Moment and tail estimates for multidimensional chaoses generated by positive random variables with logarithmically concave tails

Rafał Latała and Rafał Łochowski

Abstract. In this paper we give estimates of tails and moments of multidimensional chaoses $\sum a_{i_1,\ldots,i_d} X_{i_1}^{(1)} \cdots X_{i_d}^{(d)}$ ($a_{i_1,\ldots,i_d} \geq 0$) generated by positive random variables $X_{i_1}^{(1)}, \ldots, X_{i_d}^{(d)}$ with logarithmically concave tails. The estimates are exact up to constants depending only on the dimension d.

1. Introduction

Let $S = \sum a_{i_1,\ldots,i_d} X_{i_1}^{(1)} \cdots X_{i_d}^{(d)}$ be a random chaos of order d generated by random variables $X_i^{(r)}$. In this paper we present two sided estimates for moments and tails of S in the case when all coefficients are nonnegative and $X_i^{(r)}$ are positive r.v.'s with log-concave tails (precise definitions are given below). Many important inequalities for moments and tails of S are known, even in the more general case of U-statistics (see [2], [4], [7] and [8] for a review of some results). However bounds obtained here are much more precise with constants depending only on the order d.

The estimates for moments of linear combinations (case $d = 1$) of symmetric log-concave random variables were obtained by Gluskin and Kwapień [5]. In the paper [9] similar inequalities were given for chaoses of order 2 generated by such random variables (Gaussian chaoses of order 2 were considered much earlier in [6]). Methods of both papers can be adapted to the nonnegative case (see [10]).

However there seems to be a significant difference between chaoses of order $d \geq 3$ and $d = 2$. As far as we know two sided estimates with deterministic quantities and universal constants are not known for moments and tails of Gaussian chaoses of order 3 (Arcones and Giné [1] obtained some bounds but involving suprema of empirical processes). We hope that our paper is just a first step towards more general results.

1991 *Mathematics Subject Classification.* Primary 60E15 .
Key words and phrases. random chaos, moments.
Research partially supported by KBN grant #2 PO3A 027 22.

Let us comment on the organization of the paper. In the next section we present the necessary notation and definitions and formulate main results. Then, in Section 3, we prove moment estimates for one dimensional case and in Section 4, by induction, for the multidimensional case.

2. Notation and Main Results

Assumptions and notation. Let d and n be positive integers. By $\mathbf{i}$ we denote d-dimensional multiindex $(i_1,\ldots,i_d)$, where $1 \le i_1,\ldots,i_d \le n$ and by $\mathbf{I}$ the set of all such multiindices.

For $I \subset \{1,2,\ldots,d\}$ and $\mathbf{i} \in \mathbf{J} \subset \mathbf{I}$, let $\mathbf{i}_I$ denote $(i_k)_{k\in I}$ and $\mathbf{J}_I$ be the set of all such $\mathbf{i}_I$'s. For $I \subset \{1,2,\ldots,d\}$ let I' be the set $\{1,2,\ldots,n\}\setminus I$. With this notation we have for example $\mathbf{i}_{\{1\}'} = (i_2,\ldots,i_d)$. The symbol $\#I$ denotes the cardinality of the set I.

Henceforth $X^{(1)}_{i_1},\ldots,X^{(d)}_{i_d}$, $1 \le i_1,\ldots,i_d \le n$ be independent positive random variables with logarithmically concave tails so that

$$\mathbf{P}\left(X_i^{(r)} \ge t\right) = e^{-N_i^{(r)}(t)}, \text{ for } 1 \le r \le d, 1 \le i \le n, t \ge 0$$

where $N_i^{(r)} : \mathbb{R}_+ \to \mathbb{R}_+$ is a convex, strictly increasing function normalized in such a way that

$$N_i^{(r)}(1) = 1. \tag{2.1}$$

Let us define a family of subsets of $\mathbb{R}^n_+$, $\left\{B^r_{\mathcal{N},p}\right\}_{1\le r\le d}$, by the formula

$$B^r_{\mathcal{N},p} = \left\{x \in \mathbb{R}^n_+ : \sum_{i=1}^n N_i^{(r)}(x_i) \le p \text{ and } (x_i = 0 \text{ or } x_i \ge 1) \text{ for } i = 1,2,\ldots,n \right\}.$$

For the multidimensional table of reals, $(a_\mathbf{i})_{\mathbf{i}\in\mathbf{I}}$, we set

$$\left\|(a_\mathbf{i})_{\mathbf{i}\in\mathbf{I}}\right\|_{\mathcal{N},u} = \sup\left\{\sum_{\mathbf{i}\in\mathbf{I}} a_\mathbf{i} \prod_{r=1}^d \left(1 + x^{(r)}_{i_r}\right) : x^{(1)} \in B^1_{\mathcal{N},u},\ldots,x^{(d)} \in B^d_{\mathcal{N},u}\right\}.$$

By S we denote a random variable $S = \sum a_\mathbf{i} X^{(1)}_{i_1}\cdots X^{(d)}_{i_d}$. For $p \ge 1$ we put $\|S\|_p = (\mathbf{E}S^p)^{1/p}$.

We use letters c and C to denote universal positive constants, that may change from line to line and $c(d)$, $C(d)$ to denote positive constants, depending only on d ($c(d), C(d)$ may also differ at each occurrence). Finally relation $A \sim_d B$ means that $c(d)A \le B \le C(d)A$.

Now we are ready to formulate two main theorems.

Theorem 2.1. *There exist constants* $0 < c_1(d), C_1(d) < \infty$ *depending only on* d *such that for any* $p \ge 1$ *we have*

$$c_1(d)\left\|(a_\mathbf{i})_{\mathbf{i}\in\mathbf{I}}\right\|_{\mathcal{N},p} \le \|S\|_p \le C_1(d)\left\|(a_\mathbf{i})_{\mathbf{i}\in\mathbf{I}}\right\|_{\mathcal{N},p}.$$

We postpone the proof of Theorem 2.1 till the end of the paper and now present some applications and remarks.

