{\|X\|_\infty}{2}$, (1.4) holds if $C > 2$. In that case $x = \frac{1}{2}\|X\|_q$ for some $q \geq 1$ (since $\|X\|_1 = 1$). From the Paley-Zygmunt inequality (see Corollary 6.1) and (1.1)

$$\mathbb{P}(X \geq x) = \mathbb{P}\left(X^q \geq \frac{1}{2^q}\mathbb{E}X^q\right) \geq \left(1 - \frac{1}{2^q}\right)^2 \left(\frac{\|X\|_q}{\|X\|_{2q}}\right)^{2q}$$
$$\geq \left(1 - \frac{1}{2^q}\right)^2 \frac{1}{2^{2kq}} \geq \left(\frac{1}{2^{k+1}}\right)^{2q}. \tag{1.5}$$

Let $A = \lceil \frac{1}{k}\log(t) \rceil$. By (1.1) we get $\|X\|_{q2^A} \leq 2^{kA}\|X\|_q$. Hence, Chebyshev's inequality yields

$$\mathbb{P}(X \geq Ctx) = \mathbb{P}\left(X \geq \frac{Ct}{2}\|X\|_q\right) \leq \mathbb{P}\left(X \geq \frac{Ct}{2^{kA+1}}\|X\|_{q2^A}\right) \leq \left(\frac{2^{kA+1}}{Ct}\right)^{q2^A}.$$

We have $2^A \geq t^{\frac{1}{k}}$ and $2^{kA} \leq 2^{k\left(\frac{1}{k}\log(t)+1\right)} = t2^k$ so if $C = 8^{k+1}$ then

$$\mathbb{P}(X \geq Ctx) \leq \left(\frac{1}{4^{k+1}}\right)^{qt^{\frac{1}{k}}} = \left(\frac{1}{2^{k+1}}\right)^{2qt^{\frac{1}{k}}}. \tag{1.6}$$

The assertion follows by (1.5) and (1.6) □

In fact, the statement of Lemma 1.2.1 may be also reversed.

Remark 1.2.2. Let X be a nonnegative random variable, $\mathbb{E}X = 1$ and there exist constants $C, \beta > 0$ such that $N(Ctx) \geq t^\beta N(x)$ for $t, x \geq 1$. Then, there exists $\bar{K} = \bar{K}(C, \beta)$ such that

$$\|X\|_{2p} \leq \bar{K}\,\|X\|_p \quad \text{for } p \geq 1.$$

Proof. In this proof K means a constant which may depend on C, β and varies from line to line. Integration by parts yields

$$\mathbb{E}\left|\frac{X}{2C}\right|^{2p} = \int_0^\infty 2pt^{2p-1}e^{-N(2Ct)}dt \leq \|X\|_p^{2p} + \int_{\|X\|_p}^\infty 2pt^{2p-1}e^{-N(2Ct)}dt$$
$$\leq \|X\|_p^{2p} + \int_{\|X\|_p}^\infty 2pt^{2p-1}e^{-N(2\|X\|_p)(\frac{t}{\|X\|_p})^\beta}\,dt. \tag{1.7}$$

Let $\alpha = N(2\|X\|_p)^{\frac{1}{\beta}}/\|X\|_p$, substituting $y = \alpha t$ into (1.7) we get

$$\mathbb{E}\left|\frac{X}{2C}\right|^{2p} \leq \|X\|_p^{2p} + \frac{1}{\alpha^{2p}}\int_0^\infty 2py^{2p-1}e^{-y^\beta}dy.$$

Thus

$$\|X\|_{2p} \leq 2C\,\|X\|_p + \frac{2C}{\alpha}\left(\int_0^\infty 2py^{2p-1}e^{-y^\beta}dy\right)^{1/2p}$$
$$= 2C\,\|X\|_p + \frac{2C}{\alpha}\left(\frac{2p}{\beta}\Gamma(\frac{2p}{\beta})\right)^{1/2p} \leq 2C\,\|X\|_p + K\frac{p^{\frac{1}{\beta}}}{\alpha}. \tag{1.8}$$

By Chebyshev's inequality $N(2\|X\|_p) \geq p\ln 2$ and the assertion follows by (1.8). □

Now, we state the crucial technical lemma.

Lemma 1.2.3. *There exist C, t_0 which depend on k, a probability space with a version of X and nonnegative i.i.d. r.v's $Y_1, ..., Y_k$ with the following properties*

(i) $C(X + t_0) \geq Y_1 \cdots Y_k$,
(ii) $C(Y_1 \cdots Y_k + t_0) \geq X$,
(iii) $Y_1, \ldots, Y_k$ *have log-concave tails,*
(iv) $H(t) \leq N(t^k) \leq H(Ct)$ *for* $t \geq t_0$, *where* $H(t) = -\ln \mathbb{P}(Y_l \geq t)$,
(v) $\frac{1}{C} \leq \mathbb{E}Y_l \leq C$.

Proof. Let $M(t) = N(t^k)$. By Lemma 1.2.1 there exists C (depending on k) such that $M(C\lambda t) \geq \lambda M(t)$ for all $\lambda \geq 1,\ t \geq 1$. By Lemma 6.2 (applied with $t_0 = 1$) there exists a convex nondecreasing function H and constants $C = C(k), t_0 = t_0(k) > 0$ such that

$$H(t) \leq M(t) \leq H(Ct) \text{ for } t \geq t_0$$
$$H(t) = 0 \text{ for } t \leq t_0 \tag{1.9}$$

Let Y_i be nonnegative i.i.d, r.v's such that $\mathbb{P}(Y_l \geq t) = e^{-H(t)}$, then (iii) and (iv) hold.

Now, we verify (i) and (ii). For $t \geq \max\{1, t_0\}$ we obtain

$$\mathbb{P}\left(\prod_{l=1}^k Y_l \geq t\right) \geq \prod_{l=1}^k \mathbb{P}\left(Y_l \geq t^{\frac{1}{k}}\right) = e^{-kH(t^{\frac{1}{k}})} \geq e^{-kM(t^{\frac{1}{k}})} = e^{-kN(t)}$$
$$\geq e^{-N(Ck^k t)} = \mathbb{P}\left(X \geq Ck^k t\right), \tag{1.10}$$

where the last inequality comes from Lemma 1.2.1. Furthermore,

$$\mathbb{P}\left(\prod_{l=1}^{k} Y_l \ge C^k t\right) \le \sum_{l=1}^{k} \mathbb{P}(Y_l \ge C t^{\frac{1}{k}}) = k e^{-H(Ct^{\frac{1}{k}})} \le k e^{-M(t^{\frac{1}{k}})} = k e^{-N(t)}. \tag{1.11}$$

By Chebyshev's inequality $1 = \mathbb{E}X \ge e\mathbb{P}(X \ge e) = e^{1-N(e)}$, so $N(e) \ge 1$ and by Lemma 1.2.1 we get for $t \ge 1$, $N(Cte) \ge t^{\frac{1}{k}} N(e) \ge t^{\frac{1}{k}}$. Thus

$$\ln k - N(t) \le -\frac{1}{2} N(t) \quad \text{for } t \ge eC \max(1, 2\ln k)^k. \tag{1.12}$$

Lemma 1.2.1 also gives $\frac{N(t)}{2} \ge N(\frac{t}{2^k C})$ for $t > 2^k C$, so from (1.12) and (1.11)

$$\mathbb{P}\left(\prod_{l=1}^{k} Y_l \ge C^k t\right) \le e^{-N(\frac{t}{2^k C})} = \mathbb{P}\left(X \ge \frac{t}{2^k C}\right). \tag{1.13}$$

Inequalities (1.10) and (1.13) implies (i) and (ii). To show (v) observe that

$$(\mathbb{E}Y_l)^k = \mathbb{E}Y_1 \cdots Y_k \le C\,(\mathbb{E}X + t_0) = C(1 + t_0)$$

an by (1.9)

$$\mathbb{E}Y_l \ge t_0 > 0$$

□

1.3 Proof of Theorem 1.1.1

Let $X_i^{(r)}$, $r \le d, i \le n$ satisfy the assumptions of Theorem 1.1.1. By Lemma 1.2.3 we may assume (enlarging if necessary the probability space) that there exist independent r.v's $Y_{i,l}^{(r)}$, $l \le k, r \le d, i \le n$ such that conditions (i)-(v) of Lemma 1.2.3 hold (for $X_i^{(r)}$ and $Y_{i,l}^{(r)}$ instead of X and Y_l). Let $H_i^{(r)}(x) := -\ln \mathbb{P}\left(Y_{i,l}^{(r)} \ge x\right)$ (observe that this function does not depend on l).

Let us start with the following Proposition.

Fact 1.3.1. *For any $p \ge 1$,*

$$\left\| \sum_{1 \le i_1, \ldots, i_d \le n} a_{i_1,\ldots,i_d} \prod_{r=1}^{d} X_{i_r}^{(r)} \right\|_p \sim^{k,d} \left\| \sum_{1 \le i_1, \ldots, i_d \le n} a_{i_1,\ldots,i_d} \prod_{r=1}^{d} \prod_{l=1}^{k} Y_{i_r,l}^{(r)} \right\|_p .$$

Proof. Lemma 1.2.3 (ii) yields

$$\begin{aligned} \left\| \sum_{1 \le i_1, \ldots, i_d \le n} a_{i_1,\ldots,i_d} \prod_{r=1}^{d} X_{i_r}^{(r)} \right\|_p &\le C^d \left\| \sum_{1 \le i_1, \ldots, i_d \le n} a_{i_1,\ldots,i_d} \prod_{r=1}^{d} \prod_{l=1}^{k} \left(Y_{i_r,l}^{(r)} + t_0\right) \right\|_p \\ &\le C^d \sum_{\substack{\varepsilon_r \in \{0,1\} \\ r=1,\ldots,d}} \left\| \sum_{1 \le i_1, \ldots, i_d \le n} a_{i_1,\ldots,i_d} \prod_{r=1}^{d} \prod_{l=1}^{k} \left((Y_{i_r,l}^{(r)})^{\varepsilon_r} t_0^{1-\varepsilon_r}\right) \right\|_p . \end{aligned} \tag{1.14}$$

We have $\mathbb{E}Y_{i,l}^{(r)} \geq \frac{1}{C}$, so by Jensen's inequality for any $\varepsilon \in \{0,1\}^d$ we get,

$$\left\|\sum_{1\le i_1,\ldots,i_d\le n} a_{i_1,\ldots,i_d}\prod_{r=1}^{d}\prod_{l=1}^{k} Y_{i_r,l}^{(r)}\right\|_p \geq \left\|\sum_{1\le i_1,\ldots,i_d\le n} a_{i_1,\ldots,i_d}\prod_{r=1}^{d}\prod_{l=1}^{k} (Y_{i_r,l}^{(r)})^{\varepsilon_r}((\mathbb{E}Y_{i_r,l}^{(r)}))^{1-\varepsilon_r}\right\|_p$$
$$\geq \frac{1}{C^{kd}}\left\|\sum_{1\le i_1,\ldots,i_d\le n} a_{i_1,\ldots,i_d}\prod_{r=1}^{d}\prod_{l=1}^{k}\left((Y_{i_r,l}^{(r)})^{\varepsilon_r}(t_0)^{1-\varepsilon_r}\right)\right\|_p . \tag{1.15}$$

The lower estimate in Proposition 1.3.1 follows by (1.14) and (1.15). The proof of the upper bound is analogous. □

So to prove Theorem 1.1.1 we need to estimate $\left\|\sum a_{i_1,\ldots,i_d}\prod_{r=1}^{d}\prod_{l=1}^{k} Y_{i_r,l}^{(r)}\right\|_p$. For this purpose we will apply the following result of Latała and Łochowski.

Theorem 1.3.2 ([19, Theorem 2.1]). *Let $\{Z_i^{(1)}\},\ldots,\{Z_i^{(d)}\}$ be independent nonnegative r.v's with logarithmically concave tails and $M_i^{(r)}(t) = -\ln\left(\mathbb{P}\left(Z_i^{(r)} \geq t\right)\right)$. Let assume that $1 = \inf\{t>0 : M_i^{(r)}(t) \geq 1\}$. Then*

$$\frac{1}{C}\sup\left\{\sum_{1\le i_1,\ldots,i_d\le n} a_{i_1,\ldots,i_d}\prod_{r=1}^{d}(1+b_{i_r}^{(r)}) \;\middle|\; \left(b_i^{(r)}\right)\in T_p^{(r)}\right\}$$
$$\leq \left\|\sum_{1\le i_1,\ldots,i_d\le n} a_{i_1,\ldots,i_d} Z_{i_1}^{(1)}\ldots Z_{i_d}^{(d)}\right\|_p$$
$$\leq C\sup\left\{\sum_{1\le i_1,\ldots,i_d\le n} a_{i_1,\ldots,i_d}\prod_{r=1}^{d}(1+b_{i_r}^{(r)}) \;\middle|\; \left(b_i^{(r)}\right)\in T_p^{(r)}\right\},$$

where $T_p^{(r)} = \left\{b\in\mathbb{R}_+^n : \sum_{i=1}^n M_i^{(r)}(b_i)\leq p\right\}$.

To use the above result we need to normalize variables $Y_{i,l}^{(r)}$. Let

$$t_i^{(r)} = \inf\{t>0 \;:\; H_i^{(r)}(t)\geq 1\},\quad r\leq d,\ i\leq k.$$

Lemma 1.2.3 (v) gives $t_i^{(r)} \leq e\mathbb{E}Y_{i,l}^{(r)} \leq eC$ and by (1.9) $t_i^{(r)} \geq t_0 > 0$, thus

$$\frac{1}{C(k,d)} \leq t_i^{(r)} \leq C(k,d). \tag{1.16}$$

Theorem 1.3.2 applied to variables $Y_{i,l}^{(r)} = Y_{i,l}^{(r)}/t_i^{(r)}$ together with (1.16) gives

$$\left\|\sum_{i_1,\ldots,i_d} a_{i_1,\ldots,i_d}\prod_{r=1}^{d}\prod_{l=1}^{k} Y_{i_r,l}^{(r)}\right\|_p$$
$$\sim_{k,d}\sup\left\{\sum_{i_1,\ldots,i_d} a_{i_1,\ldots,i_d}\prod_{r=1}^{d}\prod_{l=1}^{k}\left(1+v_{i_r,l}^{(r)}\right) \;\middle|\; \left(v_{i,l}^{(r)}\right)\in D_{k,p}^{(r)},\ r=1,\ldots d\right\},$$

where

$$D_{k,p}^{(r)} = \left\{ \left(v_{i,l}\right)_{i\le n; l\le k} \in \mathbb{R}^{nk} \ : \ \sum_{i=1}^{n} H_i^{(r)}(v_{i,l}) \le p \text{ for all } l \le k \right\}.$$

Lemma 1.2.3 (iv) yields

$$\sup\left\{ \sum_{i_1,\ldots,i_d} a_{i_1,\ldots,i_d} \prod_{r=1}^{d}\prod_{l=1}^{k}\left(1+v_{i_r,l}^{(r)}\right) \mid \left(v_{i,l}^{(r)}\right) \in D_{k,p}^{(r)},\ r=1,\ldots d \right\}$$
$$\sim^{k,d} \sup\left\{ \sum_{i_1,\ldots,i_d} a_{i_1,\ldots,i_d} \prod_{r=1}^{d}\prod_{l=1}^{k}\left(1+v_{i_r,l}^{(r)}\right) \mid \left(v_{i,l}^{(r)}\right) \in B_{k,p}^{(r)}\ r=1,\ldots d \right\} =: \left\|(a_{i_1,\ldots,i_d})\right\|_{k,p}',$$

where

$$B_{k,p}^{(r)} = \left\{ \left(v_{i,l}\right)_{i\le n; l\le k} \in \mathbb{R}^{nk} \ : \ \sum_{i=1}^{n} N_i^{(r)}(v_{i,l}^{k}) \le p \text{ for all } l \le k \right\}.$$

To finish the proof of Theorem 1.1.1 we need to show that

$$\left\|(a_{i_1,\ldots,i_d})\right\|_{k,p}' \sim \left\|(a_{i_1,\ldots,i_d})\right\|_{p}. \tag{1.17}$$

First we will show this holds for $d=1$ that is

$$\sup\left\{ \sum_i b_i \prod_{l=1}^{k}(1+a_{i,l}) \mid \sum_i N_i(a_{i,l}^k) \le p \text{ for all } l\le k \right\} \sim^k \sup\left\{ \sum_i b_i(1+w_i) \mid \sum_i N_i(w_i) \le p \right\}. \tag{1.18}$$

We have

$$\sup\left\{ \sum_i b_i \prod_{l=1}^{k}(1+a_{i,l}) \mid \sum_i N_i(a_{i,l}^k) \le p \text{ for all } l\le k \right\}$$
$$\le \sum_{\substack{\varepsilon_l\in\{0,1\}\\ l=1\ldots k}} \sup\left\{ \sum_i b_i \prod_{l=1}^{k} a_{i,l}^{\varepsilon_l} \mid \sum_i N_i(a_{i,l}^k) \le p \text{ for all } l\le k \right\}.$$

So to establish the upper bound in (1.18) it is enough to prove

$$\sup\left\{ \sum_i b_i \prod_{l=1}^{k} a_{i,l}^{\varepsilon_l} \mid \sum_i N_i(a_{i,l}^k) \le p \text{ for all } l\le k \right\} \le C(k) \sup\left\{ \sum_i b_i(1+w_i) \mid \sum_i N_i(w_i) \le p \right\}$$

or equivalently (after permuting indexes) that for any $0\le k_0\le k$,

$$\sup\left\{ \sum_i b_i \prod_{l=1}^{k_0} a_{i,l} \mid \sum_i N_i(a_{i,l}^k) \le p \text{ for all } l\le k_0 \right\} \le C(k) \sup\left\{ \sum_i b_i(1+w_i) \mid \sum_i N_i(w_i) \le p \right\}. \tag{1.19}$$

Let us fix sequences $(a_{i,l})$ such that $\sum_i N_i(a_{i,l}^k) \le p$ for all $l\le k_0$. Let C be a constant from Lemma 1.2.1, define

$$w_i = \begin{cases} \frac{\prod_{l=1}^{k_0} a_{i,l}}{Ck^k} & \text{if } \prod_{l=1}^{k_0} a_{i,l} > 2Ck^k \\ 0 & \text{otherwise.} \end{cases}$$

For such w_i we have $\prod_{l=1}^{k_0} a_{i,l} \le 2Ck^k(1+w_i)$, so to establish (1.19) it is enough to check that $\sum_i N_i(w_i) \le p$. Lemma 1.2.1 yields

$$\sum_i N_i(w_i) \le \frac{1}{k}\sum_i N_i(Ck^k \cdot w_i) \le \frac{1}{k}\sum_{i:\ w_i \neq 0} N_i\left(\max\{a_{i,1},\ldots,a_{i,k_0}\}^{k_0}\right)$$
$$\le \frac{1}{k}\sum_i N_i\left(\max\{a_{i,1},\ldots,a_{i,k_0}\}^{k}\right) \le \frac{1}{k}\sum_i \sum_{l=1}^{k_0} N_i(a_{i,l}^k) \le p,$$

where the third inequality comes from the observation that $w_i \neq 0$ implies $\max\{a_{i,1},\ldots,a_{i,k_0}\} \ge 1$.

To show the lower bound in (1.18) we fix $w_i \in B_p$, choose $a_{i,1} = a_{i,2} = \ldots = a_{i,k} = w_i^{\frac{1}{k}}$ and observe that

$$\sum_i b_i \prod_{l=1}^{k}(1+a_{i,l}) = \sum_i b_i \left(1 + w_i^{\frac{1}{k}}\right)^k \ge \sum_i b_i(1+w_i).$$

We showed that (1.18) holds. Now we prove (1.17) for any d. We have

$$\left\|(a_{i_1,\ldots,i_d})\right\|'_{k,p} = \sup\left\{\sum_{i_1,\ldots,i_d} a_{i_1,\ldots,i_d} \prod_{r=1}^{d}\prod_{l=1}^{k}\left(1+v^{(r)}_{i_r,l}\right) \;\middle|\; \left(v^{(r)}_{i,l}\right) \in B^{(r)}_{k,p},\ r \le d\right\}$$
$$= \sup\left\{\sup\left\{\sum_{i_1,\ldots,i_d} a_{i_1,\ldots,i_d} \prod_{r=1}^{d}\prod_{l=1}^{k}\left(1+v^{(r)}_{i_r,l}\right) \;\middle|\; (v^{(d)}_{i,l}) \in B^{(d)}_{k,p}\right\} \;\middle|\; \forall_{r \le d-1}(v^{(r)}_{i,j}) \in B^{(r)}_{k,p}\right\}$$
$$\sim^k \sup\left\{\sum_{i_1,\ldots,i_d} a_{i_1,\ldots,i_d} \prod_{r=1}^{d-1}\prod_{l=1}^{k}\left(1+v^{(r)}_{i_r,l}\right)\left(1+w^{(d)}_{i_d}\right) \;\middle|\; (w^{(d)}) \in B^{(d)}_{p}, \forall_{r \le d-1} v^{(r)}_{i,j} \in B^{(r)}_{k,p}\right\},$$

where the last equivalence follows by (1.18). Iterating the above procedure d times we obtain (1.17).

Chapter 2

Tail and moment estimates for a class of random chaoses of order two.

The aim of this chapter is to derive two-sided bounds for moments and tails of real-valued random quadratic forms (chaoses of order two) $\sum_{i\neq j} a_{i,j} X_i X_j$ under the assumption that for any i and some $\alpha \geq 1$ we have

$$\|X_i\|_{2p} \leq \alpha \|X_i\|_p \quad \text{for any } p \geq 1. \tag{2.1}$$

Since any symmetric random variable X with log-concave tails satisfies (2.1) with $\alpha = 2$, this generalizes the previous result of Latała [17].

In the proof of main Theorem 2.1.1 below we use the same idea as in the previous chapter. We replace variables X_i by products of independent variables with log-concave tails. However, the situation is much more difficult to handle than in the nonnegative case, since in the symmetric log-concave case two-sided moment bounds are known only for chaoses of small order. Instead we first establish Gluskin-Kwapien-type bounds for moments of linear combinations, decouple quadratic forms, apply conditionally bounds for $d = 1$ and get to the point of estimating the L_p-norms of suprema of linear combinations of X_i's. Although formulas are similar as in Latała's paper [17], we cannot use his approach since our random variables do not satisfy nice dimension-free concentration inequalities. Instead we use a recent result of Latała and Strzelecka [21] and reduce the question to finding a right bound on L_1-norm of suprema. To treat this we use some ideas from [19] and [1].

2.1 Notation and main results

Unless otherwise stated, we assume that all vector are from $\mathbb{R}^N$ (we do not exclude $N = \infty$) and all matrices are real-valued. If $\mathbf{v}$ is a deterministic vector then $\|\mathbf{v}\|_r$, $r \in [1,\infty]$, is its l^r norm. We denote by $g_1, g_2, \ldots$ independent $\mathcal{N}(0,1)$ random variables and by $\varepsilon_1, \varepsilon_2, \ldots$ independent symmetric ± 1 random variables (Bernoulli sequence). We write $[n]$ for $\{1,\ldots,n\}$. If X is a r.v. then $\|X\|_p := (\mathbb{E}|X|^p)^{1/p}$. We say that a r.v. X belongs to the class $\mathbb{S}(d)$ if X is symmetric, $\|X\|_2 = 1/e$ and for every $p \geq 1$ $\|X\|_{2p} \leq 2^d \|X\|_p$ (the constant $1/e$ is chosen for technical reasons). For a sequence $(X_i)_{i\geq 1}$ we define the function $N_i^X(t) = -\ln \mathbb{P}(|X_i| \geq |t|) \in [0,\infty]$ and set

$$\hat{N}_i^X(t) := \begin{cases} t^2 & \text{for } |t| \leq 1, \\ N_i^X(t) & \text{for } |t| > 1. \end{cases} \tag{2.2}$$

Analogously we define $N_j^Y(t), \hat{N}_j^Y(t)$. The following three norms will play crucial role in this chapter:

$$\|(a_{i,j})\|_{X,Y,p} = \sup\left\{\sum_{i,j} a_{i,j}x_iy_j \;\middle|\; \sum_i \hat{N}_i^X(x_i) \le p,\ \sum_j \hat{N}_j^Y(y_j) \le p\right\},$$

$$\|(a_i)\|_{X,p} = \sup\left\{\sum_i a_ix_i \;\middle|\; \sum_i \hat{N}_i^X(x_i) \le p\right\}, \quad \|(a_j)\|_{Y,p} = \sup\left\{\sum_j a_jy_j \;\middle|\; \sum_j \hat{N}_j^Y(y_j) \le p\right\},$$

(see Lemma 2.2.1 for the proof, that they are norms).

To shorten notation, we write

$$m_p((a_{i,j})) = \|(a_{i,j})\|_{X,Y,p} + \left\|\left(\sqrt{\sum_i a_{i,j}^2}\right)_j\right\|_{Y,p} + \left\|\left(\sqrt{\sum_j a_{i,j}^2}\right)_i\right\|_{X,p},$$

$$\hat{m}_p((a_{i,j})) = \|(a_{i,j})\|_{X,X,p} + \left\|\left(\sqrt{\sum_i a_{i,j}^2}\right)_j\right\|_{X,p}.$$

By C, c we denote a universal constant which may differ at each occurrence. We also write $C(d), c(d)$ if the constants may depend on the parameter d. We write $a \sim b$ ($a \sim^d b$ resp.) if $b/C \le a \le Cb$ ($b/C(d) \le a \le C(d)b$ resp.).

Our main result is the following theorem.

Theorem 2.1.1. *Assume that $(X_i), (Y_j)$ are independent random variables from the $\mathbb{S}(d)$ class. Then for any finite matrix $(a_{i,j})$ and any $p \ge 1$,*

$$\left\|\sum_{i,j} a_{i,j}X_iY_j\right\|_p \sim^d m_p((a_{i,j})).$$

We postpone the proof of Theorem 2.1.1 till the end of this chapter and now present some corollaries. The first one shows that property (2.1) is preserved by the variable $\sum_{i,j} a_{i,j}X_iY_j$.

Corollary 2.1.2. *Under the assumptions of Theorem 2.1.1 we have*

$$\left\|\sum_{i,j} a_{i,j}X_iY_j\right\|_{2p} \le C(d)\left\|\sum_{i,j} a_{i,j}X_iY_j\right\|_p \tag{2.3}$$

Proof. Using Lemma 2.2.6 below twice we get

$$\begin{aligned}\|(a_{i,j})\|_{X,Y,2p} &= \sup\left\{\left\|\left(\sum_j a_{i,j}y_j\right)_i\right\|_{X,2p} \;\middle|\; \sum_j \hat{N}_j^Y(y_j) \le 2p\right\} \\ &\le C(d)\sup\left\{\left\|\left(\sum_j a_{i,j}y_j\right)_i\right\|_{X,p} \;\middle|\; \sum_j \hat{N}_j^Y(y_j) \le 2p\right\} \\ &= C(d)\sup\left\{\left\|\left(\sum_i a_{i,j}x_i\right)_j\right\|_{Y,2p} \;\middle|\; \sum_i \hat{N}_i^X(x_i) \le p\right\} \le C(d)\,\|(a_{i,j})\|_{X,Y,p}.\end{aligned}$$

The above estimate together with Lemma 2.2.6 and Theorem 2.1.1 yields the assertion. □

Standard arguments show how to get from moment to tail bounds.

Corollary 2.1.3. *Under the assumptions of Theorem 2.1.1 we have*

$$\mathbb{P}\left(\left|\sum_{i,j} a_{i,j} X_i Y_j\right| \geq C(d) m_p\left((a_{i,j})\right)\right) \leq e^{-p} \tag{2.4}$$

and

$$\mathbb{P}\left(\left|\sum_{i,j} a_{i,j} X_i Y_j\right| \geq c(d) m_p\left((a_{i,j})\right)\right) \geq e^{-c(d)p}. \tag{2.5}$$

Proof. The upper bound (2.4) is an immediate consequence of Chebyshev's inequality and Theorem 2.1.1. To establish the lower bound we have

$$\begin{aligned}\mathbb{P}\left(\left|\sum_{i,j} a_{i,j} X_i Y_j\right| \geq c(d) m_p\left((a_{i,j})\right)\right) &\geq \mathbb{P}\left(\left|\sum_{i,j} a_{i,j} X_i Y_j\right| \geq \frac{1}{2}\left\|\sum_{i,j} a_{i,j} X_i Y_j\right\|_p\right) \\ &\geq \left(1-\frac{1}{2^p}\right)^2 \left(\frac{\left\|\sum_{i,j} a_{i,j} X_i Y_j\right\|_p}{\left\|\sum_{i,j} a_{i,j} X_i Y_j\right\|_{2p}}\right)^{2p} \geq e^{-c(d)p},\end{aligned}$$

where the first inequality follows by Theorem 2.1.1, the second by the Paley-Zygmund inequality (cf. Corollary 6.1) and the last one by (2.3). □

We formulate undecoupled versions of Theorem 2.1.1 and Corollary 2.1.3.

Corollary 2.1.4. *Let $X_1, X_2, \ldots$ be independent r.v's from the $\mathbb{S}(d)$ class and $(a_{i,j})$ be a finite matrix such that $a_{i,i} = 0$ and $a_{i,j} = a_{j,i}$ for all i,j. Then for each $p \geq 1$,*

$$\left\|\sum_{i,j} a_{i,j} X_i X_j\right\|_p \sim^d \hat{m}_p\left((a_{i,j})\right), \tag{2.6}$$

$$\mathbb{P}\left(\left|\sum_{i,j} a_{i,j} X_i X_j\right| \geq C(d) \hat{m}_p\left((a_{i,j})\right)\right) \leq e^{-p} \tag{2.7}$$

and

$$\mathbb{P}\left(\left|\sum_{i,j} a_{i,j} X_i X_j\right| \geq c(d) \hat{m}_p\left((a_{i,j})\right)\right) \geq e^{-c(d)p}. \tag{2.8}$$

Proof. Moment estimate (2.6) is an immediate consequence of Theorem 2.1.1 and the Kwapień decoupling inequalities (Theorem 6.5).

We may derive tail bounds from the moment estimates in the similar way as in the undecoupled case. Alternatively we may use the more general decoupling result of de la Peña and Montgomery-Smith (Theorem 6.6) and get (2.7) and (2.8) from (2.4) and (2.5). □

Remark 2.1.5. A simple approximation argument shows that Theorem 2.1.1 and Corollaries 2.1.3, 2.1.4 hold for infinite square summable matrices $(a_{i,j})$.

We derive some examples from Corollary 2.1.4. Firstly we recover the special case of the Kolesko and Latała result [13, Example 3].

Example 2.1.6. *Let $X_1, X_2, \ldots$ be independent r.v's with symmetric Weibull distribution with scale parameter 1 and shape parameter $r \in (0,1]$, i.e. for any i $\mathbb{P}(|X_i| \geq t) = \exp(-t^r)$ for $t \geq 0$. Then for any $p \geq 1$ and any square summable matrix $(a_{i,j})$ such that $a_{i,i} = 0$ and $a_{i,j} = a_{j,i}$ for all i,j we have*

$$\left\| \sum_{i,j} a_{i,j} X_i X_j \right\|_p \sim^r p^{2/r} \sup_{i,j} |a_{i,j}| + p^{1/r+1/2} \sup_i \sqrt{\sum_j a_{i,j}^2}$$
$$+ p \sup \left\{ \sqrt{\sum_i \left(\sum_j a_{i,j} x_j \right)} \;\middle|\; \|x\|_2 = 1 \right\} + \sqrt{p} \sqrt{\sum_{i,j} a_{i,j}^2}. \tag{2.9}$$

Proof. By direct computation one may check that $\|X_i\|_{2p} \leq 2^{1/r} \|X_i\|_p$ (it may be also checked that (2.1) holds using Remark 1.2.2). First observe that

$$\|(v_i)\|_{X,p} \sim \sqrt{p} \|v\|_2 + p^{1/r} \sup_i |v_i|. \tag{2.10}$$

Indeed we have that $|x|^r \leq x^2$ for $|x| \geq 1$ and $|x|^r > x^2$ for $|x| < 1$ so

$$\|(v_i)\|_{X,p} \sim \sqrt{p} \|v\|_2 + p^{1/r} \sup \left\{ \sum_i v_i x_i \;\middle|\; \sum_i |x_i|^r \leq 1 \right\}.$$

Obviously $\sup\{\sum_i v_i x_i \mid \sum |x_i|^r \leq 1\} \geq \sup_i |v_i|$. Since $r \in (0,1]$ we have $\sum_i |x_i|^r \geq (\sum_i |x_i|)^{1/r}$ and as a result

$$\sup \left\{ \sum_i v_i x_i \;\middle|\; \sum |x_i|^r \leq 1 \right\} = \sup_i |v_i|$$

and (2.10) holds.

Iterating (2.10) we get

$$\|(a_{i,j})\|_{X,X,p} \sim p^{2/r} \sup_{i,j} |a_{i,j}| + p^{1/r+1/2} \sup_i \sqrt{\sum_j a_{i,j}^2} + p \sup \left\{ \sqrt{\sum_i \left(\sum_j a_{i,j} x_j \right)} \;\middle|\; \|x\|_2 = 1 \right\}. \tag{2.11}$$

The inequality (2.9) follows by Corollary 2.1.4, (2.10) and (2.11). □

Next example presents a situation when tails of X_i are neither log-concave nor log-convex, so it cannot be deduced from previous results.

Example 2.1.7. *Let $X_1, X_2, \ldots$ be i.i.d r.v's distributed as $W \mathbf{1}_{|W| \leq R}$, where $R > 1$ and W be a symmetric Weibull distribution with scale parameter 1 and shape parameter $r \in (0,1]$. Assume that $(a_{i,j})_{i,j \geq 0}$ is a square summable matrix such that $a_{i,i} = 0$, $a_{i,j} = a_{j,i}$. Denote $A_i = \sqrt{\sum_j a_{i,j}^2}$ and*

$$
\begin{aligned}
|||(a_{i,j})|||_{R,r,p} = &\sup\left\{\sum_{i,j} a_{i,j}x_iy_j \;\middle|\; \|x\|_2^2 \le p,\ \|x\|_\infty \le R,\ \|y\|_2^2 \le p,\ \|y\|_\infty \le R\right\}\\
&+\sup\left\{\sum_{i,j} a_{i,j}x_iy_j \;\middle|\; \|x\|_r^r \le p,\ \|x\|_\infty \le R,\ \|y\|_2^2 \le p,\ \|y\|_\infty \le R\right\}\\
&+\sup\left\{\sum_{i,j} a_{i,j}x_iy_j \;\middle|\; \|x\|_r^r \le p,\ \|x\|_\infty \le R,\ \|y\|_r^r \le p,\ \|y\|_\infty \le R\right\}.
\end{aligned}
$$

Then

$$
\left\|\sum_{i,j} a_{i,j}X_iX_j\right\|_p \sim^r \begin{cases} |||(a_{i,j})|||_{R,r,p} + \sqrt{p}\sqrt{\sum_{i,j} a_{i,j}^2} + p^{1/r}A_1^* & \textit{for } 1 \le p \le R^r,\\ |||(a_{i,j})|||_{R,r,p} + \sqrt{p}\sqrt{\sum_{i,j} a_{i,j}^2} + R\sum_{i\le p/R^r} A_i^* & \textit{for } R^r < p \le R^2,\\ |||(a_{i,j})|||_{R,r,p} + \sqrt{p}\sqrt{\sum_{i\ge p/R^2}(A_i^*)^2} + R\sum_{i\le p/R^r} A_i^* & \textit{for } R^2 < p, \end{cases}
$$

where (A_i^*) is a nonincreasing rearrangement of (A_i).

Proof. W.l.o.g we may assume that A_i is nonincreasing. Since $|t|^r \le t^2$ for $|t| \ge 1$ and $|t|^r \ge t^2$ for $|t| \le 1$ we obtain

$$
\|(A_i)_i\|_{X,p} \sim \sup\left\{\sum_i A_it_i \;\middle|\; \|t\|_2^2 \le p,\ |t_i| \le R\right\} + \sup\left\{\sum_i A_it_i \;\middle|\; \|t\|_r^r \le p,\ |t_i| \le R\right\} =: S_1 + S_2.
$$

Iteration of the above argument easily yields

$$
|||(a_{i,j})|||_{R,r,p} \sim \|(a_{i,j})\|_{X,X,p}. \tag{2.12}
$$

Now we will estimate S_1 and S_2 separately.

1. Obviously $S_1 = \sqrt{p}\sqrt{\sum_i A_i^2}$ for $p \le R^2$. Assume that $p > R^2$. By homogeneity and Lemma 2.2.3 below we get

$$
S_1 = R\left\{\sum_i A_it_i \;\middle|\; \sum_i t_i^2 \le \frac{p}{R^2},\ |t_i| \le 1\right\} \sim R\sum_{i\le p/R^2} A_i + \sqrt{p}\sqrt{\sum_{i\ge p/R^2} A_i^2}.
$$

2. It is easy to see that $S_2 = p^{1/r}A_1 = p^{1/r}\|(A_i)\|_\infty$ for $p \le R^r$. Assume that $p > R^r$. Since $0 < r \le 1$, by homogeneity we have

$$
S_2 = R\left\{\sum_i A_it_i \;\middle|\; \sum_i |t_i|^r \le \frac{p}{R^r},\ |t_i| \le 1\right\} \le R\left\{\sum_i A_it_i \;\middle|\; \sum_i |t_i| \le \frac{p}{R^r},\ |t_i| \le 1\right\} \le 2R\sum_{i\le p/R^r} A_i.
$$

   Moreover, by picking $t_1 = \ldots = t_{\lfloor p/R^r\rfloor} = 1$ and $t_{\lfloor p/R^r\rfloor+1} = \ldots = t_n = 0$ we obtain

$$
S_2 \ge R\sum_{i\le p/R^r} A_i.
$$

Since $R^2 \ge R^r$ we have that (recall (2.12))

$$\hat{m}_p\left((a_{i,j})\right) \sim \begin{cases} |||(a_{i,j})|||_{R,r,p} + \sqrt{p}\sqrt{\sum_i A_i^2} + p^{1/r}A_1 & \text{for } 1 \le p \le R^r \\ |||(a_{i,j})|||_{R,r,p} + \sqrt{p}\sqrt{\sum_i A_i^2} + R\sum_{i\le p/R^r} A_i & \text{for } R^r < p \le R^2 \\ |||(a_{i,j})|||_{R,r,p} + \sqrt{p}\sqrt{\sum_{i\ge p/R^2} A_i^2} + R\sum_{i\le p/R^r} A_i & \text{for } R^2 < p. \end{cases}$$

Now it is enough to observe that $\|X_1\|_2 \sim^r 1$, X_1 satisfies (2.1) with $\alpha = \alpha(r)$ (see Remark 1.2.2) and invoke Corollary 2.1.4.

□

Remark 2.1.8. In the Gaussian and Rademacher case Corollary 2.1.4 implies (see also Examples 1 and 2 in [17])

$$\left\|\sum_{i,j} a_{i,j} g_i g_j\right\|_p \sim p\,\|(a_{i,j})\|_{l^2\to l^2} + \sqrt{p}\,\|(a_{i,j})\|_2\,, \tag{2.13}$$

$$\left\|\sum_{i,j} a_{i,j} \varepsilon_i \varepsilon_j\right\|_p \sim |||(a_{i,j})|||_{1,2,p} + \sum_{i\le p} A_i^* + \sqrt{p}\sqrt{\sum_{i>p}(A_i^*)^2}, \tag{2.14}$$

where A_i^* is nonincreasing rearrangement of $A_i = \sqrt{\sum_j a_{i,j}^2}$. Neither (2.13) nor (2.14) can be expressed by a closed formula which do not involves suprema. Thus, there is no hope for any closed formulas in Examples 2.1.6 and 2.1.7.

The chapter is organized as follows. In the next section we present some technical facts used in the main proof. In Section 2.3 we establish Gluskin-Kwapień-type bounds for moments of linear combinations of X_i's. In Section 2.4 we obtain bounds for expected values of suprema and conclude the proof of Theorem 2.1.1 in Section 2.5.

2.2 Preliminary facts

We begin with three simple technical facts.

Lemma 2.2.1. *Assume that $T \subset \mathbb{R}^n$ is bounded and span$(T) = \mathbb{R}^n$. Then $\|x\| = \sup_{t\in T}|\sum_i x_i t_i|$ is a norm. In particular $\|(a_i)\|_{X,p}$, $\|(a_{i,j})\|_{X,Y,p}$ are norms.*

Proof. It is clear that $\|\cdot\|$ is well-defined, homogeneous and satisfies the triangle inequality. Now observe that $\|x\| = 0$ gives $\sum_i x_i t_i = 0$ for all $t \in span(T) = \mathbb{R}^n$, hence $x = 0$ and the first part of the assertion follows. Now let

$$\hat{T} = \left\{ (x_i y_j)_{i,j} \in \mathbb{R}^{n^2} \,\middle|\, \sum_i \hat{N}_i^X(x_i) \le p,\ \sum_j \hat{N}_j^Y(y_j) \le p \right\}.$$

Observe that $\hat{T}$ spans $\mathbb{R}^{n^2}$ and is bounded since for any r.v X, $N^X(t) = -\ln\mathbb{P}(|X| \ge t) \to \infty$ as $t \to \infty$. Analogously we prove that $\|\cdot\|_{X,p}$ is a norm. □

Fact 2.2.2. *For any $p \ge 1$, any set $T \subset \mathbb{R}^{\lfloor p\rfloor}$ which fulfills assumptions of Lemma 2.2.1 and any r.v's $Z_1, \ldots, Z_p$ we have*

$$\mathbb{E}\sup_{x\in T}\left|\sum_{i\leq p}x_iZ_i\right| \leq \left(\mathbb{E}\sup_{x\in T}\left|\sum_{i\leq p}x_iZ_i\right|^p\right)^{1/p} \leq 10\sup_{x\in T}\left\|\sum_i x_iZ_i\right\|_p.$$

Proof. Consider a norm on $\mathbb{R}^{\lfloor p\rfloor}$ given by

$$\|(y_1,\ldots,y_{\lfloor p\rfloor})\| = \sup_{x\in T}\left|\sum_{i\leq p}x_iy_i\right|.$$

Let K be the unit ball of the dual norm $\|\ \|_*$. Then $K=\mathrm{conv}(T\cup -T)$. Let M be a $1/2$ net in K (with respect to $\|\ \|_*$) of cardinality not larger than $5^{\lfloor p\rfloor}$ (M exists by standard volumetric arguments). Then for all $y\in\mathbb{R}^{\lfloor p\rfloor}$

$$\|y\| \leq 2\sup_{u\in M}\left|\sum_{i\leq p}u_iy_i\right|.$$

Thus

$$\mathbb{E}\sup_{x\in T}\left|\sum_{i\leq p}x_iZ_i\right|^p \leq 2^p\mathbb{E}\sup_{u\in M}\left|\sum_{i\leq p}u_iZ_i\right|^p.$$

Observe that

$$\left(\mathbb{E}\sup_{u\in M}\left|\sum_{i\leq p}u_iZ_i\right|^p\right)^{1/p} \leq \left(\sum_{u\in M}\mathbb{E}\left|\sum_{i\leq p}u_iZ_i\right|^p\right)^{1/p} \leq \left(5^p\sup_{u\in M}\mathbb{E}\left|\sum_{i\leq p}u_iZ_i\right|^p\right)^{1/p}$$
$$\leq 5\sup_{x\in T}\left\|\sum x_iZ_i\right\|_p.$$

□

Lemma 2.2.3. *For any $p\geq 1$ and any $a_1\geq a_2\geq\ldots\geq a_n\geq 0$, we have*

$$\frac{1}{2}\left(\sum_{i\leq p}a_i+\sqrt{p}\sqrt{\sum_{i>p}a_i^2}\right) \leq \sup\left\{\sum_i a_it_i\ \middle|\ \sum_i t_i^2\leq p,\ |t_i|\leq 1\right\} \leq \sum_{i\leq p}a_i+\sqrt{p}\sqrt{\sum_{i>p}a_i^2}.$$

Proof. Denote $M=\sup\{\sum_i a_it_i \mid \sum_i t_i^2\leq p,\ |t_i|\leq 1\}$. By choosing $t_i=1$ for $i\leq p$ and $t_i=0$ for $i>p$ we see that $M\geq\sum_{i\leq p}a_i$.

Now let $k=\lfloor p\rfloor+1$, $A=\sqrt{ka_k^2+\sum_{i>k}a_i^2}$, $t_i=\sqrt{p}a_k/A$ for $i\leq k$ and $t_i=\sqrt{p}a_i/a$ for $i>k$. Then $|t_i|\leq 1$, $\sum t_i^2=p$ and

$$M\geq\sum_i a_it_i\geq\sqrt{p}A\geq\sqrt{p}\sqrt{\sum_{i>p}a_i^2}.$$

To show the upper bound it is enough to observe that

$$\sum a_it_i\leq\|t\|_\infty\sum_{i\leq p}a_i+\|t\|_2\sqrt{\sum_{i>p}a_i^2}.$$

□

Now we slightly reformulate the crucial technical lemma from Chapter 1.

Lemma 2.2.4. *If X is from the $\mathbb{S}(d)$ class then there exists symmetric i.i.d r.v's $X^1,\dots,X^d$ on the extended probability space and a constant $t_0(d)\geq 1$ with the following properties:*

$$C(d)(|X|+1)\geq |X^1\cdots X^d| \quad \text{and } C(d)(|X^1\cdots X^d|+1)\geq |X|, \tag{2.15}$$

$$X^1,\dots,X^d \text{ have log-concave tails}, \tag{2.16}$$

$$M(t)\leq N(t^d)\leq M(C(d)t) \text{ for } t\geq t_0(d), \tag{2.17}$$

$$\text{where } M(t)=-\ln\mathbb{P}(|X^1|\geq t),$$

$$\frac{1}{C(d)}\leq \mathbb{E}|X^1|\leq C(d), \tag{2.18}$$

$$\inf\{t>0 \mid M(t)\geq 1\}=1. \tag{2.19}$$

Proof. From Lemma 1.2.3 we know that there exists symmetric i.i.d r.v's $X^1,\dots,X^d$ which satisfy (2.15)-(2.18) and $M(t)=0$ for $t<t_0(d)$ where $t_0(d)>0$ (see formula (1.9)). So

$$\inf\{t>0 \mid M(t)\geq 1\}\geq t_0(d).$$

By Chebyshev's inequality $M(3\mathbb{E}|X^1|)\geq \ln(3)>1$. Combining it with (2.18) yields

$$\inf\{t>0 \mid M(t)\geq 1\}\leq 3\mathbb{E}|X^1|\leq C(d).$$

So we have proved that

$$0<c(d)\leq \inf\{t>0 \mid M(t)\geq 1\}\leq C(d)<\infty.$$

The variables $X_i/\inf\{t>0 \mid M(t)\geq 1\}$ satisfy (2.15)-(2.19). □

Till the end of the chapter we assign to every X_i from the $\mathbb{S}(d)$ class the r.v's $X_i^1,\dots,X_i^d$ obtained by Lemma 2.2.4.

Denote for $t\in\mathbb{R}$, $M_i^X(t)=-\ln\mathbb{P}\left(|X_i^1|\geq |t|\right)\in[0,\infty]$ and

$$\hat{M}_i^X(t)=\begin{cases} t^2 & \text{for } |t|<1,\\ M_i^X(t) & \text{for } |t|\geq 1.\end{cases} \tag{2.20}$$

Observe that convexity of M_i^X and the normalization condition (2.19) imply

$$\hat{M}_i^X\left(\frac{t}{u}\right)\leq \frac{\hat{M}_i^X(t)}{u} \text{ for } u\geq 1, \tag{2.21}$$

and

$$\hat{M}_i^X(t)=M_i^X(t)\geq |t| \text{ for } |t|\geq 1. \tag{2.22}$$

We define the following technical norms (the proof that they are norms is the same as for $\|\cdot\|_{X,Y,p}$, see Lemma 2.2.1)

$$|||(a_i)|||_{X,p,1}=\sup\left\{\sum_i a_i x_i \,\middle|\, \sum_i \hat{M}_i^X(x_i)\leq p\right\} \tag{2.23}$$

and for $d>1$

$$|||(a_i)|||_{X,p,d} = \sup\left\{ \sum_i a_i x_i^1 \prod_{k=2}^{d} (1+x_i^k) \;\middle|\; \forall_{k=1,\ldots,d} \sum_i \hat{M}_i^X(x_i^k) \le p \right\} \tag{2.24}$$

Lemma 2.2.5. *For any $p \ge 1$ we have*

$$\|(a_i)\|_{X,p} \sim^d |||(a_i)|||_{X,p,d}\,.$$

Proof. Without loss of generality we can assume that a_i are nonnegative. Let $t_0(d)$ be a constant from Lemma 2.2.4. We have

$$\begin{aligned}\|(a_i)\|_{X,p} \le & \sup\left\{ \sum_i a_i b_i \mathbf{1}_{\{1>b_i\ge 0\}} \;\middle|\; \sum_i \hat{N}_i^X(b_i) \le p \right\} \\ & + \sup\left\{ \sum_i a_i b_i \mathbf{1}_{\{t_0(d)^d > b_i \ge 1\}} \;\middle|\; \sum_i \hat{N}_i^X(b_i) \le p \right\} \\ & + \sup\left\{ \sum_i a_i b_i \mathbf{1}_{\{b_i \ge t_0(d)^d\}} \;\middle|\; \sum_i \hat{N}_i^X(b_i) \le p \right\} =: I + II + III.\end{aligned}$$

The equality $\hat{N}_i^X(t) = \hat{M}_i^X(t)$ for $|t| \le 1$ implies

$$I \le \sup\left\{ \sum_i a_i x_i^1 \;\middle|\; \sum_i \hat{M}_i^X(x_i^1) \le p \right\} \le |||(a_i)|||_{X,p,d}\,.$$

Since $\|X_i\|_2 = 1/e$, Chebyshev's inequality yields $\hat{N}_i^X(s) \ge 1$ for $s \ge 1$. Hence

$$\begin{aligned} II &= \sup\left\{ \sum_{i\in I} a_i b_i \mathbf{1}_{\{t_0(d)^d > b_i \ge 1\}} \;\middle|\; \sum_{i\in I} \hat{N}_i^X(b_i) \le p,\ |I| \le \lfloor p \rfloor \right\} \le t_0(d)^d \sup_{|I|=\lfloor p\rfloor} \sum_{i\in I} a_i \\ &\le t_0(d)^d\, |||(a_i)|||_{X,p,d}\,.\end{aligned}$$

To see the last inequality it is enough to take in (2.24) $x_i^1 = \mathbf{1}_{\{i\in I\}}$ and $x_i^2 = \ldots = x_i^d = 0$. From (2.17) we obtain

$$III \le \sup\left\{ \sum_i a_i (b_i)^d \;\middle|\; \sum_i \hat{M}_i^X(b_i) \le p \right\} \le |||(a_i)|||_{X,p,d}\,,$$

where to get the last inequality we take $x_i^1 = \ldots = x_i^d = b_i$.

It remains to show

$$|||(a_i)|||_{X,p,d} \le C(d)\, \|(a_i)\|_{X,p}\,. \tag{2.25}$$

By an easy computation

$$\frac{1}{(1+C(d)t_0(d))^{d-1}}\,|||(a_i)|||_{X,p,d} \le \sup\left\{\sum_i a_i x_i^1 \prod_{k=2}^d (1+x_i^k \mathbf{1}_{\{x_i^k > C(d)t_0(d)\}}) \;\middle|\; \forall_k \sum_i \hat{M}_i^X(x_i^k) \le p\right\}$$
$$\le C(d)\sup\left\{\sum_i a_i x_i^1 \mathbf{1}_{\{0\le x_i^1 \le C(d)t_0(d)\}} \;\middle|\; \sum_i \hat{M}_i^X(x_i^1) \le p\right\}$$
$$+C(d)\sum_{\substack{I\subset[d]\\ I\neq\emptyset}} \sup\left\{\sum_i a_i \prod_{k\in I} x_i^k \mathbf{1}_{\{x_i^k > C(d)t_0(d)\}} \;\middle|\; \forall_k \sum_i \hat{M}_i^X(x_i^k) \le p\right\}. \tag{2.26}$$

Putting $y_i = x_i^1/(C(d)t_0(d))$ we see that

$$\sup\left\{\sum_i a_i x_i^1 \mathbf{1}_{\{0\le x_i^1 \le C(d)t_0(d)\}} \;\middle|\; \sum_i \hat{M}_i^X(x_i^1) \le p\right\}$$
$$= C(d)t_0(d)\sup\left\{\sum_i a_i y_i \mathbf{1}_{\{0\le y_i\le 1\}} \;\middle|\; \sum_i \hat{M}_i^X(C(d)t_0(d)y_i) \le p\right\}$$
$$\le C(d)\sup\left\{\sum_i a_i y_i \mathbf{1}_{\{0\le y_i\le 1\}} \;\middle|\; \sum_i \hat{N}_i^X(y_i) \le p\right\} \le C(d)\,\|(a_i)\|_{X,p}, \tag{2.27}$$

where the first inequality follows by the monotonicity of the functions $(\hat{M}_i^X)_i$.

Now we estimate the second term in (2.26). For a $I\subset[d],\ I\neq\emptyset$,

$$\sup\left\{\sum_i a_i \prod_{k\in I} x_i^k \mathbf{1}_{\{x_i^k > C(d)t_0(d)\}} \;\middle|\; \forall_k \sum_i \hat{M}_i^X(x_i^k) \le p\right\}$$
$$\le \sup\left\{\sum_i a_i \frac{1}{|I|}\sum_{k\in I} (x_i^k)^{|I|} \mathbf{1}_{\{x_i^k > C(d)t_0(d)\}} \;\middle|\; \forall_k \sum_i \hat{M}_i^X(x_i^k) \le p\right\}$$
$$\le \sup\left\{\sum_i a_i (x_i^1)^d \mathbf{1}_{\{x_i^1 > C(d)t_0(d)\}} \;\middle|\; \sum_i \hat{M}_i^X(x_i^1) \le p\right\}$$
$$\le C(d)\sup\left\{\sum_i a_i (x_i^1)^d \mathbf{1}_{\{x_i^1 > t_0(d)\}} \;\middle|\; \sum_i \hat{M}_i^X(C(d)x_i^1) \le p\right\}$$
$$\le C(d)\sup\left\{\sum_i a_i x_i^1 \mathbf{1}_{\{x_i^1 > (t_0(d))^d\}} \;\middle|\; \sum_i \hat{N}_i^X(x_i^1) \le p\right\} \le C(d)\,\|a_i\|_{X,p}, \tag{2.28}$$

where to get the fourth inequality we used (2.17). Estimates (2.26)-(2.28) imply (2.25). □

Lemma 2.2.6. *There exists $C=C(d)$ such that for any $p,u\ge 1$ we have*

$$\|(a_i)\|_{X,up} \le C(d)u^d\,\|(a_i)\|_{X,p}.$$

Proof. By Lemma 2.2.5 it is enough to show the following inequality

$$|||(a_i)|||_{X,up,d} \le u^d\,|||(a_i)|||_{X,p,d}.$$

The inequality (2.21) yields

$$\|\|(a_i)\|\|_{X,up,d} \le \sup\left\{ \sum_i a_i x_i^1 \prod_{k=2}^d (1+x_i^k) \,\middle|\, \forall_k \sum_i \hat{M}_i^X(\frac{x_i^k}{u}) \le p \right\}$$
$$= \sup\left\{ \sum_i a_i (u x_i^1) \prod_{k=2}^d (1+u x_i^k) \,\middle|\, \forall_k \sum_i \hat{M}_i^X(x_i^k) \le p \right\} \le u^d \|\|(a_i)\|\|_{X,p,d}\,.$$

□

2.3 Moment estimates in the one dimensional case

In this section we will show two-sided bound for moments of linear combinations of r.v's from the $\mathbb{S}(d)$ class.

Theorem 2.3.1. *Let $X_1, X_2, \ldots$ be independent, symmetric random variables from the $\mathbb{S}(d)$ class. Then for any $p \ge 1$ and any finite sequence (a_i) we have*

$$\left\| \sum_i a_i X_i \right\|_p \sim^d \|(a_i)\|_{X,p}\,. \tag{2.29}$$

Latała [15] (see Theorem 6.23) derived bounds for moment of $\sum_i a_i X_i$ in a general case. However we were not able to deduce Theorem 2.3.1 directly from it. Instead below we present a direct tedious proof of (2.29).

Since the r.v's $X_1, \ldots, X_n$ are symmetric and independent without loss of the generality we may assume that $a_i \ge 0$.

Lemma 2.3.2. *We have $\|\sum_i a_i X_i\|_p \sim^d \left\|\sum_i a_i \prod_{k=1}^d X_i^k\right\|_p$ for $p \ge 1$.*

Proof. Let (ε_i) be a Bernoulli sequence, independent of $\left\{X_i, X_j^k\right\}_{i,j\ge 1, k\le d}$. Using the Jensen inequality and (2.18) we get

$$\left\| \sum_i a_i \prod_{k=1}^d X_i^k \right\|_p = \left\| \sum_i a_i \varepsilon_i \prod_{k=1}^d |X_i^k| \right\|_p \ge \left\| \sum_i a_i \varepsilon_i \mathbb{E} \prod_{k=1}^d |X_i^k| \right\|_p \ge c(d) \left\| \sum_i a_i \varepsilon_i \right\|_p.$$

The contraction principle, Lemma 2.2.4 and the triangle inequality yield

$$\left\| \sum_i a_i X_i \right\|_p = \left\| \sum_i a_i \varepsilon_i |X_i| \right\|_p \le \left\| \sum_i a_i \varepsilon_i C(d) \left(\left| \prod_{k=1}^d X_i^k \right| + 1 \right) \right\|_p \le C(d) \left\| \sum_i a_i \prod_{k=1}^d X_i^k \right\|_p.$$

The reverse bound may be established in an analogous way. □

The intuition behind Lemma 2.3.2 is that it is easier to properly bound $\left\|\sum a_i \prod_{k=1}^d X_i^k\right\|_p$ since we replaced each "big" r.v X_i with a product of "smaller" pieces, which are easier to deal with.

Next lemma shows that Theorem 4.1 holds under the additional assumption that the support of the sum is small.

Lemma 2.3.3. *For any $p \ge 1$ we have*

$$\left\| \sum_{i\le p} a_i \prod_{k=1}^{d} X_i^k \right\|_p \sim^d \left|\!\left|\!\left| (a_i)_{i\le p} \right|\!\right|\!\right|_{X,p,d}.$$

Proof. We will proceed by an induction on d. For $d=1$ lemma holds by the Gluskin-Kwapień bound 6.22. Assume $d\ge 2$ and the assertion holds for $1,2,\dots,d-1$. First we establish the lower bound for $\left\| \sum_{i\le p} a_i \prod_{k=1}^{d} X_i^k \right\|_p$. We have

$$\begin{aligned} C(d)\left\| \sum_{i\le p} a_i \prod_{k=1}^{d} X_i^k \right\|_p &\ge \left\| \sum_{i\le p} a_i \prod_{k=1}^{d-1} X_i^k \mathbb{E}|X_i^d| \right\|_p + \left\| \left|\!\left|\!\left| (a_i X_i^d)_{i\le p} \right|\!\right|\!\right|_{X,p,d-1} \right\|_p \\ &\ge \frac{1}{C(d)} \left\| \sum_{i\le p} a_i \prod_{k=1}^{d-1} X_i^k \right\|_p + \sup\left\{ \left\| \sum_{i\le p} a_i x_i^1 \prod_{k=2}^{d-1} (1+x_i^k) X_i \right\|_p \,\middle|\, \forall_{k\le d-1} \sum_{i\le p} \hat{M}_i^X(x_i^k) \le p \right\} \\ &\ge \frac{1}{C(d)} \left(\sup\left\{ \sum_i a_i x_i^1 \prod_{k=2}^{d-1} (1+x_i^k) \,\middle|\, \forall_{k=1,\dots,d-1} \sum_i \hat{M}_i^X(x_i^k) \le p \right\} \right. \\ &\left. + \sup\left\{ \sum_i a_i x_i^1 \prod_{k=2}^{d-1} (1+x_i^k) x_i^d \,\middle|\, \forall_{k=1,\dots,d} \sum_i \hat{M}_i^X(x_i^k) \le p \right\} \right) \ge \frac{1}{C(d)} \left|\!\left|\!\left| (a_i) \right|\!\right|\!\right|_{X,p,d}, \end{aligned}$$

where the first inequality follows by Jensen's inequality and the induction assumption, the second by (2.18) and the third by the induction assumption.

Now we prove the upper bound. Using the right-continuity of M_i^X we have

$$\mathbb{P}(M_i^X(|X_i^d|)\ge t) = \mathbb{P}\left(|X_i^d| \ge \left(M_i^X\right)^{-1}(t) \right) \le e^{-t} \quad \text{for } t>0.$$

Therefore there exists nonnegative i.i.d r.v's $\mathcal{E}_1,\dots,\mathcal{E}_p$ with the density $e^{-t}\mathbf{1}_{\{t>0\}}$ such that $M_i^X(|X_i^d|)\le \mathcal{E}_i$. Since $\sum_i^p \mathcal{E}_i$ has the $\Gamma(p,1)$ distribution we obtain

$$\left\| \sum_{i\le p} M_i^X(|X_i^d|) \right\|_q \le \left\| \sum_{i\le p} \mathcal{E}_i \right\|_q \le Cq \quad \text{for } q\ge p.$$

Since M_i^X is convex, the above inequality implies for any $t\ge 1$,

$$\begin{aligned} \mathbb{P}\left(\sum_{i\le p} M_i^X\left(\frac{|X_i^d|}{Ct}\right) \ge p \right) &\le \mathbb{P}\left(\sum_{i\le p} M_i^X\left(|X_i^d|\right) \ge Ctp \right) \le (Ctp)^{-tp} \left\| \sum_{i\le p} M_i^X(|X_i^d|) \right\|_{tp}^{tp} \\ &\le e^{-tp}. \end{aligned} \tag{2.30}$$

From (2.30)

$$\mathbb{P}\left(\left|\!\left|\!\left|\left(a_iX_i^d\right)_{i\le p}\right|\!\right|\!\right|_{X,p,d-1} \ge Ct\left|\!\left|\!\left|(a_i)_{i\le p}\right|\!\right|\!\right|_{X,p,d}\right)$$
$$= \mathbb{P}\left(\left|\!\left|\!\left|\left(a_i\frac{X_i^d}{Ct}\right)_{i\le p}\right|\!\right|\!\right|_{X,p,d-1} \ge \sup\left\{\left|\!\left|\!\left|(a_i(1+x_i))_{i\le p}\right|\!\right|\!\right|_{X,p,d-1}\right\}\right)$$
$$\le \mathbb{P}\left(\sum_{i\le p}\hat{M}_i^X\left(\frac{|X_i^d|}{Ct}\right)\mathbf{1}_{\{\frac{|X_i^d|}{Ct}>1\}} \ge p\right) \le \mathbb{P}\left(\sum_{i\le p}M_i^X\left(\frac{|X_i^d|}{Ct}\right)\ge p\right) \le e^{-tp},$$

where the supremum in the second line is taken over all $x_1,\dots,x_p$, such that $\sum_i \hat{N}_i^X(x_i) \le p$. Integration by parts gives

$$\left\|\left|\!\left|\!\left|\left(a_iX_i^d\right)_{i\le p}\right|\!\right|\!\right|_{X,p,d-1}\right\|_p \le C\left|\!\left|\!\left|(a_i)_{i\le p}\right|\!\right|\!\right|_{X,p,d}. \tag{2.31}$$

By the induction assumption and (2.31),

$$\left\|\sum_{i\le p}a_i\prod_{k=1}^d X_i^k\right\|_p \le C(d)\left\|\left|\!\left|\!\left|\left(a_iX_i^d\right)_{i\le p}\right|\!\right|\!\right|_{X,p,d-1}\right\|_p \le C(d)\left|\!\left|\!\left|(a_i)_{i\le p}\right|\!\right|\!\right|_{X,p,d},$$

that concludes the proof of the induction step.

□

Remark 2.3.4. Observe that in the proof of the lower bound in Lemma 2.3.3 we have not used the condition $i \le p$.

The idea of the following lemma is taken from [19].

Lemma 2.3.5. *Let $p \ge 1$. Define*

$$T = \left\{v\in\mathbb{R}^n \,\middle|\, \sum_i \hat{M}_i^X(v_i)\le p,\ \textit{and}\ \forall_{i\le n}\ |v_i|\ge 1\ \textit{or}\ v_i=0\right\} \tag{2.32}$$

$$U = \bigcap_{l=1}^{\infty}\left\{v\in\mathbb{R}^n \,\middle|\, \hat{M}_i^X(v_i)\le l^3,\ i\in(2^lp,2^{l+1}p]\right\}\cap\left\{v\in\mathbb{R}^n \,\middle|\, \forall_{i\le 2p}\ v_i=0\right\}\cap T, \tag{2.33}$$

$$V = \bigcap_{l=1}^{\infty}\left\{v\in\mathbb{R}^n \,\middle|\, \hat{M}_i^X(v_i)> l^3\ \textit{or}\ v_i=0,\ i\in(2^lp,2^{l+1}p]\right\} \tag{2.34}$$
$$\cap\left\{v\in\mathbb{R}^n \,\middle|\, \forall_{i\le 2p}\ \hat{M}_i^X(v_i)\ge 1\ \textit{or}\ v_i=0\right\}\cap T.$$

If (a_i) is a nonincreasing nonnegative sequence then

$$\mathbb{E}\sup_{x^1,\dots,x^{d-1}\in U}\sum_i a_i\prod_{k=1}^{d-1}x_i^k|X_i^d| \le C(d)\,|||(a_i)|||_{X,p,d}, \tag{2.35}$$

$$\mathbb{E}\sup_{x^1,\dots,x^{d-2}\in T,\ x^{d-1}\in V}\sum_i a_i\prod_{k=1}^{d-1}x_i^k|X_i^d| \le C(d)\,|||(a_i)|||_{X,p,d}. \tag{2.36}$$

As we will see (in the next lemma) the main difficulty in proving Theorem 2.3.1 is the proper estimation of

$$\mathbb{E} \sup_{x^1,\dots,x^{d-1}\in T} \sum_i a_i \prod_{k=1}^{d-1} x_i^k |X_i^d|.$$

The key properties of sets U,V are that $U,V \subset T \subset U+V$ and that we can prove (2.35), (2.36) by some combinatorial arguments. The main difficulty in Lemma 2.3.5 is to figure how to decompose set T (which was done in [19]).

Proof. We begin with (2.35). For abbreviation, we define

$$Z_l := \left\{ x \in \mathbb{R}^n \;\middle|\; \sum_i x_i \le l^3 \left\lceil \frac{p}{l^3} \right\rceil,\ x_i \in \{0, l^3\} \right\}.$$

Using the fact that the sequence (a_i) is nonnegative and (2.22) we obtain

$$\begin{aligned}
&\frac{1}{C(d)}\mathbb{E} \sup_{\substack{x^k\in U\\ k=1,\dots,d-1}} \sum_i a_i \prod_{k=1}^{d-1} x_i^k \left|X_i^d\right| \le \frac{1}{C(d)} \sum_{l=1}^{\infty} \mathbb{E} \sup_{\substack{x^k\in U\\ k=1,\dots,d-1}} \sum_{i=2^l p+1}^{2^{l+1}p} a_i \prod_{k=1}^{d-1} x_i^k \left|X_i^d\right| \\
&\le \sum_{l=1}^{\infty} l^{3(d-2)} \mathbb{E}\sup \left\{ \sum_{i=2^l p+1}^{2^{l+1}p} a_i x_i \left|X_i^d\right| \;\middle|\; \sum_i x_i \le p,\ x_i = 0 \vee x_i \in [1, l^3] \right\} \\
&\le C(d) \sum_{l=1}^{\infty} l^{3(d-2)} \mathbb{E}\sup \left\{ \sum_{i=2^l p+1}^{2^{l+1}p} a_i x_i \left|X_i^d\right| \;\middle|\; x \in Z_l \right\}.
\end{aligned} \tag{2.37}$$

Using Lemma 6.16

$$\begin{aligned}
&\mathbb{E}\sup \left\{ \sum_{i=2^l p+1}^{2^{l+1}p} a_i x_i \left|X_i^d\right| \;\middle|\; x \in Z_l \right\} \\
&\le \mathbb{E}\sup \left\{ \left(\sum_{i=2^l p+1}^{2^{l+1}p} a_i x_i \left|X_i^d\right| - C \sum_{i=2^l p+1}^{2^{l+1}p} a_i x_i \right)_+ \;\middle|\; x \in Z_l \right\} + C l^3 \left\lceil \frac{p}{l^3} \right\rceil a_{2^l p+1} \\
&\le C \left(p + \ln \left| \left\{ v \in \mathbb{R}^{2^l p} \;\middle|\; v \in Z_l \right\} \right| \right) l^3 a_{2^l p+1} + C(l^3 + p) a_{2^l p+1} \le C(p + l^4) a_{2^l p+1}.
\end{aligned} \tag{2.38}$$

The last inequality follows by the simple estimate

$$\binom{2^l p}{\lceil \frac{p}{l^3} \rceil} \le \left(\frac{2^l p e}{\lceil \frac{p}{l^3} \rceil} \right)^{\lceil p/l^3 \rceil} \le C^{p+l}.$$

Combining (2.37) and (2.38) we obtain

$$\begin{aligned}
\mathbb{E} \sup_{x^k\in U,\ k=1,\dots,d-1} \sum_i a_i \prod_{k=1}^{d-1} x_i^k \left|X_i^d\right| &\le C(d) \sum_{l=1}^{\infty} l^{3d} p a_{2^l p+1} \le C(d) \sum_{l=1}^{\infty} l^{3d} p \frac{\sqrt{\sum_{i>p} a_i^2}}{\sqrt{2^{l-1}p}} \\
&\le C(d) \sqrt{p} \sqrt{\sum_{i>p} a_i^2}.
\end{aligned} \tag{2.39}$$

To finish the proof of (2.35) it is enough to observe that by Lemma 2.2.3

$$\frac{\sqrt{p}}{2}\sqrt{\sum_{i>p}a_i^2}\le \sup\left\{\sum_i a_i t_i \,\middle|\, \sum_i t_i^2\le p,\ \forall_i |t_i|\le 1\right\}\le |||(a_i)|||_{X,p,d}.$$

Now we show (2.36). Let

$$\mathcal{J}=\left\{I\subset \mathbb{N}\,\middle|\, |I|\le p,\ \forall_{l\in\mathbb{N}}\ |I\cap[2^l p+1,2^{l+1}p)|\le \frac{p}{l^3}\right\} \tag{2.40}$$

(in particular $\mathcal{J}$ contains any subset of $[2p]$ of cardinality not greater than p). Applying the inequality $\binom{n}{m}\le ((en)/m)^m$, we get an estimate of the cardinality of $\mathcal{J}$

$$|\mathcal{J}|\le 2^{2p}\prod_{l=1}^{\lfloor p^{1/3}\rfloor}\left(\frac{C2^l p}{\lfloor p/l^3\rfloor}\right)^{\lfloor p/l^3\rfloor}\le C^p\prod_{l=1}^{\infty}(2^l l^3)^{p/l^3}\le C^p. \tag{2.41}$$

Take any $I\in\mathcal{J}$. We obtain (see Lemma 2.3.3)

$$\mathbb{E}\left|\left|\left|(a_iX_i^d)_{i\in I}\right|\right|\right|_{X,p,d-1}\le C(d)\,|||(a_i)_{i\in I}|||_{X,p,d}. \tag{2.42}$$

Using the definition of V, (2.41) and (2.42)

$$\mathbb{E}\sup\left\{\sum_i a_i\prod_{k=1}^{d-1}x_i^k X_i^d \,\middle|\, \forall_{k\le d-2}\sum_i \hat{M}_i^X(x_i^k)\le p,\ x^{d-1}\in V\right\}$$
$$\le\left(\sum_{I\in\mathcal{J}}\left(\mathbb{E}\left|\left|\left|(a_iX_i^d)_{i\in I}\right|\right|\right|_{X,p,d-1}\right)^p\right)^{1/p}\le C(d)\sup_{I\in\mathcal{J}}|||(a_i)_{i\in I}|||_{X,p,d}\le C(d)\,|||(a_i)|||_{X,p,d}.$$

□

Lemma 2.3.6. *For any $d\in\mathbb{N}$ the following holds*

$$\mathbb{E}\left|\left|\left|(a_iX_i^d)\right|\right|\right|_{X,p,d-1}\le C(d)\,|||(a_i)|||_{X,p,d}. \tag{2.43}$$

Proof. Without loss of the generality we may assume that the sequence (a_i) is nonincreasing and recall that in this section a_i are nonnegative. We proceed by an induction on d. If $d=2$ then Corollary 6.18 (with $\hat{a}_{i,j}=a_i\mathbf{1}_{\{i=j\}}$) implies the assertion. Assume that (2.43) holds for any $2,3,\ldots,d-1$. Obviously,

$$\mathbb{E}\left|\left|\left|(a_iX_i^d)\right|\right|\right|_{X,p,d-1}\le \mathbb{E}\sup\left\{\sum_i a_i x_i^1\mathbf{1}_{\{|x_i^1|\le 1\}}\prod_{k=2}^{d-1}(1+x_i^k)X_i^d \,\middle|\, \forall_{k\le d-1}\sum_i \hat{M}_i^X(x_i^k)\le p\right\}$$
$$+\mathbb{E}\sup\left\{\sum_i a_i x_i^1\mathbf{1}_{\{|x_i^1|>1\}}\prod_{k=2}^{d-1}(1+x_i^k)X_i^d \,\middle|\, \forall_{k\le d-1}\sum_i \hat{M}_i^X(x_i^k)\le p\right\}=:S_1+S_2. \tag{2.44}$$

We have

$$
\begin{aligned}
S_1 &\le \mathbb{E}\sup\left\{ \sum_i a_i x_i^1 X_i^d \,\middle|\, \sum_i \hat{M}_i^X(x_i^1) \le p \right\} \\
&\quad + \sum_{\substack{I\subset([d-1]\setminus\{1\})\\ I\neq\emptyset}} \mathbb{E}\sup\left\{ \sum_i a_i \prod_{k\in I} x_i^k X_i^d \,\middle|\, \forall_{2\le k\le d-1} \sum_i \hat{M}_i^X(x_i^k) \le p \right\} \\
&\le 2^{d-2}\mathbb{E}\sup\left\{ \sum_i a_i x_i^2 \prod_{k=3}^{d-1}(1+x_i^k) X_i^d \,\middle|\, \forall_{2\le k\le d-1} \sum_i \hat{M}_i^X(x_i^k) \le p \right\} \le C(d)\, \|\!|(a_i)|\!\|_{X,p,d}\,,
\end{aligned} \tag{2.45}
$$

where the last inequality follows by the induction assumption.

Now we bound S_2. Denote

$$
Z := \left\{ \mathbf{x} = (x^1,\dots,x^{d-1}) \in \mathbb{R}^{n(d-1)} \,\middle|\, \forall_{k\le d-1} \sum_i \hat{M}_i^X(x_i^k) \le p \right\}.
$$

Since a_i are nonnegative,

$$
\begin{aligned}
S_2 &\le 2^d \mathbb{E}\sup\left\{ \sum_i a_i x_i^1 \mathbf{1}_{\{x_i^1>1\}} \prod_{k=2}^{d-1}(1+x_i^k\mathbf{1}_{\{x_i^k>1\}}) \left|X_i^d\right| \,\middle|\, \mathbf{x}\in Z \right\} \\
&\le 2^d \mathbb{E}\sup\left\{ \sum_i a_i \prod_{k=1}^{d-1} x_i^k \mathbf{1}_{\{x_i^k>1\}} \left|X_i^d\right| \,\middle|\, \mathbf{x}\in Z \right\} \\
&\quad + 2^d \sum_{\substack{I\subsetneq[d-1]\\ I\neq\emptyset}} \mathbb{E}\sup\left\{ \sum_i a_i \prod_{k\in I} x_i^k \mathbf{1}_{\{x_i^1>1\}} \left|X_i^d\right| \,\middle|\, \mathbf{x}\in Z \right\}.
\end{aligned} \tag{2.46}
$$

We can bound the second term in (2.46) by using the induction assumption. Inequalities (2.44)-(2.46) imply (2.43), provided

$$
\mathbb{E}\sup\left\{ \sum_i a_i \prod_{k=1}^{d-1} x_i^k \mathbf{1}_{\{x_i^k>1\}} \left|X_i^d\right| \,\middle|\, \mathbf{x}\in Z \right\} \le C(d)\, \|\!|(a_i)|\!\|_{X,p,d}\,. \tag{2.47}
$$

Let T,U,V be the sets defined in (2.32)-(2.34). Since $T\subset U+V$ we have

$$
\begin{aligned}
&\mathbb{E}\sup\left\{ \sum_i a_i \prod_{k=1}^{d-1} x_i^k \mathbf{1}_{\{x_i^k>1\}} \left|X_i^d\right| \,\middle|\, \mathbf{x}\in Z \right\} \\
&\le \mathbb{E}\sup_{x^1,\dots,x^{d-1}\in U} \sum_i a_i \prod_{k=1}^{d-1} x_i^k |X_i^d| + \sum_{\substack{I\subset[d-1]\\ I\neq\emptyset}} \mathbb{E}\sup_{x^i\in U,\ i\in I}\ \sup_{x^i\in V,\ i\notin I} \sum_i a_i \prod_{k=1}^{d-1} x_i^k |X_i^d| \\
&\le \mathbb{E}\sup_{x^1,\dots,x^{d-1}\in U} \sum_i a_i \prod_{k=1}^{d-1} x_i^k |X_i^d| + \left(2^{d-1}-1\right)\mathbb{E}\sup_{x^1,\dots,x^{d-2}\in T,\ x^{d-1}\in V} \sum_i a_i \prod_{k=1}^{d-1} x_i^k |X_i^d|.
\end{aligned} \tag{2.48}
$$

In the last inequality we used the symmetry and inclusions $U,V\subset T$. Now (2.47) follows by (2.48), (2.35) and (2.36). $\square$

Lemma 2.3.7. *We have*

$$\left\|\sum_i a_i \prod_{k=1}^d X_i^k\right\|_p \sim^d |||(a_i)|||_{X,p,d}\,. \tag{2.49}$$

Proof. We prove (2.49) by an induction. In case of $d=1$ it follows by Gluskin-Kwapień bound 6.22. Assume (2.49) hold for any $1,2,\dots,d-1$. The lower bound of $\left\|\sum_i a_i \prod_{k=1}^d X_i^k\right\|_p$ follows by Remark 2.3.4.

Now we prove

$$\left\|\sum_i a_i \prod_{k=1}^d X_i^k\right\|_p \le C(d)\,|||(a_i)|||_{X,p,d}\,.$$

Induction assumptions imply

$$\begin{aligned}\left\|\sum_i a_i \prod_{k=1}^d X_i^k\right\|_p &\le C(d)\left\|\left|\left|\left|(a_i X_i^d)\right|\right|\right|_{X,p,d-1}\right\|_p \le C(d)\mathbb{E}\left|\left|\left|(a_i X_i^d)\right|\right|\right|_{X,p,d-1}\\ &\quad + C(d)\sup\left\{\left\|\sum_i a_i x_i^1 \prod_{k=2}^{d-1}(1+x_i^k)X_i^d\right\|_p \,\middle|\, \forall_{k\le d-1} \sum_i \hat{M}_i^X(x_i^k)\le p\right\},\end{aligned} \tag{2.50}$$

where in the second inequality we used Theorem 6.25 (if Z is symmetric random variable with log-concave tail then $\|Z\|_{2p} \le 2\|Z\|_p$ for $p\ge 1$) . The Gluskin-Kwapień estimate (i.e the first step of the induction) gives

$$\sup\left\{\left\|\sum_i a_i x_i^1 \prod_{k=2}^{d-1}(1+x_i^k)X_i^d\right\|_p \,\middle|\, \forall_{k\le d-1} \sum_i \hat{M}_i^X(x_i^k)\le p\right\} \le C(d)\,|||(a_i)|||_{X,p,d}\,. \tag{2.51}$$

The assertion follows by (2.50), Lemma 2.3.6 and (2.51). □

Theorem 2.3.1 follows by Lemmas 2.3.2, 2.3.7 and 2.2.5.

Observe that Lemma 2.2.6 and (2.29) imply that moments of $\sum_i a_i X_i$ cannot grow too quickly, namely

$$\left\|\sum_i a_i X_i\right\|_{2p} \le C(d)\left\|\sum_i a_i X_i\right\|_p \quad \text{for } p\ge 1. \tag{2.52}$$

2.4 Estimates for suprema of processes

The essential difficulty with proving Theorem 2.1.1 is to properly bound $\mathbb{E}\sup_{x\in\hat{T}}\sum_{i,j} a_{i,j}x_iY_j$ in two cases:

1. $\hat{T} = \left\{x\in\mathbb{R}^n \,\middle|\, \sum_i \hat{M}_i^X(x_i)\le p\right\} \subset \sqrt{p}B_2^n + pB_1^n$,
2. $\hat{T} = \left\{\left(\prod_{k=1}^d x_i^k \mathbf{1}_{\{x_i^k\ge 1\}}\right)_i \in\mathbb{R}^n \,\middle|\, \forall_{k\le d}\sum_i \hat{M}_i^X(x_i^k)\le p\right\}$.

We start with two lemmas which are responsible for solving the second case. The rest of the section is devoted for developing some decomposition lemmas. They will be used to handle the first case.

Firstly, we show that Theorem 2.1.1 holds under additional assumption that the support of the sum is small.

Lemma 2.4.1. *For any $p\ge 1$, and set $I\subset[n]$, $|I|\le p$,*

$$\left\| \sum_{i \in I, j} a_{i,j} X_i Y_j \right\|_p \leq C(d) \left\| (a_{i,j})_{i \in I, j} \right\|_{X,Y,p}. \tag{2.53}$$

Proof. We have

$$\begin{aligned}
\left\| \sum_{i \in I, j} a_{i,j} X_i Y_j \right\|_p &\leq C(d) \left(\mathbb{E} \sup_{\sum_j \hat{N}_j^Y(y_j) \leq p} \left| \sum_{i \in I, j} a_{i,j} X_i y_j \right|^p \right)^{1/p} \\
&\leq C(d) \sup_{\sum_j \hat{N}_j^Y(y_j) \leq p} \left(\mathbb{E} \left| \sum_{i \in I, j} a_{i,j} X_i y_j \right|^p \right)^{1/p} \\
&\leq C(d) \sup_{\sum_j \hat{N}_j^Y(y_j) \leq p} \left(\mathbb{E} \left| \sum_{i \in I, j} a_{i,j} X_i y_j \right|^p \right)^{1/p} \leq C(d) \left\| (a_{i,j})_{i \in I, j} \right\|_{X,Y,p},
\end{aligned}$$

where the first inequality follows by a conditional application of Theorem 2.3.1, the second one by Fact 2.2.2 and the last one by Theorem 2.3.1. □

From now till the end of this section we assume (without loss of generality) the following condition

$$\text{the functions } i \to \sum_j a_{i,j}^2,\ j \to \sum_i a_{i,j}^2 \text{ are nonincreasing.} \tag{2.54}$$

Lemma 2.4.2 (cf. Lemma 2.3.6)**.** *Let* T, U, V *be the sets defined in* (2.32)-(2.34)*. Then*

$$S_1 := \mathbb{E} \sup_{x^1, \ldots, x^d \in U} \sum_{i,j} a_{i,j} \prod_{k=1}^d x_i^k Y_j \leq C(d) \left\| \left(\sqrt{\sum_i a_{i,j}^2} \right)_j \right\|_{Y,p}, \tag{2.55}$$

$$S_2 := \mathbb{E} \sup_{x^1, \ldots, x^{d-1} \in T, x^d \in V} \sum_{i,j} a_{i,j} \prod_{k=1}^d x_i^k Y_j \leq C(d) \left\| (a_{i,j})_{i,j} \right\|_{X,Y,p}. \tag{2.56}$$

Proof. We begin with (2.55). By the symmetry of $\hat{M}_i^X$,

$$S_1 = \mathbb{E} \sup_{x^1, \ldots, x^d \in U} \sum_i \prod_{k=1}^d x_i^k \left| \sum_j a_{i,j} Y_j \right| \leq \sum_{l=1}^{\infty} \mathbb{E} \sup_{x^1, \ldots, x^d \in U} \sum_{2^l p < i \leq 2^{l+1} p} \prod_{k=1}^d x_i^k \left| \sum_j a_{i,j} Y_j \right|.$$

So by (2.22)

$$
\begin{aligned}
S_1 &\le \sum_{l=1}^{\infty} \mathbb{E}\sup\left\{ \sum_{2^l p < i \le 2^{l+1} p} \prod_{k=1}^{d} x_i^k \left| \sum_j a_{i,j} Y_j \right| \,\middle|\, x^1, \ldots, x^d \in Z_l \right\} \\
&\le \sum_{l=1}^{\infty} l^{3d} \mathbb{E} \sup_{\substack{|I|=p \\ I \subset (2^l p, 2^{l+1} p]}} \sum_{i \in I} \left| \sum_j a_{i,j} Y_j \right| \le \sum_{l=1}^{\infty} l^{3d} p^{3/4} \mathbb{E} \sup_{\substack{|I|=p \\ I \subset (2^l p, 2^{l+1} p]}} \left(\sum_{i \in I} \left(\sum_j a_{i,j} Y_j \right)^4 \right)^{1/4} \\
&\le \sum_{l=1}^{\infty} l^{3d} p^{3/4} \mathbb{E} \left(\sum_{2^l p < i \le 2^{l+1} p} \left(\sum_j a_{i,j} Y_j \right)^4 \right)^{1/4},
\end{aligned}
$$

where

$$
Z_l := \left\{ x \in \mathbb{R}^n \,\middle|\, \sum_{i=2^l p+1}^{2^{l+1} p} |x_i| \le p,\ 1 \le |x_i| \le l^3 \text{ or } x_i = 0 \right\},
$$

and the third inequality follows by the Hölder inequality. Now using the Jensen inequality and (2.52)

$$
\begin{aligned}
\sum_{l=1}^{\infty} l^{3d} p^{3/4} \mathbb{E} \left(\sum_{2^l p < i \le 2^{l+1} p} \left(\sum_j a_{i,j} Y_j \right)^4 \right)^{1/4} &\le \sum_{l=1}^{\infty} l^{3d} p^{3/4} \left(\sum_{2^l p < i \le 2^{l+1} p} \left\| \sum_j a_{i,j} Y_j \right\|_4^4 \right)^{1/4} \\
&\le C(d) \sum_{l=1}^{\infty} l^{3d} p^{3/4} \left(\sum_{2^l p < i \le 2^{l+1} p} \left\| \sum_j a_{i,j} Y_j \right\|_2^4 \right)^{1/4} \\
&\le C(d) \sum_{l=1}^{\infty} l^{3d} p^{3/4} \left(\sum_{2^l p < i \le 2^{l+1} p} \left(\sum_j a_{i,j}^2 \right)^2 \right)^{1/4}.
\end{aligned} \tag{2.57}
$$

Denote $B = \sqrt{\sum_{\substack{i \ge p \\ j \ge 1}} a_{i,j}^2}$. Since $i \to \sum_j a_{i,j}^2$ is nonincreasing we have that

$$
\sum_j a_{i,j}^2 \le \frac{B^2}{i-p} \text{ for } i > p.
$$

Using the above estimate in (2.57) gives

$$
\begin{aligned}
\sum_{l=1}^{\infty} l^{3d} p^{3/4} \left(\sum_{2^l p < i \le 2^{l+1} p} \left(\sum_j a_{i,j}^2 \right)^2 \right)^{1/4} &\le \sum_{l=1}^{\infty} l^{3d} p^{3/4} \left(\sum_{i > 2^l p} \frac{B^4}{(i-p)^2} \right)^{1/4} \\
&\le \sum_{l=1}^{\infty} p^{3/4} l^{3d} \left(\frac{CB^4}{2^l p} \right)^{1/4} \le C(d) \sqrt{p} B.
\end{aligned}
$$

By (2.54) and Lemma 2.2.3 we have

$$\frac{\sqrt{p}}{2}B \leq \sup\left\{\sum_i \sqrt{\sum_j a_{i,j}^2} t_i \;\middle|\; \sum_i t_i^2 \leq p,\ |t_i| \leq 1\right\} \leq \left\|\left(\sqrt{\sum_i a_{i,j}^2}\right)_j\right\|_{Y,p}.$$

As a consequence

$$S_1 \leq C(d)\left\|\left(\sqrt{\sum_i a_{i,j}^2}\right)_j\right\|_{Y,p}.$$

Now we bound S_2. Let $\mathcal{J}$ be defined by (2.40) and let $I \in \mathcal{J}$ be arbitrary. Using conditionally (2.29), Lemma 2.2.5 and the Jensen inequality

$$\begin{aligned}
&\mathbb{E} \sup_{x^1,\ldots,x^{d-1}\in T, x^d \in V} \sum_{i\in I}\prod_{k=1}^{d} x_i^k \left|\sum_j a_{i,j} Y_j\right| \leq \mathbb{E} \sup_{x^1,\ldots,x^d\in T} \sum_{i\in I}\prod_{k=1}^{d} x_i^k \left|\sum_j a_{i,j} Y_j\right| \\
&\leq C(d)\mathbb{E}_Y \left(\mathbb{E}_X \left|\sum_{i\in I,j} a_{i,j} X_i Y_j\right|^p\right)^{1/p} \leq C(d)\left\|\sum_{i\in I,j} a_{i,j} X_i Y_j\right\|_p.
\end{aligned}$$

By Lemma 2.4.1

$$\left\|\sum_{i\in I,j} a_{i,j} X_i Y_j\right\|_p \leq C(d)\left\|(a_{i,j})_{i\in I,j}\right\|_{X,Y,p}.$$

So we conclude that

$$\mathbb{E} \sup_{x^1,\ldots,x^{d-1}\in T, x^d \in V} \sum_{i\in I}\prod_{k=1}^{d} x_i^k \left|\sum_j a_{i,j} Y_j\right| \leq C(d)\left\|(a_{i,j})_{i\in I,j}\right\|_{X,Y,p} \leq C(d)\left\|(a_{i,j})_{i,j}\right\|_{X,Y,p}. \tag{2.58}$$

By (2.41) and (2.58)

$$\begin{aligned}
S_2 &\leq C(d)\left(\sum_{I\in\mathcal{J}}\left(\mathbb{E} \sup_{x^1,\ldots,x^{d-1}\in T, x^d \in V} \sum_{i\in I}\prod_{k=1}^{d} x_i^k \left|\sum_j a_{i,j} Y_j\right|\right)^p\right)^{1/p} \\
&\leq C(d)\sup_{I\in\mathcal{J}}\left\|(a_{i,j})_{i\in I,j}\right\|_{X,Y,p} \leq C(d)\left\|(a_{i,j})_{i,j}\right\|_{X,Y,p}.
\end{aligned}$$

□

As it was announced earlier, we proceed with the study of decomposition lemmas. It is well know that (see Lemma 6.10) if $T = \bigcup_{k=1}^m T_l$ then

$$\mathbb{E}\sup_{t\in T}\sum_i t_i g_i \leq \max_{k\leq m}\mathbb{E}\sup_{t\in T_k}\sum_i t_i g_i + C\sqrt{\log(m)}\sup_{s,t\in T}\left\|\sum_i (t_i - s_i) g_i\right\|_2.$$

We will generalize this formula to any variables from the $\mathbb{S}(d)$ class.

Corollary 2.4.3. *For any $p \geq 1$ we have*

$$\mathbb{P}\left(\sup_{t\in T}\left|\sum_i t_iX_i\right| \geq C(d)\left(\mathbb{E}\sup_{t\in T}\left|\sum_i t_iX_i\right| + \sup_{t\in T}\|(t_i)\|_{X,p}\right)\right) \leq e^{-p}.$$

Proof. It is a simple consequence of Theorem 6.25, (2.29) and Chebyshev's inequality. □

Lemma 2.4.4. *Let* $T = \bigcup_{k=1}^m T_k$, $m \geq 8$. *Then*

$$\mathbb{E}\sup_{t\in T}\sum_i t_iX_i \leq C(d)\left(\max_{k\leq m}\mathbb{E}\sup_{t\in T_k}\sum_i t_iX_i + \sup_{s,t\in T}\|(t_i-s_i)_i\|_{X,\ln(m)}\right).$$

Proof. We choose any $s\in T$. Since $\mathbb{E}X_i = 0$ we have

$$\mathbb{E}\sup_{t\in T}\sum_i t_iX_i = \mathbb{E}\max_{k\leq m}\sup_{t\in T_k}\sum_i (t_i-s_i)X_i \leq \mathbb{E}\max_{k\leq m}\sup_{t\in T_k}\left|\sum_i (t_i-s_i)X_i\right|.$$

Let us denote

$$M := \max_{k\leq m}\mathbb{E}\sup_{t\in T_k}\left|\sum_i (t_i-s_i)X_i\right|.$$

Corollary 2.4.3 and the union bound yield for $u\geq 1$,

$$\mathbb{P}\left(\max_{k\leq m}\sup_{t\in T_k}\left|\sum_i (t_i-s_i)X_i\right| \geq C(d)\left[M + \|(t_i-s_i)_i\|_{X,u\ln(m)}\right]\right) \leq me^{-u\ln(m)} \leq 4^{1-u}.$$

Lemma 2.2.6 implies $\|(t_i-s_i)_i\|_{X,u\ln(m)} \leq C(d)u^d\|(t_i-s_i)_i\|_{X,\ln(m)}$. Hence

$$\mathbb{P}\left(\sup_{t\in T}\left|\sum_i (t_i-s_i)X_i\right| \geq C(d)\left[M + u^d\sup_{s,t\in T}\|(t_i-s_i)_i\|_{X,\ln(m)}\right]\right) \leq 4^{1-u}.$$

Integration by parts gives

$$\begin{aligned}\mathbb{E}\sup_{t\in T}\left|\sum_i (t_i-s_i)X_i\right| &\leq C(d)\left(M + \sup_{s,t\in T}\|(t_i-s_i)_i\|_{X,\ln(m)}\right)\\ &= C(d)\left(\max_k\mathbb{E}\sup_{t\in T_k}\left|\sum_i (t_i-s_i)X_i\right| + \sup_{s,t\in T}\|(t_i-s_i)_i\|_{X,\ln(m)}\right)\end{aligned}$$

So to finish the proof it is enough to show that

$$\frac{\mathbb{E}\sup_{t\in T_k}\left|\sum_i (t_i-s_i)X_i\right|}{C(d)} \leq \mathbb{E}\sup_{t\in T_k}\sum_i t_iX_i + \sup_{s,t\in T}\|(t_i-s_i)_i\|_{X,\ln(m)}. \tag{2.59}$$

Let $z\in T_k$. We have

$$\begin{aligned}\mathbb{E}\sup_{t\in T_k}\left|\sum_i (t_i-s_i)X_i\right| &\le \mathbb{E}\sup_{t\in T_k}\left|\sum_i (t_i-z_i)X_i\right| + \mathbb{E}\left|\sum_i (z_i-s_i)X_i\right| \\ &\le \mathbb{E}\sup_{t\in T_k}\left|\sum_i (t_i-z_i)X_i\right| + \left\|\sum_i (z_i-s_i)X_i\right\|_2 \\ &\le \mathbb{E}\sup_{t\in T_k}\left|\sum_i (t_i-z_i)X_i\right| + C(d)\left\|(t_i-s_i)_i\right\|_{X,\ln(m)}. \end{aligned} \tag{2.60}$$

The last inequality is true since (we recall $\ln(m)\ge \ln(8)>2$)

$$\left\|\sum_i (z_i-s_i)X_i\right\|_2 \le \left\|\sum_i (z_i-s_i)X_i\right\|_{\ln(m)} \le C(d)\left\|(t_i-s_i)_i\right\|_{X,\ln(m)}.$$

Let us also notice that

$$\begin{aligned}&\mathbb{E}\sup_{t\in T_k}\left|\sum_i (t_i-z_i)X_i\right| = \mathbb{E}\max\left(\left(\sup_{t\in T_k}\sum_i (t_i-z_i)X_i\right)_+, \left(\sup_{t\in T_k}\sum_i (t_i-z_i)X_i\right)_-\right) \\ &\le \mathbb{E}\left(\sup_{t\in T_k}\sum_i (t_i-z_i)X_i\right)_+ + \mathbb{E}\left(\sup_{t\in T_k}\sum_i (t_i-z_i)X_i\right)_- \\ &= 2\mathbb{E}\sup_{t\in T_k}\left(\sum_i (t_i-z_i)X_i\right)_+ = 2\mathbb{E}\sup_{t\in T_k}\sum_i (t_i-z_i)X_i,\end{aligned}$$

where in the second equality we used that X_i are symmetric and in the last one that $z\in T_k$. The above together with (2.60) imply (2.59). □

The next Theorem (together with Lemma 2.4.4) allows us to pass from the bounds on expectations of suprema of Gaussian processes developed in [1] (Theorem 6.19) to empirical processes involving general random variables with bounded fourth moments (in particular all random variables from the $\mathbb{S}(d)$ class).

Theorem 2.4.5. *Let $p\ge 1$ and $T\subset B_2^n+\sqrt{p}B_1$. There is a decomposition $T=\bigcup_{l=1}^N T_l$, $N\le e^{Cp}$ such that for every $l\le N$ and $z\in\mathbb{R}^n$ the following holds:*

$$\mathbb{E}\sup_{x\in T_l}\sum_{i,j} a_{i,j}x_i g_j z_j \le C\left(\sum_{i,j} a_{i,j}^2 z_j^4\right)^{1/4}\left(\sum_{i,j} a_{i,j}^2\right)^{1/4}. \tag{2.61}$$

Proof. Let $\alpha_z(x)=\sqrt{\sum_j z_j^2\left(\sum_i a_{i,j}x_i\right)^2}$. By the Cauchy-Schwarz inequality,

$$\alpha_z(x)\le \left(\sum_{i,j} z_j^4 a_{i,j}^2\right)^{1/4}\beta(x), \tag{2.62}$$

where

$$\beta(x)=\left(\sum_j \frac{\left(\sum_i a_{i,j}x_i\right)^4}{\sum_i a_{i,j}^2}\right)^{1/4}.$$

Let $\mathcal{E}=(\mathcal{E}_j)_j$, where $\mathcal{E}_j$ are i.i.d symmetric exponential r.v's with the density $e^{-|x|}/2$. We have

$$\mathbb{E}\beta(\mathcal{E}) \le \left(\mathbb{E}\beta(\mathcal{E})^4\right)^{1/4} \le C\left(\sum_{i,j} a_{i,j}^2\right)^{1/4}.$$

So by using Corollary 6.9 with, $\sqrt{p}$ instead of p, $\varepsilon = 1/\sqrt{p}$ and $\rho_\alpha(x,y) = \beta(x-y)$ we can decompose T into $\bigcup_{l=1}^N T_l$ in such a way that $N \le \exp(Cp)$ and

$$\forall_{l\le N} \sup_{x,\widetilde{x}\in T_l} \beta(x-\widetilde{x}) \le \frac{C}{\sqrt{p}}\left(\sum_{i,j} a_{i,j}^2\right)^{1/4}. \tag{2.63}$$

By Lemma 6.19 we obtain

$$\begin{aligned}
&\mathbb{E}\sup_{x\in T_l}\sum_{i,j} a_{i,j}x_i g_j z_j \le C\left(\sqrt{\sum_{i,j} a_{i,j}^2 z_j^2} + \sqrt{p}\sup_{x,\widetilde{x}\in T_l}\sqrt{\sum_j z_j^2\left(\sum_i a_{i,j}(x_i-\widetilde{x}_i)\right)^2}\right)\\
&\le \sqrt{\sum_{i,j} a_{i,j}^2 z_j^2} + \sqrt{p}\sup_{x,\widetilde{x}\in T_l}\left(\sum_{i,j} a_{i,j}^2 z_j^4\right)^{1/4}\beta(x-\widetilde{x}) \le C\left(\sum_{i,j} a_{i,j}^2 z_j^4\right)^{1/4}\left(\sum_{i,j} a_{i,j}^2\right)^{1/4},
\end{aligned}$$

where in the last inequality we used the Cauchy-Schwarz inequality and (2.63). □

Fact 2.4.6. *For any symmetric set $T \subset \sqrt{p}B_2^n + pB_1^n$, we have*

$$S(T) \le C(d)\left(\sup_{x\in T}\left\|\left(\sum_i a_{i,j}x_i\right)_j\right\|_{Y,p} + \left\|\left(\sqrt{\sum_i a_{i,j}^2}\right)_j\right\|_{Y,p}\right),$$

where $S(T) := \mathbb{E}\sup_{x\in T}\sum_{i,j} a_{i,j}x_i Y_j$.