Theorem 2.2. *There exist constants* $0 < c_2(d), C_2(d) < \infty$ *depending only on* d *such that for any* $t \geq 0$ *we have*

$$\mathbf{P}(S \geq C_2(d) \left\|(a_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right\|_{\mathcal{N},t}) \leq e^{-t}$$

and

$$\mathbf{P}(S \geq c_2(d) \left\|(a_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right\|_{\mathcal{N},t}) \geq \min(c_2(d), e^{-t}).$$

Proof. Notice that $\left\|(a_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right\|_{\mathcal{N},t} = \left\|(a_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right\|_{\mathcal{N},1}$ for $0 \leq t \leq 1$, so it is enough to consider only $t \geq 1$. First inequality then immediately follows by Theorem 2.1 and Chebyshev's inequality. To see the second inequality notice that for $p \geq 1$ by Theorem 2.1 we get

$$\|S\|_{2p} \leq C_1(d) \left\|(a_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right\|_{\mathcal{N},2p} \leq 4^d C_1(d) \left\|(a_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right\|_{\mathcal{N},p} \leq C(d)\|S\|_p$$

Now it is enough to notice that $\left\|(a_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right\|_{\mathcal{N},\lambda p} \leq (2\lambda)^d \left\|(a_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right\|_{\mathcal{N},p}$ for $\lambda \geq 1$ and use the Paley-Zygmund inequality as in [5]. □

Remark 2.3. We may define the norms $\left\|(a_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right\|_{\mathcal{N},u}$ in a slightly different, a little bit more natural way. To do this it is enough to observe that

$$\left\|(a_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right\|_{\mathcal{N},u} \leq \sup\left\{\sum_{\mathbf{i}\in\mathbf{I}} a_{\mathbf{i}} \prod_{r=1}^{d}\left(1 + x_{i_r}^{(r)}\right) : \sum_i N_i^{(r)}(x_i^{(r)}) \leq u, r = 1, 2, \ldots, d\right\}$$
$$\leq 2^d \left\|(a_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right\|_{\mathcal{N},u}.$$

Remark 2.4. All estimates obtained in Theorems 2.1 and 2.2 may be extended by simple approximation argument to the situation where functions $N_i^{(r)}$ are still convex but not necesserily continuous or stricly increasing on $\mathbb{R}$, i.e. they may take value 0 on some subinterval of $\mathbb{R}_+$ as well value $+\infty$ beginning from some number t_0 (it means that variable $X_i^{(r)}$ does not admit values greater than t_0). In this case condition (2.1) should be substituted by the following one

$$\inf\left\{t : N_i^{(r)}(t) \geq 1\right\} = 1.$$

Remark 2.5. Theorems 2.1 and 2.2 are also valid (with worse constants) for the undecoupled chaoses of order d, that is for random variables of the form

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

where coefficients $a_{i_1,\ldots,i_d}$ are nonnegative and positive r.v.'s X_i are independent with log-concave tails. It is an immediate consequence of the result of de la Peña and Montgomery-Smith [3] that moments and tails of S are comparable (with constants depending only on d) with moments and tails of the decoupled chaos

$$\tilde{S} = \sum_{\pi} \sum_{1 \le i_1 \le \ldots \le i_d \le n} a_{i_1,\ldots,i_d} X^{(1)}_{i_{\pi(1)}} \cdots X^{(d)}_{i_{\pi(d)}},$$

where the first sum is taken over all permutations π of the set $\{1,\ldots,d\}$ and r.v.'s $X_i^{(r)}$ are independent copies of X_i.

Example. Let us consider a special case, when all variables $X^{(1)}_{i_1},\ldots,X^{(d)}_{i_d}$, $1 \le i_1,\ldots,i_d \le n$ have the exponential distribution, i.e.

$$\mathbf{P}\left(X_i^{(r)} \ge t\right) = e^{-t} \text{ for } 1 \le r \le d, 1 \le i \le n.$$

In this case we may easily find a simpler expression for the norm $\left\|(a_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right\|_{\mathcal{N},u}$. We have

$$\begin{aligned}\left\|(a_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right\|_{\mathcal{N},u} &\le \sup\left\{\sum a_{\mathbf{i}} \prod_{r=1}^{d}\left(1+b^{(r)}_{i_r}\right) : \sum b^{(1)}_{i_1} \le u,\ldots,\sum b^{(d)}_{i_d} \le u\right\} \\ &= \sup\left\{\sum_{I\subset\{1,2,\ldots,d\}} \sum a_{\mathbf{i}} \prod_{r\in I} b^{(r)}_{i_r} : \sum b^{(1)}_{i_1} \le u,\ldots,\sum b^{(d)}_{i_d} \le u\right\}.\end{aligned}$$

So

$$\begin{aligned}\left\|(a_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right\|_{\mathcal{N},u} &\le \sum_{I\subset\{1,2,\ldots,d\}} \sup\left\{\sum a_{\mathbf{i}} \prod_{r\in I} b^{(r)}_{i_r} : \sum b^{(r)}_{i_r} \le u \text{ for } r\in I\right\} \\ &= \sum_{I\subset\{1,2,\ldots,d\}} u^{\#I} \max_{\mathbf{j}\in\mathbf{I}_I} \sum_{\mathbf{i}:\mathbf{i}_I=\mathbf{j}} a_{\mathbf{i}}. \qquad (2.2)\end{aligned}$$

On the other hand

$$\begin{aligned}\left\|(a_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right\|_{\mathcal{N},u} &\ge \max_{I\subset\{1,2,\ldots,d\}} \sup\left\{\sum a_{\mathbf{i}} \prod_{r\in I} b^{(r)}_{i_r} : \sum b^{(r)}_{i_r} \le u \text{ for } r\in I\right\} \\ &\ge 2^{-d} \sum_{I\subset\{1,2,\ldots,d\}} u^{\#I} \max_{\mathbf{j}\in\mathbf{I}_I} \sum_{\mathbf{i}:\mathbf{i}_I=\mathbf{j}} a_{\mathbf{i}}. \qquad (2.3)\end{aligned}$$

From (2.2) and (2.3) and Theorem 2.1 we get that for all $p \ge 1$

$$\begin{aligned}\left\|\sum a_{\mathbf{i}} X^{(1)}_{i_1}\cdots X^{(d)}_{i_d}\right\|_p &\sim_d \sum_{I\subset\{1,2,\ldots,d\}} p^{\#I} \max_{\mathbf{j}\in\mathbf{I}_I} \sum_{\mathbf{i}:\mathbf{i}_I=\mathbf{j}} a_{\mathbf{i}} \\ &= \sum_{i_1,\ldots,i_d} a_{i_1,\ldots,i_d} + p\max_{i_1} \sum_{i_2,\ldots,i_d} a_{i_1,\ldots,i_d} + p\max_{i_2} \sum_{i_1,i_3,\ldots,i_d} a_{i_1,\ldots,i_d} + \ldots + \\ &\quad + \ldots + p^d \max_{i_1,\ldots,i_d} a_{i_1,\ldots,i_d}.\end{aligned}$$

Remark 2.6. In the paper we do not put much emphasis on the rate of growth of constants $c_i^{-1}(d)$ and $C_i(d)$ that appear in Theorems 2.1 and 2.2. Following the presented proofs one can get $c_1^{-1}(d) \leq C^d$ and $c_2^{-1}(d), C_i(d) \leq (Cd)^d$. However there is no indication that this is their optimal rate of growth.

3. One dimensional case

In this section we will prove Theorem 2.1 for $d = 1$. It will be the starting point of an induction proof of the general case. Let us start with the following simple Lemma.