Proof. Obviously,

$$S(T) \le \mathbb{E}\sup_{x\in T}\sum_i\sum_{j\le p} a_{i,j}x_i Y_j + \mathbb{E}\sup_{x\in T}\sum_i\sum_{j>p} a_{i,j}x_i Y_j =: S_1(T) + S_2(T). \tag{2.64}$$

By Fact 2.2.2 and Theorem 2.3.1 we have

$$S_1(T) \le C\sup_{x\in T}\left\|\sum_{i,j} a_{i,j}x_i Y_j\right\|_p \le C(d)\sup_{x\in T}\left\|\left(\sum_i a_{i,j}x_i\right)_j\right\|_{Y,p}. \tag{2.65}$$

Now we bound $S_2(T)$. By Theorem 2.4.5 we may decompose T into $T = \bigcup_{l=1}^N T_l$ in such a way that $N \le \exp(Cp)$ and (2.61) holds for $T_l/\sqrt{p}$ instead of T_l and $(a_{i,j})_{i\ge 1, j>p}$ instead of $(a_{i,j})_{i,j}$. Using Lemmas 2.4.4 and 2.2.6 we get

$$\frac{S_2(T)}{C(d)} \le \max_{l\le N}\mathbb{E}\sup_{x\in T_l}\sum_i\sum_{j>p} a_{i,j}x_i Y_j + \sup_{x\in T}\left\|\left(\sum_i a_{i,j}x_i\right)_j\right\|_{Y,p}. \tag{2.66}$$

By the symmetry of Y_j's, Jensen's inequality and the contraction principle

$$\mathbb{E}\sup_{x\in T_l}\sum_i\sum_{j>p}a_{i,j}x_iY_j=\mathbb{E}\sup_{x\in T_l}\sum_i\sum_{j>p}a_{i,j}x_i\varepsilon_j|Y_j|\leq\sqrt{\frac{2p}{\pi}}\mathbb{E}\sup_{x\in T_l/\sqrt{p}}\sum_i\sum_{j>p}a_{i,j}x_ig_j|Y_j|.$$

Theorem 2.4.5 states that last term in the above formula does not exceed

$$\begin{aligned}C(d)\sqrt{p}\,\mathbb{E}\left(\sum_i\sum_{j>p}a_{i,j}^2Y_j^4\right)^{1/4}\left(\sum_i\sum_{j>p}a_{i,j}^2\right)^{1/4}&\leq C(d)\sqrt{p}\left(\sum_i\sum_{j\geq p}a_{i,j}^2\mathbb{E}Y_j^4\right)^{1/4}\\&\leq C(d)\sqrt{p}\sqrt{\sum_i\sum_{j>p}a_{i,j}^2}.\end{aligned}\tag{2.67}$$

In the last inequality we used $\|Y_j\|_4\leq 2^d\|Y_j\|_2=2^d/e$. From Lemma 2.2.3 and (2.54)

$$\frac{\sqrt{p}}{2}\sqrt{\sum_i\sum_{j>p}a_{i,j}^2}\;\leq\;\sup\left\{\sum_j\sqrt{\sum_i a_{i,j}^2}t_j\;\middle|\;\sum_j t_j^2\leq p,\;|t_j|\leq 1\right\}\;\leq\;\left\|\left(\sqrt{\sum_i a_{i,j}^2}\right)_j\right\|_{Y,p}.\tag{2.68}$$

The assertion follows by (2.64)-(2.68). □

2.5 Proof of Theorem 2.1.1

We are ready to prove the main theorem of this chapter. We begin with the lower bound.

Repeated application of (2.29) gives

$$\left\|\sum_{i,j}a_{i,j}X_iY_j\right\|_p\geq c(d)\,\|a_{i,j}\|_{X,Y,p}.\tag{2.69}$$

Symmetry of $Y_j's$, Jensen's inequality, (2.52) with $p=1$, normalization $\|X_i\|_2=1/e$ and (2.29) imply

$$\begin{aligned}\left\|\sum_{i,j}a_{i,j}X_iY_j\right\|_p&=\left\|\sum_j Y_j\left|\sum_i a_{i,j}X_i\right|\right\|_p\geq\left\|\sum_j Y_j\mathbb{E}_X\left|\sum_i a_{i,j}X_i\right|\right\|_p\\&\geq\frac{1}{C(d)}\left\|\sum_j Y_j\sqrt{\mathbb{E}_X\left(\sum_i a_{i,j}X_i\right)^2}\right\|_p\geq\frac{1}{C(d)}\left\|\left(\sum_i a_{i,j}^2\right)_j\right\|_{Y,p}.\end{aligned}\tag{2.70}$$

In the same way we show

$$\left\|\sum_{i,j}a_{i,j}X_iY_j\right\|_p\geq\frac{1}{C(d)}\left\|\sqrt{\sum_j a_{i,j}^2}\right\|_{X,p}.\tag{2.71}$$

Inequalities (2.69)-(2.71) gives the lower bound in Theorem 2.1.1.

Now we establish the upper bound. To this end we observe that Theorems 2.3.1 and 6.25 and Lemma 2.2.5 yield

$$\left\|\sum_{i,j} a_{i,j} X_i Y_j\right\|_p \le C(d)\left(\mathbb{E}\left\|\left\|\left(\sum_j a_{i,j} Y_j\right)_i\right\|\right\|_{X,p,d} + \|(a_{i,j})_{i,j}\|_{X,Y,p}\right),$$

whereas $\mathbb{E}\left\|\left\|\left(\sum_j a_{i,j} Y_j\right)_i\right\|\right\|_{X,p,d}$ is bounded by the following Proposition.

Proposition 2.5.1. *For any $d \ge 1$ we have*

$$\mathbb{E}\left\|\left\|\left(\sum_j a_{i,j} Y_j\right)_i\right\|\right\|_{X,p,d} \le C(d)\left(\|(a_{i,j})\|_{X,Y,p} + \left\|\left(\sqrt{\sum_i a_{i,j}^2}\right)_j\right\|_{Y,p}\right). \tag{2.72}$$

Proof. Since the functions $\hat{M}_i^X(t)$ are symmetric and $1+x \le 2 + x\mathbf{1}_{\{x\ge 1\}}$ we have

$$\begin{aligned}
\mathbb{E}\left\|\left\|\left(\sum_j a_{i,j} Y_j\right)_i\right\|\right\|_{X,p,d} &= \mathbb{E}\sup\left\{\sum_i x_i^1 \prod_{k=2}^d (1+x_i^k)\left|\sum_j a_{i,j} Y_j\right| \;\middle|\; \forall_{k\le d} \sum_i \hat{M}_i^X(x_i^k) \le p\right\} \\
&\le 2^{d-1}\Bigg[\mathbb{E}\sup\left\{\sum_i x_i^1 \left|\sum_j a_{i,j} Y_j\right| \;\middle|\; \sum_i \hat{M}_i^X(x_i^1) \le p\right\} \\
&\quad + \sum_{I\subset[d]} \mathbb{E}\sup\left\{\sum_i \prod_{k\in I} x_i^k \mathbf{1}_{\{|x_i^k|\ge 1\}}\left|\sum_j a_{i,j} Y_j\right| \;\middle|\; \forall_{k\in I} \sum_i \hat{M}_i^X(x_i^k) \le p\right\}\Bigg].
\end{aligned}$$

Since $\{x \in \mathbb{R}^n \mid \sum_i \hat{M}_i^X(x_i) \le p\} \subset \sqrt{p}B_2^n + pB_1^n$ (recall (2.22)) Fact 2.4.6 implies

$$\begin{aligned}
&\frac{1}{C(d)}\mathbb{E}\sup\left\{\sum_i x_i^1 \left|\sum_j a_{i,j} Y_j\right| \;\middle|\; \sum_i \hat{M}_i^X(x_i^1) \le p\right\} \\
&= \frac{1}{C(d)}\mathbb{E}\sup\left\{\sum_i x_i^1 \sum_j a_{i,j} Y_j \;\middle|\; \sum_i \hat{M}_i^X(x_i^1) \le p\right\} \\
&\le \sup\left\{\left\|\left(\sum_i a_{i,j} x_i^1\right)_j\right\|_{Y,p} \;\middle|\; \sum_i \hat{M}_i^X(x_i^1) \le p\right\} + \left\|\left(\sqrt{\sum_i a_{i,j}^2}\right)_j\right\|_{Y,p} \\
&\le C(d)\left(\|(a_i)\|_{X,Y,p} + \left\|\left(\sqrt{\sum_i a_{i,j}^2}\right)_j\right\|_{Y,p}\right).
\end{aligned}$$

If $I \subset [d]$ then

$$\begin{aligned}
&\mathbb{E}\sup\left\{\sum_i \prod_{k\in I} x_i^k \mathbf{1}_{\{|x_i^k|\ge 1\}}\left|\sum_j a_{i,j} Y_j\right| \;\middle|\; \forall_{k\in I} \sum_i \hat{M}_i^X(x_i^k) \le p\right\} \\
&\le \mathbb{E}\sup\left\{\sum_i \prod_{k=1}^d x_i^k \mathbf{1}_{\{|x_i^k|\ge 1\}}\left|\sum_j a_{i,j} Y_j\right| \;\middle|\; \forall_{k\in I} \sum_i \hat{M}_i^X(x_i^k) \le p\right\} =: S,
\end{aligned}$$

where in the inequality we choose $x_i^k = x_i^{k_0}$ for any $k \notin I$ and some $k_0 \in I$.

Let T, U, V be the sets defined in (2.32)-(2.34). Then $T \subset U + V$. So we have that

$$\begin{aligned} S &\le \mathbb{E} \sup_{x^1,\dots,x^d \in U} \sum_i \prod_{k=1}^d x_i^k \left| \sum_j a_{i,j} Y_j \right| + \sum_{\substack{I \subset [d] \\ I \neq \emptyset}} \mathbb{E} \sup_{x^i \in U \text{ for } i \in I} \sup_{x^i \in V \text{ for } i \notin I} \sum_i \prod_{k=1}^d x_i^k \left| \sum_j a_{i,j} Y_j \right| \\ &\le \mathbb{E} \sup_{x^1,\dots,x^d \in U} \sum_i \prod_{k=1}^d x_i^k \left| \sum_j a_{i,j} Y_j \right| + \left(2^d - 1\right) \mathbb{E} \sup_{x^1,\dots,x^{d-1} \in T, x^d \in V} \sum_i \prod_{k=1}^d x_i^k \left| \sum_j a_{i,j} Y_j \right|. \end{aligned} \tag{2.73}$$

The last inequality follows by the symmetry and the inclusions $U, V \subset T$. Observe that (2.55) implies

$$\mathbb{E} \sup_{x^1,\dots,x^d \in U} \sum_i \prod_{k=1}^d x_i^k \left| \sum_j a_{i,j} Y_j \right| \le C(d) \left\| \left(\sqrt{\sum_i a_{i,j}^2} \right)_j \right\|_{Y,p},$$

whereas (2.56) implies

$$\mathbb{E} \sup_{x^1,\dots,x^{d-1} \in T, x^d \in V} \sum_i \prod_{k=1}^d x_i^k \left| \sum_j a_{i,j} Y_j \right| \le C(d) \left\| (a_i) \right\|_{X,Y,p}.$$

$\square$

Part II
Vector case

Chapter 3
Hanson-Wright inequality in Banach Spaces.

3.1 Introduction and main results

The Hanson-Wright inequality gives an upper bound for tails of real quadratic forms in independent subgaussian random variables. Recall that a random variable X is called α-subgaussian if for every $t>0$, $\mathbb{P}(|X|\geq t)\leq 2\exp(-t^2/2\alpha^2)$. The Hanson-Wright inequality states that for any sequence of independent mean zero α-subgaussian random variables $X_1,\ldots,X_n$ and any symmetric real-valued matrix $A=(a_{ij})_{i,j\leq n}$ one has

$$\mathbb{P}\left(\left|\sum_{i,j=1}^{n} a_{ij}(X_iX_j-\mathbb{E}(X_iX_j))\right|\geq t\right)\leq 2\exp\left(-\frac{1}{C}\min\left\{\frac{t^2}{\alpha^4\|A\|_{\mathrm{HS}}},\frac{t}{\alpha^2\|A\|_{\mathrm{op}}}\right\}\right), \tag{3.1}$$

where in the whole chapter we use the letter C to denote universal constants which may differ at each occurrence. Estimate (3.1) was essentially established in [10] in the symmetric and in [33] in the mean zero case (in fact in both papers the operator norm of A was replaced by the operator norm of $(|a_{ij}|)$, which in general could be much bigger, proofs of (3.1) may be found in [4] and [28]).

The Hanson-Wright inequality has found numerous applications in high-dimensional probability and statistics, as well as in random matrix theory (see e.g., [32]). However in many problems one faces the need to analyze not a single quadratic form but a supremum of a collection of them or equivalently a norm of a quadratic form with coefficients in a Banach space. While in the literature there are inequalities addressing this problem (see ineq. (3.3) below), they are usually expressed in terms of quantities which themselves are troublesome to analyze. The main objective of this chapter is to provide estimates on vector-valued quadratic forms which can be applied more easily and are of optimal form.

The main step in modern proofs of the Hanson-Wright inequality is to get a bound similar to (3.1) in the Gaussian case. The extension to general subgaussian variables is then obtained with use of the by now standard tools of probability in Banach spaces, such as decoupling, symmetrization and the contraction principle. Via Chebyshev's inequality to obtain a tail estimate it is enough to bound appropriately the moments of quadratic forms in the case when $X_i=g_i$ are standard Gaussian $\mathcal{N}(0,1)$ random variables. One may in fact show that (cf. [16,17])

$$\left(\mathbb{E}\left|\sum_{i,j=1}^{n} a_{ij}(g_ig_j-\delta_{ij})\right|^p\right)^{1/p}\sim p\|A\|_{\mathrm{op}}+\sqrt{p}\|A\|_{\mathrm{HS}}, \tag{3.2}$$

where δ_{ij} is the Kronecker delta, and $\sim$ stands for a comparison up to universal multiplicative constants.

Following the same line of arguments, in order to extend the Hanson-Wright bound to the Banach space setting we first estimate moments of centered vector-valued Gaussian quadratic forms, i.e. quantities

$$\Big\|\sum_{i,j=1}^n a_{ij}(g_ig_j-\delta_{ij})\Big\|_p = \left(\mathbb{E}\Big\|\sum_{i,j=1}^n a_{ij}(g_ig_j-\delta_{ij})\Big\|^p\right)^{1/p}, \quad p\geq 1,$$

where $A=(a_{ij})_{i,j\leq n}$ is a symmetric matrix with values in a normed space $(F,\|\ \|)$. We note that (as mentioned above) there exist two-sided estimates for the moments of Gaussian quadratic forms with vector-valued coefficients. To the best of our knowledge they were obtained first in [5] and then they were reproved in various context by several authors (see e.g., [3, 22, 24]). They state that for $p\geq 1$,

$$\Big\|\sum_{ij} a_{ij}(g_ig_j-\delta_{ij})\Big\|_p \sim \mathbb{E}\Big\|\sum_{ij} a_{ij}(g_ig_j-\delta_{ij})\Big\| + \sqrt{p}\,\mathbb{E}\sup_{\|x\|_2\leq 1}\Big\|\sum_{ij} a_{ij}x_ig_j\Big\| + p\sup_{\|x\|_2\leq 1,\|y\|_2\leq 1}\Big\|\sum_{ij} a_{ij}x_iy_j\Big\|. \tag{3.3}$$

Unfortunately the second term on the right hand side of (3.3) is usually difficult to estimate. The main effort in this chapter will be to replace it by quantities which even if still involve expected values of Banach space valued random variables in many situations can be handled more easily. More precisely, we will obtain inequalities in which additional suprema over Euclidean spheres are placed outside the expectations, which reduces the complexity of the involved stochastic processes. As one of the consequences we will derive two-sided bounds in L_r spaces involving only purely deterministic quantities.

Our first observation is a simple lower bound

Proposition 3.1.1. *Let $(a_{ij})_{i,j\leq n}$ be a symmetric matrix with values in a normed space $(F,\|\cdot\|)$. Then for any $p\geq 1$ we have*

$$\Big\|\sum_{ij} a_{ij}(g_ig_j-\delta_{ij})\Big\|_p \geq \frac{1}{C}\Bigg(\mathbb{E}\Big\|\sum_{ij} a_{ij}(g_ig_j-\delta_{ij})\Big\| + \sqrt{p}\sup_{\|x\|_2\leq 1}\mathbb{E}\Big\|\sum_{i\neq j} a_{ij}x_ig_j\Big\| + \sqrt{p}\sup_{\|(x_{ij})\|_2\leq 1}\Big\|\sum_{ij} a_{ij}x_{ij}\Big\| + p\sup_{\|x\|_2\leq 1,\|y\|_2\leq 1}\Big\|\sum_{ij} a_{ij}x_iy_j\Big\|\Bigg).$$

This motivates the following conjecture.

Conjecture 3.1.2. *Under the assumptions of Proposition 3.1.1 we have*

$$\Big\|\sum_{ij} a_{ij}(g_ig_j-\delta_{ij})\Big\|_p \leq C\Bigg(\mathbb{E}\Big\|\sum_{ij} a_{ij}(g_ig_j-\delta_{ij})\Big\| + \sqrt{p}\sup_{\|x\|_2\leq 1}\mathbb{E}\Big\|\sum_{i\neq j} a_{ij}x_ig_j\Big\| + \sqrt{p}\sup_{\|(x_{ij})\|_2\leq 1}\Big\|\sum_{ij} a_{ij}x_{ij}\Big\| + p\sup_{\|x\|_2\leq 1,\|y\|_2\leq 1}\Big\|\sum_{ij} a_{ij}x_iy_j\Big\|\Bigg).$$

We are able to show that the conjectured estimate holds up to logarithmic factors.

Theorem 3.1.3. *Let $(a_{ij})_{i,j\leq n}$ be a symmetric matrix with values in a normed space $(F,\|\cdot\|)$. Then for any $p\geq 1$ the following two estimates hold*

$$\Big\|\sum_{ij}a_{ij}(g_ig_j-\delta_{ij})\Big\|_p \leq C\bigg(\ln(ep)\mathbb{E}\Big\|\sum_{ij}a_{ij}(g_ig_j-\delta_{ij})\Big\|+\sqrt{p}\sup_{\|x\|_2\leq 1}\mathbb{E}\Big\|\sum_{i\neq j}a_{ij}x_ig_j\Big\| \\ +\sqrt{p}\sup_{\|(x_{ij})\|_2\leq 1}\Big\|\sum_{ij}a_{ij}x_{ij}\Big\|+p\ln(ep)\sup_{\|x\|_2\leq 1,\|y\|_2\leq 1}\Big\|\sum_{ij}a_{ij}x_iy_j\Big\|\bigg) \tag{3.4}$$

and

$$\Big\|\sum_{ij}a_{ij}(g_ig_j-\delta_{ij})\Big\|_p \leq C\bigg(\mathbb{E}\Big\|\sum_{ij}a_{ij}(g_ig_j-\delta_{ij})\Big\|+\sqrt{p}\sup_{\|x\|_2\leq 1}\mathbb{E}\Big\|\sum_{i\neq j}a_{ij}x_ig_j\Big\| \\ +\sqrt{p}\ln(ep)\sup_{\|(x_{ij})\|_2\leq 1}\Big\|\sum_{ij}a_{ij}x_{ij}\Big\|+p\sup_{\|x\|_2\leq 1,\|y\|_2\leq 1}\Big\|\sum_{ij}a_{ij}x_iy_j\Big\|\bigg). \tag{3.5}$$

One of the main reasons behind the appearance of additional logarithmic factors is lack of good Sudakov-type estimates for Gaussian quadratic forms. Such bounds hold for linear forms and as a result we may show the following ($(g_{i,j})_{i,j\leq n}$ below denote as usual i.i.d. $\mathcal{N}(0,1)$ random variables).

Theorem 3.1.4. *Under the assumptions of Theorem 3.1.3 we have*

$$\Big\|\sum_{ij}a_{ij}(g_ig_j-\delta_{ij})\Big\|_p \leq C\bigg(\mathbb{E}\Big\|\sum_{ij}a_{ij}(g_ig_j-\delta_{ij})\Big\|+\mathbb{E}\Big\|\sum_{i\neq j}a_{ij}g_{ij}\Big\| \\ +\sqrt{p}\sup_{\|x\|_2\leq 1}\mathbb{E}\Big\|\sum_{i\neq j}a_{ij}x_ig_j\Big\|+\sqrt{p}\sup_{\|(x_{ij})\|_2\leq 1}\Big\|\sum_{ij}a_{ij}x_{ij}\Big\| \\ +p\sup_{\|x\|_2\leq 1,\|y\|_2\leq 1}\Big\|\sum_{ij}a_{ij}x_iy_j\Big\|\bigg). \tag{3.6}$$

In particular we know that Conjecture 3.1.2 holds in Banach spaces, in which Gaussian quadratic forms dominate in mean Gaussian linear forms, i.e. in Banach spaces $(F,\|\ \|)$ for which there exists a constant $\lambda<\infty$ such for any finite symmetric matrix $(a_{ij})_{i,j\leq n}$ with values in F one has

$$\mathbb{E}\Big\|\sum_{i\neq j}a_{ij}g_{ij}\Big\|\leq\lambda\mathbb{E}\Big\|\sum_{i\neq j}a_{ij}g_ig_j\Big\|. \tag{3.7}$$

It is easy to check (see Proposition 3.2.1 below) that such property holds for L_r-spaces with $\lambda=\lambda(r)\leq Cr$.

Remark 3.1.5. For non-centered Gaussian quadratic forms $S=\sum_{ij}a_{ij}g_ig_j$ one has $\|S\|_p\sim\|\mathbb{E}S\|+\|S-\mathbb{E}S\|_p$, so Proposition 3.1.1 yields

$$\left\| \sum_{ij} a_{ij} g_i g_j \right\|_p \geq \frac{1}{C} \Bigg(\mathbb{E} \left\| \sum_{ij} a_{ij} g_i g_j \right\| + \sqrt{p} \sup_{\|x\|_2 \leq 1} \mathbb{E} \left\| \sum_{i \neq j} a_{ij} x_i g_j \right\| + \sqrt{p} \sup_{\|(x_{ij})\|_2 \leq 1} \left\| \sum_{ij} a_{ij} x_{ij} \right\| + p \sup_{\|x\|_2 \leq 1, \|y\|_2 \leq 1} \left\| \sum_{ij} a_{ij} x_i y_j \right\| \Bigg).$$

and Theorem 3.1.4 implies

$$\left\| \sum_{ij} a_{ij} g_i g_j \right\|_p \leq C \Bigg(\mathbb{E} \left\| \sum_{ij} a_{ij} g_i g_j \right\| + \mathbb{E} \left\| \sum_{i \neq j} a_{ij} g_{ij} \right\| + \sqrt{p} \sup_{\|x\|_2 \leq 1} \mathbb{E} \left\| \sum_{i \neq j} a_{ij} x_i g_j \right\| + \sqrt{p} \sup_{\|(x_{ij})\|_2 \leq 1} \left\| \sum_{ij} a_{ij} x_{ij} \right\| + p \sup_{\|x\|_2 \leq 1, \|y\|_2 \leq 1} \left\| \sum_{ij} a_{ij} x_i y_j \right\| \Bigg).$$

Proposition 3.1.1 and Theorem 3.1.4 may be expressed in terms of tails.

Theorem 3.1.6. *Let $(a_{ij})_{i,j \leq n}$ be a symmetric matrix with values in a normed space $(F, \|\cdot\|)$. Then for any $t > 0$,*

$$\mathbb{P}\left(\left\| \sum_{ij} a_{ij}(g_i g_j - \delta_{ij}) \right\| \geq t + \frac{1}{C} \mathbb{E} \left\| \sum_{ij} a_{ij}(g_i g_j - \delta_{ij}) \right\| \right) \geq \frac{1}{C} \exp\left(-C \min\left\{ \frac{t^2}{U^2}, \frac{t}{V} \right\} \right),$$

where

$$U = \sup_{\|x\|_2 \leq 1} \mathbb{E} \left\| \sum_{i \neq j} a_{ij} x_i g_j \right\| + \sup_{\|(x_{ij})\|_2 \leq 1} \left\| \sum_{ij} a_{ij} x_{ij} \right\|, \tag{3.8}$$

$$V = \sup_{\|x\|_2 \leq 1, \|y\|_2 \leq 1} \left\| \sum_{ij} a_{ij} x_i y_j \right\|. \tag{3.9}$$

Moreover, for $t > C(\mathbb{E}\| \sum_{ij} a_{ij}(g_i g_j - \delta_{ij})\| + \mathbb{E}\| \sum_{i \neq j} a_{ij} g_{ij}\|)$ we have

$$\mathbb{P}\left(\left\| \sum_{ij} a_{ij}(g_i g_j - \delta_{ij}) \right\| \geq t \right) \leq 2 \exp\left(-\frac{1}{C} \min\left\{ \frac{t^2}{U^2}, \frac{t}{V} \right\} \right).$$

As a corollary we get a Hanson-Wright-type inequality for Banach space valued quadratic forms in general independent subgaussian random variables.

Theorem 3.1.7. *Let $X_1, X_2, \ldots, X_n$ be independent mean zero α-subgaussian random variables. Then for any symmetric matrix $(a_{ij})_{i,j \leq n}$ with values in a normed space $(F, \|\cdot\|)$ and $t > C\alpha^2(\mathbb{E}\| \sum_{ij} a_{ij}(g_i g_j - \delta_{ij})\| + \mathbb{E}\| \sum_{i \neq j} a_{ij} g_{ij}\|)$ we have*

$$\mathbb{P}\left(\left\| \sum_{ij} a_{ij}(X_i X_j - \mathbb{E}(X_i X_j)) \right\| \geq t \right) \leq 2 \exp\left(-\frac{1}{C} \min\left\{ \frac{t^2}{\alpha^4 U^2}, \frac{t}{\alpha^2 V} \right\} \right), \tag{3.10}$$

where U and V are as in Theorem 3.1.6.

Remark 3.1.8. It is not hard to check that in the case $F = \mathbb{R}$ we have $U \sim \|(a_{ij})\|_{\mathrm{HS}}$ and $V = \|(a_{ij})\|_{\mathrm{op}}$. Moreover,

$$\mathbb{E}\left\|\sum_{ij}a_{ij}(g_ig_j-\delta_{ij})\right\|+\mathbb{E}\left\|\sum_{i\neq j}a_{ij}g_{ij}\right\|\leq 2\|(a_{ij})\|_{\mathrm{HS}},$$

so the right hand side of (3.10) is at least 1 for $t<C'(\mathbb{E}\|\sum_{ij}a_{ij}(g_ig_j-\delta_{ij})\|+\mathbb{E}\|\sum_{i\neq j}a_{ij}g_{ij}\|)$ and sufficiently large C. Hence (3.10) holds for any $t>0$ in the real case and is equivalent to the Hanson-Wright bound.

Remark 3.1.9. Proposition 3.4.2 below shows that we may replace in all estimates above the term $\sup_{\|x\|_2\leq 1}\mathbb{E}\|\sum_{i\neq j}a_{ij}x_ig_j\|$ by $\sup_{\|x\|_2\leq 1}\mathbb{E}\|\sum_{ij}a_{ij}x_ig_j\|$.

The organization of the chapter is as follows. In the next section we discuss a few corollaries of Theorems 3.1.4 and 3.1.7. In Section 3.3 we prove Proposition 3.1.1 and show that it is enough to bound separately moments of diagonal and off-diagonal parts of chaoses. In Section 3.4 we reduce Theorems 3.1.3 and 3.1.4 to the problem of estimating means of suprema of certain Gaussian processes. In Section 3.5 we show how to bound expectations of such suprema – the main new ingredient are entropy bounds presented in Corollary 3.5.3 (derived via volumetric-type arguments). Unfortunately our entropy bounds are too weak to use the Dudley integral bound. Instead, we present a technical chaining argument (of similar type as in [16]). In the last section we conclude the proofs of main Theorems.

3.2 Consequences and extensions

3.2.1 L_r-spaces

We start with showing that L_r spaces for $r<\infty$, satisfy (3.7) with $\lambda=Cr$, so Theorem 3.1.4 implies Conjecture 3.1.2 for L_r spaces (and as a consequence the Hanson-Wright inequality). Moreover, in this case one may express all parameters without any expectations as is shown in the proposition below.

Proposition 3.2.1. *For any symmetric matrix $(a_{ij})_{i,j\leq n}$ with values in $L_r=L_r(X,\mu)$, $1\leq r<\infty$ and $x_1,\dots,x_n\in\mathbb{R}$ we have*

$$\frac{1}{C}\left\|\sqrt{\sum_{ij}a_{ij}^2}\right\|_{L_r}\leq\mathbb{E}\left\|\sum_{ij}a_{ij}g_{ij}\right\|_{L_r}\leq C\sqrt{r}\left\|\sqrt{\sum_{ij}a_{ij}^2}\right\|_{L_r}, \tag{3.11}$$

$$\frac{1}{C}\left\|\sqrt{\sum_{j}\left(\sum_{j}a_{ij}x_i\right)^2}\right\|_{L_r}\leq\mathbb{E}\left\|\sum_{ij}a_{ij}x_ig_j\right\|_{L_r}\leq C\sqrt{r}\left\|\sqrt{\sum_{j}\left(\sum_{i}a_{ij}x_i\right)^2}\right\|_{L_r}, \tag{3.12}$$

$$\frac{1}{C\sqrt{r}}\left\|\sqrt{\sum_{ij}a_{ij}^2}\right\|_{L_r}\leq\mathbb{E}\left\|\sum_{ij}a_{ij}(g_ig_j-\delta_{ij})\right\|_{L_r}\leq Cr\left\|\sqrt{\sum_{ij}a_{ij}^2}\right\|_{L_r}. \tag{3.13}$$

Proof. For any a_i's in L_r the Gaussian concentration yields

$$\mathbb{E}\left\|\sum_i a_ig_i\right\|_{L_r}\leq\left(\mathbb{E}\left\|\sum_i a_ig_i\right\|_{L_r}^r\right)^{1/r}\leq C\sqrt{r}\,\mathbb{E}\left\|\sum_i a_ig_i\right\|_{L_r}.$$

Since

$$\left(\mathbb{E}\left\|\sum_i a_i g_i\right\|_{L_r}^r\right)^{1/r} = \left(\int_X \mathbb{E}\left|\sum_i a_i(x) g_i\right|^r d\mu(x)\right)^{1/r}$$
$$= \left(\int_X \mathbb{E}|g_1|^r \Big(\sum_i a_i^2(x)\Big)^{r/2} d\mu(x)\right)^{1/r} \sim \sqrt{r}\left\|\sqrt{\sum_i a_i^2}\right\|_{L_r},$$

estimates (3.11),(3.12) follow easily. The proof of (3.13) is analogous. It is enough to observe that from Theorem 6.3

$$\mathbb{E}\left\|\sum_{ij} a_{ij}(g_i g_j - \delta_{ij})\right\|_{L_r} \le \left(\mathbb{E}\left\|\sum_{ij} a_{ij}(g_i g_j - \delta_{ij})\right\|_{L_r}^r\right)^{1/r} \le Cr\mathbb{E}\left\|\sum_{ij} a_{ij}(g_i g_j - \delta_{ij})\right\|_{L_r}$$

and (3.2) imples for any $x \in X$,

$$\frac{\sqrt{r}}{C}\sqrt{\sum_{ij} a_{ij}^2(x)} \le \left(\mathbb{E}\left|\sum_{ij} a_{ij}(x)(g_i g_j - \delta_{ij})\right|^r\right)^{1/r} \le Cr\sqrt{\sum_{ij} a_{ij}^2(x)}.$$

□

The above proposition, together with Proposition 3.1.1 and Theorems 3.1.4 and 3.1.7 immediately yield the following corollaries (in particular they imply that Conjecture 3.1.2 holds in L_r spaces with r-dependent constants)

Corollary 3.2.2. *For any symmetric matrix $(a_{ij})_{ij}$ with values in L_r and $p \ge 1$ we have*

$$\left\|\sum_{ij} a_{ij}(g_i g_j - \delta_{ij})\right\|_p \sim^r \left\|\sqrt{\sum_{ij} a_{ij}^2}\right\|_{L_r} + \sqrt{p}\sup_{\|x\|_2\le 1}\left\|\sqrt{\sum_j\Big(\sum_{i\ne j} a_{ij}x_i\Big)^2}\right\|_{L_r}$$
$$+\sqrt{p}\sup_{\|(x_{ij})\|_2\le 1}\left\|\sum_{ij} a_{ij}x_{ij}\right\|_{L_r} + p\sup_{\|x\|_2\le 1,\|y\|_2\le 1}\left\|\sum_{ij} a_{ij}x_i y_j\right\|_{L_r}.$$

The implicit constants in the estimates for moments can be taken to be equal to Cr in the upper bound and $r^{-1/2}/C$ in the lower bound.

Corollary 3.2.3. *Let $X_1, X_2, \ldots, X_n$ be independent mean zero α-subgaussian random variables. Then for any symmetric finite matrix $(a_{ij})_{i,j\le n}$ with values in $L_r = L_r(X,\mu)$, $1 \le r < \infty$ and $t > C\alpha^2 r\|\sqrt{\sum_{ij} a_{ij}^2}\|_{L_r}$ we have*

$$\mathbb{P}\left(\left\|\sum_{ij} a_{ij}(X_i X_j - \mathbb{E}(X_i X_j))\right\|_{L_r} \ge t\right) \le 2\exp\left(-\frac{1}{C}\min\left\{\frac{t^2}{\alpha^4 r U^2}, \frac{t}{\alpha^2 V}\right\}\right), \tag{3.14}$$

where

$$U = \sup_{\|x\|_2 \le 1} \left\| \sqrt{\sum_j \left(\sum_{i \ne j} a_{ij} x_i \right)^2} \right\|_{L_r} + \sup_{\|(x_{ij})\|_2 \le 1} \left\| \sum_{ij} a_{ij} x_{ij} \right\|_{L_r},$$
$$V = \sup_{\|x\|_2 \le 1, \|y\|_2 \le 1} \left\| \sum_{ij} a_{ij} x_i y_j \right\|_{L_r}.$$

3.2.2 Spaces of type 2

Recall that a normed space F is of type 2 with constant λ if for every positive integer n and $v_1, \ldots, v_n \in F$,

$$\mathbb{E}\left\| \sum_{i=1}^n v_i \varepsilon_i \right\| \le \lambda \sqrt{\sum_{i=1}^n \|v_i\|^2},$$

where $\varepsilon_1, \varepsilon_2, \ldots$ is a sequence of independent Rademacher variables.

By standard symmetrization inequalities one easily obtains that if F is of type two with constant λ then for any independent random variables X_i,

$$\mathbb{E}\left\| \sum_i a_i (X_i^2 - \mathbb{E}X_i^2) \right\| \le 2\lambda \sqrt{\sum_i \|a_i\|^2 \mathbb{E}X_i^4}$$

and if $\mathbb{E}X_i = 0$, then decoupling arguments combined with symmetrization and Khintchine-Kahane inequalities give

$$\mathbb{E}\left\| \sum_{i \ne j} a_{ij} X_i X_j \right\| \le C\lambda^2 \sqrt{\sum_{i \ne j} \|a_{ij}\|^2 \mathbb{E}X_i^2 \mathbb{E}X_j^2}.$$

Therefore, Theorem 3.1.7 gives immediately the following

Corollary 3.2.4. *Let $X_1, X_2, \ldots, X_n$ be independent mean zero α-subgaussian random variables and let F be a normed space of type two constant λ. Then for any symmetric finite matrix $(a_{ij})_{i,j \le n}$ with values in F and $t > C\lambda^2\alpha^2 \sqrt{\sum_{ij} \|a_{ij}\|^2}$ we have*

$$\mathbb{P}\left(\left\| \sum_{ij} a_{ij} (X_i X_j - \mathbb{E}(X_i X_j)) \right\| \ge t \right) \le 2\exp\left(-\frac{1}{C} \min\left\{ \frac{t^2}{\alpha^4 U^2}, \frac{t}{\alpha^2 V} \right\} \right), \tag{3.15}$$

where

$$U = \lambda \sup_{\|x\|_2 \le 1} \sqrt{\sum_j \left\| \sum_{i \ne j} a_{ij} x_i \right\|^2} + \sup_{\|(x_{ij})\|_2 \le 1} \left\| \sum_{ij} a_{ij} x_{ij} \right\|, \quad V = \sup_{\|x\|_2 \le 1, \|y\|_2 \le 1} \left\| \sum_{ij} a_{ij} x_i y_j \right\|.$$

Remark 3.2.5. We note that from Theorem 3.1.7 one can also derive similar inequalities for suprema of quadratic forms over VC-type classes of functions appearing e.g., in the analysis of randomized U-processes (cf. e.g., [7, Chapter 5.4]).

3.2.3 Random vectors with dependencies

Let us assume that $X=(X_1,\ldots,X_n)$ is an image of a standard Gaussian vector in $\mathbb{R}^n$ under an α-Lipschitz map. In particular, by the celebrated Caffarelli contraction principle [6], this is true if X has density of the form e^{-V}, where $\nabla^2 V \geq \alpha^{-2}\mathrm{Id}$. As observed by Ledoux and Oleszkiewicz [23, Corollary 1], by combining the well known comparison result due to Pisier [27] with a stochastic domination-type argument, one gets that for any smooth function $f\colon \mathbb{R}^n \to F$, and any $p\geq 1$,

$$\|f(X)-\mathbb{E}f(X)\|_p \leq \frac{\pi\alpha}{2}\|\langle \nabla f(X), G\rangle\|_p, \tag{3.16}$$

where here and subsequently G_n is a standard Gaussian vector in $\mathbb{R}^n$ independent of X and for $a\in F^n$, $b\in\mathbb{R}^n$ we denote $\langle a,b\rangle = \sum_{i=1}^n a_i b_i$. This inequality together with Theorem 3.1.4 allow us to implement a simple argument from [2] and obtain inequalities for quadratic forms and more general F-valued functions of the random vector X. Below, we will denote the second partial derivatives of f by $\partial_{ij} f$. For the sake of brevity, we will focus on moment estimates, clearly tail bounds follow from them by an application of the Chebyshev inequality.

Corollary 3.2.6. *Let X be an α-Lipschitz image of a standard Gaussian vector in $\mathbb{R}^n$ and let $f\colon\mathbb{R}^n\to F$ be a function with bounded derivatives of order two. Assume moreover that $\mathbb{E}\nabla f(X)=0$. Then for any $p\geq 2$,*

$$\begin{aligned}\|f(X)-\mathbb{E}f(X)\|_p \leq & C\alpha^2 \sup_{z\in\mathbb{R}^n}\Bigg(\mathbb{E}\Big\|\sum_{ij}\partial_{ij}f(z)(g_ig_j-\delta_{ij})\Big\| + \mathbb{E}\Big\|\sum_{i\neq j}\partial_{ij}f(z)g_{ij}\Big\| \\ & + \sqrt{p}\sup_{\|x\|_2\leq 1}\mathbb{E}\Big\|\sum_{i\neq j}\partial_{ij}f(z)x_ig_j\Big\| + \sqrt{p}\sup_{\|(x_{ij})\|_2\leq 1}\Big\|\sum_{ij}\partial_{ij}f(z)x_{ij}\Big\| \\ & + p\sup_{\|x\|_2\leq 1,\|y\|_2\leq 1}\Big\|\sum_{ij}\partial_{ij}f(z)x_iy_j\Big\|\Bigg).\end{aligned} \tag{3.17}$$

In particular if X is of mean zero, then

$$\begin{aligned}\Big\|\sum_{ij}a_{ij}(X_iX_j-\mathbb{E}(X_iX_j))\Big\|_p \leq & C\alpha^2\Bigg(\mathbb{E}\Big\|\sum_{ij}a_{ij}(g_ig_j-\delta_{ij})\Big\| + \mathbb{E}\Big\|\sum_{i\neq j}a_{ij}g_{ij}\Big\| \\ & + \sqrt{p}\sup_{\|x\|_2\leq 1}\mathbb{E}\Big\|\sum_{i\neq j}a_{ij}x_ig_j\Big\| + \sqrt{p}\sup_{\|(x_{ij})\|_2\leq 1}\Big\|\sum_{ij}a_{ij}x_{ij}\Big\| \\ & + p\sup_{\|x\|_2\leq 1,\|y\|_2\leq 1}\Big\|\sum_{ij}a_{ij}x_iy_j\Big\|\Bigg)\end{aligned} \tag{3.18}$$

and the inequality (3.10) *is satisfied.*

Proof. Let $G_n=(g_1,\ldots,g_n)$, $G_n'=(g_1',\ldots,g_n')$ be independent standard Gaussian vectors in $\mathbb{R}^n$, independent of X. By an iterated application of (3.16) (the second time conditionally on G_n) we have

$$\mathbb{E}\|f(X)-\mathbb{E}f(X)\|^p \le C^p\alpha^p\mathbb{E}\|\langle\nabla f(X),G_n\rangle\|^p \le C^{2p}\alpha^{2p}\mathbb{E}\left\|\sum_{ij}\partial_{ij}f(X)g_ig_j'\right\|^p$$
$$\le \tilde{C}^{2p}\alpha^{2p}\mathbb{E}\left\|\sum_{ij}\partial_{ij}f(X)(g_ig_j-\delta_{ij})\right\|^p,$$

where the last inequality follows by Theorem 6.4. To finish the proof of (3.17) it is now enough to apply Theorem 3.1.4 conditionally on X and replace the expectation in X by the supremum over $z\in\mathbb{R}^n$.

The inequality (3.18) follows by a direct application of (3.17). □

3.3 Lower bounds

In this part we show Proposition 3.1.1 and the lower bound in Theorem 3.1.6. We start with a simple lemma.

Lemma 3.3.1. *Let $W=\|\sum_{i\neq j}a_{ij}g_ig_j\|_p+\|\sum_i a_{ii}(g_i^2-1)\|_p$. Then for any $p\ge 1$,*

$$\frac{1}{3}W\le\left\|\sum_{ij}a_{ij}(g_ig_j-\delta_{ij})\right\|_p\le W.$$

Proof. Let $(\varepsilon_i)_i$ be a sequence of i.i.d. symmetric ±1 r.v's independent of $(g_i)_i$. We have by symmetry of g_i and Jensen's inequality,

$$\left\|\sum_{ij}a_{ij}(g_ig_j-\delta_{ij})\right\|_p=\left\|\sum_{ij}a_{ij}(\varepsilon_i\varepsilon_jg_ig_j-\delta_{ij})\right\|_p\ge\left\|\mathbb{E}_\varepsilon\sum_{ij}a_{ij}(\varepsilon_i\varepsilon_jg_ig_j-\delta_{ij})\right\|_p$$
$$=\left\|\sum_i a_{ii}(g_i^2-1)\right\|_p.$$

To conclude we use the triangle inequality in L_p and get

$$\left\|\sum_{i\neq j}a_{ij}g_ig_j\right\|_p\le\left\|\sum_{ij}a_{ij}(g_ig_j-\delta_{ij})\right\|_p+\left\|\sum_i a_{ii}(g_i^2-1)\right\|_p\le 2\left\|\sum_{ij}a_{ij}(g_ig_j-\delta_{ij})\right\|_p.$$

Adding the inequalities above yields the first estimate of the lemma. The second one follows trivially from the triangle inequality. □

Proof of Proposition 3.1.1. Obviously

$$\left\|\sum_{ij}a_{ij}(g_ig_j-\delta_{ij})\right\|_p\ge\mathbb{E}\left\|\sum_{ij}a_{ij}(g_ig_j-\delta_{ij})\right\|.$$

Moreover, denoting by $\|\cdot\|_*$ the norm in the dual of F, we have

$$\left\|\sum_{ij} a_{ij}(g_ig_j-\delta_{ij})\right\|_p \geq \sup_{\|\varphi\|_*\leq 1}\left\|\sum_{ij}\varphi(a_{ij})(g_ig_j-\delta_{ij})\right\|_p$$

$$\geq \frac{1}{C}\left(\sqrt{p}\sup_{\|\varphi\|_*\leq 1}\|(\varphi(a_{ij}))_{ij}\|_{\mathrm{HS}}+p\sup_{\|\varphi\|_*\leq 1}\|(\varphi(a_{ij}))_{ij}\|_{\mathrm{op}}\right)$$

$$=\frac{1}{C}\left(\sqrt{p}\sup_{\|(x_{ij})\|_2\leq 1}\left\|\sum_{ij}a_{ij}x_{ij}\right\|+p\sup_{\|x\|_2\leq 1,\|y\|_2\leq 1}\left\|\sum_{ij}a_{ij}x_iy_j\right\|\right),$$

where in the second inequality we used (3.2).

Lemma 3.3.1 and the decoupling Theorem of Kwapień 6.5 (see also Theorem 6.6) yield

$$\left\|\sum_{ij} a_{ij}(g_ig_j-\delta_{ij})\right\|_p \geq \frac{1}{3}\left\|\sum_{i\neq j}a_{ij}g_ig_j\right\|_p \geq \frac{1}{C}\left\|\sum_{i\neq j}a_{ij}g_ig_j'\right\|_p, \tag{3.19}$$

where $(g_i')_i$ denotes an independent copy of $(g_i)_i$.

For any finite sequence $(b_i)_i$ in $(F,\|\cdot\|)$ we have

$$\left\|\sum_i b_ig_i\right\|_p \geq \sup_{\|\varphi\|_*\leq 1}\left\|\sum_i\varphi(b_i)g_i\right\|_p = \sup_{\|\varphi\|_*\leq 1}\|(\varphi(b_i))_i\|_2\cdot\|g_1\|_p \geq \frac{\sqrt{p}}{C}\sup_{\|x\|_2\leq 1}\left\|\sum_i x_ib_i\right\|. \tag{3.20}$$

Thus, by (3.19) and the Fubini Theorem, we get

$$\left\|\sum_{ij} a_{ij}(g_ig_j-\delta_{ij})\right\|_p \geq \frac{\sqrt{p}}{C}\sup_{\|x\|_2\leq 1}\left\|\sum_{i\neq j}a_{ij}x_ig_j'\right\|_p \geq \frac{\sqrt{p}}{C}\sup_{\|x\|_2\leq 1}\mathbb{E}\left\|\sum_{i\neq j}a_{ij}x_ig_j\right\|.$$

□

3.4 Reduction to a bound on the supremum of a Gaussian process

In this section we will reduce the upper estimates of Theorems 3.1.3 and 3.1.4 to an estimate on expected value of a supremum of a certain Gaussian process. The arguments in this part of the chapter are well-known, we present them for the sake of completeness. In particular we will demonstrate the upper bounds given in (3.3).

The first lemma shows that we may easily bound the diagonal terms.

Lemma 3.4.1. *For $p\geq 1$ we have*

$$\left\|\sum_i a_{ii}(g_i^2-1)\right\|_p \leq C\left(\mathbb{E}\left\|\sum_{ij}a_{ij}(g_ig_j-\delta_{ij})\right\|+\sqrt{p}\sup_{\|(x_{ij})\|_2\leq 1}\left\|\sum_{ij}a_{ij}x_{ij}\right\|\right.$$
$$\left.+p\sup_{\|x\|_2\leq 1,\|y\|_2\leq 1}\left\|\sum_{ij}a_{ij}x_iy_j\right\|\right).$$

Proof. Let X_i be a sequence of i.i.d. standard symmetric exponential r.v's. A simple argument (cf. proof of Lemma 9.5 in [1]) shows that

$$\Big\|\sum_i a_{ii}(g_i^2-1)\Big\|_p \sim \Big\|\sum_i a_{ii}g_ig_i'\Big\|_p \sim \Big\|\sum_i a_{ii}X_i\Big\|_p, \tag{3.21}$$

the latter quantity can be bounded by Theorem 6.24, thus

$$\begin{aligned}\Big\|\sum_i a_{ii}(g_i^2-1)\Big\|_p &\sim \Big\|\sum_i a_{ii}(g_i^2-1)\Big\|_1 + \sqrt{p}\sup_{\|x\|_2\le 1}\Big\|\sum_i a_{ii}x_i\Big\| + p\sup_i\|a_{ii}\| \\ &\le C\left(\mathbb{E}\Big\|\sum_{ij} a_{ij}(g_ig_j-\delta_{ij})\Big\| + \sqrt{p}\sup_{\|(x_{ij})\|_2\le 1}\Big\|\sum_{ij} a_{ij}x_{ij}\Big\| + p\sup_{\|x\|_2\le 1,\|y\|_2\le 1}\Big\|\sum_{ij} a_{ij}x_iy_j\Big\|\right),\end{aligned}$$

where in the last inequality we used Lemma 3.3.1. □

The next proposition implies that in all our main results the term $\sup_{\|x\|_2\le 1}\mathbb{E}\|\sum_{i\neq j} a_{ij}x_ig_j\|$ can be replaced by $\sup_{\|x\|_2\le 1}\mathbb{E}\|\sum_{ij} a_{ij}x_ig_j\|$.