Lemma 3.1. *We have*

$$e^{-1} \leq \mathbf{E}X_i^{(r)} \leq 1 + e^{-1}.$$

Proof. Let us observe that by (2.1), $\mathbf{E}X_i^{(r)} \geq \mathbf{P}(X_i^{(r)} \geq 1) \geq e^{-1}$ and on the other hand by the convexity of $N_i^{(r)}$ we have

$$\mathbf{E}X_i^{(r)} \leq 1 + \int_1^\infty \exp(-N_i^{(r)}(t))dt \leq 1 + \int_1^\infty e^{-t}dt = 1 + e^{-1}.$$

□

In order to ease the notation in the rest of this section, instead of writing $X_1^{(1)}, \ldots, X_n^{(1)}$ we will write $X_1, \ldots, X_n$ and instead of writing $N_1^{(1)}, \ldots, N_n^{(1)}$ we will write $N_1, \ldots, N_n$. So in this case for a sequence (a_i) of nonnegative real numbers and $u > 0$ we have $S = \sum a_i X_i$ and

$$\|(a_i)\|_{\mathcal{N},u} = \sup\left\{\sum a_i\,(1+b_i) : \ (b_i) \in B^1_{\mathcal{N},u}\right\}.$$

We will show that for $p \geq 1$

$$\frac{1}{2e}\|(a_i)\|_{\mathcal{N},p} \leq \|S\|_p \leq 26\|(a_i)\|_{\mathcal{N},p}. \tag{3.1}$$

Proof. First we will prove the estimate from below, by the same method as in [5]. Since $\|S\|_1 \leq \|S\|_p$ by the definition of $\|(a_i)\|_{\mathcal{N},u}$ and Lemma 3.1 it is enough to show that

$$\sum a_i b_i \leq e\|S\|_p \tag{3.2}$$

for every sequence (b_i) of nonnegative reals such that $\sum N_i(b_i) \leq p$. We have

$$\begin{aligned}\|S\|_p = \left\|\sum a_i X_i\right\|_p &\geq \left(\sum a_i b_i\right)(\mathbf{P}(X_i \geq b_i : i = 1, 2, \ldots, n))^{1/p} \\ &= \left(\sum a_i b_i\right)\exp\left(-\frac{1}{p}\sum N_i(b_i)\right) \geq \frac{1}{e}\sum a_i b_i.\end{aligned}$$

So (3.2) holds.

Now we will prove the estimate from above. Let $M_i : \mathbb{R}_+ \to \mathbb{R}_+$ be the inverse of N_i restricted to $\mathbb{R}_+$. The variable X_i has distribution $M_i(\mu)$, where μ is the measure with density $e^{-x} I_{\{x \geq 0\}}$. This means in particular that

$$\mathbf{P}\left(\sum a_i X_i > t\right) = \mu^n \left(x \in \mathbb{R}_+^n : \sum a_i M_i(x_i) > t\right),$$

where μ^n is the product measure $\mu \otimes \mu \otimes \ldots \otimes \mu$ on $\mathbb{R}^n$. Let us define

$$A = \left\{x \in \mathbb{R}_+^n : \sum a_i M_i(x_i) \leq 3 \sum a_i\right\}.$$

By the Chebyshev inequality (since by Lemma 3.1, $\mathbf{E}S \leq \frac{3}{2} \sum a_i$) $\mu^n(A) \geq 1/2$ and therefore by a result of Talagrand (cf. [12] and [11] for a simpler proof) we get

$$\mu^n(A + V_s) \geq 1 - 2e^{-s},$$

where

$$V_s = \{x \in \mathbb{R}^n : \sum (|x_i| - 1)_+ = \sum \max(|x_i| - 1, 0) \leq 5s\}.$$

The last inequality is just an application of Lemma 4 from [11] for the couple $\left(\mu^n, w^{(n)}\right)$ where $w^{(n)}(x) = \sum w(x_i)$ and $w(t) = \frac{1}{5} \max(|t| - 1, 0)$. It is easy to check that the couple $\left(\mu^n, w^{(n)}\right)$ satisfies assumptions of the Lemma 4. Indeed, applying for a moment some notation from [11], it means that the couple $\left(\mu^n, w^{(n)}\right)$ has property (τ), which follows from the Proposition from page 191, the inequality $w \leq W$ and Lemma 1 from page 190 of the cited paper.

Let now $\mathbb{R}_+^n \ni x = y + z$ with $y \in A$ and $z \in V_s$. By the convexity of N_i we have $|M_i(x_i) - M_i(y_i)| \leq M_i(|x_i - y_i|)$, so we obtain

$$\begin{aligned}
\sum a_i M_i(x_i) &\leq \sum a_i M_i(y_i) + \sum a_i |M_i(x_i) - M_i(y_i)| \\
&\leq 3 \sum a_i + \sum a_i M_i(|z_i|) \\
&\leq 4 \sum a_i + \|(a_i)\|_{\mathcal{N},5s},
\end{aligned}$$

since $\sum N_i((M_i(|z_i|) - 1)_+) I_{\{M_i(|z_i|) \geq 1\}} \leq \sum (N_i(M_i(|z_i|)) - 1)_+ = \sum (|z_i| - 1)_+ \leq 5s$. So

$$\mathbf{P}\left(S > 4 \sum a_i + \|(a_i)\|_{\mathcal{N},5s}\right) \leq 2e^{-s} \tag{3.3}$$

and since $\|(a_i)\|_{\mathcal{N},\lambda u} \leq 2\lambda \|(a_i)\|_{\mathcal{N},u}$ for $\lambda \geq 1$ we have for $t \geq 1$

$$\mathbf{P}\left(S > 4 \sum a_i + 2t \|(a_i)\|_{\mathcal{N},u}\right) \leq 2e^{-tu/5}.$$

Therefore, integrating by parts, we get

$$\begin{aligned}\|S\|_p &\le 4\sum a_i + 2\|(a_i)\|_{\mathcal{N},p} + 2\|(a_i)\|_{\mathcal{N},p} \times \\ &\times \left(\int_0^\infty pt^{p-1}\mathbf{P}\left(S > 4\sum a_i + 2\,(1+t)\,\|(a_i)\|_{\mathcal{N},p}\right) dt\right)^{1/p} \\ &\le 4\sum a_i + 2\|(a_i)\|_{\mathcal{N},p}\left(1 + \left(\int_0^\infty 2pt^{p-1}e^{-tp/5}dt\right)^{1/p}\right) \\ &\le 4\sum a_i + 2\|(a_i)\|_{\mathcal{N},p}\left(1 + 5\left(2\Gamma\,(p)\,/p^{p-1}\right)^{1/p}\right) \\ &\le 4\sum a_i + 22\|(a_i)\|_{\mathcal{N},p} \le 26\|(a_i)\|_{\mathcal{N},p}.\end{aligned}$$

and the proof of (3.1) is now completed. □

Now, using (3.3), we will prove the following lemma, which will be used in the next section.

Lemma 3.2. *Let A be a finite subset of $\mathbb{R}^n_+$ and for $\mathbf{a} = (a_1, \ldots, a_n) \in A$ let us define $S_{\mathbf{a}} = \sum a_i X_i$, then the following estimate holds for any $p \ge 1$*

$$\left\|\max_{\mathbf{a}\in A}\left(S_{\mathbf{a}} - C\sum a_i\right)_+\right\|_p \le C\,(p + \ln \#A)\max_{\mathbf{a}\in A}\max_i a_i \tag{3.4}$$

Proof. By the convexity of N_i and by the condition (2.1) we have $N_i\,(x) \ge x$ for $x \ge 1$, thus we get for $s \ge 0$

$$\begin{aligned}\|(a_i)\|_{\mathcal{N},s} &\le \sum a_i + \sup\left\{\sum a_i b_i : \sum b_i \le s\right\} \\ &= \sum a_i + s\max_i a_i.\end{aligned}$$

So, applying (3.3), we obtain

$$\mathbf{P}\left(S_{\mathbf{a}} > 4\sum a_i + 5s\max_i a_i\right) \le 2e^{-s},$$

$$\mathbf{P}\left(\left(S_{\mathbf{a}} - C\sum a_i\right)_+ > Cs\max_{\mathbf{a}\in A}\max_i a_i\right) \le 2e^{-s}.$$

Thus

$$\begin{aligned}&\mathbf{P}\left(\max_{\mathbf{a}\in A}\left(S_{\mathbf{a}} - C\sum a_i\right)_+ \ge Cs\max_{\mathbf{a}\in A}\max_i a_i\right) \\ &\qquad \le \sum_{\mathbf{a}\in A}\mathbf{P}\left(\left(S_{\mathbf{a}} - C\sum a_i\right)_+ \ge Cs\max_{\mathbf{a}\in A}\max_i a_i\right) \le 2\#Ae^{-s}.\end{aligned}$$

and

$$\begin{aligned}&\mathbf{P}\left(\max_{\mathbf{a}\in A}\left(S_{\mathbf{a}} - C\sum a_i\right)_+ \ge C\,(1+s)\,(p + \ln \#A)\max_{\mathbf{a}\in A}\max_i a_i\right) \\ &\qquad\qquad \le 2\#Ae^{-(1+s)(p+\ln \#A)} \le 2e^{-sp}.\end{aligned}$$

Integrating by parts we obtain

$$\begin{aligned}\left\|\max_{\mathbf{a}\in A}\left(S_{\mathbf{a}}-C\sum a_i\right)_+\right\|_p &\leq C\left(p+\ln\#A\right)\max_{\mathbf{a}\in A}\max_i a_i\\ &+C\left(p+\ln\#A\right)\max_{\mathbf{a}\in A}\max_i a_i\left(\int_0^\infty 2pt^{p-1}e^{-pt}dt\right)^{1/p}\\ &\leq C\left(p+\ln\#A\right)\max_{\mathbf{a}\in A}\max_i a_i.\end{aligned}$$

□

4. Multidimensional case

In this section we will prove Theorem 2.1 proceeding by induction on d. The first step was already established in the previous section. So in the sequel we will assume that Theorem 2.1 holds for $d-1$ and we will show it for d. First we will show the estimate from below that is

$$\|S\|_p \geq c_1(d)\left\|(a_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right\|_{\mathcal{N},p}. \tag{4.1}$$

Proof. We will proceed in the similar way as in the proof of the lower estimate in the previous section. By the induction assumption and Lemma 3.1

$$\begin{aligned}\|S\|_p &\geq c_1(d-1)\left\|\left\|\left(\sum_{\mathbf{i}:\mathbf{i}_{\{1\}'}=\mathbf{j}} a_{\mathbf{i}}X_{i_1}^{(1)}\right)_{\mathbf{j}\in\mathbf{I}_{\{1\}'}}\right\|_{\mathcal{N},p}\right\|_p\\ &\geq c_1(d-1)\,\mathbf{E}\left\|\left(\sum_{\mathbf{i}:\mathbf{i}_{\{1\}'}=\mathbf{j}} a_{\mathbf{i}}X_{i_1}^{(1)}\right)_{\mathbf{j}\in\mathbf{I}_{\{1\}'}}\right\|_{\mathcal{N},p}\\ &\geq c_1(d-1)\sup\left\{\mathbf{E}\sum_{\mathbf{i}} a_{\mathbf{i}}X_{i_1}^{(1)}\prod_{r>1}\left(1+x_{i_r}^{(r)}\right): x^{(r)}\in B^r_{\mathcal{N},u} \text{ for } r>1\right\}\\ &\geq e^{-1}c_1(d-1)\sup\left\{\sum_{\mathbf{i}} a_{\mathbf{i}}\prod_{r>1}\left(1+x_{i_r}^{(r)}\right): x^{(r)}\in B^r_{\mathcal{N},u} \text{ for } r>1\right\}. \qquad (4.2)\end{aligned}$$

Let us now assume that (b_i) is a sequence of nonnegative reals with $\sum N_i^{(1)}(b_i)\leq p$, then again by the induction assumption

$$\|S\|_p \geq c_1(d-1) \left\| \left\| \left(\sum_{\mathbf{i}:\mathbf{i}_{\{1\}'}=\mathbf{j}} a_{\mathbf{i}} X_{i_1}^{(1)} \right)_{\mathbf{j}\in \mathbf{I}_{\{1\}'}} \right\|_{\mathcal{N},p} \right\|_p$$

$$\geq c_1(d-1) \left(\left\| \left(\sum_{\mathbf{i}:\mathbf{i}_{\{1\}'}=\mathbf{j}} a_{\mathbf{i}} b_{i_1} \right)_{\mathbf{j}\in \mathbf{I}_{\{1\}'}} \right\|_{\mathcal{N},p}^p \mathbf{P}\left(X_i^{(1)} \geq b_i,\ i=1,2,\ldots,n \right) \right)^{1/p}$$

$$\geq c_1(d-1) \left\| \left(\sum_{\mathbf{i}:\mathbf{i}_{\{1\}'}=\mathbf{j}} a_{\mathbf{i}} b_{i_1} \right)_{\mathbf{j}\in \mathbf{I}_{\{1\}'}} \right\|_{\mathcal{N},p} \exp\left(-\frac{1}{p} \sum_{i=1}^n N_i^{(1)}(b_i) \right)$$

$$\geq e^{-1} c_1(d-1) \sup \left\{ \sum_{\mathbf{i}\in\mathbf{I}} a_{\mathbf{i}} b_{i_1} \prod_{r=2}^n \left(1 + x_{i_r}^{(r)} \right) : x^{(2)} \in B^1_{\mathcal{N},u}, \ldots, x^{(d)} \in B^d_{\mathcal{N},u} \right\}.$$

Combining the above with (4.2) we get the lower estimate (4.1) with $c_1(d) = (2e)^{-1} c_1(d-1)$. □

In order to prove the estimate from above that is

$$\|S\|_p \leq C_1(d) \left\| (a_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}} \right\|_{\mathcal{N},p}. \tag{4.3}$$

we will need following two observations.