Proposition 3.4.2. *Under the assumption of Proposition 3.1.1 we have for $p\ge 1$,*

$$\begin{aligned}\sqrt{p}\sup_{\|x\|_2\le 1}\mathbb{E}\Big\|\sum_i a_{ii}x_ig_i\Big\| \le C\Bigg(&\mathbb{E}\Big\|\sum_{ij} a_{ij}(g_{ij}-\delta_{ij})\Big\| + \sqrt{p}\sup_{\|(x_{ij})\|_2\le 1}\Big\|\sum_{ij} a_{ij}x_{ij}\Big\| \\ &+ p\sup_{\|x\|_2\le 1,\|y\|_2\le 1}\Big\|\sum_{ij} a_{ij}x_iy_j\Big\|\Bigg)\end{aligned}$$

Proof. Applying (3.20) conditionally on $(g_i)_i$ and Jensen's inequality yield

$$\sqrt{p}\sup_{\|x\|_2\le 1}\mathbb{E}\Big\|\sum_i a_{ii}x_ig_i\Big\| \le \mathbb{E}\sqrt{p}\sup_{\|x\|_2\le 1}\Big\|\sum_i a_{ii}x_ig_i\Big\| \le C\Big\|\sum_i a_{ii}g_ig_i'\Big\|_p \le C\Big\|\sum_i a_{ii}(g_i^2-1)\Big\|_p,$$

where in the last line we used (3.21). The assertion follows by Lemma 3.4.1.

□

For the off-diagonal terms we use first the concentration approach.

Proposition 3.4.3. *For $p\ge 1$ we have*

$$\Big\|\sum_{i\neq j} a_{ij}g_ig_j\Big\|_p \le C\left(\mathbb{E}\Big\|\sum_{i\neq j} a_{ij}g_ig_j\Big\| + \sqrt{p}\mathbb{E}\sup_{\|x\|_2\le 1}\Big\|\sum_{i\neq j} a_{ij}x_ig_j\Big\| + p\sup_{\|x\|_2\le 1,\|y\|_2\le 1}\Big\|\sum_{ij} a_{ij}x_iy_j\Big\|\right).$$

Proof. Let

$$A := \left\{z\in\mathbb{R}^n:\ \Big\|\sum_{i\neq j} a_{ij}z_iz_j\Big\| \le 4\mathbb{E}\Big\|\sum_{i\neq j} a_{ij}g_ig_j\Big\|,\ \sup_{\|x\|_2\le 1}\Big\|\sum_{i\neq j} a_{ij}x_iz_j\Big\| \le 4\mathbb{E}\sup_{\|x\|_2\le 1}\Big\|\sum_{i\neq j} a_{ij}x_ig_j\Big\|\right\}.$$

Then $\gamma_n(A)\ge\frac{1}{2}$ by the Chebyshev inequality. Gaussian concentration gives $\gamma_n(A+tB_2^n)\ge 1-e^{-t^2/2}$ for $t\ge 0$. It is easy to check that for $z\in A+tB_2^n$ we have

$$\left\|\sum_{i\neq j} a_{ij} z_i z_j\right\| \leq 4S(t),$$

where

$$S(t) = \mathbb{E}\left\|\sum_{i\neq j} a_{ij} g_i g_j\right\| + 2t\mathbb{E}\sup_{\|x\|_2\leq 1}\left\|\sum_{ij} a_{ij} x_i g_j\right\| + t^2 \sup_{\|x\|_2\leq 1, \|y\|_2\leq 1}\left\|\sum_{i\neq j} a_{ij} x_i y_j\right\|.$$

So

$$\mathbb{P}\left(\left\|\sum_{i\neq j} a_{ij} g_i g_j\right\| > 4S(t)\right) \leq e^{-t^2/2} \quad \text{for } t \geq 0.$$

Integrating by parts we get $\|\sum_{i\neq j} a_{ij} g_i g_j\|_p \leq CS(\sqrt{p})$ for $p \geq 1$, which ends the proof. □

Observe that for any symmetric matrix the Kwapień decoupling Theorem 6.5 yields

$$\mathbb{E}\left\|\sum_{i\neq j} a_{ij} g_i g_j\right\| \sim \mathbb{E}\left\|\sum_{i\neq j} a_{ij} g_i g_j'\right\|.$$

Moreover introducing decoupled chaos enables us to release the assumptions of the symmetry of the matrix and zero diagonal.

Taking into account the above observations, Conjecture 3.1.2 reduces to the statement that for any $p \geq 1$ and any finite matrix (a_{ij}) in $(F, \|\cdot\|)$ we have

$$\mathbb{E}\sup_{\|x\|_2\leq 1}\left\|\sum_{ij} a_{ij} g_i x_j\right\| \leq C\Bigg(\frac{1}{\sqrt{p}}\mathbb{E}\left\|\sum_{ij} a_{ij} g_i g_j'\right\| + \sup_{\|x\|_2\leq 1}\mathbb{E}\left\|\sum_{ij} a_{ij} g_i x_j\right\| + \sup_{\|(x_{ij})\|_2\leq 1}\left\|\sum_{ij} a_{ij} x_{ij}\right\| + \sqrt{p}\sup_{\|x\|_2\leq 1, \|y\|_2\leq 1}\left\|\sum_{ij} a_{ij} x_i y_j\right\|\Bigg). \tag{3.22}$$

Let us rewrite (3.22) in another language. We may assume that $F = \mathbb{R}^m$ for some finite m and $a_{ij} = (a_{ijk})_{k\leq m}$. Let $T = B_{F^*}$ be the unit ball in the dual space F^*. Then (3.22) takes the following form.

Conjecture 3.4.4. *Let $p \geq 1$. Then for any triple indexed matrix $(a_{ijk})_{i,j\leq n, k\leq m}$ and bounded nonempty set $T \subset \mathbb{R}^m$ we have*

$$\mathbb{E}\sup_{\|x\|_2\leq 1, t\in T}\left|\sum_{ijk} a_{ijk} g_i x_j t_k\right| \leq C\Bigg(\frac{1}{\sqrt{p}}\mathbb{E}\sup_{t\in T}\left|\sum_{ijk} a_{ijk} g_i g_j' t_k\right| + \sup_{\|x\|_2\leq 1}\mathbb{E}\sup_{t\in T}\left|\sum_{ijk} a_{ijk} g_i x_j t_k\right| + \sup_{t\in T}\left(\sum_{ij}\left(\sum_k a_{ijk} t_k\right)^2\right)^{1/2} + \sqrt{p}\sup_{\|x\|_2\leq 1, t\in T}\left(\sum_i\left(\sum_{jk} a_{ij} x_j t_k\right)^2\right)^{1/2}\Bigg). \tag{3.23}$$

Obviously it is enough to show this for finite sets T.

3.5 Estimating suprema of Gaussian processes

To estimate the supremum of a centered Gaussian process $(G_v)_{v\in V}$ one needs to study the distance on V given by $d(v,v') := (\mathbb{E}|G_v - G_{v'}|^2)^{1/2}$ (cf. [30]). In the case of the Gaussian process from Conjecture 3.4.4 this distance is defined on $B_2^n \times T \subset \mathbb{R}^n \times \mathbb{R}^m$ by the formula

$$d_A((x,t),(x',t')) := \left(\sum_i \left(\sum_{jk} a_{ijk}(x_j t_k - x'_j t'_k)\right)^2\right)^{1/2} = \alpha_A(x\otimes t - x'\otimes t'),$$

where $x\otimes t = (x_j t_k)_{j,k} \in \mathbb{R}^{nm}$ and α_A is a norm on $\mathbb{R}^{nm}$ given by

$$\alpha_A(y) := \left(\sum_i \left(\sum_{jk} a_{ijk} y_{jk}\right)^2\right)^{1/2},$$

(as in Conjecture 3.4.4 in this section we do not assume that the matrix $(a_{ijk})_{ijk}$ is symmetric or that it has 0 on the generalized diagonal).

Let

$$B((x,t),d_A,r) = \left\{(x',t') \in \mathbb{R}^n \times T \colon\ \alpha_A(x\otimes t - x'\otimes t') \le r\right\}$$

be the closed ball in d_A with center at (x,t) and radius r.

Observe that

$$\operatorname{diam}(B_2^n \times T, d_A) \sim \sup_{\|x\|_2\le 1, t\in T} \left(\sum_i \left(\sum_{jk} a_{ijk} x_j t_k\right)^2\right)^{1/2}.$$

Now we try to estimate entropy numbers $N(B_2^n \times T, d_A, \varepsilon)$ for $\varepsilon > 0$ (recall that $N(S,\rho,\varepsilon)$ is the smallest number of closed balls with the diameter ε in metric ρ that cover set S). To this end we first introduce some notation. For a nonempty bounded set S in $\mathbb{R}^m$ let

$$\beta_{A,S}(x) := \mathbb{E}\sup_{t\in S}\left|\sum_{ijk} a_{ijk} g_i x_j t_k\right|, \quad x\in\mathbb{R}^n.$$

Observe that $\beta_{A,S}$ is a norm on $\mathbb{R}^n$. Moreover, by the classical Sudakov minoration (cf. Theorem 6.7) for any $x\in\mathbb{R}^n$ there exists a set $S_{x,\varepsilon} \subset S$ of cardinality at most $\exp(C\varepsilon^{-2})$ such that

$$\forall_{t\in S}\ \exists_{t'\in S_{x,\varepsilon}}\ \alpha_A(x\otimes(t-t')) \le \varepsilon\beta_{A,S}(x).$$

For a finite set $S\subset\mathbb{R}^m$ and $\varepsilon > 0$ define a measure $\mu_{\varepsilon,S}$ on $\mathbb{R}^n \times S$ in the following way

$$\mu_{\varepsilon,S}(C) := \int_{\mathbb{R}^n} \sum_{t\in S_{x,\varepsilon}} \delta_{(x,t)}(C)\, d\gamma_{n,\varepsilon}(x),$$

where $\gamma_{n,\varepsilon}$ is the distribution of the vector εG_n (recall that G_n is the standard Gaussian vector in $\mathbb{R}^n$). Since S is finite, we can choose sets $S_{x,\varepsilon}$ in such a way that there are no problems with measurability.

To bound $N(B_2^n \times T, d_A, \varepsilon)$ we need two lemmas.

Lemma 3.5.1. [16, Lemma 1] *For any norms α_1,α_2 on $\mathbb{R}^n$, $y\in B_2^n$ and $\varepsilon>0$,*

$$\gamma_{n,\varepsilon}(x\colon\ \alpha_1(x-y)\le 4\varepsilon\mathbb{E}\alpha_1(G_n),\ \alpha_2(x)\le 4\varepsilon\mathbb{E}\alpha_2(G_n)+\alpha_2(y))\ge \frac{1}{2}\exp(-\varepsilon^{-2}/2).$$

Lemma 3.5.2. *For any finite set S in $\mathbb{R}^m$, any $(x,t)\in B_2^n\times S$ and $\varepsilon>0$ we have*

$$\mu_{\varepsilon,S}\left(B\left((x,t),d_A,r(\varepsilon)\right)\right)\ge \frac{1}{2}\exp(-\varepsilon^{-2}/2),$$

where

$$r(\varepsilon)=r(A,S,x,t,\varepsilon)=4\varepsilon^2\mathbb{E}\beta_{A,S}(G_n)+\varepsilon\beta_{A,S}(x)+4\varepsilon\mathbb{E}\alpha_A(G_n\otimes t).$$

Proof. Let

$$U=\left\{x'\in\mathbb{R}^n\colon\ \beta_{A,S}(x')\le 4\varepsilon\mathbb{E}\beta_{A,S}(G_n)+\beta_{A,S}(x),\alpha_A((x-x')\otimes t)\le 4\varepsilon\mathbb{E}\alpha_A(G_n\otimes t)\right\}.$$

For any $x'\in U$ there exists $t'\in S_{x',\varepsilon}$ such that $\alpha_A(x'\otimes(t-t'))\le\varepsilon\beta_{A,S}(x')$. By the triangle inequality

$$\alpha_A(x\otimes t-x'\otimes t')\le\alpha_A((x-x')\otimes t)+\alpha_A(x'\otimes(t-t'))\le r(\varepsilon).$$

Thus, by Lemma 3.5.1, $\mu_{\varepsilon,S}\left(B\left((x,t),d_A,r(\varepsilon)\right)\right)\ge\gamma_{n,\varepsilon}(U)\ge\frac{1}{2}\exp(-\varepsilon^{-2}/2)$. □

Having Lemma 3.5.2 we can estimate the entropy numbers by a version of the usual volumetric argument.

Corollary 3.5.3. *For any $\varepsilon>0$, $U\subset B_2^n$ and $S\subset\mathbb{R}^m$,*

$$N\left(U\times S,d_A,8\varepsilon^2\mathbb{E}\beta_{A,S}(G_n)+2\varepsilon\sup_{x\in U}\beta_{A,S}(x)+8\varepsilon\sup_{t\in S}\mathbb{E}\alpha_A(G_n\otimes t)\right)\le\exp(C\varepsilon^{-2})\qquad(3.24)$$

and for any $\delta>0$,

$$\sqrt{\ln N(U\times S,d_A,\delta)}\le C\left(\delta^{-1}\left(\sup_{x\in U}\beta_{A,S}(x)+\sup_{t\in S}\mathbb{E}\alpha_A(G_n\otimes t)\right)+\delta^{-1/2}(\mathbb{E}\beta_{A,S}(G_n))^{1/2}\right).$$

Proof. Let $r=4\varepsilon^2\mathbb{E}\beta_{A,S}(G_n)+\varepsilon\sup_{x\in U}\beta_{A,S}(x)+4\varepsilon\sup_{t\in S}\mathbb{E}\alpha_A(G_n\otimes t)$ and $N=N(U\times S,d_A,2r)$. Then there exist points $(x_i,t_i)_{i=1}^N$ in $U\times S$ such that $d_A((x_i,t_i),(x_j,t_j))>2r$. To show (3.24) we consider two cases.

If $\varepsilon>2$ then

$$\begin{aligned}2r&\ge 4\sup_{x\in U}\beta_{A,S}(x)\ge 4\sup_{(x,t)\in U\times S}\mathbb{E}\left|\sum_{ijk}a_{ijk}g_it_jx_k\right|\\&=4\sqrt{\frac{2}{\pi}}\sup_{(x,t)\in U\times S}\left(\sum_i\left(\sum_{jk}a_{ijk}t_jx_k\right)^2\right)^{1/2}\ge\operatorname{diam}(U\times S,d_A)\end{aligned}$$

so $N=1\le\exp(C\varepsilon^{-2})$.

If $\varepsilon<2$, note that the balls $B((x_i,t_i),d_A,r)$ are disjoint and, by Lemma 3.5.2, each of these balls has $\mu_{\varepsilon,S}$ measure at least $\frac{1}{2}\exp(-\varepsilon^{-2}/2)\ge\exp(-5\varepsilon^{-2})$. On the other hand we obviously have $\mu_{\varepsilon,S}(\mathbb{R}^n\times S)\le\exp(C\varepsilon^{-2})$. Comparing the upper and lower bounds on $\mu_{\varepsilon,S}(\mathbb{R}^n\times S)$ gives (3.24) in this case.

The second estimate from the assertion is an obvious consequence of the first one. □

Remark 3.5.4. The classical Dudley's bound on suprema of Gaussian processes (see Corollary 6.14) gives

$$\mathbb{E}\sup_{\|x\|_2\le 1, t\in T}\left|\sum_{ijk} a_{ijk}g_i x_j t_k\right| \le C\int_0^{\mathrm{diam}(B_2^n\times T, d_A)} \sqrt{\ln N(B_2^n\times T, d_A,\delta)}d\delta.$$

Observe that

$$\int_0^{\mathrm{diam}(B_2^n\times T,d_A)} \delta^{-1/2}(\mathbb{E}\beta_{A,T}(G_n))^{1/2}d\delta = 2\sqrt{\mathrm{diam}(B_2^n\times T,d_A)\mathbb{E}\beta_{A,T}(G_n)}$$
$$\le \frac{1}{\sqrt{p}}\mathbb{E}\beta_{A,T}(G_n)+\sqrt{p}\mathrm{diam}(B_2^n\times T,d_A)$$

appears on the right hand side of (3.23). Unfortunately the other term in the estimate of $\sqrt{\ln N(B_2^n\times T,d_A,\delta)}$ is not integrable. The remaining part of the proof is devoted to improve on Dudley's bound.

We will now continue along the lines of [16]. We will need in particular to partition the set T into smaller pieces T_i such that $\sup_{t,s\in T_i}\mathbb{E}\alpha_A(G_n\otimes(t-t'))$ is small on each piece. To this end we apply the following Sudakov-type estimate for chaoses, derived by Talagrand ([31] or [30, Section 8.2]).

Theorem 3.5.5. *Let $\mathcal{A}$ be a subset of n by n real valued matrices and d_2, d_∞ be distances associated to the Hilbert-Schmidt and operator norms respectively. Then*

$$\varepsilon \ln^{1/4} N(\mathcal{A},d_2,\varepsilon)\le C\mathbb{E}\sup_{a\in\mathcal{A}}\sum_{ij}a_{ij}g_ig_j' \quad \textit{for } \varepsilon>0$$

and

$$\varepsilon \ln^{1/2} N(\mathcal{A},d_2,\varepsilon)\le C\mathbb{E}\sup_{a\in\mathcal{A}}\sum_{ij}a_{ij}g_ig_j' \quad \textit{for } \varepsilon> C\sqrt{\mathrm{diam}(\mathcal{A},d_\infty)\mathbb{E}\sup_{a\in\mathcal{A}}\sum_{ij}a_{ij}g_ig_j'}.$$

To make the notation more compact let for $T\subset\mathbb{R}^m$ and $V\subset\mathbb{R}^n\times\mathbb{R}^m$,

$$\begin{aligned}
s_A(T) &:= \mathbb{E}\beta_{A,T}(G_n)=\mathbb{E}\sup_{t\in T}\left|\sum_{ijk}a_{ijk}g_ig_j't_k\right|,\\
F_A(V) &:= \mathbb{E}\sup_{(x,t)\in V}\sum_{ijk}a_{ijk}g_ix_jt_k\\
\varDelta_{A,\infty}(T) &:= \sup_{\|x\|_2\le1,\|y\|_2\le1,t,t'\in T}\left|\sum_{ijk}a_{ijk}x_iy_j(t_k-t_k')\right|,\\
\varDelta_A(V) &:= \mathrm{diam}(V,d_A)=\sup_{(x,t),(x',t')\in V}\alpha_A(x\otimes t-x'\otimes t').
\end{aligned}$$

Corollary 3.5.6. *Let T be a subset of $\mathbb{R}^m$. Then for any $r>0$ there exists a decomposition $T-T=\bigcup_{i=1}^N T_i$ such that, $N\le e^{Cr}$ and*

$$\sup_{t,t'\in T_i}\mathbb{E}\alpha_A(G_n\otimes(t-t'))\le\min\left\{r^{-1/4}s_A(T), r^{-1/2}s_A(T)+C\sqrt{s_A(T)\varDelta_{A,\infty}(T)}\right\}.$$

Proof. We use Theorem 3.5.5 with $\mathcal{A}=\{(\sum_k a_{ijk}t_k)_{ij}:\ t\in T-T\}$. It is enough to observe that

$$\mathbb{E}\sup_{b\in\mathcal{A}}\left|\sum_{ij} b_{ij}g_ig_j'\right| = s_A(T-T)\le 2s_A(T),\quad \operatorname{diam}(\mathcal{A},d_\infty)=2\Delta_{A,\infty}(T)$$

and

$$\mathbb{E}\alpha_A(G_n\otimes t-t')\le \left\|\left(\sum_k a_{ijk}(t_k-t_k')\right)_{ij}\right\|_{HS}.$$

□

On the other hand the dual Sudakov minoration (cf. Theorem 6.8) yields the following

Corollary 3.5.7. *Let U be a subset of B_2^n. Then for any $r>0$ there exists a decomposition $U=\bigcup_{i=1}^N U_i$ such that $N\le e^{Cr}$ and*

$$\sup_{x,x'\in U_i}\beta_{A,T}(x-x')\le r^{-1/2}s_A(T).$$

Putting the above two corollaries together with Corollary 3.5.3 we get the following decomposition of subsets $B_2^n\times T$.

Corollary 3.5.8. *Let $V\subset\mathbb{R}^n\times\mathbb{R}^m$ be such that $V-V\subset B_2^n\times(T-T)$. Then for $r\ge 1$ we may find a decomposition $V=\bigcup_{i=1}^N((x_i,t_i)+V_i)$ such that $N\le e^{Cr}$ and for each $1\le i\le N$,*
i) $(x_i,t_i)\in V$, $V_i-V_i\subset V-V$, $V_i\subset B_2^n\times(T-T)$,
ii) $\sup_{(x,t)\in V_i}\beta_{A,T}(x)\le r^{-1/2}s_A(T)$,
iii) $\sup_{(x,t)\in V_i}\mathbb{E}\alpha_A(G_n\otimes t)\le\min\left\{r^{-1/4}s_A(T),r^{-1/2}s_A(T)+C\sqrt{s_A(T)\Delta_{A,\infty}(T)}\right\}$,
iv) $\Delta_A(V_i)\le\min\left\{r^{-3/4}s_A(T),r^{-1}s_A(T)+r^{-1/2}\sqrt{s_A(T)\Delta_{A,\infty}(T)}\right\}$.

Proof. The assertion is invariant under translations of the set V thus we may assume that $(0,0)\in V$ and so $V\subset V-V\subset B_2^n\times(T-T)$. By Corollaries 3.5.6 and 3.5.7 we may decompose $B_2^n=\bigcup_{i=1}^{N_1}U_i$, $T-T=\bigcup_{i=1}^{N_2}T_i$ in such a way that $N_1,N_2\le e^{Cr}$ and

$$\sup_{x,x'\in U_i}\beta_{A,T}(x-x')\le r^{-1/2}s_A(T),$$
$$\sup_{t,t'\in T_i}\mathbb{E}\alpha_A(G_n\otimes(t-t'))\le\min\left\{r^{-1/4}s_A(T),r^{-1/2}s_A(T)+C\sqrt{s_A(T)\Delta_{A,\infty}(T)}\right\}.$$

Let $V_{ij}:=V\cap(U_i\times T_j)$. If $V_{ij}\ne\emptyset$ we take any point $(x_{ij},y_{ij})\in V_{ij}$ and using Corollary 3.5.3 with $\varepsilon=r^{-1/2}/C$ we decompose

$$V_{ij}-(x_{ij},y_{ij})=\bigcup_{k=1}^{N_3}V_{ijk}$$

in such a way that $N_3\le e^{Cr}$ and

$$\begin{aligned}\Delta_A(V_{ijk})&\le\frac{1}{C}\left(r^{-1}s_A(T)+r^{-1/2}\sup_{x'\in U_i}\beta_{A,T}(x'-x_{ij})+r^{-1/2}\sup_{y'\in T_j}\mathbb{E}\alpha_A(G_n\otimes(y'-y_{ij}))\right)\\&\le\min\left\{r^{-3/4}s_A(T),r^{-1}s_A(T)+r^{-1/2}\sqrt{s_A(T)\Delta_{A,\infty}(T)}\right\}.\end{aligned}$$

The final decomposition is obtained by relabeling of the decomposition $V=\bigcup_{ijk}((x_{ij},y_{ij})+V_{ijk})$.
□

Remark 3.5.9. We may also use a trivial bound in iii):

$$\sup_{(x,t)\in V_i} \mathbb{E}\alpha_A(G_n\otimes t) \le \sup_{t,t'\in T} \mathbb{E}\alpha_A(G_n\otimes(t-t')) \le 2\sup_{t\in T}\mathbb{E}\alpha_A(G_n\otimes t),$$

this will lead to the following bound in iv):

$$\Delta_A(V_i) \le r^{-1}s_A(T) + r^{-1/2}\sup_{t\in T}\mathbb{E}\alpha_A(G_n\otimes t).$$

Remark 3.5.10. By using Sudakov minoration (cf. Theorem 6.7) instead of Theorem 3.5.5 we may decompose the set $T = \bigcup_{i=1}^N T_i$, $N \le \exp(Cr)$ in such a way that

$$\forall_{i\le N} \sup_{t,t'\in T_i} \mathbb{E}\alpha_A(G_n\otimes(t-t')) \le r^{-1/2}\mathbb{E}\sup_{t\in T}\sum_{ijk} a_{ijk}g_{ij}t_k.$$

This will lead to the following bounds in iii) and iv):

$$\sup_{(x,t)\in V_i} \mathbb{E}\alpha_A(G_n\otimes t) \le r^{-1/2}\mathbb{E}\sup_{t\in T}\sum_{ijk} a_{ijk}g_{ij}t_k$$

$$\Delta_A(V_i) \le r^{-1}\left(\mathbb{E}\sup_{t\in T}\sum_{ijk} a_{ijk}g_{ij}t_k + s_A(T)\right).$$

Lemma 3.5.11. *Let V be a subset of $B_2^n\times(T-T)$. Then for any $(y,s)\in\mathbb{R}^n\times\mathbb{R}^m$ we have*

$$F_A(V+(y,s)) \le F_A(V) + 2\beta_{A,T}(y) + C\mathbb{E}\alpha_A(G_n\otimes s).$$

Proof. We have

$$F_A(V+(y,s)) \le F_A(V) + \mathbb{E}\sup_{(x,t)\in V}\sum_{ijk} a_{ijk}g_iy_jt_k + \mathbb{E}\sup_{(x,t)\in V}\sum_{ijk} a_{ijk}g_ix_js_k.$$

Obviously,

$$\mathbb{E}\sup_{(x,t)\in V}\sum_{ijk} a_{ijk}g_iy_jt_k \le \mathbb{E}\sup_{t,t'\in T}\left|\sum_{ijk} a_{ijk}g_iy_j(t_k-t'_k)\right| \le 2\beta_{A,T}(y).$$

Moreover,

$$\mathbb{E}\sup_{(x,t)\in V}\sum_{ijk} a_{ijk}g_ix_js_k \le \left(\mathbb{E}\sup_{x\in B_2^n}\left|\sum_{ijk} a_{ijk}g_ix_js_k\right|^2\right)^{1/2} = \left(\sum_{ij}\left(\sum_k a_{ijk}s_k\right)^2\right)^{1/2}$$
$$= (\mathbb{E}\alpha_A(G_n\otimes s)^2)^{1/2} \le C\mathbb{E}\alpha_A(G_n\otimes s),$$

where in the second inequality we used Theorem 6.3. □

Proposition 3.5.12. *For any nonempty finite set T in $\mathbb{R}^m$ and $p\ge 1$ we have*

$$F_A(B_2^n \times T) \le C\Bigg(\frac{\ln(ep)}{\sqrt{p}} s_A(T) + \sup_{\|x\|_2\le 1} \beta_{A,T}(x) + \sup_{t\in T} \mathbb{E}\alpha_A(G_n\otimes t) + \ln(ep)\sqrt{p}\Delta_A(B_2^n\times T)\Bigg), \quad (3.25)$$

$$F_A(B_2^n \times T) \le C\Bigg(\frac{1}{\sqrt{p}} s_A(T) + \sup_{\|x\|_2\le 1} \beta_{A,T}(x) + \ln(ep)\sup_{t\in T} \mathbb{E}\alpha_A(G_n\otimes t) + \sqrt{p}\Delta_A(B_2^n\times T)\Bigg), \quad (3.26)$$

$$F_A(B_2^n \times T) \le C\Bigg(\frac{1}{\sqrt{p}} s_A(T) + \frac{1}{\sqrt{p}}\mathbb{E}\sup_{t\in T}\sum a_{ijk}g_{ij}t_k + \sup_{\|x\|_2\le 1} \beta_{A,T}(x) + \sup_{t\in T} \mathbb{E}\alpha_A(G_n\otimes t) + \sqrt{p}\Delta_A(B_2^n\times T)\Bigg). \quad (3.27)$$

Proof. First we prove (3.25) Let $l_0 \in \mathbb{N}$ be such that $2^{l_0-1} \le p < 2^{l_0}$. Define

$$\Delta_0 := \Delta_A(B_2^n\times T), \quad \tilde{\Delta}_0 := \sup_{x\in B_2^n} \beta_{A,T}(x) + \sup_{t\in T}\mathbb{E}\alpha_A(G_n\otimes t),$$

$$\Delta_l = 2^{-3l/4}p^{-3/4}s_A(T), \quad \tilde{\Delta}_l = 2^{-l/4}p^{-1/4}s_A(T), \quad l > l_0.$$

and for $1 \le l \le l_0$,

$$\Delta_l := 2^{-l}p^{-1}s_A(T) + 2^{-l/2}p^{-1/2}\sqrt{s_A(T)\Delta_{A,\infty}(T)},$$
$$\tilde{\Delta}_l := 2^{-l/2}p^{-1/2}s_A(T) + C\sqrt{s_A(T)\Delta_{A,\infty}(T)}.$$

Let for $l = 0,1,\dots$ and $m = 1,2,\dots$

$$c(l,m) := \sup\big\{F_A(V):\ V - V \subset B_2^n \times (T-T), \#V \le m, \Delta_A(V) \le \Delta_l, \sup_{(x,t)\in V}(\beta_{A,T}(x) + \mathbb{E}\alpha_A(G_n\otimes t)) \le 2\tilde{\Delta}_l\big\}.$$

Obviously $c(l,1) = 0$. We will show that for $m > 1$ and $l \ge 0$ we have

$$c(l,m) \le c(l+1,m-1) + C\left(2^{l/2}\sqrt{p}\Delta_l + \tilde{\Delta}_l\right). \quad (3.28)$$

To this end take any set V as in the definition of $c(l,m)$ and apply to it Corollary 3.5.8 with $r = 2^{l+1}p$ to obtain decomposition $V = \bigcup_{i=1}^N((x_i,t_i)+V_i)$. We may obviously assume that all V_i have smaller cardinality than V. Conditions i)-iv) from Corollary 3.5.8 easily imply that $F_A(V_i) \le c(l+1,m-1)$.

Lemma 6.10 yields

$$F_A(V) = F_A\left(\bigcup_i((x_i,t_i)+V_i)\right) \le C\sqrt{\ln N}\Delta_A(V) + \max_i F_A((x_i,t_i)+V_i).$$

Estimate (3.28) follows since

$$\sqrt{\ln N}\varDelta_A(V) \le C2^{l/2}\sqrt{p}\varDelta_l$$

and for each i by Lemma 3.5.11 we have (recall that $(x_i,t_i)\in V$)

$$F_A((x_i,t_i)+V_i) \le F_A(V)+2\beta_{A,T}(x_i)+C\mathbb{E}\alpha_A(G_n\otimes t_i) \le c(l+1,m-1)+C\tilde{\varDelta}_l.$$

Hence

$$\begin{aligned} c(0,m) &\le C\left(\sum_{l=0}^{\infty}2^{l/2}\sqrt{p}\varDelta_l+\sum_{l=0}^{\infty}\tilde{\varDelta}_l\right)\\ &\le C\left(\sqrt{p}\varDelta_0+\tilde{\varDelta}_0+\frac{1}{\sqrt{p}}s_A(T)+l_0\sqrt{s_A(T)\varDelta_{A,\infty}(T)}+2^{-l_0/4}p^{-1/4}s_A(T)\right). \end{aligned}$$

Since $\ln_2 p < l_0 \le \ln_2 p+1$ and $\sqrt{s_A(T)\varDelta_{A,\infty}(T)} \le \frac{1}{\sqrt{p}}s_A(T)+\sqrt{p}\varDelta_{A,\infty}(T)$ and clearly $\varDelta_{A,\infty}(T) \le \varDelta_A(B_2^n\times T)$ we get for all $m\ge 1$,

$$c(0,m) \le C\left(\frac{\ln(ep)}{\sqrt{p}}s_A(T)+\sup_{\|x\|_2\le 1}\beta_{A,T}(x)+\sup_{t\in T}\mathbb{E}\alpha_A(G_n\otimes t)+\ln(ep)\sqrt{p}\varDelta_A(B_2^n\times T)\right).$$

To conclude the proof of (3.25) it is enough to observe that

$$F_A(B_2^n\times T)=2F_A\left(\frac{1}{2}B_2^n\times T\right) \le 2\sup_{m\ge 1}c(0,m).$$

The proofs of (3.26) and (3.27) are the same as the proof of (3.25). The only difference is that for $1\le l\le l_0$ we change the definitions of $\varDelta_l$, $\tilde{\varDelta}_l$ and we use Remarks 3.5.9 and 3.5.10 respectively. In the first case we take

$$\begin{aligned} \varDelta_l &:= 2^{-l}p^{-1}s_A(T)+2^{-l/2}p^{-1/2}\sup_{t\in T}\mathbb{E}\alpha_A(G_n\otimes t)\\ \tilde{\varDelta}_l &:= 2^{-l/2}p^{-1/2}s_A(T)+\sup_{t\in T}\mathbb{E}\alpha_A(G_n\otimes t), \end{aligned}$$

while in the second

$$\begin{aligned} \varDelta_l &:= 2^{-l}p^{-1}\left(s_A(T)+\mathbb{E}\sup_{t\in T}\sum_{ijk}a_{ijk}g_{ij}t_k\right)\\ \tilde{\varDelta}_l &= 2^{-l/2}p^{-1/2}\left(s_A(T)+\mathbb{E}\sup_{t\in T}\sum_{ijk}a_{ijk}g_{ij}t_k\right). \end{aligned}$$

□

Remark 3.5.13. Note that the proof of Proposition 3.5.12 gives in fact the following estimate for any set $U\subset B_2^n\times T$. For any $p\ge 1$,

$$F_A(U) \le C\left(\frac{1}{\sqrt{p}}s_A(T)+\frac{1}{\sqrt{p}}\mathbb{E}\sup_{t\in T}\sum a_{ijk}g_{ij}t_k+\sup_{\|x\|_2\le 1}\beta_{A,T}(x)+\sup_{t\in T}\mathbb{E}\alpha_A(G_n\otimes t)+\sqrt{p}\varDelta_A(U)\right).$$

Indeed, as in the proof above one can reduce the problem to the case of $U\subset\frac{1}{2}B_2^n\times T$ and then it is enough to set $\varDelta_0=\varDelta_A(U)$ and add the condition $V\subset U$ to the definition of $c(l,m)$.

3.6 Proofs of main results

Proof of Theorems 3.1.3 and 3.1.4. By Lemmas 3.3.1, 3.4.1 and Proposition 3.4.3 we need only to establish (3.4)-(3.6) with $\|\sum_{ij} a_{ij}(g_i g_j - \delta_{ij})\|_p$ replaced by $\sqrt{p}\mathbb{E}\sup_{\|x\|_2\le 1}\|\sum_{i\neq j} a_{ij} g_i x_j\|$. We may assume that $F=\mathbb{R}^m$ and $a_{ii}=0$, so taking for T the unit ball in the dual space F^* we have

$$\Big\|\sum_{i\neq j} a_{ij} g_i x_j\Big\| = \sup_{t\in T}\sum_{ijk} a_{ijk} g_i x_j t_k.$$

Then, using the notation introduced in Section 3.5,

$$\mathbb{E}\sup_{\|x\|_2\le 1}\Big\|\sum_{i\neq j} a_{ij} g_i x_j\Big\| = F_A(B_2^n\times T),\quad \mathbb{E}\Big\|\sum_{ij} a_{ij} g_i x_j\Big\| = \beta_{A,T}(x),$$

$$\sup_{\|(x_{ij})\|_2\le 1}\Big\|\sum_{ij} a_{ij} x_{ij}\Big\| = \sup_{t\in T}(\mathbb{E}\alpha_A^2(G_n\otimes t))^{1/2}\sim \sup_{t\in T}\mathbb{E}\alpha_A(G_n\otimes t),$$

$$\sup_{\|x\|_2\le 1,\|y\|_2\le 1}\Big\|\sum_{ij} a_{ij} x_i y_j\Big\|\sim \varDelta_A(B_2^n\times T),$$

$$\mathbb{E}\Big\|\sum_{i\neq j} a_{ij} g_{ij}\Big\| = \mathbb{E}\sup_{t\in T}\sum_{ijk} a_{ijk} g_{ij} t_k \quad\text{and}\quad \mathbb{E}\Big\|\sum_{i\neq j} a_{ij} g_i g_j\Big\|\sim s_A(T),$$

where the last estimate follows by decoupling. We conclude the proof invoking Proposition 3.5.12. □

Proof of Theorem 3.1.6. Let $S=\|\sum_{ij} a_{ij}(g_i g_j-\delta_{ij})\|$. By the Paley-Zygmund inequality (see Corollary 6.1) and comparison of moments of Gaussian quadratic forms (see Theorem 6.3) we have for $p\ge 1$,

$$\mathbb{P}\left(S\ge \frac{1}{2}(\mathbb{E}S^p)^{1/p}\right) = \mathbb{P}\left(S^p\ge \frac{1}{2^p}\mathbb{E}S^p\right)\ge \left(1-\frac{1}{2^p}\right)^2\frac{(\mathbb{E}S^p)^2}{\mathbb{E}S^{2p}}\ge C_1^{-2p}.$$

So, to prove the lower bound on tails of S it is enough to use Proposition 3.1.1 and substitute $p=1+C\min\{t^2/U^2,t/V\}$.

To derive the upper bound we use Theorem 3.1.4, estimate $\mathbb{P}(S\ge e\|S\|_p)\le e^{-p}$ for $p\ge 1$ and make an analogous substitution. □

Proof of Theorem 3.1.7. Recall that for $r>0$ the ψ_r-norm of a random variable Y is defined as

$$\|Y\|_{\psi_r} = \inf\Big\{a>0\colon \mathbb{E}\exp\Big(\Big(\frac{|Y|}{a}\Big)^r\Big)\le 2\Big\}. \tag{3.29}$$

(formally for $r<1$ this is a quasi-norm, but it is customary to use the name ψ_r-norm for all r). By [2, Lemma 5.4] if k is a positive integer and $Y_1,\ldots,Y_n$ are symmetric random variables such that $\|Y\|_{\psi_{2/k}}\le M$, then

$$\left\|\sum_{i=1}^{n} a_i Y_i\right\|_p \le C_k M \left\|\sum_{i=1}^{n} a_i g_{i1}\cdots g_{ik}\right\|_p, \tag{3.30}$$

where g_{ik} are i.i.d. standard Gaussian variables (we remark that the lemma in [2] is stated only for $F=\mathbb{R}$ but its proof, based on contraction principle, works in any normed space).

To prove the theorem we will again establish a moment bound and then combine it with Chebyshev's inequality. Similarly as in the Gaussian setting we will treat the diagonal and off-diagonal part separately. Let $\varepsilon_1,\ldots,\varepsilon_n$ be a sequence of i.i.d. Rademacher variables independent of X_i's. For $p \ge 1$ we have

$$\left\|\sum_i a_{ii}(X_i^2 - \mathbb{E}X_i^2)\right\|_p \le 2\left\|\sum_i a_{ii}\varepsilon_i X_i^2\right\|_p \le C\alpha^2 \left\|\sum_i a_{ii} g_i g_i'\right\|_p,$$

where in the first inequality we used symmetrization and in the second one (3.30) together with the observation $\|\varepsilon_i X_i^2\|_{\psi_1} \le C\alpha^2$ (which can be easily proved by integration by parts).

Now by Lemma 6.13 (see also proof of [1, Lemma 9.5]),

$$\left\|\sum_i a_{ii} g_i g_i'\right\|_p \le C\left\|\sum_i a_{ii}\varepsilon_i g_i^2\right\|_p \le 2C\left\|\sum_i a_{ii}(g_i^2-1)\right\|_p,$$

and thus

$$\left\|\sum_i a_{ii}(X_i^2 - \mathbb{E}X_i^2)\right\|_p \le C\alpha^2 \left\|\sum_i a_{ii}(g_i^2-1)\right\|_p. \tag{3.31}$$

The estimate of the off-diagonal part is analogous, the only additional ingredient is decoupling. Denoting $(X_i')_{i=1}^n$ an independent copy of the sequence $(X_i)_{i=1}^n$ and by $(\varepsilon_i)_{i=1}^n$, $(\varepsilon_i')_{i=1}^n$ (resp. $(g_i)_{i=1}^n$, $(g_i')_{i=1}^n$) independent sequences of Rademacher (resp. standard Gaussian) random variables, we have

$$\left\|\sum_{i\neq j} a_{ij} X_i X_j\right\|_p \sim \left\|\sum_{i\neq j} a_{ij} X_i X_j'\right\|_p \sim \left\|\sum_{i\neq j} a_{ij}\varepsilon_i X_i \varepsilon_i' X_j'\right\|_p \le C\alpha^2 \left\|\sum_{i\neq j} a_{ij} g_i g_j'\right\|_p \sim \alpha^2 \left\|\sum_{i\neq j} a_{ij} g_i g_j\right\|_p, \tag{3.32}$$

where in the first and last inequality we used decoupling, the second one follows from iterated conditional application of symmetrization inequalities and the third one from iterated conditional application of (3.30) (note that by integration by parts we have $\|\varepsilon_i X_i\|_{\psi_2} \le C\alpha$).

Combining inequalities (3.31) and (3.32) with Lemma 3.3.1 we obtain

$$\left\|\sum_{ij} a_{ij}(X_i X_j - \mathbb{E}X_i X_j)\right\|_p \le C\alpha^2 \left\|\sum_{ij} a_{ij}(g_i g_j - \delta_{ij})\right\|_p.$$

To finish the proof of the theorem it is now enough to invoke moment estimates of Theorem 3.1.4 and use Chebyshev's inequality in L_p. □

Chapter 4
Moments of Gaussian chaoses in Banach spaces

4.1 Introduction

In this chapter we study Gaussian chaoses of order $d \in \mathbb{N}$, i.e. random variables of the form

$$S = \sum_{1 \le i_1,\dots,i_d \le n} a_{i_1,\dots,i_d} g_{i_1} \cdots g_{i_d},$$

where $(a_{i_1,\dots,i_d})_{i_1,\dots,i_d}$ is a multi-indexed symmetric array with values in a Banach space $(F, \|\cdot\|)$.

We establish upper bounds for moments, which for some classes of Banach spaces, including L_q spaces, can be reversed (up to constants depending only on d and the Banach space, but not on n or $a_{i_1,\dots,i_d}$). As a corollary we are able to deduce moment and tail bounds for homogeneous polynomials in i.i.d. symmetric exponential random variables and arbitrary polynomials in Gaussian random variables.

In the sequel we will consider mainly decoupled chaoses $S' = \sum_{i_1,\dots,i_d} a_{i_1,\dots,i_d} g^1_{i_1} \cdots g^d_{i_d}$, where $(g^k_i)_{i,k\in\mathbb{N}}$ are independent $\mathcal{N}(0,1)$ variables - under some natural assumptions, moments and tails of S, S' are comparable with constants depending only on d (cf. Theorems 6.5, 6.6).

For $d = 1$ and any $p \ge 1$ the Gaussian concentration yields easily

$$\Big\|\sum_i a_i g_i\Big\|_p = \Big(\mathbb{E}\Big\|\sum_i a_i g_i\Big\|^p\Big)^{1/p} \sim \mathbb{E}\Big\|\sum_i a_i g_i\Big\| + \sup_{\varphi\in F^*, \|\varphi\|\le 1} \Big\|\sum_i \varphi(a_i) g_i\Big\|_p$$
$$\sim \mathbb{E}\Big\|\sum_i a_i g_i\Big\| + \sqrt{p} \sup_{x\in B_2^n} \Big\|\sum_i a_i x_i\Big\|,$$

where F^* is the dual space and $\sim$ stands for a comparison up to universal multiplicative constants.

We recall that, for chaoses of order 2 it can be shown that

$$\Big\|\sum_{i,j} a_{ij} g_i g'_j\Big\|_p \sim \mathbb{E}\Big\|\sum_{i,j} a_{ij} g_i g'_j\Big\| + \sqrt{p}\,\mathbb{E} \sup_{x\in B_2^n} \Big(\Big\|\sum_{i,j} a_{ij} g_i x_j\Big\| + \Big\|\sum_{i,j} a_{ij} x_i g_j\Big\|\Big)$$
$$+ p \sup_{x,y\in B_2^n} \Big\|\sum_{i,j} a_{ij} x_i y_j\Big\|. \qquad (4.1)$$

If $d > 2$ then Gaussian concentration yields

$$\left\|\sum_{i_1,\ldots,i_d} a_{i_1,\ldots,i_d} g^1_{i_1}\cdots g^d_{i_d}\right\|_p \sim^d \sum_{J\subset[d]} p^{d/2}\mathbb{E}\sup\left\|\sum_{i_1,\ldots,i_d} a_{i_1,\ldots,i_d}\prod_{j\in J} x^j_{i_j}\prod_{j\in[d]\setminus J} g^j_{i_j}\right\|, \tag{4.2}$$

where the supremum is taken over $x^1,\ldots,x^n$ from the Euclidean unit sphere and $\sim_a$ stands for comparison up to constants depending only on the parameter a. The above formula (so in particular (4.1)) gives the precise dependence on p, but unfortunately involves suprema of certain stochastic processes, which are hard to estimate. Assume that $d=2$ and we are interested in the upper bound only. Then the problematic term $\mathbb{E}\sup_{x\in B_2^n}\left(\left\|\sum_{i<j} a_{ij}g_ix_j\right\|+\left\|\sum_{i<j} a_{ij}x_ig_j\right\|\right)$, can be replaced by quantities that can be handled more easily (cf. Chapter 3). Namely for any $p\geq 1$ we have (cf. Theorems 3.1.4, 6.5, 6.6)

$$\begin{aligned} C^{-1}\left\|\sum_{i,j} a_{i,j}g_ig'_j\right\|_p \leq\ & \mathbb{E}\left\|\sum_{i,j} a_{i,j}g_ig'_j\right\|+\mathbb{E}\left\|\sum_{i,j} a_{i,j}g_{ij}\right\| \\ & + p^{1/2}\left(\sup_{x\in B_2^n}\mathbb{E}\left\|\sum_{i,j} a_{ij}g_ix_j\right\|+\sup_{x\in B_2^{n^2}}\left\|\sum_{i,j} a_{ij}x_{ij}\right\|\right) \\ & + p\sup_{x,y\in B_2^n}\left\|\sum_{i,j} a_{ij}x_iy_j\right\|. \end{aligned}$$

As we mentioned in the previous chapter this upper bound turns out to be two-sided in a certain class of Banach spaces containing L_q spaces. This motivates the question of obtaining similar results for arbitrary d. We give an answer to it in Theorem 4.2.1. We rely on the techniques developed in [16]. In particular the heart of the proof is estimation of the expectation of the supremum of a certain Gaussian process.

The chapter is organized as follows. In the next section we set up the notation and formulate the main results including the pivotal upper bound (4.5) for moments of Gaussian chaoses in arbitrary Banach space. In Section 4.3 we reformulate (4.5) in an equivalent way and derive the entropy bounds. In Section 4.4 we use these entropy bounds to estimate expectation of supremum of a certain Gaussian process. In Section 4.5 we prove the bound (4.5) and then we deduce the remaining claims from it.

4.2 Notation and main results

We write $[n]$ for the set $\{1,\ldots,n\}$. Throughout the chapter C will denote an absolute constant which may differ at each occurrence. Accordingly $C(\alpha)$ stands for constants depending on the parameter α. By A we typically denote a finite multi-indexed matrix $(a_{i_1,\ldots,i_d})_{1\leq i_1,\ldots,i_d\leq n}$ of order d with values in a normed space $(F,\|\cdot\|)$. If $\mathbf{i}=(i_1,\ldots,i_d)\in[n]^d$ and $I\subset[d]$, then we define $i_I:=(i_j)_{j\in I}$.

If U is a finite set then $|U|$ stands for its cardinality and by $\mathcal{P}(U)$ we denote a family of (unordered) partitions of U into nonempty, pairwise disjoint sets. Note that if $U=\emptyset$ then $\mathcal{P}(U)$ consists only of the empty partition $\emptyset$.

With a slight abuse of notation we write $(\mathcal{P},\mathcal{P}')\in\mathcal{P}(U)$ if $\mathcal{P}\cup\mathcal{P}'\in\mathcal{P}(U)$ and $\mathcal{P}\cap\mathcal{P}'=\emptyset$.

Let $\mathcal{P}=\{I_1,\ldots,I_k\},\ \mathcal{P}'=\{J_1,\ldots,J_m\}$ be such that $(\mathcal{P},\mathcal{P}')\in\mathcal{P}([d])$. Then we define

$$\|A\|_{\mathcal{P}',\mathcal{P}} := \sup\left\{\mathbb{E}\left\|\sum_{i_1,\ldots,i_d} a_{i_1,\ldots,i_d}\prod_{r=1}^{k} x^r_{i_{I_r}}\prod_{l=1}^{m} g^l_{i_{J_l}}\right\| \;\Big|\; \forall_{r\le k}\sum_{i_{I_r}}\left(x^r_{i_{I_r}}\right)^2\le 1\right\}, \tag{4.3}$$

$$|||A|||_{\mathcal{P}} := \sup\left\{\mathbb{E}\left\|\sum_{i_1,\ldots,i_d} a_{i_1,\ldots,i_d}\prod_{r=1}^{k} x^r_{i_{I_r}}\prod_{l\in[d]\setminus(\bigcup\mathcal{P})} g^l_{i_l}\right\| \;\Big|\; \forall_{r\le k}\sum_{i_{I_r}}\left(x^r_{i_{I_r}}\right)^2\le 1\right\}. \tag{4.4}$$

We do not exclude the situation that $\mathcal{P}'$ or $\mathcal{P}$ is an empty partition. In the first case $\|A\|_{\mathcal{P}',\mathcal{P}} = |||A|||_{\mathcal{P}}$ is defined in non-probabilistic terms. In particular for $d=3$ we have

$$\|A\|_{\emptyset,\{1,2,3\}} = |||A|||_{\{1,2,3\}} = \sup_{\sum_{i,j,k} x_{ijk}^2\le 1}\left\|\sum a_{ijk}x_{ijk}\right\|,$$

$$\|A\|_{\{\{1,2\},\{3\}\},\emptyset} = \mathbb{E}\left\|\sum_{i,j,k} a_{ijk}g_{ij}g'_k\right\|,$$

$$\|A\|_{\{1\},\{\{2\},\{3\}\}} = |||A|||_{\{2\},\{3\}} = \sup_{\sum_j x_j^2\le 1,\sum y_k^2\le 1}\mathbb{E}\left\|\sum_{i,j,k} a_{ijk}g_i x_j y_k\right\|,$$

$$\|A\|_{\{\{1\},\{2\},\{3\}\},\emptyset} = |||A|||_{\emptyset} = \mathbb{E}\left\|\sum_{ijk} a_{ijk}g_i g'_j g'''_k\right\|,$$

$$\|A\|_{\{\{1\},\{3\}\},\{\{2\}\}} = |||A|||_{\{2\}} = \sup_{\sum_j x_j^2\le 1}\mathbb{E}\left\|\sum_{i,j,k} a_{ijk}g_i x_j g'_l\right\|,$$

$$\|A\|_{\emptyset,\{\{1\},\{2,3\}\}} = |||A|||_{\{1\},\{2,3\}} = \sup_{\sum_i x_i^2\le 1,\sum_{jk} y_{jk}^2\le 1}\left\|\sum_{ijk} a_{ijk}x_i y_{jk}\right\|.$$

The main result is the following moment estimate of the variable S'.

Theorem 4.2.1. *Assume that $A=(a_{i_1,\ldots,i_d})_{i_1,\ldots,i_d}$ is a finite matrix with values in a normed space $(F,\|\cdot\|)$. Then for any $p\ge 1$,*

$$C^{-1}(d)\sum_{J\subset[d]}\sum_{\mathcal{P}\in\mathcal{P}(J)} p^{|\mathcal{P}|/2}|||A|||_{\mathcal{P}} \le \left\|\sum_{i_1,\ldots,i_d} a_{i_1,\ldots,i_d}g^1_{i_1}\cdots g^d_{i_d}\right\|_p \le C(d)\sum_{(\mathcal{P},\mathcal{P}')\in\mathcal{P}([d])} p^{|\mathcal{P}|/2}\|A\|_{\mathcal{P}',\mathcal{P}}. \tag{4.5}$$

The lower bound in (4.5) motivates the following conjecture (we leave it to the reader to verify that in general Banach spaces it is impossible to reverse the upper bound even if $d=2$).

Conjecture 4.2.2. *Under the assumption of Theorem 4.2.1 we have*

$$\left\|\sum_{i_1,\ldots,i_d} a_{i_1,\ldots,i_d}g^1_{i_1}\cdots g^d_{i_d}\right\|_p \le C(d)\sum_{J\subset[d]}\sum_{\mathcal{P}\in\mathcal{P}(J)} p^{|\mathcal{P}|/2}|||A|||_{\mathcal{P}}. \tag{4.6}$$

Remark 4.2.3. Unfortunately we are able to show (4.6) only for $d=2$ and with an additional factor $\ln p$ (cf. Chapter 3). It is likely that by a modification of our proof one can show (4.6) for arbitrary d with an additional factor $(\ln p)^{C(d)}$.

By a standard application of Chebyshev's inequality, Theorem 4.2.1 can be expressed in term of tails.

Theorem 4.2.4. *Under the assumptions of Theorem 4.2.1 the following two inequalities hold. For any* $t > C(d)\sum_{\mathcal{P}'\in\mathcal{P}([d])}\|A\|_{\mathcal{P}',\{\emptyset\}}$

$$\mathbb{P}\left(\left\|\sum_{i_1,\ldots,i_d} a_{i_1,\ldots,i_d} g^1_{i_1}\cdots g^d_{i_d}\right\| \geq t\right) \leq 2\exp\left(-\frac{1}{C(d)}\min_{\substack{(\mathcal{P},\mathcal{P}')\in\mathcal{P}([d])\\ |\mathcal{P}|>0}}\left(\frac{t}{\|A\|_{\mathcal{P}',\mathcal{P}}}\right)^{2/|\mathcal{P}|}\right),$$

and for any $t \geq 0$,

$$\begin{aligned}\mathbb{P}&\left(\left\|\sum_{i_1,\ldots,i_d} a_{i_1,\ldots,i_d} g^1_{i_1}\cdots g^d_{i_d}\right\| \geq C(d)^{-1}\mathbb{E}\left\|\sum_{i_1,\ldots,i_d} a_{i_1,\ldots,i_d} g^1_{i_1}\cdots g^d_{i_d}\right\| + t\right)\\ &\geq \frac{1}{C(d)}\exp\left(-C(d)\min_{\emptyset\neq J\subset[d]}\min_{\mathcal{P}\in\mathcal{P}(J)}\left(\frac{t}{\|\!|A|\!\|_{\mathcal{P}}}\right)^{2/|\mathcal{P}|}\right).\end{aligned}$$

In view of (4.2) and [16] it is clear that to prove Theorem 4.2.1 one needs to estimate suprema of some Gaussian processes. The next statement is the key element of the proof of the upper bound in (4.5).

Theorem 4.2.5. *Under the assumptions of Theorem 4.2.1 we have for any* $p \geq 1$

$$\mathbb{E}\sup_{(x^2,\ldots,x^d)\in(B_2^n)^{d-1}}\left\|\sum_{i_1,\ldots,i_d} a_{i_1,\ldots,i_d} g_{i_1}\prod_{k=2}^d x^k_{i_k}\right\| \leq C(d)\sum_{(\mathcal{P},\mathcal{P}')\in\mathcal{P}([d])} p^{\frac{|\mathcal{P}|+1-d}{2}}\|A\|_{\mathcal{P}',\mathcal{P}}. \tag{4.7}$$

We postpone proofs of the above results till Section 4.5 and discuss now some of their consequences.

4.2.1 Two-sided estimates in special classes of Banach spaces

We will start by introducing a class of normed spaces for which the estimate (4.5) is two-sided. To this end we restrict our attention to normed spaces $(F,\|\cdot\|)$ which satisfy the following condition: there exists a constant $K = K(F)$ such that for any $n\in\mathbb{N}$ and any matrix $(b_{i,j})_{i,j\leq n}$ with values in F,

$$\mathbb{E}\left\|\sum_{i,j} b_{i,j} g_{i,j}\right\| \leq K\mathbb{E}\left\|\sum_{i,j} b_{i,j} g_i g'_j\right\|. \tag{4.8}$$

Remark 4.2.6. By considering $n = 1$ it is easy to see that $K \geq \sqrt{\pi/2} > 1$.

A simple inductive argument and (4.8) yield that for any $d,n\in\mathbb{N}$ and any F-valued matrix $(b_{i_1,\ldots,i_d})_{i_1,\ldots,i_d\leq n}$,

$$\mathbb{E}\left\|\sum_{\mathbf{i}} b_{\mathbf{i}} g_{\mathbf{i}}\right\| \leq K^{d-1}\mathbb{E}\left\|\sum_{\mathbf{i}} b_{\mathbf{i}} g^1_{i_1}\cdots g^d_{i_d}\right\|, \tag{4.9}$$

where we recall that $\mathbf{i}=(i_1,\dots,i_d)\in[n]^d$. It turns out that under the condition (4.8) our bound (4.5) is actually two-sided.

Proposition 4.2.7. *Assume that* $(F,\|\cdot\|)$ *satisfies* (4.8) *and* $(\mathcal{P},\mathcal{P}')\in\mathcal{P}([d])$. *Then*

$$\|A\|_{\mathcal{P}',\mathcal{P}} \leq K^{|\bigcup\mathcal{P}'|-|\mathcal{P}'|}\|\!|A|\!\|_{\mathcal{P}}.$$

Proof. Let $\mathcal{P}'=(J_1,\dots,J_k)$, $\mathcal{P}=(I_1,\dots,I_m)$. The proof is by induction on $r=|\{l\ :\ |J_l|\geq 2\}|$. If $r=1$ then conditional application of (4.9) implies the assertion. Assume that the statement holds for $r>1$ and $|\{l\ :\ |J_l|\geq 2\}|=r+1$. Without loss of generality $|J_1|\geq 2$. Combining Fubini's Theorem with (4.9) we obtain

$$\begin{aligned}
\|A\|_{\mathcal{P}',\mathcal{P}} &= \sup\left\{\mathbb{E}^{(G^2,\dots,G^m)}\mathbb{E}^{G^1}\left\|\sum_{\mathbf{i}} a_{\mathbf{i}}\prod_{r=1}^m x^r_{i_{I_r}} g^1_{i_{J_1}}\prod_{l=2}^k g^l_{i_{J_l}}\right\| \ \Big|\ \forall_{r\leq m}\sum_{i_{I_r}}\left(x^r_{i_{I_r}}\right)^2\leq 1\right\}\\
&\leq K^{|J_1|-1}\sup\left\{\mathbb{E}^{(G^2,\dots,G^m)}\mathbb{E}^{G'}\left\|\sum_{\mathbf{i}} a_{\mathbf{i}}\prod_{r=1}^m x^r_{i_{I_r}}\prod_{j\in J_1}(g')^j_{i_j}\prod_{l=2}^k g^l_{i_{J_l}}\right\| \ \Big|\ \forall_{r\leq m}\sum_{i_{I_r}}\left(x^r_{i_{I_r}}\right)^2\leq 1\right\}\\
&\leq K^{|\bigcup\mathcal{P}'|-|\mathcal{P}'|}\|\!|A|\!\|_{\mathcal{P}},
\end{aligned}$$

where $G^l=(g^l_{i_{I_l}})_{i_{I_l}}$, $G'=((g')^j_{i_j})_{j\in J_1, i_{J_1}}$ and in the last inequality we used the induction assumption. □

The following is an obvious consequence of Proposition 4.2.7 and Theoerems 4.2.1, 4.2.4.

Corollary 4.2.8. *For any normed space* $(F,\|\cdot\|)$ *satisfying* (4.8) *and any* $p\geq 1$, *we have*

$$\left\|\sum_{i_1,\dots,i_d} a_{i_1,\dots,i_d} g^1_{i_1}\cdots g^d_{i_d}\right\|_p \leq C(d)K^{d-1}\sum_{J\subset[d]}\sum_{\mathcal{P}\in\mathcal{P}([J])} p^{|\mathcal{P}|/2}\|\!|A|\!\|_{\mathcal{P}},$$

and for $t>C(d)K^{d-1}\mathbb{E}\left\|\sum_{\mathbf{i}} a_{\mathbf{i}} g^1_{i_1}\cdots g^d_{i_d}\right\|$,

$$\mathbb{P}\left(\left\|\sum_{i_1,\dots,i_d} a_{i_1,\dots,i_d} g^1_{i_1}\cdots g^d_{i_d}\right\|\geq t\right)\leq \exp\left(-\frac{1}{C(d)}K^{2-2d}\min_{\emptyset\neq J\subset[d]}\min_{\mathcal{P}\in\mathcal{P}(J)}\left(\frac{t}{\|\!|A|\!\|_{\mathcal{P}}}\right)^{2/|\mathcal{P}|}\right).$$

Thanks to infinite divisibility of Gaussian variables, the above corollary can be in fact generalized to arbitrary polynomials in Gaussian variables.

Theorem 4.2.9. *Let* $(F,\|\cdot\|)$ *be a Banach space. If* G *is a standard Gaussian vector in* $\mathbb{R}^n$ *and* $f\colon\mathbb{R}^n\to F$ *is a polynomial of degree* D, *then for all* $p\geq 2$,

$$\|f(G)-\mathbb{E}f(G)\|_p \geq C(D)^{-1}\Big(\mathbb{E}\|f(G)-\mathbb{E}f(G)\| + \sum_{1\leq d\leq D}\sum_{\emptyset\neq J\subset[d]}\sum_{\mathcal{P}\in\mathcal{P}(J)} p^{\frac{|\mathcal{P}|}{2}}\|\!|\mathbb{E}\nabla^d f(G)|\!\|_{\mathcal{P}}\Big) \tag{4.10}$$

and for all $t>0$,

$$\mathbb{P}\Big(\|f(G)-\mathbb{E}f(G)\| \geq C(D)^{-1}(\mathbb{E}\|f(G)-\mathbb{E}f(G)\|+t)\Big) \geq \frac{1}{C(D)}\exp\Big(-C(D)\eta_f(t)\Big), \tag{4.11}$$

where

$$\eta_f(t) = \min_{1\leq d\leq D}\min_{\emptyset\neq J\subset[d]}\min_{\mathcal{P}\in\mathcal{P}(J)}\Big(\frac{t}{\|\|\mathbb{E}\nabla^d f(G)\|\|_{\mathcal{P}}}\Big)^{2/|\mathcal{P}|}.$$

Moreover, if F satisfies (4.8)*, then for all $p\geq 1$,*

$$\|f(G)-\mathbb{E}f(G)\|_p \leq C(D)K^{D-1}\Big(\mathbb{E}\|f(G)-\mathbb{E}f(G)\| + \sum_{1\leq d\leq D}\sum_{\emptyset\neq J\subset[d]}\sum_{\mathcal{P}\in\mathcal{P}(J)} p^{\frac{|\mathcal{P}|}{2}}\|\|\mathbb{E}\nabla^d f(G)\|\|_{\mathcal{P}}\Big) \tag{4.12}$$

and for all $t\geq C(D)K^{D-1}\mathbb{E}\|F(G)-\mathbb{E}f(G)\|$,

$$\mathbb{P}\Big(\|f(G)-\mathbb{E}f(G)\|\geq t\Big) \leq 2\exp\Big(-C(D)^{-1}K^{2-2D}\eta_f(t)\Big). \tag{4.13}$$

4.2.2 L_q spaces

It turns out that L_q spaces satisfy (4.8) and as a result estimate (4.5) is two-sided. Moreover, as is shown in Lemma 4.2.11 below, in this case one may express all the parameters without any expectations. For the sake of brevity, we will focus on moment estimates, clearly tail bounds follow from them by standard arguments (cf. the proof of Theorem 4.2.4).

Proposition 4.2.10. *Space $L_q(X,\mu)$ satisfies* (4.8) *with $K = Cq$.*

Proof. Observe that (4.8) reads as

$$\mathbb{E}\left(\int_X \left|\sum_{i,j} b_{ij}(x)g_{ij}\right|^q d\mu(x)\right)^{1/q} \leq K\mathbb{E}\left(\int_X \left|\sum_{i,j} b_{ij}(x)g_i g_j'\right|^q d\mu(x)\right)^{1/q},$$

where $(b_{ij})_{ij}$ is a matrix with values in $L_q(X,d\mu)$. Jensen's inequality and Fubini's theorem imply

$$\begin{aligned}\mathbb{E}\left(\int_X \left|\sum_{i,j} b_{ij}(x)g_{ij}\right|^q d\mu(x)\right)^{1/q} &\leq \left(\int_X \mathbb{E}\left|\sum_{i,j} b_{ij}(x)g_{ij}\right|^q d\mu(x)\right)^{1/q} \\ &\leq C\sqrt{q}\left(\int_X\left(\sum_{i,j} b_{ij}^2(x)\right)^{q/2} d\mu(x)\right)^{1/q}.\end{aligned}$$

On the other hand Theorem 6.3 yields

$$C q\mathbb{E}\left(\int_X\left|\sum_{i,j} b_{ij}(x) g_i g_j'\right|^q d\mu(x)\right)^{1/q} \geq \left(\int_X \mathbb{E}\left|\sum_{i,j} b_{ij}(x) g_i g_j'\right|^q d\mu(x)\right)^{1/q}$$
$$\geq C^{-1}\sqrt{q}\left(\int_X\left(\sum_{i,j} b_{ij}^2(x)\right)^{q/2} d\mu(x)\right)^{1/q},$$

where in the last inequality we used Corollary 6.21. □

For a multi-indexed matrix A of order d with values in $L_q(X,\mu)$ and $J\subset[d]$, $\mathcal{P}=(I_1,\ldots,I_k)\in\mathcal{P}([J])$ we define

$$|||A|||_{\mathcal{P}}^{L_q} = \sup\left\{\left\|\sqrt{\sum_{i_{[d]\setminus J}}\left(\sum_{i_J} a_{\mathbf{i}}\prod_{r=1}^k x^r_{i_{I_r}}\right)^2}\right\|_{L_q} \;\middle|\; \forall_{r\leq k}\sum_{i_{I_r}}\left(x^r_{i_{I_r}}\right)^2\leq 1\right\}.$$

For $J=[d]$ and $\mathcal{P}\in\mathcal{P}(J)$ we obviously have $|||A|||_{\mathcal{P}}^{L_q}=|||A|||_{\mathcal{P}}$. The following lemma asserts that for general J the corresponding two norms are comparable.

Lemma 4.2.11. *For any $J\subsetneq[d]$ and $\mathcal{P}=(I_1,\ldots,I_k)\in\mathcal{P}(J)$ and any multi-indexed matrix A of order d with values in $L_q(X,d\mu)$ we have*

$$C(d)^{-1}q^{\frac{1-d+|J|}{2}}|||A|||_{\mathcal{P}}^{L_q}\leq |||A|||_{\mathcal{P}}\leq C(d)q^{\frac{d-|J|}{2}}|||A|||_{\mathcal{P}}^{L_q}.$$

Proof. By Jensen's inequality and Corollary 6.21 we get

$$|||A|||_{\mathcal{P}}\leq\sup\left\{\left(\int_X\mathbb{E}\left|\sum_{\mathbf{i}} a_{\mathbf{i}}(x)\prod_{j\in[d]\setminus J} g^j_{i_j}\prod_{r=1}^k x^r_{i_{I_r}}\right|^q d\mu(x)\right)^{1/q} \;\middle|\; \forall_{r\leq k}\sum_{i_{I_r}}\left(x^r_{i_{I_r}}\right)^2\leq 1\right\}$$
$$\leq C(d)q^{\frac{d-|J|}{2}}\sup\left\{\left\|\sqrt{\sum_{i_{[d]\setminus J}}\left(\sum_{i_J} a_{\mathbf{i}}\prod_{r=1}^k x^r_{i_{I_r}}\right)^2}\right\|_{L_q} \;\middle|\; \forall_{r\leq k}\sum_{i_{I_r}}\left(x^r_{i_{I_r}}\right)^2\leq 1\right\}.$$

On the other hand Theorem 6.3 and Corollary 6.21 yield

$$|||A|||_{\mathcal{P}}\geq C(d)^{-1}q^{\frac{|J|-d}{2}}\sup\left\{\left(\mathbb{E}\left\|\sum_{\mathbf{i}} a_{\mathbf{i}}(x)\prod_{j\in[d]\setminus J} g^j_{i_j}\prod_{r=1}^k x^r_{i_{I_r}}\right\|_{L_q}^q\right)^{1/q} \;\middle|\; \forall_{r\leq k}\sum_{i_{I_r}}\left(x^r_{i_{I_r}}\right)^2\leq 1\right\}$$
$$\geq C(d)^{-1}q^{\frac{1-d+|J|}{2}}\sup\left\{\left\|\sqrt{\sum_{i_{[d]\setminus J}}\left(\sum_{i_J} a_{\mathbf{i}}\prod_{r=1}^k x^r_{i_{I_r}}\right)^2}\right\|_{L_q} \;\middle|\; \forall_{r\leq k}\sum_{i_{I_r}}\left(x^r_{i_{I_r}}\right)^2\leq 1\right\}.$$

□

Theorem 4.2.12. *Let $A=(a_{i_1,\ldots,i_d})_{i_1,\ldots,i_d}$ be a matrix with values in $L_q(X,\mu)$. Then for any $p\geq 1$ we have*

$$\frac{1}{C(d)}q^{\frac{1-d}{2}}\sum_{J\subset[d]}\sum_{\mathcal{P}\in\mathcal{P}([J])}p^{\frac{|\mathcal{P}|}{2}}|||A|||_{\mathcal{P}}^{L_q}\leq\left\|\sum_{i_1,\ldots,i_d}a_{i_1,\ldots,i_d}g^1_{i_1}\cdots g^d_{i_d}\right\|_p$$
$$\leq C(d)q^{3d/2-1}\sum_{J\subset[d]}\sum_{\mathcal{P}\in\mathcal{P}([J])}p^{\frac{|\mathcal{P}|}{2}}|||A|||_{\mathcal{P}}^{L_q}.$$

Proof. This is an obvious consequence of Theorem 4.2.1, Propositions 4.2.7, 4.2.10 and Lemma 4.2.11. □

Using Theorem 4.2.9 we can generalize the above result to general polynomials

Theorem 4.2.13. *Let G be standard Gaussian vector in $\mathbb{R}^n$ and $f\colon\mathbb{R}^n\to L_q(X,\mu)$ a polynomial of degree D. Then for $p\geq 1$, we have*

$$\frac{1}{C(D)}\sum_{d=0}^{D}q^{(1-d)/2}\sum_{\emptyset\neq J\subset[d]}\sum_{\mathcal{P}\in\mathcal{P}([J])}p^{\frac{|\mathcal{P}|}{2}}|||\mathbb{E}\nabla^d f(G)|||_{\mathcal{P}}^{L_q}\leq\|f(G)-\mathbb{E}f(G)\|_p$$
$$\leq C(D)q^{D-1}\sum_{d=1}^{D}q^{d/2}\sum_{J\subset[d]}\sum_{\mathcal{P}\in\mathcal{P}([J])}p^{\frac{|\mathcal{P}|}{2}}|||\mathbb{E}\nabla^d f(G)|||_{\mathcal{P}}^{L_q}.$$

4.2.3 Exponential variables

Theorem 4.2.12 together with Lemma 6.13 allows us to obtain inequalities for chaoses based on the i.i.d standard symmetric exponential random variables (i.e., variables with density $2^{-1}\exp(-|x|/2)$) which are denoted by $(E^i_j)_{i,j\in\mathbb{N}}$ below. Similar as in the previous Section we concentrate only on the moment estimates.

Proposition 4.2.14. *Let $A=(a_{i_1,\ldots,i_d})_{i_1,\ldots,i_d}$ be a matrix with values in $L_q(X,\mu)$. Then for any $p\geq 1,\ q\geq 2$ we have*

$$\left\|\sum_{\mathbf{i}}a_{\mathbf{i}}\prod_{r=1}^{d}E^r_{i_r}\right\|_p\sim^{d,q}\sum_{I\subset[d]}\sum_{J\subset[d]\setminus I}\sum_{\mathcal{P}([d]\setminus(I\cup J))}p^{|I|+|\mathcal{P}|/2}\max_{i_I}|||(a_{i_1,\ldots,i_d})_{i_{I^c}}|||_{\mathcal{P}}^{L_q}.$$

One can take $C^{-1}(d)q^{1/2-d}$ in the lower bound and $C(d)q^{3d-1}$ in the upper bound.

Example 4.2.15. If $d=2$ then Proposition 4.2.14 reads as

$$\left\|\sum_{ij} a_{i,j} E_i^1 E_j^2\right\|_p \sim^q p^2 \max_{i,j} \|a_{i,j}\|_{L_q} + p^{3/2}\left(\max_i \sup_{x\in B_2^n}\left\|\sum_j a_{i,j}x_j\right\|_{L_q} + \max_j \sup_{x\in B_2^n}\left\|\sum_i a_{i,j}x_i\right\|_{L_q}\right)$$
$$+ p\left(\max_{x,y\in B_2^n}\left\|\sum a_{i,j}x_iy_j\right\|_{L_q} + \max_i\left\|\sqrt{\sum_j a_{i,j}^2}\right\|_{L_q} + \max_j\left\|\sqrt{\sum_i a_{i,j}^2}\right\|_{L_q}\right)$$
$$+ p^{1/2}\left(\sup_{x\in B_2^n}\left\|\sqrt{\sum_i\left(\sum_j a_{i,j}x_j\right)^2}\right\|_{L_q} + \sup_{x\in B_2^n}\left\|\sqrt{\sum_j\left(\sum_i a_{i,j}x_i\right)^2}\right\|_{L_q}\right)$$
$$+ \left\|\sqrt{\sum_{ij} a_{i,j}^2}\right\|_{L_q}.$$

Proof of Proposition 4.2.14. Lemma 6.13 implies

$$\left\|\sum_{\mathbf{i}} a_{\mathbf{i}} \prod_{r=1}^{d} E_{i_r}^r\right\|_p \sim^d \left\|\sum_{i_1,\ldots,i_{2d}} \hat{a}_{i_1,\ldots,i_{2d}} \prod_{r=1}^{2d} g_{i_r}^r\right\|_p, \tag{4.14}$$

where

$$\hat{a}_{i_1,\ldots,i_{2d}} = a_{i_1,\ldots,i_d}\mathbf{1}_{\{i_1=i_{d+1},\ldots,i_d=i_{2d}\}}.$$

Let $\hat{A} = (\hat{a}_{i_1,\ldots,i_{2d}})_{i_1,\ldots,i_{2d}}$. Theorem 4.2.12 and (4.14) yields

$$C^{-1}(d)q^{1/2-d}\sum_{J\subset[2d]}\sum_{\mathcal{P}\in\mathcal{P}([J])} p^{\frac{|\mathcal{P}|}{2}}|||\hat{A}|||_{\mathcal{P}}^{L_q} \le \left\|\sum_{\mathbf{i}} a_{\mathbf{i}}\prod_{r=1}^{d} E_{i_r}^r\right\|_p \le C(d)q^{3d-1}\sum_{J\subset[2d]}\sum_{\mathcal{P}\in\mathcal{P}([J])} p^{\frac{|\mathcal{P}|}{2}}|||\hat{A}|||_{\mathcal{P}}^{L_q} \tag{4.15}$$

Observe that $\sum_{J\subset[2d]}\sum_{\mathcal{P}\in\mathcal{P}([J])} p^{\frac{|\mathcal{P}|}{2}}|||\hat{A}|||_{\mathcal{P}}^{L_q}$ can be expressed in terms of the matrix A by using a different language. Consider a finite sequence $\mathcal{M} = (J, I_1, \ldots, I_k)$ of subsets of $[d]$, such that $J \cup I_1 \cup \ldots \cup I_k = [d]$, $I_1, \ldots, I_k \neq \emptyset$ and each number $m \in [d]$ belongs to at most two of the sets $J, I_1, \ldots, I_k$. Denote the family of all such sequences by $\mathcal{M}([d])$. For $\mathcal{M} = (J, I_1, \ldots, I_k)$ set $|\mathcal{M}| = k+1$ and

$$\langle A\rangle_{\mathcal{M}}^{L_q} := \sup\left\{\left\|\sqrt{\sum_{i_J}\left(\sum_{i_{[d]\setminus J}} a_{\mathbf{i}}\prod_{r=1}^{k} x_{i_{I_r}}^r\right)^2}\right\|_{L_q} \mid \forall_{r\le k}\sum_{i_{I_r}}\left(x_{i_{I_r}}^r\right)^2 \le 1\right\},$$

where we do not exclude that $J = \emptyset$. By an easy verification

$$\sum_{J\subset[2d]}\sum_{\mathcal{P}\in\mathcal{P}([J])} p^{\frac{|\mathcal{P}|}{2}}|||\hat{A}|||_{\mathcal{P}}^{L_q} \sim^d \sum_{\mathcal{M}\in\mathcal{M}([d])} p^{\frac{|\mathcal{M}|-1}{2}}\langle A\rangle_{\mathcal{M}}^{L_q}. \tag{4.16}$$

It is enough to prove that

$$\sum_{\mathcal{M}\in\mathcal{M}([d])} p^{\frac{|\mathcal{M}|-1}{2}}\langle A\rangle_{\mathcal{M}}^{L_q} \sim^d \sum_{\mathcal{M}\in\mathcal{C}} p^{\frac{|\mathcal{M}|-1}{2}}\langle A\rangle_{\mathcal{M}}^{L_q}, \tag{4.17}$$

where

$$\mathcal{C} = \Bigg\{ \mathcal{M} = (J, I_1, \ldots, I_k) \in \mathcal{M}([d]) \mid J \cap \left(\bigcup_{r=1}^{k} I_r \right) = \emptyset,$$

$$\forall_{r,m \le k} I_m \cap I_r \neq \emptyset \Rightarrow (|I_r| = |I_m| = 1,\ I_r = I_m) \Bigg\}.$$

Indeed assume that (4.17) holds and choose $\mathcal{M} = (J, I_1, \ldots, I_k) \in \mathcal{C}$.

Let $I = \{i \mid \exists_{r<m\le k}\ \{i\} = I_r = I_m\}$. Then $J \cap I = \emptyset$ and we have

$$\left(\langle A\rangle_{\mathcal{M}}^{L_q}\right)^q = \sup\Bigg\{ \int_X \left(\sum_{i_J} \left(\sum_{i_{J^c}} a_{i_1,\ldots,i_d} \prod_{r\in I} y^r_{i_r} x^r_{i_r} \prod_{\substack{r\le k\\ I_r\cap I=\emptyset}} x^r_{i_{I_r}} \right)^2 \right)^{q/2} d\mu(x) \ \Big| $$

$$\forall_{1\le r\le k} \sum_{i_{I_r}} (x^r_{i_{I_r}})^2 \le 1,\ \sum_{i_r} (y^r_{i_r})^2 \le 1 \Bigg\}$$

$$= \max_{i_I} \sup \Bigg\{ \int_X \left(\sum_{i_J} \left(\sum_{i_{J^c\setminus I}} a_{i_1,\ldots,i_d} \prod_{\substack{r\le k\\ I_r\cap I=\emptyset}} x^r_{i_{I_r}} \right)^2 \right)^{q/2} d\mu(x) \ \Big| \ \forall_{1\le r\le k} \sum_{i_{I_r}} (x^r_{i_{I_r}})^2 \le 1 \Bigg\}$$

$$= \max_{i_I} \left(\| (a_{i_1,\ldots,i_d})_{i_{I^c}} \|^{L_q}_{\{I_r\,:\, I_r\cap I=\emptyset\}} \right)^q =: \max_{i_I} \left(\| (a_{i_1,\ldots,i_d})_{i_{I^c}} \|^{L_q}_{\mathcal{P}} \right)^q, \tag{4.18}$$

where in the second equality we used the fact that $(y^r_{i_r} x^r_{i_r})_{i_r} \in B^n_1$ together with convexity and homogeneity of the norm

$$\|(f_{i_J})_{i_J}\|_{L_q(\ell_2)} = \left(\int_X \left(\sum_{i_J} f^2_{i_J} \right)^{q/2} \right)^{1/q}.$$

By combining the above with (4.15)-(4.17) we conclude the assertion.

The proof is completed by showing that

$$\sum_{\mathcal{M}\in\mathcal{M}([d])} p^{\frac{|\mathcal{M}|-1}{2}} \langle A\rangle_{\mathcal{M}}^{L_q} \le C(d) \sum_{\mathcal{M}\in\mathcal{C}} p^{\frac{|\mathcal{M}|-1}{2}} \langle A\rangle_{\mathcal{M}}^{L_q}$$

(the second inequality in (4.17) is trivial), which will be done in two steps. Let us fix $\mathcal{M} = (J, I_1, \ldots, I_k) \in \mathcal{M}([d])$.

1. Assume first that $J \cap (\bigcup_{i=1}^k I_i) \neq \emptyset$.
   Without loss of the generality $1 \in J \cap I_1$. Denote $\hat{I}_1 = I_1 \setminus \{1\}$ and for any matrix $(x^1_{i_I})_{i_I}$ such that $\sum_{i_{I_1}} (x^1_{i_{I_1}})^2 \le 1$, set $(b^2_{i_1})_{i_1} := (\sum_{i_{I_1\setminus\{1\}}} (x^1_{i_{I_1}})^2)_{i_1} \in B^n_1$. Observe that for any $f_1, \ldots, f_n \in L^q(X, d\mu)$ the function

$$[0, +\infty)^n \ni v \to \int_X \left(\sum_i f_i^2(x) v_i \right)^{q/2} d\mu(x)$$

is convex (recall that $q \geq 2$). Therefore we have

$$\left(\langle A\rangle_{\mathcal{M}}^{L_q}\right)^q = \sup\left\{\int_X \left(\sum_{i_J} b_{i_1}^2 \left(\sum_{i_{J^c}} a_{i_1,\ldots,i_d} \frac{x^1_{i_{I_1}}}{b_{i_1}} \prod_{r=2}^k x^r_{i_{I_r}}\right)^2\right)^{q/2} d\mu(x) \;\middle|\; \forall_{1\leq r\leq k} \sum_{i_{I_r}} (x^r_{i_{I_r}})^2 \leq 1\right\}$$

$$\leq \max_{i_1} \sup\left\{\int_X \left(\sum_{i_{J\setminus\{1\}}} \left(\sum_{i_{J^c}} a_{i_1,\ldots,i_d} \frac{x^1_{i_{I_1}}}{b_{i_1}} \prod_{r=2}^k x^r_{i_{I_r}}\right)^2\right)^{q/2} d\mu(x) \;\middle|\; \forall_{1\leq r\leq k} \sum_{i_{I_r}} (x^r_{i_{I_r}})^2 \leq 1\right\}$$

$$\leq \max_{i_1} \sup\left\{\int_X \left(\sum_{i_{J\setminus\{1\}}} \left(\sum_{i_{J^c}} a_{i_1,\ldots,i_d} y_{i_{\hat{I}_1}} \prod_{r=2}^k x^r_{i_{I_r}}\right)^2\right)^{q/2} d\mu(x) \;\middle|\; \sum_{i_{\hat{I}_1}} (y_{i_{\hat{I}_1}})^2 \leq 1,\ \forall_{1\leq r\leq k} \sum_{i_{I_r}} (x^r_{i_{I_r}})^2 \leq 1\right\}.$$

If $\hat{I}_1 \neq \emptyset$ let $\mathcal{M}' = (J\setminus\{1\}, \{1\}, \{1\}, \hat{I}_1, I_2, \ldots, I_k)$, otherwise set $\mathcal{M}' = (J\setminus\{1\}, \{1\}, \{1\}, I_2, \ldots, I_k)$. By the same argument as was used for the second equality of (4.18) we obtain that the right-hand side above equals $\langle A\rangle_{\mathcal{M}'}^{L_q}$, which gives

$$\langle A\rangle_{\mathcal{M}}^{L_q} \leq \langle A\rangle_{\mathcal{M}'}^{L_q}.$$

Observe that

$$p^{(|\mathcal{M}|-1)/2} \langle A\rangle_{\mathcal{M}}^{L_q} \leq p^{(|\mathcal{M}|-1)/2} \langle A\rangle_{\mathcal{M}'}^{L_q} \leq p^{(|\mathcal{M}'|-1)/2} \langle A\rangle_{\mathcal{M}'}^{L_q}.$$

By iterating this argument we obtain that $p^{(|\mathcal{M}|-1)/2} \langle A\rangle_{\mathcal{M}}^{L_q} \leq p^{(|\mathcal{M}''|-1)/2} \langle A\rangle_{\mathcal{M}''}^{L_q}$ for some $\mathcal{M}'' = (J'', I_1'', \ldots, I_m'')$ such that $J'' \cap (\bigcup_{r=1}^m I_r'') = \emptyset$.

2. Assume that for some $r, m \leq k$ $I_r \cap I_m \neq \emptyset$ and $|I_r| \geq 2$ or $|I_m| \geq 2$.
Without loss of the generality $1 \in I_1 \cap I_2$ and $|I_1| \geq 2$. Clearly,

$$\left(\langle A\rangle_{\mathcal{M}}^{L_q}\right)^q = \sup\left\{\int_X \left(\sum_{i_J} \left(\sum_{i_{J^c}} a_{i_1,\ldots,i_d} b_{i_1} c_{i_1} \frac{x^1_{i_{I_1}}}{b_{i_1}} \frac{x^2_{i_{I_2}}}{c_{i_1}} \prod_{r=3}^k x^r_{i_{I_r}}\right)^2\right)^{q/2} d\mu(x) \;\middle|\; \forall_{1\leq r\leq k} \sum_{i_{I_r}} (x^r_{i_{I_r}})^2 \leq 1\right\},$$

where $(b_{i_1})_{i_1} := \left(\sqrt{\sum_{i_{I_1\setminus\{1\}}} (x^1_{i_{I_1}})^2}\right)_{i_1}$, $(c_{i_1})_{i_1} := \left(\sqrt{\sum_{i_{I_2\setminus\{1\}}} (x^2_{i_{I_2}})^2}\right)_{i_1} \in B_2^n$. Since

$$\forall_{i_1} \sum_{i_{I_1\setminus\{1\}}} \left(\frac{x^1_{i_{I_1}}}{b_{i_1}}\right)^2 \leq 1, \quad \sum_{i_{I_2\setminus\{1\}}} \left(\frac{x^2_{i_{I_2}}}{b_{i_1}}\right)^2 \leq 1$$

we obtain similarly as in step 1,

$$p^{(|\mathcal{M}|-1)/2} \langle A\rangle_{\mathcal{M}}^{L_q} \leq p^{(|\mathcal{M}|-1)/2} \langle A\rangle_{\mathcal{M}'}^{L_q} \leq p^{(|\mathcal{M}'|-1)/2} \langle A\rangle_{\mathcal{M}'}^{L_q}$$

where $\mathcal{M}' = (J, \{1\}, \{1\}, I_1 \setminus \{1\}, I_2 \setminus \{1\}, I_3, \ldots, I_k)$ if $I_2 \setminus \{1\} \neq \emptyset$ and $\mathcal{M}' = (J, \{1\}, \{1\}, I_1 \setminus \{1\}, I_3, \ldots, I_k)$ otherwise. An iteration of this argument shows that indeed one can assume that $\mathcal{M}$ satisfies the implication $I_m \cap I_r \neq \emptyset \Rightarrow (|I_r| = |I_m| = 1,\ I_r = I_m)$.

Combining Steps 1 and 2 we obtain that for any $\mathcal{M} \in \mathcal{M}([d])$ there exists $\mathcal{M}' \in \mathcal{C}$ such that $p^{(|\mathcal{M}|-1)/2} \langle A \rangle_{\mathcal{M}}^{L_q} \leq p^{(|\mathcal{M}'|-1)/2} \langle A \rangle_{\mathcal{M}'}^{L_q}$ which yields (4.17). □

4.3 Reformulation of Theorem 4.2.5 and entropy estimates

Let us rewrite Theorem 4.2.5 in a different language. We may assume that $F = \mathbb{R}^m$ for some finite m and $a_{i_1,\ldots,i_d} = (a_{i_1,\ldots,i_d,i_{d+1}})_{i_{d+1} \leq m}$. For this reason $\mathbf{i} \in [n]^d \times [m]$ from now on. Let $T = B_{F^*}$ be the unit ball in the dual space F^*.

In this setup we have

$$\mathbb{E} \sup_{(x^2,\ldots,x^d) \in (B_2^n)^{d-1}} \left\| \sum_{\mathbf{i}} a_{\mathbf{i}} g_{i_1} \prod_{k=2}^{d} x_{i_k}^k \right\| = \mathbb{E} \sup_{(x^2,\ldots,x^d) \in (B_2^n)^{d-1}} \sup_{t \in T} \sum_{\mathbf{i}} a_{\mathbf{i}} g_{i_1} \prod_{k=2}^{d} x_{i_k}^k t_{i_{d+1}}$$

$$\|A\|_{\mathcal{P}',\mathcal{P}} = \sup \left\{ \mathbb{E} \sup_{t \in T} \sum_{\mathbf{i}} a_{\mathbf{i}} \prod_{r=1}^{k} x_{i_{I_r}}^r \prod_{s=1}^{l} g_{i_{J_s}}^s t_{i_{d+1}} \;\middle|\; \forall_{j=1\ldots,k} \sum_{i_{I_j}} \left(x_{i_{I_j}}^{(j)} \right)^2 = 1 \right\},$$

where $\mathcal{P} = (I_1, \ldots, I_k), \mathcal{P}' = (J_1, \ldots, J_l)$, $(\mathcal{P}', \mathcal{P}) \in \mathcal{P}([d])$.

To make the notation more compact we define

$$s_k(A) = \sum_{\substack{(\mathcal{P},\mathcal{P}') \in \mathcal{P}([d]) \\ |\mathcal{P}|=k}} \|A\|_{\mathcal{P}',\mathcal{P}}.$$

To prove Theorem 4.2.5 it suffices to show the following.

Theorem 4.3.1. *For any $p \geq 1$ we have*

$$\mathbb{E} \sup_{(x^2,\ldots,x^d,t) \in (B_2^n)^{d-1} \times T} \sum_{\mathbf{i}} a_{\mathbf{i}} g_{i_1} \prod_{k=2}^{d} x_{i_k}^k t_{i_{d+1}} \leq C(d) \sum_{k=0}^{d} p^{\frac{k+1-d}{2}} s_k(A). \tag{4.19}$$

To estimate the supremum of a centered Gaussian process $(G_v)_{v \in V}$ one needs to study the distance on V given by $d(v,v') := (\mathbb{E}|G_v - G_{v'}|^2)^{1/2}$ (cf. [30]). In the case of the Gaussian process from (4.19) this distance is defined on $(B_2^n)^{d-1} \times T \subset \mathbb{R}^{n(d-1)} \times \mathbb{R}^m$ by the formula

$$\begin{aligned}
&\rho_A((x^2,\ldots,x^d,t),(y^2,\ldots,y^d,t')) \\
&:= \left(\sum_{i_1} \left(\sum_{i_2,\ldots,i_{d+1}} a_{i_1,\ldots,i_{d+1}} \left(\prod_{k=2}^{d} y_{i_k}^k t'_{i_{d+1}} - \prod_{k=2}^{d} x_{i_k}^k t_{i_{d+1}} \right) \right) \right)^{1/2} \\
&= \alpha_A^d \left(\left(\bigotimes_{k=2}^{d} x^k \right) \otimes t - \left(\bigotimes_{k=2}^{d} y^k \right) \otimes t' \right),
\end{aligned} \tag{4.20}$$

where $\left(\bigotimes_{k=2}^{d} x^k \right) \otimes t = (x_{i_2}^2 \cdots x_{i_d}^d t_{i_{d+1}})_{i_2,\ldots,i_{d+1}} \in \mathbb{R}^{n^{d-1}m}$ and α_A is a norm on $\mathbb{R}^{n^{d-1}m}$ given by

$$\alpha_A^d(\mathbf{x}) := \sqrt{\sum_{i_1}\left(\sum_{i_{[d+1]\setminus\{1\}}} a_{\mathbf{i}}\mathbf{x}_{i_{[d+1]\setminus\{1\}}}\right)^2}. \tag{4.21}$$

Now we try to estimate entropy numbers $N(U,\rho_A,\varepsilon)$ for $\varepsilon>0$ and $U\subset(B_2^n)^{d-1}\times T$ (recall that $N(S,\varphi,\varepsilon)$ is the minimal number of closed balls with the diameter ε in metric φ that cover set S). To this end we first introduce some notation. From now on $G_n=(g_1,\dots,g_n)$ and $G_n^i=(g_1^i,\dots,g_n^i)$ stand for independent standard Gaussian vectors in $\mathbb{R}^n$. For $\varepsilon>0$, $U=\{(x^2,\dots,x^d,t)\in U\}\subset(\mathbb{R}^n)^{d-1}\times T$ we set

$$W_d^U(\alpha,\varepsilon) := \sum_{k=1}^{d-1}\varepsilon^k\sum_{I\subset\{2,\dots,d\}:|I|=k} W_I^U(\alpha), \tag{4.22}$$

where

$$W_I^U(\alpha) := \sup_{(x^2,\dots,x^{d+1},t)\in U}\mathbb{E}\alpha_A^d\left(\left(\bigotimes_{k=2}^d\left(x^k(1-\mathbf{1}_I(k))+G^k\mathbf{1}_I(k)\right)\right)\otimes t\right).$$

We define a norm β_A^d on $\mathbb{R}^{n^{d-1}}$ by

$$\beta_A^d(\mathbf{y}) := \mathbb{E}\sup_{t\in T}\sum_{\mathbf{i}} a_{\mathbf{i}}g_{i_1}\mathbf{y}_{i_{[d]\setminus\{1\}}}t_{i_{d+1}}. \tag{4.23}$$

Following (4.22) we denote

$$V_d^U(\beta,\varepsilon) := \sum_{k=0}^{d-1}\varepsilon^{k+1}\sum_{I\subset\{2,\dots,d\}:|I|=k} V_I^U(\beta), \tag{4.24}$$

where

$$V_I^U(\beta) := \sup_{(x^2,\dots,x^d,t)\in U}\mathbb{E}\beta_A^d\left(\bigotimes_{k=2}^d\left(x^k(1-\mathbf{1}_I(k))+G^k\mathbf{1}_I(k)\right)\right).$$

In particular

$$V_d^U(\beta,\varepsilon)\ge\varepsilon\cdot V_\emptyset^U(\beta)=\varepsilon\cdot\sup_{(x^2,\dots,x^d,t)\in U}\beta_A^d\left(\bigotimes_{k=2}^d x^k\right). \tag{4.25}$$

Observe that by the classical Sudakov minoration (Theorem 6.7), for any $(x^k)\in\mathbb{R}^n$, $k=2,\dots,d$ there exists $T_{\bigotimes x^k,\varepsilon}$ such that $|T_{\bigotimes x^k,\varepsilon}|\le\exp(C\varepsilon^{-2})$ and

$$\forall_{t\in T}\exists_{t'\in T_{\bigotimes x^k,\varepsilon}}\ \alpha_A^d\left(\bigotimes_{k=2}^d x^k\otimes(t-t')\right)\le\varepsilon\beta_A^d\left(\bigotimes_{k=2}^d x^k\right).$$

We define a measure $\mu_{\varepsilon,T}^d$ on $\mathbb{R}^{(d-1)n}\times T$ by the formula

$$\mu_{\varepsilon,T}^d(C) := \int_{\mathbb{R}^{(d-1)n}}\sum_{t\in T_{\bigotimes x^k,\varepsilon}}\mathbf{1}_C\big((x^2,\dots,x^d,t)\big)d\gamma_{(d-1)n,\varepsilon}((x^k)_{k=2,\dots,d}),$$

where $\gamma_{n,t}$ is the distribution of $tG_n=t(g_1,\dots,g_n)$.

To bound $N(U,\rho_A,\varepsilon)$ for $\varepsilon>0$ and $U\subset(B_2^n)^{d-1}\times T$ we need two lemmas.

Lemma 4.3.2. [16, Lemma 2] *For any* $\boldsymbol{x} = (x^1, \dots, x^d) \in (B_2^n)^d$*, norm* α' *on* $\mathbb{R}^{n^d}$ *and* $\varepsilon > 0$ *we have*

$$\gamma_{dn,\varepsilon}\left(B_{\alpha'}(\boldsymbol{x}, r(\varepsilon,\alpha'))\right) \geq 2^{-d}\exp(-d\varepsilon^{-2}/2),$$

where

$$B_{\alpha'}(\boldsymbol{x}, r(\varepsilon,\alpha')) = \left\{\boldsymbol{y} = (y^1,\dots,y^d) \in (\mathbb{R}^n)^d \;\middle|\; \alpha'\left(\bigotimes_{k=1}^d x^k - \bigotimes_{k=1}^d y^k\right) \leq r(\varepsilon,\alpha')\right\},$$

and

$$r(\varepsilon,\alpha') = \sum_{k=1}^d \varepsilon^k \sum_{I\subset[d]:\ |I|=k} \mathbb{E}\alpha'\left(\bigotimes_{k=1}^d \left(x^k(1-\boldsymbol{1}_{k\in I}) + G^k \boldsymbol{1}_{k\in I}\right)\right).$$

Lemma 4.3.3. *For any* $(\boldsymbol{x},t) = (x^2,\dots,x^d,t) \in (B_2^n)^{d-1}\times T$ *and* $\varepsilon > 0$ *we have*

$$\mu_{\varepsilon,T}^d\left(B\left((\boldsymbol{x},t),\rho_A, W_d^{\{(\boldsymbol{x},t)\}}(\alpha,8\varepsilon) + V_d^{\{(\boldsymbol{x},t)\}}(\beta,8\varepsilon)\right)\right) \geq C^{-d}\exp\left(-C(d)\varepsilon^{-2}\right).$$

Proof. Fix $(\mathbf{x},t) \in (B_2^n)^{d-1}\times T$, $\varepsilon > 0$ and consider

$$U = \left\{(y^2,\dots,y^d) \in (\mathbb{R}^n)^{d-1}:\ \alpha_A^d\left(\left(\bigotimes_{k=2}^d x^k - \bigotimes_{k=2}^d y^k\right)\otimes t\right) + \varepsilon\beta_A^d\left(\bigotimes_{k=2}^d x^k - \bigotimes_{k=2}^d y^k\right) \right.$$
$$\left. \leq W_d^{\{(\mathbf{x},t)\}}(\alpha,4\varepsilon) + V_d^{\{(\mathbf{x},t)\}}(\beta,4\varepsilon)\right\}.$$

For any $(y^2,\dots,y^d) \in U$ there exists $t' \in T_{\bigotimes y^k,\varepsilon}$ such that

$$\alpha_A^d\left(\bigotimes_{k=2}^d y^k \otimes (t-t')\right) \leq \varepsilon\beta_A^d\left(\bigotimes_{k=2}^d y^k\right).$$

By the triangle inequality,

$$\begin{aligned}
&\alpha_A^d\left(\bigotimes_{k=2}^d x^k\otimes t - \bigotimes_{k=2}^d y^k \otimes t'\right)\\
&\quad\leq \alpha_A^d\left(\left(\bigotimes_{k=2}^d x^k - \bigotimes_{k=2}^d y^k\right)\otimes t\right) + \alpha_A^d\left(\bigotimes_{k=2}^d y^k\otimes(t-t')\right)\\
&\quad\leq \alpha_A^d\left(\left(\bigotimes_{k=2}^d x^k - \bigotimes_{k=2}^d y^k\right)\otimes t\right) + \varepsilon\beta_A^d\left(\bigotimes_{k=2}^d x^k - \bigotimes_{k=2}^d y^k\right) + \varepsilon\beta_A^d\left(\bigotimes_{k=2}^d x^k\right)\\
&\quad\leq W_d^{\{(\mathbf{x},t)\}}(\alpha,4\varepsilon) + 2V_d^{\{(\mathbf{x},t)\}}(\beta,4\varepsilon) \leq W_d^{\{(\mathbf{x},t)\}}(\alpha,8\varepsilon) + V_d^{\{(\mathbf{x},t)\}}(\beta,8\varepsilon),
\end{aligned}$$

where in the third inequality we used (4.25). Thus,

$$\mu_{\varepsilon,T}^d\left(B\left((\mathbf{x},t),\rho_A,W_d^{\{(\mathbf{x},t)\}}(\alpha,8\varepsilon) + V_d^{\{(\mathbf{x},t)\}}(\beta,8\varepsilon)\right)\right) \geq \gamma_{(d-1)n,\varepsilon}(U) \geq C^{-d}\exp(-C(d)\varepsilon^{-2}),$$

where the last inequality follows from Lemma 4.3.2 applied to the norm $\alpha_A^d + \varepsilon\beta_A^d$.