1. We may rearrange indexes $i_1, \ldots, i_d$ in such a way that for any $1 \leq r \leq d$

$$i < j \Rightarrow \sum_{\mathbf{i}:i_r=i} a_{\mathbf{i}} \geq \sum_{\mathbf{i}:i_r=j} a_{\mathbf{i}}. \tag{4.4}$$

2. We may split $\mathbf{I}$ into sum of disjoint sets $\mathbf{I}_r, 1 \leq r \leq d$, in the following way.

   $$\mathbf{I}_r = \{\mathbf{i} \in \mathbf{I} : i_r > \max(i_1, \ldots, i_{r-1}) \text{ and } i_r \geq \max(i_{r+1}, \ldots, i_d)\}.$$

   Assuming that (4.3) holds for $d-1$, we may write

$$\|S\|_p = \left\| \sum_{r=1}^d \sum_{\mathbf{i}\in\mathbf{I}_r} a_{\mathbf{i}} X_{i_1}^{(1)} \cdots X_{i_d}^{(d)} \right\|_p \leq \sum_{r=1}^d \left\| \sum_{\mathbf{i}\in\mathbf{I}_r} a_{\mathbf{i}} X_{i_1}^{(1)} \cdots X_{i_d}^{(d)} \right\|_p$$

$$\leq C_1(d-1) \sum_{r=1}^d \left\| \left\| \left(\sum_{\mathbf{i}:\mathbf{i}_{\{r\}'}=\mathbf{j}} a_{\mathbf{i}} X_{i_r}^{(r)} \right)_{\mathbf{j}\in(\mathbf{I}_r)_{\{r\}'}} \right\|_{\mathcal{N},p} \right\|_p.$$

So to show (4.3) it is sufficient to get for $r = 1, \ldots, d$

$$\left\| \left\| \left(\sum_{\mathbf{i}:\mathbf{i}_{\{r\}'}=\mathbf{j}} a_{\mathbf{i}} X^{(r)}_{i_r} \right)_{\mathbf{j}\in(\mathbf{I}_r)_{\{r\}'}} \right\|_{\mathcal{N},p} \right\|_p \leq C(d) \left\| (a_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}} \right\|_{\mathcal{N},p}. \tag{4.5}$$

Proof. We will consider only the summand corresponding to $r = 1$, since the proof for other values of r is entirely similar. Recall that $C(d)$ denotes the constant that may differ at each occurrence and that depends only on d (in particular it may depend on values of $c_1(d-1)$ and $C_1(d-1)$).

To ease the notation let us set $\mathbf{J}_1 = (\mathbf{I}_1)_{\{1\}'}$ and additionally define for $I \subset \{1, \ldots, d\}$

$$\left\| (a_{\mathbf{j}})_{\mathbf{j}\in\mathbf{I}_I} \right\|_{\mathcal{N},u} = \sup \left\{ \sum_{\mathbf{j}\in\mathbf{I}_I} a_{\mathbf{j}} \prod_{r\in I} \left(1 + x^{(r)}_{j_r}\right) : x^{(r)} \in B^r_{\mathcal{N},u} \text{ for } r \in I \right\}.$$

We have

$$\begin{aligned} &\left\| \left(\sum_{\mathbf{i}:\mathbf{i}_{\{1\}'}=\mathbf{j}} a_{\mathbf{i}} X^{(1)}_{i_1} \right)_{\mathbf{j}\in\mathbf{J}_1} \right\|_{\mathcal{N},p} \\ &\qquad = \sup_{x^{(r)}\in B^r_{\mathcal{N},p}, 2\leq r\leq d} \sum_{\mathbf{j}\in\mathbf{J}_1} \left(\sum_{\mathbf{i}\in\mathbf{I}_1:\mathbf{i}_{\{1\}'}=\mathbf{j}} a_{\mathbf{i}} X^{(1)}_{i_1} \right) \left(1 + x^{(2)}_{j_2}\right) \cdots \left(1 + x^{(d)}_{j_d}\right) \end{aligned}$$

and for $x^{(r)} \in \mathbb{R}^n_+$, $2 \leq r \leq d$,

$$\left(1 + x^{(2)}_{j_2}\right) \cdots \left(1 + x^{(d)}_{j_d}\right) \leq \sum_{2\leq q\leq d} \prod_{\substack{r=2 \\ r\neq q}}^{d} \left(1 + x^{(r)}_{j_r}\right) + x^{(2)}_{j_2} \cdots x^{(d)}_{j_d}.$$

Therefore

$$\begin{aligned} &\left\| \left(\sum_{\mathbf{i}:\mathbf{i}_{\{1\}'}=\mathbf{j}} a_{\mathbf{i}} X^{(1)}_{i_1} \right)_{\mathbf{j}\in\mathbf{J}_1} \right\|_{\mathcal{N},p} \\ &\quad \leq \sum_{2\leq q\leq d} \sup_{x^{(r)}\in B^r_{\mathcal{N},p}, 2\leq r\leq d} \sum_{\mathbf{j}\in\mathbf{J}_1} \left(\sum_{\mathbf{i}\in\mathbf{I}_1:\mathbf{i}_{\{1\}'}=\mathbf{j}} a_{\mathbf{i}} X^{(1)}_{i_1} \right) \prod_{\substack{r=2, \\ r\neq q}}^{n} \left(1 + x^{(r)}_{j_r}\right) \\ &\qquad\qquad + \sup_{x^{(r)}\in B^r_{\mathcal{N},p}, 2\leq r\leq d} \sum_{\mathbf{j}\in\mathbf{J}_1} \left(\sum_{\mathbf{i}\in\mathbf{I}_1:\mathbf{i}_{\{1\}'}=\mathbf{j}} a_{\mathbf{i}} X^{(1)}_{i_1} \right) x^{(2)}_{j_2} \cdots x^{(d)}_{j_d}. \end{aligned}$$

By the inductive hypothesis, for any $2 \le q \le d$,

$$\left\| \sup_{x^{(r)} \in B^r_{\mathcal{N},p}, 2\le r\le d} \sum_{\mathbf{j}\in \mathbf{J}_1} \left(\sum_{\mathbf{i}\in\mathbf{I}_1 : \mathbf{i}_{\{1\}'}=\mathbf{j}} a_{\mathbf{i}} X^{(1)}_{i_1} \right) \prod_{\substack{r=2\\ r\neq q}}^n \left(1 + x^{(r)}_{j_r}\right) \right\|_p$$

$$= \left\| \sup_{x^{(r)} \in B^r_{\mathcal{N},p}, 2\le r\le d} \sum_{\mathbf{j}\in (\mathbf{I}_1)_{\{q\}'}} \left(\sum_{\mathbf{i}\in\mathbf{I}_1 : \mathbf{i}_{\{q\}'}=\mathbf{j}} a_{\mathbf{i}} \right) X^{(1)}_{j_1} \prod_{\substack{r=2\\ r\neq q}}^n \left(1 + x^{(r)}_{j_r}\right) \right\|_p$$

$$\le c_1^{-1}(d-1) \left\| \sum_{\mathbf{j}\in (\mathbf{I}_1)_{\{q\}'}} \left(\sum_{\mathbf{i}\in\mathbf{I}_1 : \mathbf{i}_{\{q\}'}=\mathbf{j}} a_{\mathbf{i}} \right) \prod_{\substack{r=1\\ r\neq q}}^n X^{(r)}_{j_r} \right\|_p$$

$$\le C(d) \left\| \left(\sum_{\mathbf{i}\in\mathbf{I}_1 : \mathbf{i}_{\{q\}'}=\mathbf{j}} a_{\mathbf{i}} \right)_{\mathbf{j}\in (\mathbf{I}_1)_{\{q\}'}} \right\|_{\mathcal{N},p} \le C(d) \left\| (a_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}} \right\|_{\mathcal{N},p}.$$

So it suffices to show that
(4.6)

$$\| \sup_{x^{(r)} \in B^r_{\mathcal{N},p}, 2\le r\le d} \sum_{\mathbf{j}\in \mathbf{J}_1} (\sum_{\mathbf{i}\in\mathbf{I}_1 : \mathbf{i}_{\{1\}'}=\mathbf{j}} a_{\mathbf{i}} X^{(1)}_{i_1}) x^{(2)}_{j_2} \cdots x^{(d)}_{j_d} \|_p \le C(d) \left\| (a_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}} \right\|_{\mathcal{N},p}.$$

Let us define for $2 \le r \le d$ the following sets

$$U^{(r)} = \left\{ x^{(r)} \in B^r_{\mathcal{N},p} : \ N^{(r)}_i \left(x^{(r)}_i \right) \le 1 \text{ if } i < p \right\} \cap$$
$$\cap \bigcap_{k=1}^{\infty} \left\{ x^{(r)} \in B^{(r)}_{\mathcal{N},p} : N^{(r)}_i \left(x^{(r)}_i \right) \le k^3 \text{ if } i \in \left[2^{k-1}p; 2^k p\right) \right\},$$

$$V^{(r)} = \left\{ x^{(r)} \in B^r_{\mathcal{N},p} : x^{(r)}_i = 0 \text{ or } N^{(r)}_i \left(x^{(r)}_i \right) > 1 \text{ if } i < p \right\} \cap$$
$$\cap \bigcap_{k=1}^{\infty} \left\{ x^{(r)} \in B^r_{\mathcal{N},p} : x^{(r)}_i = 0 \text{ or } N^{(r)}_i \left(x^{(r)}_i \right) > k^3 \text{ if } i \in \left[2^{k-1}p; 2^k p\right) \right\}.$$

For every $x^{(r)} \in B^r_{\mathcal{N},p}$ there exist $y^{(r)} \in U^{(r)}, z^{(r)} \in V^{(r)}$ such that $x^{(r)} = y^{(r)} + z^{(r)}$. Let $\mathcal{J}$ denote a family of subsets of indices defined in the following way:

$$\mathcal{J} = \left\{ I \subseteq \{1, 2, \ldots\} : \#I \le p, \#I \cap \left[2^{k-1}p; 2^k p\right) \le \frac{p}{k^3} \text{ for } k = 1, 2, \ldots \right\}.$$

Applying the inequality $\sum_{k\le m} \binom{n}{k} \le \left(\frac{Cn}{m}\right)^m$, we get an estimate of the cardinality of $\mathcal{J}$.

$$\#\mathcal{J} \le 2^p \prod_{k=1}^{\infty} \left(\frac{C 2^{k-1} p}{p/k^3} \right)^{p/k^3} \le C^p \prod_{k=1}^{\infty} \left(2^{k-1} k^3\right)^{p/k^3} \le C^p. \tag{4.7}$$

Let us notice that for I such that $\#I \le p$ and $2 \le q \le d$

$$\left\| \sup_{x^{(r)} \in B^r_{\mathcal{N},u}, 2 \le r \le d} \sum_{\mathbf{i} \in \mathbf{I}_1 : i_q \in I} a_{\mathbf{i}} X^{(1)}_{i_1} x^{(2)}_{i_2} \cdots x^{(d)}_{i_d} \right\|_p$$

$$\le c_1^{-1}(d-1) \left\| \sum_{\mathbf{i} \in \mathbf{I}_1 : i_q \in I} a_{\mathbf{i}} X^{(1)}_{i_1} X^{(2)}_{i_2} \cdots X^{(d)}_{i_d} \right\|_p$$

$$\le c_1^{-1}(d-1) \left\| \sup_{x^{(r)} \in B^r_{\mathcal{N},u}, r \in \{q\}'} \sum_{\mathbf{i} \in \mathbf{I}_1 : i_q \in I} a_{\mathbf{i}} \left(1 + x^{(1)}_{i_1}\right) \cdots X^{(q)}_{i_q} \cdots \left(1 + x^{(d)}_{i_d}\right) \right\|_p$$

$$\le C(d) \left\| (a_{\mathbf{i}})_{\mathbf{i} \in \mathbf{I}_1} \right\|_{\mathcal{N},p}. \tag{4.8}$$

The last estimate follows by integration by parts from the following two remarks.

- The function

$$x^{(q)} \longmapsto \left\| \sup_{x^{(r)} \in B^r_{\mathcal{N},u}, r \in \{q\}'} \sum_{\mathbf{i} \in \mathbf{I}_1 : i_q \in I} a_{\mathbf{i}} (1 + x^{(1)}_{i_1}) \cdots x^{(q)}_{i_q} \cdots (1 + x^{(d)}_{i_d}) \right\|_p$$

  is positive homogeneous on $\mathbb{R}^n_+$.
- The random variable $\sum_{i_q \in I} N^{(q)}_{i_q} \left(X^{(q)}_{i_q}\right)$ has the distribution $\Gamma(\#I, 1)$ and therefore (recall that $\#I \le p$)

$$\mathbf{P}\left(\sum_{i_q \in I} N^{(q)}_{i_q} \left(\frac{1}{1+t} X^{(q)}_{i_q} \right) > p \right) \le \mathbf{P}\left(\sum_{i_q \in I} N^{(q)}_{i_q} \left(X^{(q)}_{i_q} \right) > (1+t)\, p \right) \le e^{-Ctp}.$$