□

Corollary 4.3.4. *For any* $\varepsilon,\delta > 0$ *and* $U \subset (B_2^n)^{d-1}\times T$ *we have*

$$N\left(U,\rho_A,W_d^U(\alpha,\varepsilon)+V_d^U(\beta,\varepsilon)\right)\le \exp(C(d)\varepsilon^{-2}) \tag{4.26}$$

and

$$\sqrt{\log N(U,\rho_A,\delta)}\le C(d)\left(\sum_{k=1}^{d-1}\left(\sum_{\substack{I\subset\{2,\dots,d\}\\|I|=k}}W_I^U(\alpha)\right)^{\frac{1}{k}}\delta^{-\frac{1}{k}}+\sum_{k=0}^{d-1}\left(\sum_{\substack{I\subset\{2,\dots,d\}\\|I|=k}}V_I^U(\beta)\right)^{\frac{1}{k+1}}\delta^{-\frac{1}{k+1}}\right). \tag{4.27}$$

Proof. It suffices to show (4.26), since it easily implies (4.27). Consider first $\varepsilon\le 8$. Obviously, $W_d^U(\alpha,\varepsilon)+V_d^U(\beta,\varepsilon)\ge \sup_{(\mathbf{x},t)\in U}(W_d^{\{(\mathbf{x},t)\}}(\alpha,\varepsilon)+V_d^{\{(\mathbf{x},t)\}}(\beta,\varepsilon))$. Therefore, by Lemma 4.3.3 (applied with $\varepsilon/16$) we have for any $(\mathbf{x},t)\in U$,

$$\mu_{\varepsilon,T}^d\left(B\left((\mathbf{x},t),\rho_A,W_d^U(\alpha,\varepsilon/2)+V_d^U(\beta,\varepsilon/2)\right)\right)\ge C^{-d}\exp\left(-C(d)\varepsilon^{-2}\right). \tag{4.28}$$

Suppose that there exist $(\mathbf{x}_1,t_1),\dots,(\mathbf{x}_N,t_N)\in U$ such that $\rho_A((\mathbf{x}_i,t_i),(\mathbf{x}_j,t_j))>W_d^U(\alpha,\varepsilon)+V_d^U(\beta,\varepsilon)\ge 2W_d^U(\alpha,\varepsilon/2)+2V_d^U(\beta,\varepsilon/2)$ for $i\ne j$. Then the sets

$$B\left((\mathbf{x}_i,t_i),\rho_\alpha,W_d^U(\alpha,\varepsilon/2)+V_d^U(\beta,\varepsilon/2)\right)\quad i=1,2\dots,N$$

are disjoint, so by (4.28), we obtain $N\le C^d\exp(C(d)\varepsilon^{-2})\le\exp(C(d)\varepsilon^{-2})$.

If $\varepsilon\ge 8$ then

$$\begin{aligned}W_d^U(\alpha,\varepsilon)+V_d^U(\beta,\varepsilon)&\ge 8\sup_{(x^2,\dots,x^d,t)\in U}\mathbb{E}\left|\sum_{\mathbf{i}}a_{\mathbf{i}}g_{i_1}\prod_{k=2}^d x_{i_k}^k t_{i_{d+1}}\right|\\&=\sqrt{\frac{128}{\pi}}\sup_{(x^2,\dots,x^d,t)\in U}\left(\sum_{i_1}\left(\sum_{i_2,\dots,i_{d+1}}a_{\mathbf{i}}\prod_{k=2}^d x_{i_k}^k t_{i_{d+1}}\right)^2\right)^{1/2}\ge\operatorname{diam}(U,\rho_A).\end{aligned}$$

So, $N(U,\rho_A,W_d^U(\alpha,\varepsilon)+V_d^U(\beta,\varepsilon))=1\le\exp(\varepsilon^{-2})$. □

Remark 4.3.5. The classical Dudley's bound on suprema of Gaussian processes (see Corollary 6.14) gives

$$\mathbb{E}\sup_{(x^2,\dots,x^d,t)\in(B_2^n)^{d-1}\times T}\sum_{\mathbf{i}}a_{\mathbf{i}}g_{i_1}\prod_{k=2}^d x_{i_k}^k t_{i_{d+1}}\le C\int_0^{\Delta}\sqrt{\log N((B_2^n)^{d-1}\times T,\rho_A,\delta)}d\delta,$$

where Δ is equal to diameter of the set $(B_2^n)^{d-1}\times T$ in the metric ρ_A. Unfortunately the entropy bound derived in Corollary 4.3.4 involve a non integrable term δ^{-1}. The remaining part of the proof of Theorem 4.3.1 is devoted to improve on Dudley's bound.

For $\mathbf{x},\mathbf{y}\in(\mathbb{R}^n)^{d-1}$ we define (note that $\hat{\alpha}_A$ is a norm on $(\mathbb{R}^n)^{d-1}=\mathbb{R}^{(d-1)n}$)

$$\hat{\alpha}_A((x^2,\dots,x^d)) := \sum_{j=2}^{d} \sum_{\substack{(\mathcal{P},\mathcal{P}')\in\mathcal{P}([d]\setminus\{j\})\\ |\mathcal{P}|=d-2}} \Big\| \sum_{i_j} a_{\mathbf{i}} x^j_{i_j} \Big\|_{\mathcal{P}',\mathcal{P}}$$
$$= \sum_{j=2}^{d} \sum_{\mathcal{P}\in\mathcal{P}([d]\setminus\{j\},d-2)} \Big\| \sum_{i_j} a_{\mathbf{i}} x^j_{i_j} \Big\|_{\emptyset,\mathcal{P}} + \sum_{j=2}^{d} \sum_{j\neq k=1}^{d} \sum_{\mathcal{P}\in\mathcal{P}([d]\setminus\{j,k\},d-2)} \Big\| \sum_{i_j} a_{\mathbf{i}} x^j_{i_j} \Big\|_{\{k\},\mathcal{P}}.$$

Proposition 4.3.6. *For any $d+1\geq 4$, $\varepsilon>0$ and $U\subset (B_2^n)^{d-1}\times T$,*

$$N\left(U,\rho_A,\sum_{k=0}^{d-2}\varepsilon^{d-k}s_k(A)+\varepsilon\sup_{(x^2,\dots,x^d,t)\in U}\hat{\alpha}_A((x^2,\dots,x^d))\right)\leq \exp(C(d)\varepsilon^{-2}).$$

Proof. Since $U\subset (B_2^n)^{d-1}\times T$, Jensen's inequality yields for $I\subset\{2,\dots,d\}$,

$$\begin{aligned}
W_I^U(\alpha) &= \sup_{(x^2,\dots,x^d,t)\in U} \mathbb{E}\alpha_A^d\left(\left(\bigotimes_{k=2}^{d}\left(x^k(1-\mathbf{1}_I(k))+G^k\mathbf{1}_I(k)\right)\right)\otimes t\right)\\
&\leq \sup_{(x^2,\dots,x^d,t)\in U}\sqrt{\mathbb{E}\sum_{i_1}\left(\sum_{i_2,\dots,i_{d+1}} a_{\mathbf{i}}\prod_{k\in I} g^k_{i_k}\prod_{k\in[d]\setminus(I\cup\{1\})} x^k_{i_k} t_{i_{d+1}}\right)^2}\\
&= \sup_{(x^2,\dots,x^d,t)\in U}\sqrt{\sum_{i_{I\cup\{1\}}}\left(\sum_{i_{[d+1]\setminus(I\cup\{1\})}} a_{\mathbf{i}}\prod_{k\in[d]\setminus(I\cup\{1\})} x^k_{i_k} t_{i_{d+1}}\right)^2}\\
&\leq \|A\|_{\emptyset,\{I\cup\{1\},\{k\}:k\in[d]\setminus(I\cup\{1\})\}} \leq s_{d-|I|}(A).
\end{aligned}\tag{4.29}$$

By estimating a little more accurately in the second inequality in (4.29) we obtain for $2\leq j\leq d$,

$$\begin{aligned}
W_{\{j\}}^U(\alpha) &\leq \sup_{(x^2,\dots,x^d,t)\in U}\sum_{\substack{2\leq l\leq d\\ l\neq j}}\sup_{(y^2,\dots,y^d)\in(B_2^n)^{d-1}}\sqrt{\sum_{i_1,i_j}\left(\sum_{i_{[d+1]\setminus\{1,j\}}} a_{\mathbf{i}} x^l_{i_l}\prod_{\substack{2\leq k\leq d\\ k\neq j,l}} y^k_{i_k} t_{i_{d+1}}\right)^2}\\
&\leq \sup_{(x^2,\dots,x^d,t)\in U}\sum_{l=2}^{d}\sum_{\substack{\mathcal{P}\in\mathcal{P}([d]\setminus\{l\})\\ |\mathcal{P}|=d-2}}\Big\|\sum_{i_l} a_{\mathbf{i}} x^l_{i_l}\Big\|_{\emptyset,\mathcal{P}}.
\end{aligned}\tag{4.30}$$

Observe that (4.30) is not true for $d+1=3$ (cf. Remark 4.3.7). The definition of V_I^U and the inclusion $U\subset(B_2^n)^{d-1}\times T$ yield

$$V_I^U(\beta)\leq \|A\|_{\{\{1\},\{i\}\ i\in I\},\{\{k\}:k\in[d]\setminus(I\cup\{1\})\}}\leq s_{d-|I|-1}(A)\quad\text{for } I\neq\emptyset \tag{4.31}$$

and

$$V_{\emptyset}^{U}(\beta) \leq \sup_{(x^2,\ldots,x^d,t)\in U} \sum_{l=2}^{d} \sup_{(y^2,\ldots,y^d)\in (B_2^n)^{d-1}} \mathbb{E}\sup_{t\in T}\sum a_{\mathbf{i}} g_{i_1} x_{i_l}^l \prod_{\substack{2\leq k\leq d\\ k\neq l}} y_{i_k}^k t_{i_{d+1}}$$

$$\leq \sup_{(x^2,\ldots,x^d,t)\in U} \sum_{l=2}^{d} \left\| \sum_{i_j} a_{\mathbf{i}} x_{i_l}^l \right\|_{\{1\},\{\{k\}:\ k\in[d]\setminus\{1,l\}\}}. \tag{4.32}$$

Inequalities (4.29)-(4.32) imply that

$$\begin{aligned}
&W_d^U(\alpha,\varepsilon)+V_d^U(\beta,\varepsilon)\\
&=\sum_{k=2}^{d-1}\varepsilon^k \sum_{I\subset\{2,\ldots,d\}:|I|=k} W_I^U(\alpha)+\sum_{k=1}^{d-1}\varepsilon^{k+1}\sum_{I\subset\{2,\ldots,d\}:|I|=k} V_I^U(\beta)+\varepsilon\left(\sum_{j=2}^{d} W_{\{j\}}^U+V_{\emptyset}^U(\beta)\right)\\
&\leq C(d)\left(\sum_{k=0}^{d-2}\varepsilon^{d-k}s_k(A)\right)+C(d)\varepsilon \sup_{(x^2,\ldots,x^d,t)\in U}\left(\sum_{l=2}^{d}\sum_{\substack{(\mathcal{P},\mathcal{P}')\in\mathcal{P}([d]\setminus\{l\})\\ |\mathcal{P}|\geq d-2}}\left\|\sum_{i_l} a_{\mathbf{i}} x_{i_l}^l\right\|_{\mathcal{P}',\mathcal{P}}\right)
\end{aligned}$$

So the assertion is a simple consequence of Corollary 4.3.4. □

Remark 4.3.7. Proposition 4.3.6 is not true for $d+1=3$. The problem arises in (4.30) - for $d=2$ there are too few indexes for existence of a dependency on $(x^2,\ldots,x^d)$, since two indexes appear in the first summation and one is reserved for dependency on t. This is the main reason why proofs for chaoses of order $d=2$ (cf. Chapter 3) have a different nature than for higher order chaoses.

4.4 Proof of Theorem 4.3.1

We will prove Theorem 4.3.1 by induction on $d+1$ so by the order of the matrix A. To this end we need to amplify the induction thesis. For $U\subset(\mathbb{R}^n)^d$ we define

$$F_A(U)=\mathbb{E}\sup_{(x^2,\ldots,x^{d+1})\in U}\sum_{i_1,\ldots,i_{d+1}} a_{i_1,\ldots,i_{d+1}} g_{i_1}\prod_{k=2}^{d+1} x_{i_k}^k.$$

Theorem 4.4.1. *For any $U\subset(B_2^n)^{d-1}\times T$ and any $p\geq 1$*

$$F_A(U)\leq C(d)\left(\sqrt{p}\Delta_A(U)+\sum_{k=0}^{d-1}p^{\frac{k+1-d}{2}}s_k(A)\right), \tag{4.33}$$

where

$$\Delta_A(U)=\sup_{(x^2,\ldots,x^d,t),(y^2,\ldots,y^d,t')\in U}\rho_A((x^2,\ldots,x^d,t),(y^2,\ldots,y^d,t'))=\operatorname{diam}(A,\rho_\alpha).$$

Clearly it is enough to prove Theorem 4.4.1 for finite sets U. Observe that

$$\Delta_A((B_2^n)^{d-1}\times T)\leq 2\|A\|_{\{\emptyset\},\{\{j\}:j\in[d]\}}=2s_d(A),$$

thus Theorem 4.4.1 implies Theorem 4.3.1. We will prove (4.33) by induction on $d+1$, but first we will show several consequences of the theorem. In the next three lemmas, we shall assume that Theorem 4.4.1 (and thus also Theorem 4.2.5) holds for all matrices of order smaller than $d+1$.

Lemma 4.4.2. *Let $p \geq 1$, $l \geq 0$ and $d+1 \geq 4$. Then*

$$N\left((B_2^n)^{d-1}, \hat{\rho}_A, 2^{-l}\sum_{k=0}^{d-1} p^{\frac{k+1-d}{2}} s_k(A)\right) \leq \exp(C(d)2^{2l}p),$$

where $\hat{\rho}_A$ is the related distance with the $\hat{\alpha}_A$ norm.

Proof. Note that $\hat{\alpha}_A$ is a norm on $(\mathbb{R}^n)^{d-1}$ and that

$$\mathbb{E}\hat{\alpha}_A\left(G^2,\dots,G^d\right) = \sum_{j=2}^{d} \sum_{\substack{(\mathcal{P},\mathcal{P}')\in\mathcal{P}([d]\setminus\{j\})\\ |\mathcal{P}|=d-2}} \mathbb{E}\left\|\sum_{i_j} a_{\mathbf{i}} g_{i_j}\right\|_{\mathcal{P}',\mathcal{P}}.$$

Up to a permutation of the indexes we have two possibilities

$$\left\|\sum_{i_j} a_{\mathbf{i}} g_{i_j}\right\|_{\mathcal{P}',\mathcal{P}} = \begin{cases} \left\|\sum_{i_j} a_{\mathbf{i}} g_{i_j}\right\|_{\emptyset,\{\{1,2\},\{\{l\}:\ 3\leq l\leq d,\ l\neq j\}\}} & \text{or} \\ \left\|\sum_{i_j} a_{\mathbf{i}} g_{i_j}\right\|_{\{\{1\}\},\{\{l\}:2\leq l\leq d,\ l\neq j\}} & . \end{cases} \tag{4.34}$$

First assume that $\left\|\sum_{i_j} a_{\mathbf{i}} g_{i_j}\right\|_{\mathcal{P}',\mathcal{P}} = \left\|\sum_{i_j} a_{\mathbf{i}} g_{i_j}\right\|_{\emptyset,\{\{1,2\},\{\{l\}:\ 3\leq l\leq d,\ l\neq j\}\}}$. In this case

$$\left\|\sum_{i_j} a_{\mathbf{i}} g_{i_j}\right\|_{\emptyset,\{\{1,2\},\{\{l\}:\ 3\leq l\leq d,\ l\neq j\}\}} = \left\|\sum_{i_1} b_{i_1,\dots,i_d} g_{i_1}\right\|_{\emptyset,\{\{2\},\dots,\{d-1\}\}}$$

for an appropriately chosen matrix $B = b_{i_1,\dots,i_d}$ (we treat a pair of indices $\{1,2\}$ as a single index and renumerate the indices in such a way that j,$\{1,2\}$ and $d+1$ would become 1,2 and d respectively).
Clearly,

$$\sum_{\substack{(\mathcal{P}',\mathcal{P})\in\mathcal{P}([d-1])\\ |\mathcal{P}|=k}} \|B\|_{\mathcal{P}',\mathcal{P}} = \sum_{\substack{(\mathcal{P}',\mathcal{P})\in\mathcal{C}\\ |\mathcal{P}|=k}} \|A\|_{\mathcal{P}',\mathcal{P}} \leq \sum_{\substack{(\mathcal{P}',\mathcal{P})\in\mathcal{P}([d])\\ |\mathcal{P}|=k}} \|A\|_{\mathcal{P}',\mathcal{P}} = s_k(A), \tag{4.35}$$

where $\mathcal{C} \subset \mathcal{P}([d])$ is the set of partitions which do not separate 1 and 2.
Thus, Theorem 4.3.1 applied to the matrix B of order d yields

$$\begin{aligned}\mathbb{E}\left\|\sum_{i_j} a_{\mathbf{i}} g_{i_j}\right\|_{\emptyset,\{\{1,2\},\{\{l\}:\ 3\leq l\leq d,\ l\neq j\}\}} &= \mathbb{E}\left\|\sum_{i_1} b_{i_1,\dots,i_d} g_{i_1}\right\|_{\emptyset,\{\{2\},\dots,\{d-1\}\}} \\ &\leq C(d) \sum_{(\mathcal{P}',\mathcal{P})\in\mathcal{P}([d-1])} p^{\frac{|\mathcal{P}|+2-d}{2}} \|B\|_{\mathcal{P}',\mathcal{P}} \leq C(d)\sum_{k=0}^{d-1} p^{\frac{k+2-d}{2}} s_k(A). \end{aligned} \tag{4.36}$$

Now assume that $\left\|\sum_{i_j} a_{\mathbf{i}} g_{i_j}\right\|_{\mathcal{P}',\mathcal{P}} = \left\|\sum_{i_j} a_{\mathbf{i}} g_{i_j}\right\|_{\{\{1\}\},\{\{l\}:2\leq l\leq d,\ l\neq j\}}$ and observe that

$$\begin{aligned}\mathbb{E}\left\|\sum_{i_j} a_{\mathbf{i}} g_{i_j}\right\|_{\{1\},\{\{l\}:2\le l\le d,\ l\neq j\}} &= \mathbb{E}^g \sup_{x^l\in B_2^n,\ 2\le l\le d,\ l\neq j} \mathbb{E}^{g'} \sup_{t\in T}\sum_{\mathbf{i}} a_{\mathbf{i}} g'_{i_1} g_{i_j} \prod_{2\le l\le d,\ l\neq j} x^l_{i_l} t_{i_{d+1}} \\ &= \mathbb{E} \sup_{x^l\in B_2^n,\ 2\le l\le d,\ l\neq j} \sup_{m\in\mathcal{M}} \sum_{\mathbf{i}} a_{\mathbf{i}} g_{i_j} \prod_{2\le l\le d,\ l\neq j} x^l_{i_l} m_{i_1,i_{d+1}} \\ &= \mathbb{E} \sup_{x^l\in B_2^n,\ 2\le l\le d-1} \sup_{m\in\widetilde{\mathcal{M}}} \sum_{i_1,\ldots,i_d} d_{i_1,\ldots,i_d} g_{i_1} \prod_{l=2}^{d-1} x^l_{i_l} m_{i_d},\end{aligned}$$

where $D=(d_{i_1,\ldots,i_d})_{i_1,\ldots,i_d}$ is an appropriately chosen matrix of order d, the set $\mathcal{M}\subset\mathbb{R}^n\otimes\mathbb{R}^m$ satisfies

$$\mathbb{E}\sup_{t\in T}\sum_{i,j} b_{i,j} g_i t_j = \sup_{m\in\mathcal{M}}\sum_{i,j} b_{i,j} m_{i,j} \text{ for any matrix } (b_{i,j})_{i\le n, j\le m},$$

and $\widetilde{\mathcal{M}}$ corresponds to $\mathcal{M}$ under a natural identification of $\mathbb{R}^n\otimes\mathbb{R}^m$ with $\mathbb{R}^{nm}$. Applying Theorem 4.3.1 to the matrix D of order d gives

$$\begin{aligned}\mathbb{E}\left\|\sum_{i_j} a_{\mathbf{i}} g_{i_j}\right\|_{\{1\},\{\{l\}:2\le l\le d,\ l\neq j\}} &= \mathbb{E} \sup_{x^l\in B_2^n,\ 2\le l\le d-1} \sup_{m\in\widetilde{\mathcal{M}}} \sum_{i_1,\ldots,i_d} d_{i_1,\ldots,i_d} g_{i_1} \prod_{l=2}^{d-1} x^l_{i_l} m_{i_d} \\ \le C(d) \sum_{(\mathcal{P}',\mathcal{P})\in\mathcal{P}([d-1])} p^{\frac{|\mathcal{P}|+2-d}{2}} \|D\|^{\widetilde{\mathcal{M}}}_{\mathcal{P}',\mathcal{P}} &\le C(d) \sum_{(\mathcal{P}',\mathcal{P})\in\mathcal{P}([d])} p^{\frac{|\mathcal{P}|+2-d}{2}} \|A\|_{\mathcal{P}',\mathcal{P}} \\ = C(d)\sum_{k=0}^{d-1} p^{\frac{k+2-d}{2}} s_k(A), \qquad (4.37)\end{aligned}$$

where $\|D\|^{\widetilde{\mathcal{M}}}_{\mathcal{P}',\mathcal{P}}$ is defined in the same manner as $\|A\|_{\mathcal{P}',\mathcal{P}}$ but the supremum is taken over the set $\widetilde{\mathcal{M}}$ instead of T. The second inequality in (4.37) can be justified analogously as (4.35). Combining (4.34),(4.36), (4.37) and the dual Sudakov ineuqality (cf. Theorem 6.8 and note that $(B_2^n)^{d-1}\subseteq\sqrt{d-1}B_2^{n(d-1)}$) we obtain

$$\begin{aligned}N\left((B_2^n)^{d-1},\hat\rho_A,\varepsilon\sum_{k=0}^{d-1} p^{\frac{k+2-d}{2}} s_k(A)\right) &\le N\Big((B_2^n)^{d-1},\hat\rho_A, C(d)^{-1}\varepsilon\mathbb{E}\hat\alpha_A(G^2,\ldots,G^n)\Big) \\ &\le \exp(C(d)\varepsilon^{-2}).\end{aligned}$$

It is now enough to choose $\varepsilon=(\sqrt{p}2^l)^{-1}$. $\square$

From now on for $U\subseteq(\mathbb{R}^n)^d$ we denote

$$\hat\alpha_A(U) = \sup_{(x^2,\ldots,x^d,t)\in U} \hat\alpha_A\Big((x^2,\ldots,x^d)\Big).$$

Lemma 4.4.3. *Suppose that* $d+1\ge 4$, $\boldsymbol{y}=(y^2,\ldots,y^d)\in(B_2^n)^{d-1}$ *and* $U\subset(B_2^n)^{d-1}\times T$. *Then for any* $p\ge1$ *and* $l\ge0$, *we can find a decomposition*

$$U=\bigcup_{j=1}^N U_j,\ N\le\exp(C(d)2^{2l}p)$$

such that for each $j\le N$,

$$F_A((\boldsymbol{y},0)+U_j) \le F_A(U_j) + C(d)\left(\hat{\alpha}_A(\boldsymbol{y}) + \hat{\alpha}_A(U) + 2^{-l}\sum_{k=0}^{d-2} p^{\frac{k+1-d}{2}} s_k(A)\right) \tag{4.38}$$

and

$$\Delta_A(U_j) \le 2^{-l}p^{-1/2}\hat{\alpha}_A(U) + 2^{-2l}\sum_{k=0}^{d-2} p^{\frac{k-d}{2}} s_k(A). \tag{4.39}$$

Proof. Fix $\mathbf{y} \in (B_2^n)^{d-1}$ and $U \subset (B_2^n)^{d-1} \times T$. For $\mathbf{x} = (x^2,\dots,x^d,t), \tilde{\mathbf{x}} = (\tilde{x}^2,\dots,\tilde{x}^d,t') \in (\mathbb{R}^n)^d$, $S \subset (\mathbb{R}^n)^d$ and $I \subset \{2,\dots,d\}$ we define

$$\rho_A^{\mathbf{y},I}(\mathbf{x},\tilde{\mathbf{x}}) := \sqrt{\sum_{i_1}\left(\sum_{i_2,\dots,i_{d+1}} a_{\mathbf{i}} \prod_{k\in I} y_{i_k}^k \left(t_{i_{d+1}} \prod_{\substack{2\le j\le d\\ j\notin I}} x_{i_j}^j - t'_{i_{d+1}} \prod_{\substack{2\le j\le d\\ j\notin I}} \tilde{x}_{i_j}^j\right)\right)^2},$$

$$\Delta_A^{\mathbf{y},I}(S) := \sup\left\{\rho_A^{\mathbf{y},I}(\mathbf{x},\tilde{\mathbf{x}}) : \mathbf{x},\tilde{\mathbf{x}} \in S\right\}$$

and

$$F_A^{\mathbf{y},I}(S) := \mathbb{E} \sup_{(x^2,\dots,x^d,t)\in S} \sum_{\mathbf{i}} a_{\mathbf{i}} g_{i_1} \prod_{k\in I} y_{i_k}^k \left(\prod_{\substack{2\le j\le d\\ j\notin I}} x_{i_j}^j\right) t_{i_{d+1}}.$$

If $I = \{2,\dots,d\}$ then for $S \subset (B_2^n)^{d-1} \times T$ we have

$$\begin{aligned} F_A^{\mathbf{y},\{2,\dots,d\}}(S) &\le \mathbb{E}\sup_{t\in T}\sum_{\mathbf{i}} a_{\mathbf{i}} g_{i_1} \prod_{k=2}^{d} y_{i_k}^k t_{i_{d+1}} \le \sup_{(x^2,\dots,x^{d-1})\in(B_2^n)^{d-2}} \mathbb{E}\sup_{t\in T}\sum_{\mathbf{i}} a_{\mathbf{i}} g_{i_1}\left(\prod_{j=2}^{d-1} x_{i_j}^j\right) y_{i_d}^d t_{i_{d+1}} \\ &= \left\|\sum_{i_d} a_{\mathbf{i}} y_{i_d}^d\right\|_{\{1\},\{\{k\}\,:\,k=2,\dots,d-1\}} \le \hat{\alpha}_A(\mathbf{y}). \end{aligned} \tag{4.40}$$

If $I \ne \emptyset, \{2,\dots,d\}$ then Theorem 4.4.1 applied to the matrix

$$A(\mathbf{y},I) := \left(\sum_{i_I} a_{\mathbf{i}} \prod_{k\in I} y_{i_k}^k\right)_{\mathbf{i}_{[d+1]\setminus I}}$$

of order $d-|I|+1 < d+1$ gives for any $S \subset (B_2^n)^{d-1} \times T$ and $q \ge 1$,

$$F_A^{\mathbf{y},I}(S) \le C(d-|I|)\left(q^{1/2}\Delta_A^{\mathbf{y},I}(S) + \sum_{k=0}^{d-|I|-1} q^{\frac{k+1-d+|I|}{2}} s_k(A(\mathbf{y},I))\right).$$

For any $2 \le k \le d$, $y^k \in B_2^n$ thus $s_k(A(\mathbf{y},I)) \le s_{k+|I|}(A)$ for $k < d-|I|-1$ and $s_{d-|I|-1}(A(\mathbf{y},I)) \le \hat{\alpha}_A(\mathbf{y})$. Hence,

$$F_A^{\mathbf{y},I}(S) \le C(d-|I|)\left(q^{1/2}\Delta_A^{\mathbf{y},I}(S) + \hat{\alpha}_A(\mathbf{y}) + \sum_{k=0}^{d-2} q^{\frac{k+1-d}{2}} s_k(A)\right). \tag{4.41}$$

By the triangle inequality,

$$F_A((y^2,\dots,y^d,0)+S)-F_A(S)\le \sum_{\emptyset\neq I\subset\{2,\dots,d\}} F_A^{\mathbf{y},I}(S).$$

Combining (4.40) and (4.41) we obtain for $S\subset (B_2^n)^{d-1}\times T$ and $q\ge 1$,

$$\begin{aligned}&F_A((y^2,\dots,y^d,0)+S)\\ &\qquad\le F_A(S)+C(d)\left(\hat{\alpha}_A(\mathbf{y})+\sum_{\emptyset\neq I\subsetneq\{2,\dots,d\}} q^{1/2}\Delta_A^{\mathbf{y},I}(S)+\sum_{k=0}^{d-2} q^{\frac{k+1-d}{2}}s_k(A)\right).\end{aligned}\tag{4.42}$$

Fix $I\neq\{2,\dots,d\}$, $|I|<d-2$ (we do not exclude $I=\emptyset$). Taking supremum over $\mathbf{y}\in(B_2^n)^{d-1}$ we conclude that

$$\sup_{(x^2,\dots,x^d,t)\in U}\hat{\alpha}_{A(\mathbf{y},I)}(\{x^k\ :\ k\in\{2,\dots,d\}\setminus I)\le \sup_{(x^2,\dots,x^d,t)\in U}\hat{\alpha}_A((x^2,\dots,x^d)).$$

Recall also that $s_k(A(\mathbf{y},I))\le s_{k+|I|}(A)$, thus we may apply $2^{d-1}-d$ times Proposition 4.3.6 with $\varepsilon=2^{-l}p^{-1/2}$ and find a decomposition $U=\bigcup_{j=1}^{N_1}U'_j$, $N_1\le\exp(C(d)2^{2l}p)$ such that for each j and $I\subset\{2,\dots,d\}$ with $|I|<d-2$,

$$\Delta_A^{\mathbf{y},I}(U'_j)\le 2^{-l}p^{-1/2}\hat{\alpha}_A(U)+2^{-2l}\sum_{k=0}^{d-2}p^{\frac{k-d}{2}}s_k(A).\tag{4.43}$$

If $|I|=d-2$ then the distance $\rho_A^{\mathbf{y},I}$ corresponds to a norm $\alpha^2_{A(\mathbf{y},I)}$ on $\mathbb{R}^{nm}$ given by

$$\alpha^2_{A(\mathbf{y},I)}(\mathbf{x})=\sqrt{\sum_{i_1}\left(\sum_{i_2,\dots,i_{d+1}} a_{\mathbf{i}}x_{i_{\{j,d+1\}}}\prod_{k\in I}y^k_{i_k}\right)^2},$$

where j is defined by the condition $\{1,j\}=[d]\setminus I$ (cf. (4.20) and (4.21)). Recall the definitions (4.22), (4.23) and (4.24) and note that (denoting by $\widetilde{U}$ the projection of U onto the j-th and $(d+1)$-th coordinate)

$$\begin{aligned}W_2^{\widetilde{U}}(\alpha^2_{A(\mathbf{y},I)}(\mathbf{x}),\varepsilon)&=\varepsilon\sup_{(x^2,\dots,x^d,t)\in U}\mathbb{E}\sqrt{\sum_{i_1}\left(\sum_{i_2,\dots,i_{d+1}}a_{\mathbf{i}}g_{i_j}t_{i_{d+1}}\prod_{k\in I}y^k_{i_k}\right)^2}\\ &\le\varepsilon\sup_{(x^2,\dots,x^d,t)\in U}\sqrt{\sum_{i_1,i_j}\left(\sum_{\mathbf{i}_I}a_{\mathbf{i}}\prod_{k\in I}y^k_{i_k}t_{i_{d+1}}\right)^2}\le\varepsilon\hat{\alpha}(\mathbf{y}).\end{aligned}\tag{4.44}$$

where we again used that $y^k\in B_2^n$, $U\subset(B_2^n)^{d-1}\times T$.

We also have

$$V_{\{j\}}^{\widetilde{U}}(\beta^2_{A(\mathbf{y},I)})=\mathbb{E}\sup_{t\in T}\sum_{\mathbf{i}}a_{\mathbf{i}}g^1_{i_1}g^2_{i_j}t_{i_k}\prod_{k\in I}y^k_{i_k}\le s_{d-2}(A)$$

and

$$V_{\emptyset}^{\widetilde{U}}(\beta^2_{A(\mathbf{y},I)}) = \sup_{(x^2,\ldots,x^d,t)\in U} \mathbb{E}\sup_{t'\in T}\sum_{\mathbf{i}} a_{\mathbf{i}} g_{i_1} x^j_{i_j} t'_{i_{d+1}} \prod_{k\in I} y^k_{i_k} \le \hat{\alpha}(\mathbf{y}).$$

Thus

$$V_{2}^{\widetilde{U}}(\beta^2_{A(\mathbf{y},I)},\varepsilon) \le \varepsilon\hat{\alpha}(\mathbf{y}) + \varepsilon^2 s_{d-2}(A).$$

Taking $\varepsilon = 2^{-l-1}p^{-1/2}$ and combining the above estimate with (4.44) and Corollary 4.3.4 (applied $d-1$ times) we obtain a partition $U = \bigcup_{j=1}^{N_2} U''_j$ with $N_2 \le \exp(C(d)2^{2l}p)$ and

$$\Delta_A^{\mathbf{y},I}(U''_j) \le 2^{-l}p^{-1/2}\hat{\alpha}_A(\mathbf{y}) + 2^{-2l}p^{-1}s_{d-2}(A) \tag{4.45}$$

for any $I \subset \{2,\ldots,d\}$ with $|I| = d-2$ and $j \le N_2$.

Intersecting this partition with the partition into the sets U'_i given by (4.43), we obtain a partition $U = \bigcup_{i=1}^{N} U_i$ with $N \le N_1N_2 \le \exp(C(d)2^{2l}p)$ and such that for every $i \le N$ there exist $j \le N_1$ and $l \le N_2$ such that $U_i \subset U'_j \cap U''_l$.

Inequality (4.38) follows by (4.42) with $q = 2^{2l}p$, (4.43) and (4.45). Observe that (4.39) follows by (4.43) for $I = \emptyset$. □

Lemma 4.4.4. *Suppose that U is a finite subset of $(B_2^n)^{d-1}\times T$, with $|U|\ge 2$ and $U-U\subset (B_2^n)^{d-1}\times(T-T)$. Then for any $p\ge 1$ there exist finite sets $U_i \subset (B_2^n)^{d-1}\times T$ and $(\boldsymbol{y}_i,t_i)\in U$, $i=1,\ldots,N$ such that*

(i) $2 \le N \le \exp(C(d)2^{2l}p)$,
(ii) $U = \bigcup_{i=1}^{N}((\boldsymbol{y}_i,0)+U_i)$, $\quad (U_i-U_i)\subset U-U$, $\quad |U_i| \le |U|-1$,
(iii) $\Delta_A(U_i) \le 2^{-2l}\sum_{k=0}^{d-1} p^{\frac{k-d}{2}} s_k(A)$,
(iv) $\hat{\alpha}_A(U_i) \le 2^{-l}\sum_{k=0}^{d-1} p^{\frac{k+1-d}{2}} s_k(A)$,
(v) $F_A((\boldsymbol{y}_i,0)+U_i) \le F_A(U_i) + C(d)\left(\hat{\alpha}_A(U) + 2^{-l}\sum_{k=0}^{d-1} p^{\frac{k+1-d}{2}} s_k(A)\right)$.

Proof. By Lemma 4.4.2 we get

$$(B_2^n)^{d-1} = \bigcup_{i=1}^{N_1} B_i,\ N_1 \le \exp(C(d)2^{2l}p),$$

where the diameter of the sets B_i in the norm $\hat{\alpha}$ satisfies

$$\operatorname{diam}(B_i,\hat{\alpha}_A) \le 2^{-l}\sum_{k=0}^{d-1} p^{\frac{k+1-d}{2}} s_k(A).$$

Let $U_i = U\cap(B_i\times T)$. Selecting arbitrary $(\mathbf{y}_i,t_i)\in U_i$ (we can assume that these sets are nonempty) and using Lemma 4.4.3 we decompose $U_i - (\mathbf{y}_i,0)$ into $\bigcup_{j=1}^{N_2} U_{ij}$ in such a way that $N_2 \le \exp(C(d)2^{2l}p)$,

$$\begin{aligned} F_A((\mathbf{y}_i,0)+U_{ij}) &\le F_A(U_{ij}) + C(d)\left(\hat{\alpha}_A(\mathbf{y}_i) + \hat{\alpha}_A(U_i-(\mathbf{y}_i,0)) + 2^{-l}\sum_{k=0}^{d-2} p^{\frac{k+1-d}{2}} s_k(A)\right) \\ &\le F_A(U_{ij}) + C(d)\left(\hat{\alpha}_A(\mathbf{y}_i) + \operatorname{diam}(B_i,\hat{\alpha}_A) + 2^{-l}\sum_{k=0}^{d-1} p^{\frac{k+1-d}{2}} s_k(A)\right) \\ &\le F_A(U_{ij}) + C(d)\left(\hat{\alpha}_A(U) + 2^{-l}\sum_{k=0}^{d-1} p^{\frac{k+1-d}{2}} s_k(A)\right) \end{aligned}$$

and

$$\varDelta_A(U_{ij}) \le 2^{-l}p^{-1/2}\hat{\alpha}_A(U_i-(\mathbf{y}_i,0))+2^{-2l}\sum_{k=0}^{d-2}p^{\frac{k-d}{2}}s_k(A) \le 2^{-2l}\sum_{k=0}^{d-1}p^{\frac{k-d}{2}}s_k(A).$$

We take the decomposition $U=\bigcup_{i,j}((\mathbf{y}_i,0)+U_{ij})$. We have $N=N_1N_2\le \exp(C(d)2^{2l}p)$. Without loss of generality we can assume $N\ge 2$ and $|U_{i,j}|\le |U|-1$. Obviously, $U_{ij}-U_{ij}\subset U_i-U_i\subset U-U$ and $\hat{\alpha}_A(U_{ij})\le \hat{\alpha}_A(U_i-(\mathbf{y}_i,0))\le 2^{-l}\sum_{k=0}^{d-1}p^{\frac{k+1-d}{2}}s_k(A)$. A relabeling of the obtained decomposition concludes the proof. □

Proof of Theorem 4.4.1. In the case of $d+1=3$ Theorem 4.4.1 is proved in Chapter 3 (see Remark 3.5.13). Assuming (4.33) to hold for all matrices of order $\{3,4,\dots,d\}$, we will prove it for matrices of order $d+1\ge 4$. Let $U\subset(\mathbb{R}^n)^d$ and let us put $\varDelta_0=\varDelta_A(U)$, $\hat{\varDelta}_0=\hat{\alpha}_A\left((B_2^n)^{d-1}\times T\right)\le C(d)s_{d-1}(A)$,

$$\varDelta_l := 2^{2-2l}\sum_{k=0}^{d-1}p^{\frac{k-d}{2}}s_k(A),\ \ \hat{\varDelta}_l := 2^{1-l}\sum_{k=0}^{d-1}p^{\frac{k+1-d}{2}}s_k(A) \text{ for } l\ge 1.$$

Suppose first that $U\subset(\frac{1}{2}(B_2^n)^{d-1})\times T$ and define

$$c_U(r,l) := \sup\Big\{F_A(S):\ \ S\subset (B_2^n)^{d-1}\times T,\ S-S\subset U-U,\ |S|\le r,\ \varDelta_A(S)\le \varDelta_l,\ \hat{\alpha}_A(S)\le \hat{\varDelta}_l\Big\}.$$

Note that any subset $S\subset U$ satisfies $\varDelta_A(S)\le\varDelta_0$ and $\hat{\alpha}_A(S)\le\hat{\varDelta}_0$, therefore,

$$c_U(r,0)\ge \sup\{F_A(S): S\subset U,\quad |S|\le r\}. \tag{4.46}$$

We will now show that for $r\ge 2$,

$$c_U(r,l)\le c_U(r-1,l+1)+C(d)\left(\hat{\varDelta}_l+2^l\sqrt{p}\varDelta_l+2^{-l}\sum_{k=0}^{d-1}p^{\frac{k+1-d}{2}}s_k(A)\right). \tag{4.47}$$

Indeed, let us take $S\subset(B_2^n)^{d-1}\times T$ as in the definition of $c_U(r,l)$. Then by Lemma 4.4.4 we may find a decomposition $S=\bigcup_{i=1}^N((\mathbf{y}_i,0)+S_i)$ satisfying $(i)-(v)$ with U, U_i replaced by S, S_i. Hence, by Lemma 6.10 we have

$$\begin{aligned} F_A(S) &\le C\sqrt{\log N}\varDelta_A(S)+\max_i F_A((\mathbf{y}_i,0)+S_i)\\ &\le C(d)\left(\hat{\alpha}_A(S)+2^l\sqrt{p}\varDelta_l+2^{-l}\sum_{k=0}^{d-1}p^{\frac{k+1-d}{2}}s_k(A)\right)+\max_i F_A(S_i). \end{aligned}\tag{4.48}$$

We have $\varDelta_A(S_i)\le\varDelta_{l+1}$, $\hat{\alpha}_A(S_i)\le\hat{\varDelta}_{l+1}$, $S_i-S_i\subset S-S\subset U-U$ and $|S_i|\le|S|-1\le r-1$, thus $\max_i F_A(S_i)\le c_U(r-1,l+1)$ and (4.48) yields (4.47). Since $c_U(1,l)=0$, (4.47) yields

$$c_U(r,0)\le C(d)\sum_{l=0}^{\infty}\left(\hat{\varDelta}_l+2^l\sqrt{p}\varDelta_l+2^{-l}\sum_{k=0}^{d-1}p^{\frac{k+1-d}{2}}s_k(A)\right).$$

For $U\subset(\frac{1}{2}(B_2^n)^{d-1})\times T$, we have by (4.46)

$$F_A(U) = \sup\{F_A(S) : S \subset T, |S| < \infty\} \le \sup_r c_U(r,0) \le C(d)\left(\sqrt{p}\Delta_A(U) + \sum_{k=0}^{d-1} p^{\frac{k+1-d}{2}} s_k(A)\right).$$

Finally, if $U \subset (B_2^n)^{d-1} \times T$, then $U' := \{(\mathbf{y}/2, t)\colon (\mathbf{y}, t) \in U\} \subset (\frac{1}{2}(B_2^n)^{d-1}) \times T$ and $\Delta_A(U') = 2^{1-d}\Delta_A(U)$, hence,

$$F_A(U) = 2^{d-1}F_A(U') \le C(d)\left(\sqrt{p}\Delta_A(U) + \sum_{k=0}^{d-1} p^{\frac{k+1-d}{2}} s_k(A)\right).$$

□

4.5 Proofs of main results

We return to the notation used Section 4.2. In particular $\mathbf{i} \in [n]^d$ in this section (instead of $[n]^d \times [m]$ as we had in the two previous sections).

4.5.1 Proofs of Theorems 4.2.1 and 4.2.4

Proof of Theorem 4.2.1. We start with the lower bound. Fix $J \subset [d]$, $\mathcal{P} \in \mathcal{P}([d] \setminus J)$ and observe that

$$\begin{aligned}\left\|\sum_{\mathbf{i}} a_{\mathbf{i}} \prod_{k=1}^d g_{i_k}^k\right\|_p &\ge \left(\mathbb{E}^{(G^j):j\in J} \sup_{\substack{\varphi\in F^* \\ \|\varphi\|\le 1}} \mathbb{E}^{(G^j):j\in[d]\setminus J}\left|\varphi\left(\sum_{\mathbf{i}} a_{\mathbf{i}} \prod_{k=1}^d g_{i_k}^k\right)\right|^p\right)^{1/p} \\ &\ge C^{-1}(d)\left(\mathbb{E}^{(g^j):j\in J} p^{\frac{p|\mathcal{P}|}{2}} \left\|\left\|\left(\sum_{i_J} a_{\mathbf{i}} \prod_{j\in J} g_{i_j}^j\right)_{i_{[d]\setminus J}}\right\|\right\|_{\mathcal{P}}^p\right)^{1/p} \\ &\ge C^{-1}(d)p^{\frac{|\mathcal{P}|}{2}}\mathbb{E}\left\|\left\|\left(\sum_{i_J} a_{\mathbf{i}} \prod_{j\in J} g_{i_j}^j\right)_{i_{[d]\setminus J}}\right\|\right\|_{\mathcal{P}} = C^{-1}(d)p^{\frac{|\mathcal{P}|}{2}} \|\!|A|\!\|_{\mathcal{P}},\end{aligned}$$

where F^* is the dual space and in the second inequality we used Theorem 6.27.

The upper bound will be proved by an induction on d. For $d+1 = 3$ it is showed in Theorem 3.1.4 (see Theorem 6.4). Suppose that $d+1 \ge 4$ and the estimate holds for matrices of order $\{2,3,\ldots,d\}$. By the induction assumption, we have

$$\left\|\sum_{\mathbf{i}} a_{\mathbf{i}} \prod_{k=1}^d g_{i_k}^k\right\|_p \le C(d) \sum_{(\mathcal{P},\mathcal{P}')\in\mathcal{P}([d-1])} p^{\frac{|\mathcal{P}|}{2}} \left\|\left\|\sum_{i_d} a_{\mathbf{i}} g_{i_d}\right\|_{\mathcal{P}',\mathcal{P}}\right\|_p. \tag{4.49}$$

Since $\|\cdot\|_{\mathcal{P}',\mathcal{P}}$ is a norm Lemma 6.20 yields

$$\left\| \left\| \sum_{i_d} a_{\mathbf{i}} g_{i_d} \right\|_{\mathcal{P}',\mathcal{P}} \right\|_p \le C\mathbb{E} \left\| \sum_{i_d} a_{\mathbf{i}} g_{i_d} \right\|_{\mathcal{P}',\mathcal{P}} + C\sqrt{p}\, \|A\|_{\mathcal{P}',\mathcal{P}\cup\{d\}}. \tag{4.50}$$

Choose $\mathcal{P} = (I_1 \ldots, I_k), \mathcal{P}' = (J_1, \ldots, J_m)$ and denote $J = \bigcup \mathcal{P}'$. By the definition of $\|A\|_{\mathcal{P}',\mathcal{P}}$ we have

$$\begin{aligned} \left\| \sum_{i_d} a_{\mathbf{i}} g_{i_d} \right\|_{\mathcal{P}',\mathcal{P}} &= \sup \left\{ \mathbb{E}^{(G^1,\ldots,G^m)} \left\| \sum_{\mathbf{i}} a_{\mathbf{i}} x^1_{i_{I_1}} \cdot \ldots \cdot x^k_{i_{I_k}} \prod_{l=1}^m g^l_{i_{J_l}} g^d_{i_d} \right\| \;\Big|\; \forall_{j=1\ldots,k} \sum_{i_{I_j}} \left(x^{(j)}_{i_{I_j}} \right)^2 = 1 \right\} \\ &= \sup \left\{ \left|\!\left|\!\left| \left(\sum_{i_{[d]\setminus J}} a_{\mathbf{i}} x^1_{i_{I_1}} \cdot \ldots \cdot x^k_{i_{I_k}} g^d_{i_d} \right)_{i_J} \right|\!\right|\!\right| \;\Big|\; \forall_{j=1\ldots,k} \sum_{i_{I_j}} \left(x^{(j)}_{i_{I_j}} \right)^2 = 1 \right\}, \end{aligned} \tag{4.51}$$

where $G^l = (g_{i_{J_l}})_{i_{J_l}}$ and $|||\cdot|||$ is a norm on $F^{n^{|J|}}$ given by

$$|||(a_{i_J})_{i_J}||| = \mathbb{E} \left\| \sum_{i_J} a_{i_J} \prod_{l=1}^m g^l_{i_{J_l}} \right\|.$$

Theorem 4.2.5 implies that

$$\begin{aligned} &\mathbb{E} \sup \left\{ \left|\!\left|\!\left| \left(\sum_{i_{[d]\setminus J}} a_{\mathbf{i}} x^1_{i_{I_1}} \cdot \ldots \cdot x^k_{i_{I_k}} g^d_{i_d} \right)_{i_J} \right|\!\right|\!\right| \;\Big|\; \forall_{j=1\ldots,k} \sum_{i_{I_j}} \left(x^{(j)}_{i_{I_j}} \right)^2 = 1 \right\} \\ &\le C(k) \sum_{(\mathcal{R}',\mathcal{R}) \in \mathcal{P}([d]\setminus J)} p^{\frac{|\mathcal{R}|-k}{2}} |||A|||_{\mathcal{R}',\mathcal{R}} = C(k) \sum_{(\mathcal{R}',\mathcal{R}) \in \mathcal{P}([d]\setminus J)} p^{\frac{|\mathcal{R}|-k}{2}} \|A\|_{\mathcal{R}'\cup\mathcal{P}',\mathcal{R}} \\ &\le C(k) \sum_{(\mathcal{R}',\mathcal{R}) \in \mathcal{P}([d])} p^{\frac{|\mathcal{R}|-k}{2}} \|A\|_{\mathcal{R}',\mathcal{R}} \end{aligned}$$

where $|||A|||_{\mathcal{R}',\mathcal{R}}$ is defined as $\|A\|_{\mathcal{R}',\mathcal{R}}$ but under the expectation occurs the norm $|||\cdot|||$.