From (4.8) and (4.7) it follows that

$$\left\| \sup_{x^{(r)} \in B^r_{\mathcal{N},u}, 2 \le r \le d, x^{(q)} \in V^{(q)}} \sum_{\mathbf{i} \in \mathbf{I}_1} a_{\mathbf{i}} X^{(1)}_{i_1} x^{(2)}_{i_2} \cdots x^{(d)}_{i_d} \right\|_p$$

$$\le \left\| \sup_{I \in \mathcal{J}} \sup_{x^{(r)} \in B^r_{\mathcal{N},u}, 2 \le r \le d} \sum_{\mathbf{i} \in \mathbf{I}_1 : i_q \in I} a_{\mathbf{i}} X^{(1)}_{i_1} x^{(2)}_{i_2} \cdots x^{(d)}_{i_d} \right\|_p$$

$$\le \left(\sum_{I \in \mathcal{J}} \left\| \sup_{x^{(r)} \in B^r_{\mathcal{N},u}, 2 \le r \le d} \sum_{\mathbf{i} \in \mathbf{I}_1 : i_q \in I} a_{\mathbf{i}} X^{(1)}_{i_1} x^{(2)}_{i_2} \cdots x^{(d)}_{i_d} \right\|_p^p \right)^{1/p}$$

$$\le C(d) \left((\#\mathcal{J}) \left\| (a_{\mathbf{i}})_{\mathbf{i} \in \mathbf{I}_1} \right\|^p_{\mathcal{N},p} \right)^{1/p} \le C(d) \left\| (a_{\mathbf{i}})_{\mathbf{i} \in \mathbf{I}_1} \right\|_{\mathcal{N},p}. \tag{4.9}$$

Let C_0 be the constant given by Lemma 3.2. Since

$$\sup_{x^{(r)}\in U^{(r)},2\le r\le d}\sum_{\mathbf{i}\in\mathbf{I}_1} a_{\mathbf{i}}C_0 x^{(2)}_{i_2}\cdots x^{(d)}_{i_d}\le C_0\left\|(a_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right\|_{\mathcal{N},p}$$

it follows by (4.9) that in order to prove (4.6) it is enough to show

$$\left\|\sup_{x^{(r)}\in U^{(r)},2\le r\le d}\sum_{\mathbf{i}\in\mathbf{I}_1} a_{\mathbf{i}}(X^{(1)}_{i_1}-C_0)x^{(2)}_{i_2}\cdots x^{(d)}_{i_d}\right\|_p\le C(d)\left\|(a_{\mathbf{i}})_{\mathbf{i}\in\mathbf{I}}\right\|_{\mathcal{N},p}. \tag{4.10}$$

Now notice that since $N^{(r)}_i(x)\ge x$ for $x\ge 1, 1\le r\le d, 1\le i\le n$, then for $a\ge 1$ and arbitrary numbers $t_1,t_2,\ldots,t_n$

$$\begin{aligned}&\sup\left\{\sum_{\mathbf{i}\in\mathbf{I}_1} a_{\mathbf{i}}t_{i_1}x^{(2)}_{i_2}\cdots x^{(d)}_{i_d}:\ x^{(r)}\in B^r_{\mathcal{N},u},\left\|\left(N^{(r)}_{i_r}\left(x^{(r)}_{i_r}\right)\right)\right\|_\infty\le a, 2\le r\le d\right\}\\ &\le\sup\left\{\sum_{\mathbf{i}\in\mathbf{I}_1} a_{\mathbf{i}}t_{i_1}x^{(2)}_{i_2}\cdots x^{(d)}_{i_d}:\sum x^{(r)}_{i_r}\le p,\left\|x^{(r)}\right\|_\infty\le a, 2\le r\le d\right\}\\ &\le a^{d-1}\max_{\#I^r\le\lceil p/a\rceil,2\le r\le d}\ \sum_{\mathbf{i}\in\mathbf{I}_1:i_r\in I^r,2\le r\le d} a_{\mathbf{i}}t_{i_1}\end{aligned}$$

(above and below we use a standard convention $\sum_{\mathbf{i}\in\emptyset} b_{\mathbf{i}}=0$). Hence by the definition of $\mathbf{I}_1$ and $U^{(r)}$ we get

$$\begin{aligned}&\left\|\sup_{x^{(r)}\in U^{(r)},2\le r\le d}\sum_{\mathbf{i}\in\mathbf{I}_1} a_{\mathbf{i}}(X^{(1)}_{i_1}-C_0)x^{(2)}_{i_2}\cdots x^{(d)}_{i_d}\right\|_p\\ &\le\sum_{l\ge1}\left\|\sup_{x^{(r)}\in U^{(r)},2\le r\le d}\ \sum_{\mathbf{i}\in\mathbf{I}_1:2^{l-1}p\le i_1<2^l p} a_{\mathbf{i}}(X^{(1)}_{i_1}-C_0)x^{(2)}_{i_2}\cdots x^{(d)}_{i_d}\right\|_p\\ &\le\sum_{1\le l\le p^{1/3}} l^{3(d-1)}\left\|\max_{\substack{I^r\subset\{1,\ldots,2^l p-1\}\\ \#I^r\le\lceil p/l^3\rceil,2\le r\le d}}\ \sum_{\substack{\mathbf{i}\in\mathbf{I}_1:2^{l-1}p\le i_1<2^l p,\\ i_r\in I^r,2\le r\le d}} a_{\mathbf{i}}(X^{(1)}_{i_1}-C_0)\right\|_p \qquad (4.11)\\ &+p^{d-1}\sum_{l>p^{1/3}}\left\|\max_{\mathbf{j}\in\mathbf{I}_1:2^{l-1}p\le j_1<2^l p}\left(\sum_{\substack{\mathbf{i}\in\mathbf{I}_1:\mathbf{i}_{\{1\}'}=\mathbf{j}_{\{1\}'}\\ 2^{l-1}p\le i_1<2^l p}} a_{\mathbf{i}}X^{(1)}_{i_1}-C_0\right)_+\right\|_p. \qquad (4.12)\end{aligned}$$

Applying Lemma 3.2 we will estimate (4.11) and (4.12). For $l \le p^{1/3}$ we have

$$\# \left\{ (I^r)_{1\le r\le d} : I^r \subset \left\{1,\ldots,2^l p-1\right\}, \# I^r \le \lceil p/l^3 \rceil \text{ for } 1\le r\le d \right\}$$
$$\le \left(\left(\frac{C2^l p}{p/l^3} \right)^{2p/l^3} \right)^d \le C^{dp/l^2} \le C^{dp}. \tag{4.13}$$