The above and (4.51) yield

$$\mathbb{E} \left\| \sum_{i_d} a_{\mathbf{i}} g_{i_d} \right\|_{\mathcal{P}',\mathcal{P}} \le C(k) \sum_{(\mathcal{R}',\mathcal{R}) \in \mathcal{P}([d])} p^{\frac{|\mathcal{R}|-k}{2}} \|A\|_{\mathcal{R}',\mathcal{R}}. \tag{4.52}$$

Since $|\mathcal{P}| = k$ the proof follows from (4.49),(4.50) and (4.52)

□

Proof of Theorem 4.2.4. Let $S = \left\| \sum a_{\mathbf{i}} g^1_{i_1} \cdots g^d_{i_d} \right\|$. Chebyshev's inequality and Theorem 4.2.1 yield

$$\mathbb{P} \left(S \ge C(d) \sum_{(\mathcal{P},\mathcal{P}') \in \mathcal{P}([d])} p^{|\mathcal{P}|/2} \|A\|_{\mathcal{P}',\mathcal{P}} \right) \le e^{-p}. \tag{4.53}$$

It is enough to substitute $t = C(d) \sum_{\substack{(\mathcal{P},\mathcal{P}') \in \mathcal{P}([d]) \\ |\mathcal{P}| \ge 1}} p^{|\mathcal{P}|/2} \|A\|_{\mathcal{P}',\mathcal{P}}$ and observe that

$$p \geq \frac{1}{C(d)} \min_{\substack{(\mathcal{P},\mathcal{P}')\in\mathcal{P}([d])\\ |\mathcal{P}|>0}} \left(\frac{t}{\|A\|_{\mathcal{P}',\mathcal{P}}}\right)^{2/|\mathcal{P}|}.$$

On the other hand by the Paley-Zygmund inequality (cf. Corollary 6.1) we get for $p \geq 2$,

$$\mathbb{P}\left(S \geq C^{-1}(d) \sum_{J\in[d]} \sum_{\mathcal{P}\in\mathcal{P}(J)} p^{|\mathcal{P}|/2} \|\!|A|\!\|_{\mathcal{P}}\right) \geq \mathbb{P}\left(S^p \geq \frac{1}{2^p}\mathbb{E}S^p\right) \geq \left(1-\frac{1}{2^p}\right)^2 \frac{(\mathbb{E}S^p)^2}{\mathbb{E}S^{2p}} \geq e^{-C(d)p},$$

where in the last inequality we used Theorem 6.3. The inequality follows by the same argument as for the upper bound. □

4.5.2 Proof of Theorem 4.2.9

Theorem 4.2.9 will be a corollary to our main results on decoupled homogeneous chaos and the following general statement.

Proposition 4.5.1. *Let F be a Banach space and let $f\colon \mathbb{R}^n \to F$ be a polynomial of degree D. Then for $p \geq 1$,*

$$\|f(G) - \mathbb{E}F(G)\|_p \sim_D \sum_{d=1}^{D} \Big\| \sum_{i_1,\ldots,i_d=1}^{n} a^d_{i_1,\ldots,i_d} g^1_{i_1} \cdots g^d_{i_d} \Big\|_p,$$

where the d-indexed F-valued matrices $A_d = (a_{i_1,\ldots,i_d})_{i_1,\ldots,i_d \leq n}$ are defined as $A_d = \mathbb{E}\nabla^k f(G)$.

Indeed, using the above proposition reduces (4.10) of Theorem 4.2.9 to the lower estimate given in Theorem 4.2.1, while (4.12) is reduced to Corollary 4.2.8. The tail bounds (4.11) and (4.13) can be then obtained by Chebyshev's and Paley-Zygmund inequalities as in the proof of Theorem 4.2.4.

The overall strategy of the proof of Proposition 4.5.1 is similar to the one used in [2] to obtain the real valued case of Theorem 4.2.9. It relies on a reduction of inequalities for general polynomials of degree D to estimates for decoupled chaoses of degree $d = 1, \ldots, D$. To this end we will approximate general polynomials by tetrahedral ones and split the latter into homogeneous parts of different degrees, which via decoupling inequalities are comparable with their decoupled counterparts. The splitting may at first appear crude but it turns out that up to constants depending on D one can in fact invert the triangle inequality, which is formalized in the following result due to Kwapień (see [14, Lemma 2]).

Theorem 4.5.2. *If $X = (X_1,\ldots,X_n)$ where X_i are independent symmetric random variables, Q is a multivariate tetrahedral polynomial of degree D with coefficients in a Banach space E and Q_d is its homogeneous part of degree d, then for any symmetric convex function $\Phi\colon E \to \mathbb{R}_+$ and any $d \in \{0,1,\ldots,D\}$,*

$$\mathbb{E}\Phi(Q_d(X)) \leq \mathbb{E}\Phi(C_d Q(X)).$$

It will be convenient to have the polynomial f represented as a combination of multivariate Hermite polynomials:

$$f(x_1,\ldots,x_n) = \sum_{d=0}^{D} \sum_{\mathbf{d}\in\Delta^n_d} a_{\mathbf{d}} h_{d_1}(x_1)\cdots h_{d_n}(x_n), \tag{4.54}$$

where

$$\Delta^n_d = \{\mathbf{d} = (d_1,\ldots,d_n) \colon \forall_{k\in[n]}\ d_k \geq 0 \text{ and } d_1 + \cdots + d_n = d\}$$

and $h_m(x) = (-1)^m e^{x^2/2} \frac{d^m}{dx^m} e^{-x^2/2}$ is the m-th Hermite polynomial. In what follows, we will use the following notation. For a set I, by $I^{\underline{k}}$ we will denote the set of all one-to-one sequences of length k with values in I. For an F-valued d-indexed matrix $A = (a_{i_1,\ldots,i_d})_{i_1,\ldots,i_d \le n}$ and $x \in \mathbb{R}^{n^d}$ we will denote

$$\langle A, x\rangle = \sum_{i_1,\ldots,i_d} a_{i_1,\ldots,i_d} x_{i_1,\ldots,i_d}.$$

Let $(W_t)_{t\in[0,1]}$ be a standard Brownian motion. Consider standard Gaussian random variables $g = W_1$ and, for any positive integer N,

$$g_{j,N} = \sqrt{N}(W_{\frac{j}{N}} - W_{\frac{j-1}{N}}), \quad j = 1,\ldots,N.$$

For any $d \ge 0$, we have the following representation of $h_d(g) = h_d(W_1)$ as a multiple stochastic integral (see [12, Example 7.12 and Theorem 3.21]),

$$h_d(g) = d! \int_0^1 \int_0^{t_d} \cdots \int_0^{t_2} dW_{t_1} \cdots dW_{t_{d-1}} dW_{t_d}.$$

Approximating the multiple stochastic integral leads to

$$h_d(g) = d! \lim_{N\to\infty} N^{-d/2} \sum_{1\le j_1<\cdots<j_d\le N} g_{j_1,N}\cdots g_{j_d,N} = \lim_{N\to\infty} N^{-d/2} \sum_{j\in[N]^{\underline{d}}} g_{j_1,N}\cdots g_{j_d,N}, \quad (4.55)$$

where the limit is in $L^2(\Omega)$ (see [12, Theorem 7.3. and formula (7.9)]) and actually the convergence holds in any L^p (see [12, Theorem 3.50]).

Now, consider n independent copies $(W_t^{(i)})_{t\in[0,1]}$ of the Brownian motion ($i = 1,\ldots,n$) together with the corresponding Gaussian random variables: $g^{(i)} = W_1^{(i)}$ and, for $N \ge 1$,

$$g_{j,N}^{(i)} = \sqrt{N}(W^{(i)}_{\frac{j}{N}} - W^{(i)}_{\frac{j-1}{N}}), \quad j = 1,\ldots,N.$$

Let also

$$G^{(n,N)} = (g^{(1)}_{1,N},\ldots,g^{(1)}_{N,N},\ g^{(2)}_{1,N},\ldots,g^{(2)}_{N,N},\ \ldots,\ g^{(n)}_{1,N},\ldots,g^{(n)}_{N,N}) = (g^{(i)}_{j,N})_{(i,j)\in[n]\times[N]}$$

be a Gaussian vector with $n\times N$ coordinates. We identify here the set $[nN]$ with $[n]\times[N]$ via the bijection $(i,j) \leftrightarrow (i-1)N + j$. We will also identify the sets $([n]\times[N])^d$ and $[n]^d \times [N]^d$ in a natural way. For $d \ge 0$ and $\mathbf{d} \in \Delta_d^n$, let

$$I_{\mathbf{d}} = \{ i \in [n]^d \colon \forall_{l\in[n]}\ \#i^{-1}(\{l\}) = d_l \},$$

and define a d-indexed matrix $B_{\mathbf{d}}^{(N)}$ of n^d blocks each of size N^d as follows: for $i \in [n]^d$ and $j \in [N]^d$,

$$\big(B_{\mathbf{d}}^{(N)}\big)_{(i,j)} = \begin{cases} \frac{d_1!\cdots d_n!}{d!} N^{-d/2} & \text{if } i \in I_{\mathbf{d}} \text{ and } (i,j) := \big((i_1,j_1),\ldots,(i_d,j_d)\big) \in ([n]\times[N])^{\underline{d}}, \\ 0 & \text{otherwise.} \end{cases}$$

Proof of Proposition 4.5.1. Assume that f is of the form (4.54), By [2, Lemma 4.3], for any $p > 0$,

$$\big\langle B_{\mathbf{d}}^{(N)}, (G^{(n,N)})^{\otimes d}\big\rangle \xrightarrow{N\to\infty} h_{d_1}(g^{(1)})\cdots h_{d_n}(g^{(n)}) \quad \text{in } L^p(\Omega),$$

which together with the triangle inequality implies that

$$\lim_{N\to\infty}\Big\|\sum_{d=1}^{D}\Big\langle\sum_{\mathbf{d}\in\Delta_d^n}a_{\mathbf{d}}B_{\mathbf{d}}^{(N)},\big(G^{(n,N)}\big)^{\otimes d}\Big\rangle\Big\|_p=\big\|f(G)-\mathbb{E}f(G)\big\|_p$$

for any $p>0$, where $G=(g^{(1)},\dots,g^{(n)})$ and we interpret multiplication of an element of F and a real valued d indexed matrix in a natural way. Thus, by Theorem 4.5.2 and the triangle inequality we obtain

$$\begin{aligned}C_D^{-1}\lim_{N\to\infty}\sum_{d=1}^{D}\Big\|\Big\langle\sum_{\mathbf{d}\in\Delta_d^n}a_{\mathbf{d}}B_{\mathbf{d}}^{(N)},\big(G^{(n,N)}\big)^{\otimes d}\Big\rangle\Big\|_p&\\ &\le\|f(G)-\mathbb{E}f(G)\|_p\\ &\le C_D\lim_{N\to\infty}\sum_{d=1}^{D}\Big\|\Big\langle\sum_{\mathbf{d}\in\Delta_d^n}a_{\mathbf{d}}B_{\mathbf{d}}^{(N)},\big(G^{(n,N)}\big)^{\otimes d}\Big\rangle\Big\|_p.\end{aligned}$$

Denote by $G^{(n,N,1)},\dots,G^{(n,N,D)}$ independent copies of $G^{(n,N)}$. By decoupling inequalities we have

$$\Big\|\Big\langle\sum_{\mathbf{d}\in\Delta_d^n}a_{\mathbf{d}}B_{\mathbf{d}}^{(N)},\big(G^{(n,N)}\big)^{\otimes d}\Big\rangle\Big\|_p\sim_d\Big\|\Big\langle\sum_{\mathbf{d}\in\Delta_d^n}a_{\mathbf{d}}B_{\mathbf{d}}^{(N)},G^{(n,N,1)}\otimes\cdots\otimes G^{(n,N,d)}\Big\rangle\Big\|_p \tag{4.56}$$

(recall that the matrices $B_{\mathbf{d}}^{(N)}$ have zeros on generalized diagonals and so do their linear combinations).

To finish the proof it is therefore enough to show that for any $d\le D$,

$$\lim_{N\to\infty}\Big\|\Big\langle\sum_{\mathbf{d}\in\Delta_d^n}a_{\mathbf{d}}B_{\mathbf{d}}^{(N)},G^{(n,N,1)}\otimes\cdots\otimes G^{(n,N,d)}\Big\rangle\Big\|_p=\frac{1}{d!}\|\langle A_d,G_1\otimes\cdots\otimes G_d\rangle\|_p \tag{4.57}$$

where $G_1,\dots,G_D$ are independent copies of G.

Fix $d\ge 1$. For any $\mathbf{d}\in\Delta_d^n$ define a symmetric d-indexed matrix $(b_{\mathbf{d}})_{i\in[n]^d}$ as

$$(b_{\mathbf{d}})_i=\begin{cases}\frac{d_1!\cdots d_n!}{d!} & \text{if } i\in I_{\mathbf{d}},\\ 0 & \text{otherwise.}\end{cases}$$

and a symmetric d-indexed matrix $(\tilde{B}_{\mathbf{d}}^{(N)})_{(i,j)\in([n]\times[N])^d}$ as

$$(\tilde{B}_{\mathbf{d}}^{(N)})_{(i,j)}=N^{-d/2}(b_{\mathbf{d}})_i\quad\text{for all } i\in[n]^d \text{ and } j\in[N]^d.$$

Using the convolution properties of Gaussian distributions one easily obtains

$$\Big\|\Big\langle\sum_{\mathbf{d}\in\Delta_d^n}a_{\mathbf{d}}\tilde{B}_{\mathbf{d}}^{(N)},G^{(n,N,1)}\otimes\cdots\otimes G^{(n,N,d)}\rangle\Big\|_p=\Big\|\Big\langle\sum_{\mathbf{d}\in\Delta_d^n}a_{\mathbf{d}}(b_{\mathbf{d}})_{i\in[n]^d},G_1\otimes\cdots\otimes G_d\Big\rangle\Big\|_p \tag{4.58}$$

On the other hand, for any $\mathbf{d}\in\Delta_d^n$, the matrices $\tilde{B}_{\mathbf{d}}^{(N)}$ and $B_{\mathbf{d}}^{(N)}$ differ at no more than $|I_{\mathbf{d}}|\cdot|([N]^d\setminus[N]^{\underline{d}})|$ entries. Thus

$$\Big\| a_{\mathbf{d}} \Big\langle \tilde{B}^{(N)}_{\mathbf{d}} - B^{(N)}_{\mathbf{d}}, G^{(n,N,1)} \otimes \cdots \otimes G^{(n,N,d)} \Big\rangle \Big\|_p \\ \leq p^{\frac{d}{2}} \|a_{\mathbf{d}}\| \cdot \Big\| \Big\langle \tilde{B}^{(N)}_{\mathbf{d}} - B^{(N)}_{\mathbf{d}}, G^{(n,N,1)} \otimes \cdots \otimes G^{(n,N,d)} \Big\rangle \Big\|_2 \\ = p^{\frac{d}{2}} \|a_{\mathbf{d}}\| \cdot \sqrt{|I_{\mathbf{d}}| \Big(\frac{d_1! \cdots d_n!}{d!} \Big)^2 N^{-d} \Big(N^d - \frac{N!}{(N-d)!} \Big)} \longrightarrow 0$$

as $N \to \infty$, where in the inequality we used Theorem 6.3.

Together with the triangle inequality and (4.58) this gives

$$\lim_{N\to\infty} \Big\| \Big\langle \sum_{\mathbf{d}\in\Delta^n_d} a_{\mathbf{d}} B^{(N)}_{\mathbf{d}}, G^{(n,N,1)} \otimes \cdots \otimes G^{(n,N,d)} \Big\rangle \Big\|_p = \Big\| \Big\langle \sum_{\mathbf{d}\in\Delta^n_d} a_{\mathbf{d}} (b_{\mathbf{d}})_{i\in[n]^d}, G_1 \otimes \cdots \otimes G_d \Big\rangle \Big\|_p. \tag{4.59}$$

Finally, we have

$$\mathbb{E}\nabla^d f(G) = d! \sum_{\mathbf{d}\in\Delta^n_d} a_{\mathbf{d}} (b_{\mathbf{d}})_{i\in[n]^d}. \tag{4.60}$$

Indeed, using the identity on Hermite polynomials, $\frac{d}{dx} h_k(x) = k h_{k-1}(x)$ $(k \geq 1)$, we obtain $\mathbb{E} \frac{d^l}{dx^l} h_k(g) = k! \mathbf{1}_{k=l}$ for $k, l \geq 0$, and thus, for any $d, l \leq D$ and $\mathbf{d} \in \Delta^n_l$,

$$\Big(\mathbb{E}\nabla^d h_{d_1}(g^{(1)}) \cdots h_{d_n}(g^{(n)}) \Big)_i = d! (b_{\mathbf{d}})_i \mathbf{1}_{d=l} \quad \text{for each } i \in [n]^d.$$

Now (4.60) follows by linearity. Combining it with (4.59) yields (4.57) and ends the proof. □

Chapter 5
Moments estimates for some types of chaos in the Banach spaces

The purpose of this chapter is to investigate moments of random quadratic forms (chaoses of order two) $S = \sum_{i \neq j} a_{ij} X_i X_j$ under the assumption that for any $X_1, X_2, \ldots$ are independent symmetric r.v's with LCT and that $(a_{ij})_{ij}$ is a matrix with values in a Banach space $(F, \|\cdot\|)$. By standard arguments (discussed in the introduction and previous chapters) one may deduce from moments estimates, bounds on tails $\mathbb{P}(\|S\| \geq t),\ t > 0$.

In the sequel we will only consider decoupled chaoses $S' = \sum_{ij} a_{ij} X_i Y_j$, where $X_1, X_2, \ldots$, $Y_1, Y_2, \ldots$ are symmetric independent r.v's with LCT - under some natural assumptions, moments and tails of S, S' (in the particular case when $Y_1, Y_2, \ldots$ is an independent copy of $X_1, X_2, \ldots$) are comparable with numerical constants (see Theorems 6.5, 6.6).

Our main result, Theorem 5.1.3 presents two-sided bound of $\|S'\|_p$ in the case when $(a_{ij})_{ij}$ is a matrix with values in L_q space, under additional assumption that r.v's $Y_1, Y_2, \ldots$ are subgaussian. It is an attempt to generalize Latała's result [17] (he studied moments of S' in the case $F = \mathbb{R}$). Theorem 5.1.3 extends also moments bounds for Gaussian chaoses with values in L_q spaces presented in the Chapter 3. We suspect that our estimates hold in L_q spaces without the additional subgaussanity assumption but we are able to show it only in the case $q = 2$ (cf. Theorem 5.1.6).

One may also ask what are two-sided bounds on moments of S' in arbitrary Banach space, or more precisely is it possible to reverse the inequality (5.9). We suspect a positive answer here, however we think that this is a difficult problem. However, we provide upper bounds for moments of $\|S'\|_p$ in an arbitrary Banach space (Theorem 5.1.5 below). The usefulness of this inequality lies in the fact that it can be reversed in a certain class of a Banach spaces (including L_q spaces).

We use ideas introduced in [1] and Chapter 3. First we generalize Proposition 3.5.12 (estimation of suprema of the Gaussian processes) to the more general sets. Then following Chapter 3 we try to decompose a "big" set $U \times T \subset (B_2^n + \sqrt{p} B_1^n) \times \mathbb{R}^m$ into a sum of "not to many" sets T_l which are "small". By small we mean that $\mathbb{E} \sup_{t \in T_l} Y_t$ is small, where $(Y_t)_{t \in T_l}$ is a stochastic process based on symmetric, independent r.v's with LCT. We were unable to find an equivalent of [1, Corollary 7.3] (what would give such decomposition), but we noticed that under the additional subgaussanity assumption one can use a different decomposition.

The chapter is organized as follows. In Section 5.1 we present the notation and formulate the main results. In Section 5.2 we reduce the problem of bounding moments to the problem of bounding expectation of suprema of a stochastic process, formulate the problems in an equivalent way and we prove the moments bounds in the Hilbert space. In Section 5.3 we prove crucial bounds on expectations of suprema of a Gaussian process. In section 5.4 we present proofs of the main results.

5.1 Notation and main results

Let $(X_i)_i, (Y_j)_j$ be independent, symmetric random variables with LCT. We define the functions

$$N_i^X(t) = -\ln \mathbb{P}(|X_i| \geq |t|) \in [0,\infty],\ M_j^Y(t) = -\ln \mathbb{P}(|Y_j| \geq |t|) \in [0,\infty]$$

(observe that N_i^X, M_j^Y are convex). We assume that r.v's are normalized in such a way that

$$\inf\left\{t \geq 0,\ N_i^X(t) \geq 1\right\} = \inf\left\{t \geq 0,\ M_i^Y(t) \geq 1\right\} = 1. \tag{5.1}$$

We set

$$\hat{N}_i^X(t) = \begin{cases} t^2 & \text{for } |t| \leq 1 \\ N_i^X(t) & \text{for } |t| > 1 \end{cases}, \quad \hat{M}_j^Y(t) = \begin{cases} t^2 & \text{for } |t| \leq 1 \\ M_i^Y(t) & \text{for } |t| > 1 \end{cases}.$$

Observe that convexity of N_i^X, M_j^Y and the normalization condition (5.1) imply that

$$\hat{N}_i^X(t) = N_i^X(t) \geq t, \quad \hat{M}_j^Y(t) = M_i^Y(t) \geq t \quad \text{for } t \geq 1, \tag{5.2}$$

$$1/e \leq \mathbb{E}X_i^2, \mathbb{E}Y_j^2 \leq 1 + 4/e \leq 3, \quad \mathbb{E}X_i^4, \mathbb{E}Y_j^4 \leq 1 + 64/e. \tag{5.3}$$

The first formula is clear, to prove the second it is enough to observe that

$$1/e \leq \mathbb{P}(|X_i^2| \geq 1) \leq \mathbb{E}X_i^2 \leq 1 + \int_1^\infty 2xe^{-x}dx \leq 1 + 4/e \leq 3$$

(we prove the bounds for $\mathbb{E}X_i^4$ analogously).

Let $(a_{ij})_{ij}$ be $\mathbb{R}$ valued matrix and $(a_i)_i \in \mathbb{R}^n$. The following three norms will play crucial role in this chapter:

$$\|(a_{ij})_{ij}\|_{X,Y,p} = \sup\left\{\sum_{ij} a_{ij}x_iy_j \;\middle|\; \sum_i \hat{N}_i^X(x_i) \leq p,\ \sum_j \hat{M}_j^Y(y_j) \leq p\right\}, \tag{5.4}$$

$$\|(a_i)\|_{X,p} = \sup\left\{\sum_i a_ix_i \;\middle|\; \sum_i \hat{N}_i^X(x_i) \leq p\right\},\quad \|(a_j)\|_{Y,p} = \sup\left\{\sum_j a_jy_j \;\middle|\; \sum_j \hat{M}_j^Y(y_j) \leq p\right\}. \tag{5.5}$$

Since $\hat{N}_i^X(t/u) \leq \hat{N}_i^X(t)/u$ for $u \geq 1$ (it follows form the convexity of N_i^X and the normalization condition (5.1)) we have that

$$\|(a_i)_i\|_{X,2p} \leq C\,\|(a_i)_i\|_{X,p}, \tag{5.6}$$

and the same holds for $\|(a_j)_j\|_{Y,p}$. The inequality (5.6) has far-reaching consequences.

Fact 5.1.1. *Assume that $X_1, X_2, \ldots, Y_1, Y_2, \ldots$ are symmetric r.v's with LCT, for any i,j $a_{ij} \in F$, where $(F, \|\cdot\|)$ is a Banach space. Then for any $p \geq 1$,*

$$\left\|\sum_{ij} a_{ij}X_iY_j\right\|_{2p} \leq C\left\|\sum_{ij} a_{ij}X_iY_j\right\|_p.$$

Proof. Let $a_1, a_2, \ldots \in F$. From Theorem 6.24 below and (5.6) it follows that

$$\begin{aligned}\left\|\sum_i a_iX_i\right\|_{2p} &\lesssim \mathbb{E}\left\|\sum a_iX_i\right\| + \sup_{\varphi\in B^*(F)} \|(\varphi(a_i))_i\|_{X,2p} \\ &\lesssim \mathbb{E}\left\|\sum a_iX_i\right\| + \sup_{\varphi\in B^*(F)} \|\varphi((a_i))_i\|_{X,p} \lesssim \left\|\sum_i a_iX_i\right\|_p.\end{aligned} \tag{5.7}$$

Conditionally using twice (5.7), first time to a Banach space $(F^\infty, |||\cdot|||)$, where $|||(a_i)_i||| = \|\sum a_i X_i\|_{2p}$ and then to $(F, \|\cdot\|)$ we conclude that

$$\left\|\sum_{ij} a_{ij} X_i Y_j\right\|_{2p} \lesssim \left(\mathbb{E}^Y \left[\left(\mathbb{E}^X \left\|\sum_{ij} a_{ij} X_i Y_j\right\|^{2p}\right)^{1/2p}\right]^p\right)^{1/p} \lesssim \left\|\sum_{ij} a_{ij} X_i Y_j\right\|_p.$$

□

Let

$$B_p^X = \left\{x \in \mathbb{R}^n : \sum_i \hat{N}_i^X(x_i) \le p\right\}, \quad B_p^Y = \left\{y \in \mathbb{R}^n : \sum_j \hat{M}_j^Y(y_j) \le p\right\}$$

be the unit balls in the dual norms to the $\|\cdot\|_{X,p}, \|\cdot\|_{Y,p}$ so that

$$\|(a_i)\|_{X,p} = \sup_{x \in B_p^X} \sum_i a_i x_i, \ \|(a_j)\|_{Y,p} = \sup_{y \in B_p^Y} \sum_j a_j y_j.$$

Observe that (5.2) implies that

$$B_p^X, B_p^Y \subset \sqrt{p} B_2^n + p B_1^n. \tag{5.8}$$

In this chapter by $(g_i), (\mathcal{E}_i)$ we denote independent random variables with standard Gaussian and exponential distribution (i.e the distribution with the density $1/2\exp(-|x|/2)$). Here and subsequently G_n stands for $(g_1, \ldots, g_n)$ and E_n for $(\mathcal{E}_1, \ldots, \mathcal{E}_n)$. For a random vector X with values in a Banach space $(F, \|\cdot\|)$ we set $\|X\|_p = (\mathbb{E}\|X\|^p)^{1/p}$. By $\mathbb{E}^X, \mathbb{E}^Y$ we mean integration with respect to $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$ respectively.

We say that a sequence of r.v's $X_1, X_2, \ldots$ is subgaussian with constant γ if $X_1, X_2, \ldots$ are independent and for any i $\mathbb{E}X_i = 0$, $\mathbb{P}(|X_i| \ge t) \le 2\exp(-t^2/(2\gamma^2))$.

In this chapter write $a \lesssim b$ (resp. $a \lesssim^\alpha b$) if $a \le Cb$ (resp. $a \le C(\alpha)b$). If $(F, \|\cdot\|)$ is a Banach space then $(F^*, \|\cdot\|_*)$ stands for its dual space and $B^*(F) = \{\varphi \in F^* : \ \|\varphi\|_* \le 1\}$ for the unit ball in the dual space.

Our first observation is a simple lower bound.

Proposition 5.1.2. *Assume that $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$ are independent random variables with LCT such that* (5.1) *holds. Let $(a_{ij})_{ij}$ be an F-valued matrix, where $(F, \|\cdot\|)$ is a Banach space. Then for any $p \ge 1$ we have*

$$\begin{aligned}\left\|\sum_{ij} a_{ij} X_i Y_j\right\|_p \gtrsim \mathbb{E}\left\|\sum_{ij} a_{ij} X_i Y_j\right\| + \sup_{x \in B_p^X} \mathbb{E}\left\|\sum_{ij} a_{ij} x_i Y_j\right\| + \sup_{y \in B_p^Y} \mathbb{E}\left\|\sum_{ij} a_{ij} X_i y_j\right\| \\ + \sup_{\varphi \in B^*(F)} \left\|\left(\sqrt{\sum_j \varphi(a_{ij})^2}\right)_i\right\|_{X,p} + \sup_{\varphi \in B^*(F)} \|\varphi(a_{ij})\|_{X,Y,p}.\end{aligned} \tag{5.9}$$

We are able to reverse the inequality (5.9) in L_q spaces under additional subgaussian assumption on $(Y_j)_j$.

Theorem 5.1.3. *Assume that $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$ are independent random variables with LCT such that* (5.1) *holds. Assume also that the sequence $Y_1, Y_2, \ldots$ is subgaussian with constant γ. Let $(a_{ij})_{ij}$ be an F-valued matrix, where $F = L_q(T, d\mu)$. Then for any $p \ge 1$ we have*

$$\left\|\sum_{ij} a_{ij}X_iY_j\right\|_p \sim^{q,\gamma} \mathbb{E}\left\|\sum_{ij} a_{ij}X_iY_j\right\|_{L_q} + \sup_{x\in B_p^X}\mathbb{E}\left\|\sum_{ij} a_{ij}x_iY_j\right\|_{L_q} + \sup_{y\in B_p^Y}\mathbb{E}\left\|\sum_{ij} a_{ij}X_iy_j\right\|_{L_q}$$
$$+ \sup_{\varphi\in B^*(L_q)}\left\|\left(\sqrt{\sum_j \varphi(a_{ij})^2}\right)_i\right\|_{X,p} + \sup_{\varphi\in B^*(L_q)} . \|(\varphi(a_{ij}))_{ij}\|_{X,Y,p} \tag{5.10}$$

Remark 5.1.4. The formula (5.10) can by simplified, but we intended to present it in a way which leads itself to generalizations. By standard arguments (cf. Proposition 3.2.1, Fact 5.1.1) it can be showed that

$$\mathbb{E}\left\|\sum_{ij} a_{ij}X_iY_j\right\|_{L_q} \sim^q \left\|\sqrt{\sum_{ij} a_{ij}^2}\right\|_{L_q}, \quad \sup_{x\in B_p^X}\mathbb{E}\left\|\sum_{ij} a_{ij}x_iY_j\right\|_{L_q} \sim^q \sup_{x\in B_p^X}\left\|\sqrt{\sum_j\left(\sum_i a_{ij}x_i\right)^2}\right\|_{L_q},$$

$$\sup_{y\in B_p^Y}\mathbb{E}\left\|\sum_{ij} a_{ij}X_iy_j\right\|_{L_q} \sim^q \sup_{y\in B_p^Y}\left\|\sqrt{\sum_i\left(\sum_j a_{ij}y_j\right)^2}\right\|_{L_q}.$$

The upper bound in (5.10) is a consequence of the following inequality.

Theorem 5.1.5. *Let $(F,\|\cdot\|)$ be an arbitrary Banach space and $(a_{ij})_{ij}$ an F-valued matrix. Assume that $(X_i)_i, (Y_j)_j$ are symmetric r.v's with LCT such that (5.1) holds. Additionally assume that Y_j are subgaussian with constant γ. Then for any $p \geq 1$ we have*

$$\left\|\sum_{ij} a_{ij}X_iY_j\right\|_p \lesssim \mathbb{E}\left\|\sum_{ij} a_{ij}X_iY_j\right\| + \gamma\left(\mathbb{E}\left\|\sum_{ij} a_{ij}g_{ij}\right\| + \mathbb{E}\left\|\sum_{ij} a_{ij}g_i\mathcal{E}_j\right\|\right) + \sup_{x\in B_p^X}\mathbb{E}\left\|\sum_{ij} a_{ij}x_iY_j\right\|$$
$$+ \sup_{y\in B_p^Y}\mathbb{E}\left\|\sum_{ij} a_{ij}X_iy_j\right\| + \sup_{\varphi\in B^*(F)}\left\|\left(\sqrt{\sum_j \varphi(a_{ij})^2}\right)_i\right\|_{X,p} + \sup_{\varphi\in B^*(F)} \|\varphi(a_{ij})\|_{X,Y,p}. \tag{5.11}$$

In the Hilbert spaces the assumption about the subgaussianity can be relaxed.

Theorem 5.1.6. *Assume that $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$ are independent random variables with LCT such that (5.1) holds. Let $(a_{ij})_{ij}$ be an F =-valued matrix, where $(F,\|\cdot\|)$ is a Hilbert space. Then for any $p \geq 1$ we have*

$$\left\|\sum_{ij} a_{ij}X_iY_j\right\|_p \sim \mathbb{E}\left\|\sum_{ij} a_{ij}X_iY_j\right\| + \sup_{x\in B_p^X}\mathbb{E}\left\|\sum_{ij} a_{ij}x_iY_j\right\| + \sup_{y\in B_p^Y}\mathbb{E}\left\|\sum_{ij} a_{ij}X_iy_j\right\|$$
$$+ \sup_{\varphi\in B^*(F)}\left\|\left(\sqrt{\sum_j \varphi(a_{ij})^2}\right)_i\right\|_{X,p} + \sup_{\varphi\in B^*(F)} \|\varphi(a_{ij})\|_{X,Y,p}. \tag{5.12}$$

5.2 Reduction to a bound on the supremum of a certain stochastic process

We begin with reducing Theorem 5.1.5 and the upper estimates in Theorem 5.1.6 to an estimate on expected value of a supremum of a certain stochastic process.

Lemma 5.2.1. *Under the assumption of Theorem 5.1.6 with "Hilbert space" replaced by "Banach space" we have*

$$\left\|\sum_{ij} a_{ij}X_iY_j\right\|_p \lesssim \mathbb{E}\left\|\sum_{ij} a_{ij}X_iY_j\right\| + \sup_{y\in B_p^Y}\mathbb{E}\left\|\sum_{ij} a_{ij}X_iy_j\right\| + \mathbb{E}\sup_{x\in B_p^X}\left\|\sum_{ij} a_{ij}x_iY_j\right\| + \sup_{\varphi\in B^*(F)}\|(\varphi(a_{ij}))_{ij}\|_{X,Y,p}$$

Proof. By conditionally applying Theorem 6.24 we obtain

$$\left\|\sum_{ij} a_{ij}X_iY_j\right\|_p \lesssim \left\|\mathbb{E}^X\left\|\sum_{ij} a_{ij}X_iY_j\right\|\right\|_p + \left\|\sup_{\varphi\in B^*(F)}\left\|\left(\sum_j \varphi(a_{ij})Y_j\right)_i\right\|_{X,p}\right\|_p$$

Since $y\to\mathbb{E}^X\left\|\sum_{ij} a_{ij}X_iy_j\right\|$ is a norm by again applying Theorem 6.24 we get

$$\left\|\mathbb{E}^X\left\|\sum_{ij} a_{ij}X_iY_j\right\|\right\|_p \lesssim \mathbb{E}\left\|\sum_{ij} a_{ij}X_iY_j\right\| + \sup_{y\in B_p^Y}\mathbb{E}\left\|\sum_{ij} a_{ij}X_iy_j\right\|.$$

Again Theorem 6.24 yield

$$\left\|\sup_{\varphi\in B^*(F)}\left\|\left(\sum_j \varphi(a_{ij})Y_j\right)_i\right\|_{X,p}\right\|_p \lesssim \mathbb{E}\sup_{x\in B_p^X}\left\|\sum_{ij} a_{ij}x_iY_j\right\| + \sup_{\varphi\in B^*(F)}\|(\varphi(a_{ij}))_{ij}\|_{X,Y,p}.$$

□

So in order to prove Theorem 5.1.5 and the upper bound in Theorem 5.1.6 it is convenient to establish upper bounds on $\mathbb{E}\sup_{x\in B_p^X}\left\|\sum_{ij} a_{ij}x_iY_j\right\|$. To this end we reformulate the language we use. Observe that it suffices to show Theorems 5.1.3-5.1.6 when we sum up after $i,j\le n$. We may also assume that $F=\mathbb{R}^m$ so that $a_{ij}=(a_{ijk})_{k\le m}$ and $\|x\|=\sup_{t\in T}\sum_{i\le m} t_ix_i$ where $T\subset\mathbb{R}^m$ is the unit ball in the dual space. In this language (5.11) becomes

$$\left\|\sup_{t\in T}\sum_{ijk}a_{ijk}X_iY_jt_k\right\|_p \lesssim \mathbb{E}\sup_{t\in T}\sum_{ijk}a_{ijk}X_iY_jt_k+\gamma\left(\mathbb{E}\sup_{t\in T}\sum_{ijk}a_{ijk}g_{ij}t_k+\mathbb{E}\sup_{t\in T}\sum_{ijk}a_{ijk}g_i\mathcal{E}_jt_k\right)$$
$$+\sup_{x\in B_p^X}\mathbb{E}\sup_{t\in T}\sum_{ijk}a_{ijk}x_iY_jt_k+\sup_{y\in B_p^Y}\mathbb{E}\sup_{t\in T}\sum_{ijk}a_{ijk}X_iy_jt_k$$
$$+\sup_{t\in T}\left\|\left(\sqrt{\sum_j\left(\sum_k a_{ijk}t_k\right)^2}\right)_i\right\|_{X,p}+\sup_{t\in T}\left\|\left(\sum_k a_{ijk}t_k\right)_{ij}\right\|_{X,Y,p}. \tag{5.13}$$

We are now in a position to show Theorem 5.1.6.

Proof of Theorem 5.1.6. The lower bound follows by Proposition 5.1.2. We will prove the upper bound. W.l.o.g (changing $(a_{ij})_{ij}$ if necessary) we may assume that the unit ball in the dual space is $T=B_2^m=\{x\in\mathbb{R}^m:\ \|x\|_2\le 1\}$. Using the language introduced above the upper bound in (5.12) becomes

$$\left\|\sup_{t\in B_2^m}\sum_{ijk}a_{ijk}X_iY_jt_k\right\|_p\lesssim \mathbb{E}\sup_{t\in B_2^m}\sum_{ijk}a_{ijk}X_iY_jt_k+\sup_{x\in B_p^X}\mathbb{E}\sup_{t\in B_2^m}\sum_{ijk}a_{ijk}x_iY_jt_k$$
$$+\sup_{y\in B_p^Y}\mathbb{E}\sup_{t\in B_2^m}\sum_{ijk}a_{ijk}X_iy_jt_k+\sup_{t\in B_2^m}\left\|\left(\sqrt{\sum_j\left(\sum_k a_{ijk}t_k\right)^2}\right)_i\right\|_{X,p}$$
$$+\sup_{t\in B_2^m}\left\|\left(\sum_k a_{ijk}t_k\right)_{ij}\right\|_{X,Y,p}. \tag{5.14}$$

Applying Lemma 5.2.1, to show (5.14) it is enough to prove that

$$\mathbb{E}\sup_{t\in B_2^m,x\in B_p^X}\sum_{ijk}a_{ijk}x_iY_jt_k\lesssim \mathbb{E}\sup_{t\in B_2^m}\sum_{ijk}a_{ijk}X_iY_jt_k+\sup_{x\in B_p^X}\mathbb{E}\sup_{t\in B_2^m}\sum_{ijk}a_{ijk}x_iY_jt_k$$
$$+\sup_{t\in B_2^m}\left\|\left(\sqrt{\sum_j\left(\sum_k a_{ijk}t_k\right)^2}\right)_i\right\|_{X,p}+\sup_{t\in B_2^m}\left\|\left(\sum_k a_{ijk}t_k\right)_{ij}\right\|_{X,Y,p}. \tag{5.15}$$

Since $B_p^X\subset\sqrt{p}B_2^n+pB_1^n$ (recall (5.8)) we take the decomposition $(p^{-1/2}B_p^X)\times B_2=\bigcup_{l=1}^N((z^l,s^l)+T_l)$ obtained from Corollary 6.12 with $Z_i=Y_i$. Applying Lemma 6.11 (recall (5.6)) yields

$$\mathbb{E}\sup_{x\in B_p^X,t\in B_2^m}\sum_{ijk}a_{ijk}x_iY_jt_k\lesssim\sqrt{p}\max_l\mathbb{E}\sup_{(x,t)\in(z^l,s^l)+T_l}\sum_{ijk}a_{ijk}x_iY_jt_k$$
$$+\sup_{t\in B_2^m,x\in B_p^X,y\in B_p^Y}\sum_{ijk}a_{ijk}x_iy_jt_k. \tag{5.16}$$

Because $\mathbb{E}Y_j=0$ and $\mathbb{E}Y_j^4\le C$ (recall (5.3) and Corollary 6.12) we conclude

$$\begin{aligned}\mathbb{E}\sup_{(x,t)\in(z^l,s^l)+T_l}\sum_{ijk}a_{ijk}x_iY_jt_k &\le \mathbb{E}\sup_{(x,t)\in T_l}\sum_{ijk}a_{ijk}x_iY_jt_k\\ &\quad+\mathbb{E}\sup_{(x,t)\in T_l}\sum_{ijk}a_{ijk}x_iY_js_k^l+\mathbb{E}\sup_{(x,t)\in T_l}\sum_{ijk}a_{ijk}z_i^lY_jt_k\\ &\lesssim p^{-1/2}\sqrt{\sum_{ijk}a_{ijk}^2}+\mathbb{E}\sup_{(x,t)\in T_l}\sum_{ijk}a_{ijk}x_iY_js_k^l+\mathbb{E}\sup_{(x,t)\in T_l}\sum_{ijk}a_{ijk}z_i^lY_jt_k. \end{aligned}\tag{5.17}$$

Using Jensen's inequality, (5.3) and Fact 5.1.1 with $F=l_2^m$ and $p=1$ we obtain

$$\sqrt{\sum_{ijk}a_{ijk}^2}\le C\mathbb{E}\sup_{t\in B_2^m}\sum_{ijk}a_{ijk}X_iY_jt_k. \tag{5.18}$$

Corollary 6.12 assures, that $(z^l,s^l)\in(p^{-1/2}B_p^X)\times B_2^m$ so that $T_l\subset 2(p^{-1/2}B_p^X)\times 2B_2^m$. Applying Corollary 6.18 we get

$$\begin{aligned}&\mathbb{E}\sup_{(x,t)\in T_l}\sum_{ijk}a_{ijk}x_iY_js_k^l\le 2p^{-1/2}\sup_{t\in B_2^m}\mathbb{E}\sup_{x\in B_p^X}\sum_{ijk}a_{ijk}x_iY_jt_k\\ &\le Cp^{-1/2}\left(\sup_{t\in B_2^m,x\in B_p^X,y\in B_p^Y}\sum_{ijk}a_{ijk}x_iy_jt_k+\sup_{t\in B_2^m,\ x\in B_p^X}\sum_i\sqrt{\sum_j\left(\sum_k a_{ijk}t_k\right)^2}x_i\right)\\ &=Cp^{-1/2}\left(\sup_{t\in B_2^m}\left\|\left(\sum_k a_{ijk}t_k\right)_{ij}\right\|_{X,Y,p}+\sup_{t\in B_2^m}\left\|\left(\sqrt{\sum_j\left(\sum_k a_{ijk}t_k\right)^2}\right)_i\right\|_{X,p}\right).\end{aligned}\tag{5.19}$$

Finally, since $z^l\in p^{-1/2}B_p^X$

$$\mathbb{E}\sup_{(x,t)\in T_l}\sum_{ijk}a_{ijk}z_i^lY_jt_k\le p^{-1/2}\sup_{x\in B_p^X}\mathbb{E}\sup_{t\in B_2^m}\sum_{ijk}a_{ijk}x_iY_jt_k. \tag{5.20}$$

Formulas (5.16)-(5.20) imply (5.15). □

5.3 Expectation of suprema of a certain Gaussian processes

The main result of this section is Proposition 5.3.8 in which we estimate the expectation of the supremum of a Gaussian process $(G_{(x,t)})_{(x,t)\in V}$, where $V\subset B_p^X\times T$ and $G_{(x,t)}=\sum_{ijk}a_{ijk}g_ix_jt_k$. To estimate such quantity one needs to study the distance on $B_p^X\times T$ given by

$$d_A((x,t),(x',t'))=(\mathbb{E}|G_{(x,t)}-G_{(x',t')}|^2)^{1/2}=\alpha_A(x\otimes t-x'\otimes t'),$$

where $x\otimes t=(x_j\cdot t_k)_{j\le n,k\le m}\in\mathbb{R}^{nm}$ and α_A is a norm on $\mathbb{R}^{nm}$ defined by the formula

$$\alpha_A((x)_{jk})=\sqrt{\sum_{i=1}^n\left(\sum_{j=1}^n\sum_{k=1}^m a_{ijk}x_{jk}\right)^2}.$$

We use the scheme introduced in Chapter 3. In order to proceed we need some entropy estimates (recall that $N(S,\rho,\varepsilon)$ is the smallest number of closed balls with the diameter ε in metric ρ that cover the set S). The crucial idea is to consider measure $\mu_{n,\varepsilon}$, which is the distribution of of $\varepsilon(G_n+E_n)$, $\varepsilon>0$ (we recall E_n is the exponential vector in $\mathbb{R}^n$).

Let $S\subset\mathbb{R}^m$. By the classical Sudakov minoration (cf. Theorem 6.7) for any $x\in\mathbb{R}^m$ there exists a set $S_{x,\varepsilon}\subset S$ of cardinality at most $\exp(C\varepsilon^{-2})$ such that

$$\forall_{s\in S}\exists_{s'\in S_{x,\varepsilon}}\ \alpha_A(x\otimes(s-s'))\leq\varepsilon\mathbb{E}\sup_{s\in S}\left|\sum_{ijk}a_{ijk}g_ix_js_k\right|=:\varepsilon\beta_{A,S}(x).$$

Observe that $\beta_{A,S}(x)$ is a norm. We define the following measure on $\mathbb{R}^n\times S$:

$$\hat{\mu}_{\varepsilon,S}(C)=\int_{\mathbb{R}^n}\sum_{t\in S_{x,\varepsilon}}\delta_{x,t}(C)d\mu_{n,\varepsilon}.$$

To bound $N(U,d_A,\varepsilon)$, for $U\subset B_p^X\times T$ we need two lemmas.