Applying the ordering (4.4) we obtain

$$\max_{\substack{I^r\subset\left\{1,\ldots,2^l p-1\right\}\\ \# I^r\le\lceil p/l^3\rceil, 2\le r\le d}} \max_{2^{l-1}p\le i<2^l p} \sum_{\substack{\mathbf{i}\in\mathbf{I}_1: i_1=i,\\ i^{(r)}\in I^r, 2\le r\le d}} a_{\mathbf{i}}$$
$$\le \max_{i\ge 2^{l-1}p} \sum_{\mathbf{i}\in\mathbf{I}: i_1=i} a_{\mathbf{i}} \le \frac{1}{2^{l-1}p}\sum_{\mathbf{i}\in\mathbf{I}} a_{\mathbf{i}}. \tag{4.14}$$

From (3.4), (4.13) and (4.14) we get

$$\left\| \max_{\substack{I^r\subset\left\{1,\ldots,2^l p-1\right\}\\ \# I^r\le\lceil p/l^3\rceil, 2\le r\le d}} \sum_{\substack{\mathbf{i}\in\mathbf{I}_1: 2^{l-1}p\le i_1<2^l p,\\ i_r\in I^r, 2\le r\le d}} a_{\mathbf{i}} \left(X^{(1)}_{i_1} - C_0 \right) \right\|_p$$
$$\le C\left(p+\ln C^{dp}\right) \max_{\substack{I^r\subset\left\{1,\ldots,2^l p-1\right\}\\ \# I^r\le\lceil p/l^3\rceil, 2\le r\le d}} \max_{2^{l-1}p\le i<2^l p} \sum_{\substack{\mathbf{i}\in\mathbf{I}_1: i_1=i,\\ i^{(r)}\in I^r, 2\le r\le d}} a_{\mathbf{i}}$$
$$\le Cpd\frac{1}{2^{l-1}p}\sum_{\mathbf{i}\in\mathbf{I}} a_{\mathbf{i}} \le \frac{Cd}{2^{l-1}}\sum_{\mathbf{i}\in\mathbf{I}} a_{\mathbf{i}}.$$

So

$$\sum_{1\le l\le p^{1/3}} l^{3(d-1)} \left\| \max_{\substack{I^r\subset\left\{1,\ldots,2^l p-1\right\},\\ \# I^r\le\lceil p/l^3\rceil, 2\le r\le d}} \sum_{\substack{\mathbf{i}\in\mathbf{I}_1: i_1=i,\\ i^{(r)}\in I^r, 2\le r\le d}} a_{\mathbf{i}}(X^{(1)}_{i_1} - C_0) \right\|_p$$
$$\le \sum_{1\le l\le p^{1/3}} l^{3(d-1)} \frac{Cd}{2^{l-1}} \sum_{\mathbf{i}\in\mathbf{I}} a_{\mathbf{i}} \le C(d)\sum_{\mathbf{i}\in\mathbf{I}} a_{\mathbf{i}}. \tag{4.15}$$

For $l > p^{1/3}$ we similarly get

$$\max_{\mathbf{i}\in\mathbf{I}_1: 2^{l-1}p\le i_1<2^l p} a_{\mathbf{i}} \le \frac{1}{2^{l-1}p}\sum_{\mathbf{i}\in\mathbf{I}} a_{\mathbf{i}}$$

and (since $\# \left\{ \mathbf{j} \in \mathbf{I}_1 : 2^{l-1}p \le j_1 < 2^l p \right\} \le (2^l p)^d$)

$$\left\| \max_{\mathbf{j}\in\mathbf{I}_1 : 2^{l-1}p \le j_1 < 2^l p} \left(\sum_{\substack{\mathbf{i}\in\mathbf{I}_1 : \mathbf{i}_{\{1\}'} = \mathbf{j}_{\{1\}'} \\ 2^{l-1}p \le i_1 < 2^l p}} a_{\mathbf{i}} (X^{(1)}_{i_1} - C_0) \right)_+ \right\|_p \le \frac{Cld}{2^{l-1}} \sum_{\mathbf{i}\in\mathbf{I}} a_{\mathbf{i}}.$$

So we obtain

$$p^{d-1} \sum_{l > p^{1/3}} \left\| \max_{\mathbf{j}\in\mathbf{I}_1 : 2^{l-1}p \le j_1 < 2^l p} \left(\sum_{\substack{\mathbf{i}\in\mathbf{I}_1 : \mathbf{i}_{\{1\}'} = \mathbf{j}_{\{1\}'} \\ 2^{l-1}p \le i_1 < 2^l p}} a_{\mathbf{i}} (X^{(1)}_{i_1} - C_0) \right)_+ \right\|_p \tag{4.16}$$

$$\le p^{d-1} \sum_{l > p^{1/3}} \frac{Cld}{2^{l-1}} \sum_{\mathbf{i}\in\mathbf{I}} a_{\mathbf{i}} \le C(d) \sum_{\mathbf{i}\in\mathbf{I}} a_{\mathbf{i}}. \tag{4.17}$$

Finally, collecting (4.11), (4.12), (4.15) and (4.17) we get the desired estimate (4.10).

□

References

[1] M. Arcones and E. Giné, *On decoupling, series expansions, and tail behavior of chaos processes*, J. Theoret. Probab. **6** (1993), 101-122.

[2] V. H. de la Peña and E. Giné, *Decoupling: From Dependence to Independence*, Springer Verlag, New York 1999.

[3] V. H. de la Peña and S. Montgomery-Smith, *Bounds for the tail probabilities of U-statistics and quadratic forms*, Bull. Amer. Math. Soc. **31** (1994), 223–227.

[4] E. Giné, R. Latała and J. Zinn, *Exponential and Moment Inequalities for U-statistics*, High Dimensional Probability II, 13–38, Progr. Probab. 47, Birkhauser, Boston 2000.

[5] E. D. Gluskin and S. Kwapień, *Tail and moment estimates for sums of independent random variables with logarithmically concave tails*, Studia Math. **114** (1995), 303-309.

[6] 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.

[7] R. Ibragimov and Sh. Sharakhmetov, *Analogues of Khintchin, Marcinkiewicz-Zygmund and Rosenthal inequalities for symmetric statistics*, Scand. J. Statist. **26** (1999), 621-623.

[8] M. Klass and K. Nowicki *A symmetrization-desymmetrization procedure for uniformly good approximations of expectations involving arbitrary sums of generalized U-statistics*, Ann. Probab. **28** (2000), 1884–1907.

[9] R. Latała, *Tail and moment estimates for some type of chaos*, Studia Math. **135**, 39–53.

[10] R. Łochowski, *Estimation of moments of sums of independent random variables and of two dimensional chaos*, M. Sc. thesis, Warsaw University 2000 (in Polish.)

[11] B. Maurey, *Some deviation inequalities*, Geometric and Functional Analysis **1** (1991), 188-197.

[12] M. Talagrand *A new isoperimetric inequality and the concentration of measure phenomenon*, Israel Seminar (GAFA), Springer Verlag Lecture Notes in Math. 1469 (1991), 94-124.

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