Lemma 5.3.1. *For any $\varepsilon,p>0$, norms α_1,α_2 on $\mathbb{R}^n$ and $y\in B_2^n+\sqrt{p}B_1^n$ we have*

$$\mu_{n,\varepsilon}\left(x:\ \alpha_1(x-y)\leq C\varepsilon\mathbb{E}\alpha_1(E_n),\ \alpha_2(x)\leq C\mathbb{E}\alpha_2(E_n)+\alpha_2(y)\right)\geq 1/4\exp\left(-\varepsilon^{-2}/2-\sqrt{p}\varepsilon^{-1}\right).$$

Proof. Since for any norm α on $\mathbb{R}^n$ we have $\mathbb{E}\alpha(G_n)\leq 3\mathbb{E}\alpha(E_n)$ (cf. [1, Lemma 5.6]), the assertion is a simple consequence of [1, Lemma 5.3]. □

Lemma 5.3.2. *For any finite set $S\subset\mathbb{R}^m$, $p>0$, $(x,t)\in(B_2^n+\sqrt{p}B_1^n)\times S$ and $\varepsilon>0$ we have*

$$\hat{\mu}_{\varepsilon,S}(B((x,t),d_A,r(\varepsilon)))\geq\frac{1}{4}\exp(-\varepsilon^{-2}/2-\sqrt{p}\varepsilon^{-1}),$$

where

$$B((x,t),d_A,r(\varepsilon))=\{(x',t')\in\mathbb{R}^n\times S:\ \alpha_A(x\otimes t-x'\otimes t')\leq r(\varepsilon)\}$$

and

$$r(\varepsilon)=C\left(\varepsilon^2\mathbb{E}\beta_{A,S}(E_n)+\varepsilon\beta_{A,S}(x)+\varepsilon\mathbb{E}\alpha_A(E_n\otimes t)\right).$$

Proof. Let us fix $(x,t)\in B_2^n\times S$ and $\varepsilon>0$. Set

$$U=\{x'\in\mathbb{R}^n:\ \beta_{A,S}(x')\leq C\varepsilon\mathbb{E}\beta_{A,S}(E_n)+\beta_{A,S}(x),\ \alpha_A((x'-x)\otimes t)\leq C\varepsilon\mathbb{E}\alpha_A(E_n\otimes t)\}.$$

For any $x'\in U$ there exists $t'\in S'_{x,\varepsilon}$ such that $\alpha_A(x'\otimes(t-t'))\leq\varepsilon\beta_{A,S}(x')$. By the triangle inequality

$$\alpha_A(x\otimes t-x'\otimes t')\leq\alpha_A((x-x')\otimes t)+\alpha_A(x'\otimes(t-t'))\leq r(\varepsilon).$$

Thus, by Lemma 5.3.1, $\hat{\mu}_{\varepsilon,S}\left(B\left((x,t),d_A,r(\varepsilon)\right)\right)\geq\mu_{n,\varepsilon}(U)\geq 1/4\exp(-\varepsilon^{-2}/2-\sqrt{p}\varepsilon^{-1})$. □

Corollary 5.3.3. *For any $p,\varepsilon>0$, $U\subset B_2^n+\sqrt{p}B_1^n$, and $S\in\mathbb{R}^m$*

$$N\left(U\times S,d_A,\varepsilon^2\mathbb{E}\beta_{A,S}(E_n)+\varepsilon\sup_{x\in U}\beta_{A,S}(x)+\varepsilon\sup_{t\in S}\mathbb{E}\alpha_A(E_n\otimes t)\right)\leq 4\exp(C\varepsilon^{-2}+C\sqrt{p}\varepsilon^{-1}). \tag{5.21}$$

Proof. Let $r=\varepsilon^2\mathbb{E}\beta_{A,S}(E_n)+\varepsilon\sup_{x\in U}\beta_{A,S}(x)+\varepsilon\sup_{t\in S}\mathbb{E}\alpha_A(E_n\otimes t)$ and $N=N(U\times S,d_A,r)$. Then there exist points $(x_i,t_i)_{i=1}^N$ in $U\times S$ such that $d_A((x_i,t_i),(x_j,t_j))>r$. Note that the balls $B((x_i,t_i),d_A,r/2)$ are disjoint and, by Lemma 5.3.2, each of these balls has $\hat{\mu}_{\varepsilon,S}$ measure at least

$1/4\exp(-C\varepsilon^{-2}/2 - C\sqrt{p}\varepsilon^{-1})$. On the other hand we obviously have $\hat{\mu}_{\varepsilon,S}(\mathbb{R}^n \times S) \le \exp(C\varepsilon^{-2})$ what implies $N \le 4\exp(C\varepsilon^{-2} + C\sqrt{p}\varepsilon^{-1})$. □

Similarly as in Chapter 3 we need to decompose $B_2^n + \sqrt{p}B_1^n$ and T into smaller pieces $B_2^n + \sqrt{p}B_1^n = \bigcup_i U_i$, $T = \bigcup_i T_i$ in such a way that $\sup_{x,x'\in U_i} \beta_{A,T}(x-x')$, $\sup_{t,t'\in T_i} \mathbb{E}\alpha_A(E_n \otimes (t-t'))$ is small on each piece. To make the notation more compact let for $T \subset \mathbb{R}^m$ and $V \subset \mathbb{R}^n \times \mathbb{R}^m$,

$$\begin{aligned}
s_A(T) &:= \mathbb{E}\sup_{t\in T}\Big|\sum_{ijk} a_{ijk} g_{ij} t_k\Big| + \mathbb{E}\sup_{t\in T}\Big|\sum_{ijk} a_{ijk} g_i \mathcal{E}_j t_k\Big|, \\
F_A(V) &:= \mathbb{E}\sup_{(x,t)\in V}\sum_{ijk} a_{ijk} g_i x_j t_k, \\
\varDelta_A(V) &:= \operatorname{diam}(V, d_A) = \sup_{(x,t),(x',t')\in V} \alpha_A(x\otimes t - x'\otimes t').
\end{aligned}$$

An obvious consequence of the classic Sudakov minoration (cf. Theorem 6.7) is the following statement.

Lemma 5.3.4. *For any $p \ge 1$ there exists decomposition $T - T = \bigcup_{i=1}^N T_i$ such that $N \le \exp(Cp)$ and for any $i \le N$,*

$$\sup_{t,t'\in T_i} \mathbb{E}\alpha_A(E_n \otimes (t-t')) \le \sup_{t,t'\in T_i} \sqrt{\sum_{ij}\Big(\sum_k a_{ijk}(t_k - t'_k)\Big)^2} \le \frac{1}{\sqrt{p}}\mathbb{E}\sup_{t\in T}\sum a_{ijk} g_{ij} t_k \le \frac{1}{\sqrt{p}} s_A(T).$$

Lemma 5.3.5. *For any $p \ge 1$ there exists decomposition of $B_2^n + \sqrt{p}B_1^n = \bigcup_{i=1}^N U_i$ such that $N \le \exp(Cp)$ and for any $i \le N$,*

$$\sup_{x,x'\in U_i} \beta_{A,T}(x-x') \le \frac{1}{\sqrt{p}}\mathbb{E}\sup_{t\in T}\Big|\sum a_{ijk} g_i \mathcal{E}_j t_k\Big| \le \frac{1}{\sqrt{p}} s_A(T).$$

Proof. It is enough to use Corollary 6.9 with $\varepsilon = p^{-1/2}/C$ and the norm $x \to \beta_{A,T}(x)$. □

Lemma 5.3.6. *Let V be a subset of $(B_2^n + \sqrt{p}B_1^n) \times (T-T)$ and $(x,t) \in \mathbb{R}^n \times \mathbb{R}^m$. Then there exists a decomposition $V = \bigcup_{i=1}^N V_i$ such that $N \le \exp(C2^{2l}p)$ and for any $i \le N$*

$$F_A(V_i + (x,t)) \le F_A(V_i) + \beta_{A,T}(x) + C\mathbb{E}\alpha_A(E_n \otimes t)$$

and

$$\varDelta_A(V_i) \le \frac{1}{2^{2l}p} s_A(T) + \frac{1}{2^l\sqrt{p}} \sup_{(y,s)\in V}\left(\beta_{A,T}(y) + \mathbb{E}\alpha_A(E_n \otimes s)\right).$$

Proof. Corollary 6.9 with $\varepsilon = p^{-1/2}$ and $\alpha = \alpha_A$ and Corollary 5.3.3 with $\varepsilon = 2^{-l}p^{-1/2}$ yield decomposition $V = \bigcup_{i\le N} V_i$, $N \le \exp(Cp2^{2l})$ such that

$$\begin{aligned}
&\sup_{(y,s),(y',s')\in V_i} \alpha_A((y_j - y'_j)\otimes t) \le \frac{C}{\sqrt{p}}\mathbb{E}\alpha_A(E_n \otimes t), \qquad (5.22)\\
&\varDelta_A(V_i) \le \frac{1}{2^{2l}p} s_A(T) + \frac{1}{2^l\sqrt{p}} \sup_{(y,s)\in V}\left(\beta_{A,T}(y) + \mathbb{E}\alpha_A(E_n \otimes s)\right).
\end{aligned}$$

Since $\mathbb{E}\sum_{ijk} g_i x_j t_k = 0$ we have

$$F_A(V_i+(x,t)) \le F_A(V_i)+\beta_{A,T}(x)+\mathbb{E}\sup_{(y,s)\in V_i}\sum_{ijk} a_{ijk}g_iy_jt_k.$$

From Lemma 6.17 and (5.22) we obtain

$$\begin{aligned}\mathbb{E}\sup_{(y,s)\in V_i}\sum_{ijk} a_{ijk}g_iy_jt_k &\lesssim \sqrt{\sum_{ij}\left(\sum_k a_{ijk}t_k\right)^2}+\sqrt{p}\sup_{(y,s),(y',s')\in V_i}\alpha_A((y_j-y_j')\otimes t)\\ &\le C\mathbb{E}\alpha_A(E_n\otimes t),\end{aligned}$$

where in the last line we used (5.7) with $p=1$ and $F=l_2^{n^2}$ and (5.22). The assertion easily follows. □

Corollary 5.3.7. *Let $V\subset (B_2^n+\sqrt{p}B_1^n)\times(T-T)$ be such that $V-V\subset(B_2^n+\sqrt{p}B_1^n)\times(T-T)$. Then there exists decomposition $V=\bigcup_{i=1}^N((x_i,t_i)+V_i)$ such that $N\le\exp(C2^{2l}p)$ and for each $i\le N$*

i) $(x_i,t_i)\in V$, $V_i-V_i\subset V-V$, $V_i\subset(B_2^n+\sqrt{p}B_1^n)\times(T-T)$,
ii) $\sup_{(y,s)\in V_i}(\beta_{A,T}(y)+\mathbb{E}\alpha_A(E_n\otimes s))\le\frac{2}{2^l\sqrt{p}}s_A(T)$,
iii) $\Delta_A(V_i)\le\frac{C}{2^{2l}p}s_A(T)$,
iv) $F_A(V_i+(x_i,t_i))\le F_A(V_i)+\beta_{A,T}(x_i)+C\mathbb{E}\alpha_A(E_n\otimes t_i)$.

Proof. By Lemmas 5.3.4 and 5.3.5 we can find decompositions $B_2^n+\sqrt{p}B_1^n=\bigcup_{i=1}^{N_1}U_i$, $T-T=\bigcup_{i=1}^{N_2}T_i$ such that $N_1,N_2\le\exp(C2^{2l}p)$ and

$$\sup_{x,x'\in U_i}\beta_{A,T}(x-x')\le\frac{1}{2^l\sqrt{p}}s_A(T),\quad \sup_{s,s'\in T_i}\mathbb{E}\alpha_A(E_n\times(s-s'))\le\frac{1}{2^l\sqrt{p}}s_A(T).$$

Let $V_{ij}=V\cap(U_i\times T_j)$. If $V_{ij}\ne\emptyset$ we take any point $(x_{ij},t_{ij})\in V_{ij}$ and using Lemma 5.3.6 we decompose

$$V_{ij}-(x_{ij},t_{ij})=\bigcup_{k=1}^{N_3}V_{ijk}$$

in such a way that $N_3\le\exp(C2^{2l}p)$,

$$F_A(V_{ijk}+(x_{ij},t_{ij}))\le F_A(V_{ijk})+\beta_{A,T}(x_{ij})+C\mathbb{E}\alpha_A(E_n\otimes t_{ij})$$

and

$$\begin{aligned}\Delta_A(V_{ijk})&\le\frac{1}{2^{2l}p}s_A(T)+\frac{1}{2^l\sqrt{p}}\left(\sup_{(y,s)\in V_{ijk}}\beta_{A,T}(y)+\sup_{(y,s)\in V_{ijk}}\mathbb{E}\alpha_A(E_n\otimes s)\right)\\ &\le\frac{1}{2^{2l}p}s_A(T)+\frac{1}{2^l\sqrt{p}}\left(\sup_{y,y'\in U_i}\beta_{A,T}(y-y')+\sup_{s,s'\in T_j}\mathbb{E}\alpha_A(E_n\otimes(s-s'))\right)\\ &\lesssim\frac{1}{2^{2l}p}s_A(T)\end{aligned}$$

The final decomposition is obtained by relabeling of the decomposition $V=\bigcup_{ijk}((x_{ij},t_{ij})+V_{ijk})$ □

Proposition 5.3.8. *For any $p\ge1$ any nonempty $T\subset\mathbb{R}^m$ and $W\subset(B_2^n+\sqrt{p}B_1^n)\times(T-T)$ we have*

$$F_A(W) \lesssim \frac{1}{\sqrt{p}} s_A(T) + \sup_{(x,t)\in W} \beta_{A,T}(x) + \sup_{(x,t)\in W} \mathbb{E}\alpha_A(E_n \otimes t) + \sqrt{p}\Delta_A(W).$$

Proof. W.l.o.g we may assume that W is finite and $W \subset 1/2((B_2^n + \sqrt{p}B_1^n) \times T)$. We define

$$\Delta_0 := \Delta_A(W), \quad \tilde{\Delta}_0 := \sup_{(x,t)\in W} \beta_{A,T}(x) + \sup_{(x,t)\in W} \mathbb{E}\alpha_A(E_n \otimes t)$$
$$\Delta_l := 2^{-2l} p^{-1} s_A(T), \quad \tilde{\Delta}_l := 2^{-l} p^{-1/2} s_A(T) \quad \text{for } l = 1, 2, \ldots.$$

Let for $l = 0, 1\ldots$ and $k = 1, 2, \ldots$

$$c(l,k) := \sup\{F_A(V) : V \subset (B_2^n + \sqrt{p}B_1^n) \times (T - T),\ V - V \subset W - W,\ |V| \le k,$$
$$\Delta_A(V) \le \Delta_l,\ \sup_{(x,t)\in V} (\beta_{A,T}(x) + \mathbb{E}\alpha_A(E_n \otimes t)) \le C\tilde{\Delta}_l\}.$$

Obviously $c(l,1) = 0$. We will show that for $k > 1$ and $l > 0$ we have

$$c(l,k) \le c(l+1, k-1) + C\left(2^l \sqrt{p}\Delta_l + \tilde{\Delta}_l\right). \tag{5.23}$$

To this end take any set V as in the definition of $c(l,k)$ and apply to it Corollary 5.3.7 to obtain decomposition $V = \bigcup_{i=1}^N ((x_i,t_i) + V_i)$. Conditions $i) - iv)$ from Corollary 5.3.7 easily imply that

$$\max_{i\le N} F(V_i) \le c(l+1, k-1).$$

Lemma 6.10 yields

$$F_A(V) = F_A\left(\bigcup_{i=1}^N (V_i + (x_i,t_i))\right) \le C\sqrt{\log N}\Delta_A(V) + \max_i F(V_i + (x_i,t_i)).$$

Since $N \le \exp(C2^{2l}p)$ (cf. Corollary 5.3.7) from the definition of $c(l,k)$ we obtain

$$\sqrt{\log N}\Delta_A(V) \le C2^l\sqrt{p}\Delta_l$$

and for each i by Corollary 5.3.7 we have (recall that $(x_i,t_i) \in V$)

$$F(V_i + (x_i,t_i)) \le F(V_i) + \beta_{A,T}(x_i) + C\mathbb{E}\alpha_A(E_n \otimes t_i) \le F_A(V_i) + C\tilde{\Delta}_l.$$

So we have proved (5.23). It implies that for any k we have

$$c(l,k) \lesssim \sum_{k=0}^{\infty} \left(2^{k/2}\sqrt{p}\Delta_k + C\tilde{\Delta}_k\right)$$
$$\lesssim \frac{1}{\sqrt{p}} s_A(T) + \sup_{(x,t)\in W} \beta_{A,T}(x) + \sup_{(x,t)\in W} \mathbb{E}\alpha_A(E_n \otimes t) + \sqrt{p}\Delta_A(W)$$

As a result

$$F(W) \le \sup_k c(0,k) \lesssim \frac{1}{\sqrt{p}} s_A(T) + \sup_{(x,t)\in W} \beta_{A,T}(x) + \sup_{(x,t)\in W} \mathbb{E}\alpha_A(E_n \otimes t) + \sqrt{p}\Delta_A(W).$$

□

5.4 Proofs of the main results

We begin with the proof of the simple lower bound.

Proof of Proposition 5.1.2. Using Theorem 6.26 we obtain

$$\left\|\sum_{ij} a_{ij}X_iY_j\right\|_p \geq \sup_{\varphi\in B^*(F)}\left\|\sum_{ij}\varphi(a_{ij})X_iY_j\right\|_p$$
$$\gtrsim \sup_{\varphi\in B^*(F)}\left\|(\varphi(a_{ij}))_{ij}\right\|_{X,Y,p} + \sup_{\varphi\in B^*(F)}\left\|\left(\sqrt{\sum_j \varphi(a_{ij})^2}\right)_i\right\|_{X,p}.$$

Jensen's inequality and Theorem 6.22 yield

$$\left\|\sum_{ij} a_{ij}X_iY_j\right\|_p \geq \left(\mathbb{E}^Y \sup_{\varphi\in B^*(F)}\mathbb{E}^X\left|\sum_{ij}\varphi(a_{ij})X_iY_j\right|^p\right)^{1/p} \gtrsim \left\|\sup_{\varphi\in B^*(F),x\in B_p^X}\sum_{ij}\varphi(a_{ij})x_iY_j\right\|_p$$
$$\geq \sup_{x\in B_p^X}\mathbb{E}\left\|\sum_{ij}a_{ij}x_iY_j\right\|.$$

Analogously $\left\|\sum_{ij}a_{ij}X_iY_j\right\|_p \gtrsim \sup_{y\in B_p^Y}\mathbb{E}\left\|\sum_{ij}a_{ij}X_iy_j\right\|$. To finish the proof it is enough to observe that $\left\|\sum_{ij}a_{ij}X_iY_j\right\|_p \geq \left\|\sum_{ij}a_{ij}X_iY_j\right\|_1$.

□

Now we prove Theorem 5.1.5 and then we will deduce from it Theorem 5.1.3. As w mentioned, there is a problem when one wants to prove moments bounds in L_q spaces. The reason is the lack of a counterpart of Corollary 6.12 and we have not been able to derive it. However, with an additional assumption about subgaussianity it is enough to use the decomposition obtained in Corollary 5.3.7

Proof of Theorem 5.1.5. We will prove inequality (5.13) equivalent to (5.11). Let us take the decomposition from Corollary 5.3.7 with $l=1$ so that $(p^{-1/2}B_p^X)\times T = \bigcup_{l=1}^N((x^l,t^l)+T_l)$, $N\leq\exp(Cp)$ and T_l satisfies $i)-iv)$ (recall the inclusion (5.8)). From Lemma 6.11 we obtain

$$\mathbb{E}\sup_{(z,s)\in B_p^X\times T}\sum_{ijk}a_{ijk}z_iY_js_k \lesssim p^{1/2}\max_l \mathbb{E}\sup_{(z,s)\in(x^l,t^l)+T_l}\sum_{ijk}a_{ijk}z_iY_js_k$$
$$+\sup\left\{\sum_{ijk}a_{ijk}z_iy_js_k \mid s\in T, \sum_i \hat{N}_i^X(z_i)\leq p,\ \sum_j \hat{M}_j^Y(y_j)\leq Cp\right\}$$
$$\lesssim p^{1/2}\max_l\mathbb{E}\sup_{(z,s)\in(x^l,t^l)+T_l}\sum_{ijk}a_{ijk}z_iY_js_k + \sup_{s\in T}\left\|\left(\sum_k a_{ijk}s_k\right)\right\|_{X,Y,p}, \tag{5.24}$$

where in the last inequality we used (5.6). By the triangle inequality

$$\mathbb{E}\sup_{(z,s)\in(x^l,t^l)+T_l}\sum_{ijk}a_{ijk}z_iY_js_k \le \mathbb{E}\sup_{(z,s)\in T_l}\sum_{ijk}a_{ijk}z_iY_js_k + \mathbb{E}\sup_{(z,s)\in T_l}\sum_{ijk}a_{ijk}x_i^lY_js_k$$
$$+\mathbb{E}\sup_{(z,s)\in T_l}\sum_{ijk}a_{ijk}z_iY_jt_k^l. \tag{5.25}$$

Corollary 5.3.7 ensures that $(x^l,t^l)\in(p^{-1/2}B_p^X)\times T$ so

$$\mathbb{E}\sup_{(z,s)\in T_l}\sum_{ijk}a_{ijk}x_i^lY_js_k \le p^{-1/2}\sup_{z\in B_p^X}\mathbb{E}\sup_{s\in T}\sum_{ijk}a_{ijk}z_iY_js_k. \tag{5.26}$$

Since, $T_l\subset 2((p^{-1/2}B_p^X\times T))$ Corollary 6.18 yields

$$\mathbb{E}\sup_{(z,s)\in T_l}\sum_{ijk}a_{ijk}z_iY_jt_k^l \le 2p^{-1/2}\sup_{s\in T}\mathbb{E}\sup_{z\in B_p^X}\sum_{ijk}a_{ijk}z_iY_js_k$$
$$\lesssim p^{-1/2}\left(\sup_{s\in T}\left\|\left(\sum_k a_{ijk}s_k\right)_{ij}\right\|_{X,Y,p}+\sup_{s\in T}\left\|\left(\sqrt{\sum_j\left(\sum_k a_{ijk}s_k\right)^2}\right)_i\right\|_{X,p}\right). \tag{5.27}$$

Random variables $Y_1,\ldots,Y_n$ are subgaussian with constant γ so Theorem 6.15 (it is easy to check that processes $(\sum_{ijk}a_{ijk}z_iY_js_k)_{(z,s)\in T_l}$ and $(\sum_{ijk}a_{ijk}z_ig_js_k)_{(z,s)\in T_l}$ satisfy its assumptions)

$$\mathbb{E}\sup_{(z,s)\in T_l}\sum_{ijk}a_{ijk}z_iY_js_k \le C\gamma\mathbb{E}\sup_{(z,s)\in T_l}\sum_{ijk}a_{ijk}z_ig_js_k.$$

Using Proposition 5.3.8 and the properties $i)-iv)$ of Corollary 5.3.7 we obtain

$$\mathbb{E}\sup_{(x,t)\in T_l}\sum_{ijk}a_{ijk}x_ig_jt_k \lesssim p^{-1/2}\left(\mathbb{E}\sup_{s\in T}\sum_{ijk}a_{ijk}g_i\mathcal{E}_js_k+\mathbb{E}\sup_{s\in T}\sum_{ijk}a_{ijk}g_{ij}s_k\right)$$
$$+\sup_{(z,s)\in T_l}\beta_{A,S}(z)+\sup_{(z,s)\in T_l}\mathbb{E}\alpha_A(E_n\otimes s)+p^{1/2}\varDelta_A(T_l)\lesssim p^{-1/2}s_A(T). \tag{5.28}$$

The assertion follows from (5.24)-(5.28). □

Proof of Theorem 5.1.3. It is enough to prove the upper bound (the lower bound follows from Proposition 5.1.2). Observe that it is a consequence of Theorem 5.1.5 provided

$$\mathbb{E}\left\|\sum_{ij}a_{ij}g_{ij}\right\|_{L_q}+\mathbb{E}\left\|\sum_{ij}a_{ij}g_i\mathcal{E}_j\right\|_{L_q}\le C(q)\mathbb{E}\left\|\sum_{ij}a_{ij}X_iY_j\right\|_{L_q}. \tag{5.29}$$

By iteration of Fact 5.1.1 (and using it in the second inequality for $q<2$), Fubini's Theorem, Jensen's inequality and (5.3) we get

$$\mathbb{E}\left\|\sum_{ij} a_{ij}X_iY_j\right\|_{L_q} \geq C(q)^{-1}\left(\mathbb{E}\left\|\sum_{ij} a_{ij}X_iY_j\right\|_{L_q}^q\right)^{1/q} = C(q)^{-1}\left(\int_S \mathbb{E}\left|\sum_{ij} a_{ij}(s)X_iY_j\right|^q d\mu(s)\right)^{1/q}$$

$$\geq C(q)^{-1}\left(\int_S\left(\mathbb{E}\left(\sum_{ij} a_{ij}(s)X_iY_j\right)^2\right)^{q/2} d\mu(s)\right)^{1/q} \geq C(q)^{-1}\left\|\sqrt{\sum_{ij} a_{ij}^2}\right\|_{L_q}.$$

Jensen's inequality, Fubini's Theorem and iteration of Fact 5.1.1 yield

$$\mathbb{E}\left\|\sum_{ij} a_{ij}g_i\mathcal{E}_j\right\|_{L_q} \leq \left(\int_S \mathbb{E}\left|\sum_{ij} a_{ij}(s)g_i\mathcal{E}_j\right|^q d\mu(s)\right)^{1/q}$$

$$\leq C(q)\left(\int_S\left(\mathbb{E}\left(\sum_{ij} a_{ij}(s)g_i\mathcal{E}_j\right)^2\right)^{q/2} d\mu(s)\right)^{1/q} \leq C(q)\left\|\sqrt{\sum_{ij} a_{ij}^2}\right\|_{L_q}.$$

Analogously we show that $\mathbb{E}\left\|\sum_{ij} a_{ij}g_{ij}\right\|_{L_q} \leq C(q)\left\|\sqrt{\sum_{ij} a_{ij}^2}\right\|_{L_q}$ what ends the proof of (5.29).

□

Chapter 6
Appendix

Here we gather results from previous work used in this dissertation. As always C (resp. $C(\alpha)$) stands for a numerical constant (resp. constant that depends only on some parameter α) which may differ at each occurrence. We use the notation $a \sim b$ (resp. $a \sim^{\alpha} b$) to denote that $C^{-1}a \le b \le Ca$ (resp. $C(\alpha)^{-1}a \le b \le C(\alpha)a$). We apply the convention that whenever $X_1, X_2, \dots$ and $Y_1, Y_2, \dots$ are random variables, then $N_i^X(t) = -\ln \mathbb{P}(|X_i| \ge t)$, $M_i^Y(t) = -\ln \mathbb{P}(|Y_i| \ge t)$,

$$\hat{N}_i^X(t) = \begin{cases} t^2 & \text{for } |t| < 1 \\ N_i^X(t) & \text{for } |t| \ge 1 \end{cases}, \qquad \hat{M}_i^Y(t) = \begin{cases} t^2 & \text{for } |t| < 1 \\ M_i^Y(t) & \text{for } |t| \ge 1 \end{cases},$$

$$\|(a_i)_i\|_{X,p} = \sup\left\{\sum_i a_i x_i \;\middle|\; \sum_i \hat{N}_i^X(a_i) \le p\right\}, \quad \|(a_i)_i\|_{Y,p} = \sup\left\{\sum_i a_i y_i \;\middle|\; \sum_i \hat{M}_i^Y(a_i) \le p\right\},$$

$$\|(a_{ij})_{ij}\|_{X,Y,p} = \sup\left\{\sum_{ij} a_{ij} x_i y_j \;\middle|\; \sum_i \hat{N}_i^X(a_i) \le p,\ \sum_j \hat{M}_j^Y(y_j) \le p\right\}$$

where $(a_i)_i \in \mathbb{R}^n$ and $(a_{ij})_{ij}$ is a real-valued matrix.

Corollary 6.1 (Paley-Zygmund's inequality [7, Corollary 3.3.2])**.** *Let X be a nonnegative random variable. Then, for all $0 \le \lambda \le 1$,*

$$\mathbb{P}(X \ge \lambda \mathbb{E}X) \ge (1-\lambda)^2 \frac{(\mathbb{E}X)^2}{\mathbb{E}X^2}.$$

Lemma 6.2. [20, Lemma 3.5] *Let a nondecreasing function $f : \mathbb{R}_+ \to \mathbb{R}_+$ satisfy*

$$f(C\lambda t) \ge \lambda f(t), \quad \textit{for } \lambda \ge 1,\ t \ge t_0,$$

where $t_0 \ge 0$ is a constant. Then there is a convex function $g : \mathbb{R}_+ \to \mathbb{R}_+$ such that

$$g(t) \le f(t) \le g(c^2 t), \quad \textit{for } t \ge Ct_0,$$

and $g(t) = 0$ for $t \in [0, ct_0]$.

Theorem 6.3 (Hypercontractivity of Gaussian chaoses)**.** *Let*

$$S = a + \sum_{i_1} a_{i_1} g_{i_1} + \sum_{i_1,i_2} a_{i_1,i_2} g_{i_1} g_{i_2} + \ldots + \sum_{i_1,\ldots,i_d} a_{i_1,\ldots,i_d} g_{i_1} \cdots g_{i_d},$$

be a non-homogeneous Gaussian chaos of order d with values in a Banach space $(F, \|\cdot\|)$. Then for any $1 \le p < q < \infty$, we have

$$\left(\mathbb{E}\|S\|^q\right)^{1/q} \leq C(d)\left(\frac{q}{p}\right)^{d/2}\left(\mathbb{E}\|S\|^p\right)^{1/p}.$$

Proof. It is an immediate consequence of [7, Theorem 3.2.10] and Hölder's inequality. □

In the below statement δ_{ij} stands for the Kronecker delta.

Theorem 6.4 ([3, Theorem 2.2] for $m=2$)**.** *Let $(a_{ij})_{ij}$ be a symmetric matrix with values in a Banach space $(F,\|\cdot\|)$. Then for any $t \geq 0$, we have*

$$2^{-14}\mathbb{P}\left(\frac{1}{\sqrt{2}}\left\|\sum_{ij} a_{ij}(g_i g_j - \delta_{ij})\right\| \geq t\right) \leq \mathbb{P}\left(\left\|\sum_{ij} a_{ij} g_i g'_j\right\| \geq t\right)$$
$$\leq 2\mathbb{P}\left(\sqrt{2}\left\|\sum_{ij} a_{ij}(g_i g_j - \delta_{ij})\right\| \geq t\right),$$

where g_i, g'_i are i.i.d. $\mathcal{N}(0,1)$ variables.

Theorem 6.5 (Kwapień decoupling bounds [14, Theorem 2])**.** *Let $(F,\|\cdot\|)$ be a Banach space and $(X^i_j)_{i\leq d, j\leq n}, (X_j)_{j\leq n}$ are independent symmetric random variables such that*

$$\forall_{j\leq n}\ X^1_j, X^2_j, \ldots, X^d_j, X_j \text{ are i.i.d.}$$

Assume that array $(a_{i_1,\ldots,i_d})_{i_1,\ldots,i_d}$ with values in F is tetrahedral ($i_k = i_l$ for $k \neq l$, $k,l \leq n$ implies $a_{i_1,\ldots,i_d} = 0$) and symmetric ($a_{i_1,\ldots,i_d} = a_{i_{\pi(1)},\ldots,i_{\pi(d)}}$ for all permutations π of $[d]$). Then for any convex function $\phi : F \to \mathbb{R}_+$

$$\mathbb{E}\phi\left(\frac{1}{C}\sum_{i_1,\ldots,i_d} a_{\mathbf{i}} X^1_{i_1}\cdots X^d_{i_d}\right) \leq \mathbb{E}\phi\left(\sum_{i_1,\ldots,i_d} a_{\mathbf{i}} X_{i_1}\cdot\ldots\cdot X_{i_d}\right)$$
$$\leq \mathbb{E}\phi\left(C\sum_{i_1,\ldots,i_d} a_{\mathbf{i}} X^1_{i_1}\cdots X^d_{i_d}\right).$$

Theorem 6.6 (de la Peña and Montgomery-Smith decoupling bounds [8, Theorem 1])**.** *Under the condition stated in 6.5, but without the assumption that the r.v's are symmetric, we have for any $t \geq 0$,*

$$\frac{1}{C}\mathbb{P}\left(\left\|\sum_{i_1,\ldots,i_d} a_{\mathbf{i}} X^1_{i_1}\cdots X^d_{i_d}\right\| \geq Ct\right) \leq \mathbb{P}\left(\left\|\sum_{i_1,\ldots,i_d} a_{\mathbf{i}} X_{i_1}\cdot\ldots\cdot X_{i_d}\right\| \geq t\right)$$
$$\leq C\mathbb{P}\left(\left\|\sum_{i_1,\ldots,i_d} a_{\mathbf{i}} X^1_{i_1}\cdots X^d_{i_d}\right\| \geq \frac{t}{C}\right).$$

We recall that the entropy number $N(S,\rho,\varepsilon)$ is the smallest number of closed balls with the diameter ε in metric ρ that cover the set S

Theorem 6.7 (Sudakov minoration [29])**.** *For any set $T \subset \mathbb{R}^n$ and $\varepsilon > 0$ we have*

$$\varepsilon\sqrt{\ln N(T,d_2,\varepsilon)} \leq C\mathbb{E}\sup_{t\in T}\sum_i t_i g_i,$$

where d_2 is the Euclidean distance.

Theorem 6.8 (Dual Sudakov minoration [24, formula (3.15)]). *Let α be a norm on $\mathbb{R}^n$ and ρ_α be a distance on $\mathbb{R}^n$ defined by $\rho_\alpha(x,y)=\alpha(x-y)$ for $x,y\in\mathbb{R}^n$. Then for any $\varepsilon>0$ we have*

$$\varepsilon\sqrt{\ln N(B_2^n,\rho_\alpha,\varepsilon)}\le C\mathbb{E}\alpha(G_n).$$

Corollary 6.9 (Corollary 5.7 from [1] with $d=1$). *Let $\mathcal{E}_1,\dots,\mathcal{E}_n$ be i.i.d random variables with the density $1/2\exp(-|t|)$, α be a norm on $\mathbb{R}^n$ and ρ_α be a distance on $\mathbb{R}^n$ defined by $\rho_\alpha(x,y)=\alpha(x-y)$. Then for any $p>0$, $T\subset B_2^n+pB_1^n$, $\varepsilon\in(0,1]$,*

$$N(T,\rho_\alpha,C\varepsilon\mathbb{E}\alpha(\mathcal{E}_1,\dots,\mathcal{E}_n))\le\exp\left(\varepsilon^{-2}+p\varepsilon^{-1}\right).$$

Lemma 6.10. [16, Lemma 3] *Let $(G_t)_{t\in T}$ be a centered Gaussian process and $T=\bigcup_{l=1}^m T_l$, $m\ge1$. Then*

$$\mathbb{E}\sup_{t\in T}G_t\le\max_{l\le m}\mathbb{E}\sup_{t\in T_l}G_t+C\sqrt{\ln(m)}\sup_{t,t'\in T}\sqrt{\mathbb{E}(G_t-G_{t'})^2}.$$

Lemma 6.11. [1, Lemma 5.10] *Let $X_1,\dots,X_n$ be independent, symmetric r.v's with log-concave tails which fulfill the normalization condition*

$$\forall_{i\le n}\ \inf\{t>0:N_i^X(t)\ge1\}=1. \tag{6.1}$$

Then for any sets $T_1,\dots,T_k\subset\mathbb{R}^n$ we have

$$\mathbb{E}\sup_{t\in\bigcup_{l=1}^k T_l}\sum_i t_iX_i\le C\left(\max_{l\le k}\mathbb{E}\sup_{t\in t_l}\sum_i t_iX_i+\sup_{t,t'\in\bigcup_{l=1}^k T_l}\left\|t-t'\right\|_{X,C\ln(k)}\right).$$

Corollary 6.12. [1, Corollary 7.3] *Let $(a_{ijk})_{ijk}$ be a real-valued matrix and $Z_1,\dots,Z_n$ be independent mean zero r.v's. Then for any $p\ge1$ and $T\subset(B_2^n+\sqrt{p}B_1^n)\times(B_2^n+\sqrt{p}B_1^n)$ there exists a decomposition $T=\bigcup_{l=1}^N((x_l,y_l)+T_l)$ such that $N\le\exp(Cp)$, $(x_l,y_l)\in T$ and for every l,*

$$\mathbb{E}\sup_{(x,y)\in T_l}\sum_{ijk}a_{ijk}Z_ix_jy_k\le\frac{C}{\sqrt{p}}\sqrt{\sum_{ijk}a_{ijk}^2}\max_i\|Z_i\|_4.$$

Lemma 6.13. [1, Lemma 9.5] *Let $Y_i^{(1)}$ be independent standard symmetric exponential variables (variables with density $1/2\exp(-|t|)$) and $Y_i^{(2)}=g_i^2$, $Y_i^{(3)}=g_ig_i'$, where g_i,g_i' are i.i.d. $\mathcal{N}(0,1)$ variables and ε_i – i.i.d. Rademacher variables independent of $(Y^{(1)})$, $(Y^{(2)})$, $(Y^{(3)})$. Then for any Banach space $(F,\|\cdot\|)$ and any vectors $v_1,\dots,v_n\in F$ the quantities*

$$\mathbb{E}\left\|\sum_i v_i\varepsilon_iY_i^{(j)}\right\|,\quad j=1,2,3,$$

are comparable up to universal multiplicative factors.

Corollary 6.14 (Dudley's bound [7, Corollary 5.1.6]). *Let $(G_t)_{t\in T}$, be a centered Gaussian process and let $d_G(s,t)=\sqrt{\mathbb{E}(G_t-G_s)^2}$, $s,t\in T$. Then we have*

$$\mathbb{E}\sup_{t\in T}|G_t|\le C\int_0^{\mathrm{diam}(T,d_G)}\sqrt{\ln(N(T,d_G,\varepsilon))}d\varepsilon.$$

Theorem 6.15. [24, Theorem 12.16] *Let $(G_t)_{t\in T}$ be a centered Gaussian process and $(Y_t)_{t\in T}$ be a process such that*

$$\forall_{t,t'\in T}\,\|Y_t-Y_{t'}\|_{\psi_2}\le\gamma\,\|G_t-G_{t'}\|_2\quad (cf.\ (3.29)).$$

Then

$$\mathbb{E}\sup_{t\in T}Y_t\le C\gamma\mathbb{E}\sup_{t\in T}G_t.$$

Lemma 6.16. [19, Lemma 3.2] *Assume that nonnegative r.v's $X_1,\dots,X_n$ have LCT and satisfy the normalization condition* (6.1)*. Let T be a finite subset of $(\mathbb{R}_+)^n$. Then for any $p\ge 1$ we have*

$$\left\|\max_{(t_1,\dots,t_n)\in T}\max\left(\sum_i t_iX_i-C\sum_i t_i,0\right)\right\|_p\le C\,(p+\ln|T|)\max_{(t_1,\dots,t_n)\in T}\max_i t_i.$$

Lemma 6.17. [1, Lemma 6.3] *For any real-valued matrix $(a_{ij})_{ij}$, $p\ge 1$ and $T\subset B_2^n+\sqrt{p}B_1^n$ we have*

$$\mathbb{E}\sup_{x\in T}\sum_{ij}a_{ij}x_ig_j\le C\left(\sqrt{\sum_{ij}a_{ij}^2}+\sqrt{p}\cdot\sup_{x,x'\in T}\sqrt{\sum_j\left(\sum_i a_{ij}(x_i-x_i')\right)^2}\right).$$

Corollary 6.18. [17, Corollary 3] *Assume that $X_1,\dots,X_n$ and $Y_1,\dots,Y_n$ satisfy the assumptions of Lemma 6.11. Then for any $p\ge 1$ and any real-valued matrix $(a_{ij})_{ij}$ we have*

$$\mathbb{E}\left\|\left(\sum_i a_{ij}X_i\right)_j\right\|_{Y,p}\le C\left(\|(a_{ij})\|_{X,Y,p}+\left\|\left(\sqrt{\sum_i a_{ij}^2}\right)_j\right\|_{Y,p}\right).$$

Lemma 6.19. [1, Lemma 6.3] *For any real-valued matrix $(a_{ij})_{ij}$, any $T\subset B_2^n+pB_1^n$ and $p\ge 1$,*

$$\mathbb{E}\sup_{x\in T}\sum_{i,j}a_{ij}x_ig_j\le C\left(\sqrt{\sum_{i,j}a_{ij}^2}+p\cdot\sup_{x,x'\in T}\sqrt{\sum_j\left(\sum_i a_{ij}(x_i-x_i')\right)^2}\right).$$

Lemma 6.20. [16, Lemma 4] *Let G be a Gaussian variable in a normed space $(F,\|\cdot\|)$. Then for any $p\ge 2$*

$$\frac{1}{C}\left(\|G\|_1+\sqrt{p}\sup_{\substack{\varphi\in F^*\\ \|\varphi\|_*\le 1}}\mathbb{E}|\varphi(G)|\right)\le\|G\|_p\le\|G\|_1+C\sqrt{p}\sup_{\substack{\varphi\in F^*\\ \|\varphi\|_*\le 1}}\mathbb{E}|\varphi(G)|,$$

where $(F^*,\|\cdot\|_*)$ is the dual space to $(F,\|\cdot\|)$.

Corollary 6.21. *Assume that for any $i_1,\dots,i_d$, $a_{i_1,\dots,i_d}\in\mathbb{R}$. Then for all $p\ge 1$*

$$\frac{1}{C(d)}\sqrt{p}\sqrt{\sum_{i_1,\dots,i_d}a_{i_1,\dots,i_d}^2}\le\left\|\sum_{i_1,\dots,i_d}a_{i_1,\dots,i_d}g_{i_1}^1\cdots g_{i_d}^d\right\|_p\le C(d)p^{d/2}\sqrt{\sum_{i_1,\dots,i_d}a_{i_1,\dots,i_d}^2}.$$

Proof. It is an easy consequence of [16, Theorem 1] □

Theorem 6.22 was formulated in [9] in a slightly different manner. The below formulation can be found for instance in [17] (Theorem 2 therein).

Theorem 6.22 (Gluskin-Kwapień estimate)**.** *Under the assumption of Lemma 6.11 for any $p\ge 1$, $a_1,\dots a_n\in\mathbb{R}$, we have*

$$\left\|\sum_i a_i X_i\right\| \sim \|(a_i)\|_{X,p}.$$

Theorem 6.23 (Latała bound on L_p-norms of linear forms [15])**.** *Let $X_1, X_2, \dots$ be independent nonnegative random variables or independent symmetric random variables. Then*

$$\left\|\sum_i X_i\right\|_p \sim \inf\left\{t > 0:\ \sum_i \ln\left(\mathbb{E}\left(1 + \frac{X_i}{t}\right)^p\right) \leq p\right\}.$$

Theorem 6.24. [18, Theorem 1] *Let $a_1, \dots, a_n \in F$ where $(F, \|\cdot\|)$ is a Banach space. Assume that $X_1, \dots, X_n$ fulfill the assumption of Lemma 6.11. Then for any $p \geq 1$ we have*

$$\left\|\sum_i a_i X_i\right\|_p \sim \mathbb{E}\left\|\sum_i a_i X_i\right\| + \sup_{\varphi \in B^*(F)} \|(\varphi(a_i))_i\|_{X,p}.$$

Theorem 6.25. [21, Theorem 1.1] *Assume that $X_1, \dots, X_n$ are independent r.v's such that*

$$\forall_i \forall_{p \geq 1} \|X_i\|_{2p} \leq \alpha \|X_i\|_p \text{ for some } \alpha > 0.$$

Then for any non empty set $T \subset \mathbb{R}^n$ and $p \geq 1$ we have

$$\left\|\sup_{t \in T}\left|\sum_i t_i X_i\right|\right\|_p \leq C(d)\left(\mathbb{E}\sup_{t \in T}\left|\sum_i t_i X_i\right| + \sup_{t \in T}\left\|\sum_i t_i X_i\right\|_p\right).$$

Theorem 6.26. [17, Theorem 1] *Let $(a_{ij})_{ij}$ be a real-valued matrix and $X_1, \dots, X_n$, $Y_1, \dots, Y_n$, be independent r.v's which satisfy the assumptions of Lemma 6.11. Then for any $p \geq 1$ we have*

$$\left\|\sum_{ij} a_{ij} X_i Y_j\right\|_p \sim \|(a_{ij})_{ij}\|_{X,Y,p} + \left\|\left(\sqrt{\sum_j a_{ij}^2}\right)_i\right\|_{X,p} + \left\|\left(\sqrt{\sum_i a_{ij}^2}\right)_j\right\|_{Y,p}.$$

In the statement below we use the notation introduced in Section 4.2

Theorem 6.27. [16, Theorem 1] *For any real-valued matrix $(a_{i_1,\dots,i_d})_{i_1,\dots,i_d}$ and $p \geq 2$, we have*

$$\left\|\sum_{\boldsymbol{i}} a_{\boldsymbol{i}} \prod_{j=1}^d g^j_{i_j}\right\|_p \sim^d \sum_{\substack{\mathcal{P} \in \mathcal{P}([d]) \\ \mathcal{P} = (I_1, \dots, I_k)}} p^{|\mathcal{P}|/2} \sup\left\{\sum_{\boldsymbol{i}} a_{\boldsymbol{i}} \prod_{j=1}^k x^j_{i_{I_j}} \ \middle|\ \left\|(x^k_{i_{I_k}})_{i_{I_k}}\right\|_2 \leq 1\right\}.$$

References

[1] Radosław Adamczak and Rafał Latała, *Tail and moment estimates for chaoses generated by symmetric random variables with logarithmically concave tails*, Ann. Inst. Henri Poincaré Probab. Stat. **48** (2012), no. 4, 1103–1136.

[2] Radosław Adamczak and Paweł Wolff, *Concentration inequalities for non-Lipschitz functions with bounded derivatives of higher order*, Probab. Theory Related Fields **162** (2015), no. 3-4, 531–586.

[3] Miguel A. Arcones and Evarist Giné, *On decoupling, series expansions, and tail behavior of chaos processes*, J. Theoret. Probab. **6** (1993), no. 1, 101–122.

[4] Franck Barthe and Emanuel Milman, *Transference principles for log-Sobolev and spectral-gap with applications to conservative spin systems*, Comm. Math. Phys. **323** (2013), no. 2, 575–625.

[5] Christer Borell, *On the Taylor series of a Wiener polynomial*, Seminar Notes on multiple stochastic integration, polynomial chaos and their integration, Case Western Reserve Univ., Cleveland, (1984).

[6] Luis A. Caffarelli, *Monotonicity properties of optimal transportation and the FKG and related inequalities*, Comm. Math. Phys. **214** (2000), no. 3, 547–563.

[7] Víctor H. de la Peña and Evarist Giné, *Decoupling*, Probability and its Applications (New York), Springer-Verlag, New York, 1999. From dependence to independence; Randomly stopped processes. U-statistics and processes. Martingales and beyond.

[8] Victor H. de la Peña and S. J. Montgomery-Smith, *Decoupling inequalities for the tail probabilities of multivariate U-statistics*, Ann. Probab. **23** (1995), no. 2, 806–816.

[9] E. D. Gluskin and S. Kwapień, *Tail and moment estimates for sums of independent random variables with logarithmically concave tails*, Studia Math. **114** (1995), no. 3, 303–309.

[10] D. L. Hanson and F. T. Wright, *A bound on tail probabilities for quadratic forms in independent random variables*, Ann. Math. Statist. **42** (1971), 1079–1083.

[11] P. Hitczenko, S. J. Montgomery-Smith, and K. Oleszkiewicz, *Moment inequalities for sums of certain independent symmetric random variables*, Studia Math. **123** (1997), no. 1, 15–42.

[12] Svante Janson, *Gaussian Hilbert spaces*, Cambridge Tracts in Mathematics, vol. 129, Cambridge University Press, Cambridge, 1997.

[13] Konrad Kolesko and Rafał Latała, *Moment estimates for chaoses generated by symmetric random variables with logarithmically convex tails*, Statist. Probab. Lett. **107** (2015), 210–214.

[14] Stanisław Kwapień, *Decoupling inequalities for polynomial chaos*, Ann. Probab. **15** (1987), no. 3, 1062–1071.

[15] Rafał Latała, *Estimation of moments of sums of independent real random variables*, Ann. Probab. **25** (1997), no. 3, 1502–1513.

[16] Rafał Latała, *Estimates of moments and tails of Gaussian chaoses*, Ann. Probab. **34** (2006), no. 6, 2315–2331.

[17] Rafał Latała, *Tail and moment estimates for some types of chaos*, Studia Math. **135** (1999), no. 1, 39–53.

[18] Rafał Latała, *Tail and moment estimates for sums of independent random vectors with logarithmically concave tails*, Studia Math. **118** (1996), no. 3, 301–304.

[19] Rafał Latała and Rafał Łochowski, *Moment and tail estimates for multidimensional chaos generated by positive random variables with logarithmically concave tails*, Stochastic inequalities and applications, Progr. Probab., vol. 56, Birkhäuser, Basel, 2003, pp. 77–92.

[20] Rafał Latała and Tomasz Tkocz, *A note on suprema of canonical processes based on random variables with regular moments*, Electron. J. Probab. **20** (2015), no. 36, 17.

[21] Rafał Latała and Marta Strzelecka, *Comparison of weak and strong moments for vectors with independent coordinates*, Mathematika **64** (2018), no. 1, 211–229.

[22] Michel Ledoux, *A note on large deviations for Wiener chaos*, Séminaire de Probabilités, XXIV, 1988/89, Lecture Notes in Math., vol. 1426, Springer, Berlin, 1990, pp. 1–14.

[23] Michel Ledoux and Krzysztof Oleszkiewicz, *On measure concentration of vector-valued maps*, Bull. Pol. Acad. Sci. Math. **55** (2007), no. 3, 261–278.

[24] Michel Ledoux and Michel Talagrand, *Probability in Banach spaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], vol. 23, Springer-Verlag, Berlin, 1991.

[25] Rafał M. Łochowski, *Oszacowania momentów i ogonów wieloliniowych form losowych, PHD dissertation in Polish*, Warsaw 2004.

[26] Rafał M. Łochowski, *Moment and tail estimates for multidimensional chaoses generated by symmetric random variables with logarithmically concave tails*, Approximation and probability, Banach Center Publ., vol. 72, Polish Acad. Sci. Inst. Math., Warsaw, 2006, pp. 161–176.

[27] Gilles Pisier, *Probabilistic methods in the geometry of Banach spaces*, Probability and analysis (Varenna, 1985), Lecture Notes in Math., vol. 1206, Springer, Berlin, 1986, pp. 167–241.

[28] Mark Rudelson and Roman Vershynin, *Hanson-Wright inequality and sub-Gaussian concentration*, Electron. Commun. Probab. **18** (2013), no. 82, 9. pp.

[29] V. N. Sudakov, *Gauss and Cauchy measures and ε-entropy*, Dokl. Akad. Nauk SSSR **185** (1969), 51–53 (Russian).

[30] Michel Talagrand, *Upper and lower bounds for stochastic processes*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics], vol. 60, Springer, Heidelberg, 2014. Modern methods and classical problems.

[31] Michel Talagrand, *Sudakov-type minoration for Gaussian chaos processes*, Israel J. Math. **79** (1992), no. 2-3, 207–224.

[32] Roman Vershynin, *High-dimensional probability*, Cambridge Series in Statistical and Probabilistic Mathematics, vol. 47, Cambridge University Press, Cambridge, 2018.

[33] F. T. Wright, *A bound on tail probabilities for quadratic forms in independent random variables whose distributions are not necessarily symmetric*, Ann. Probability **1** (1973), no. 6, 1068–1070.