ESTIMATES OF MOMENTS AND TAILS OF GAUSSIAN CHAUSES

BY RAFAL LATAŁA

Warsaw University

We derive two-sided estimates on moments and tails of Gaussian chaoses, that is, random variables of the form $\sum a_{i_1,\ldots,i_d} g_{i_1} \cdots g_{i_d}$, where $g_i$ are i.i.d. $\mathcal{N}(0,1)$ r.v.’s. Estimates are exact up to constants depending on $d$ only.

1. Introduction. The purpose of this paper is to give precise bounds on moments and tails of Gaussian chaoses of order $d$, that is, random variables of the form $S = \sum_{i_1 < i_2 < \cdots < i_d} a_{i_1,\ldots,i_d} g_{i_1} \cdots g_{i_d}$. In the sequel, we will only consider decoupled chaoses $\tilde{S} = \sum a_i g_{i_1}^{(1)} \cdots g_{i_d}^{(d)}$, where $g_i^{(k)}$ are independent standard $\mathcal{N}(0,1)$ normal random variables and $(a_i) = (a_{i_1,\ldots,i_d})_{1 \leq i_1,\ldots,i_d \leq N}$ is a finite multi-indexed matrix—under natural symmetry assumptions, moments and tails of $S$ and $\tilde{S}$ are comparable with constants depending only on $d$ (cf. [5]).

For $d = 1$, we obviously have for $p \geq 2$,

$$\left\| \sum_i a_i g_i \right\|_p = \left( \sum_i a_i^2 \right)^{1/2} \left\| g \right\|_p \sim \sqrt{p} \left( \sum_i a_i^2 \right)^{1/2}. \tag{1}$$

For the chaoses of order 2, we have for any finite rectangular matrix $(a_{ij})$ and $p \geq 2$,

$$\left\| \sum_{ij} a_{ij} g_{i}^{(1)} g_{j}^{(2)} \right\|_p \sim \sqrt{p} \left\| (a_{ij}) \right\|_{\{1,2\}} + p \left\| (a_{ij}) \right\|_{\{1\}\{2\}}, \tag{2}$$

where $\left\| (a_{ij}) \right\|_{\{1,2\}} := \left\| (a_{ij}) \right\|_{\text{HS}} = (\sum_{ij} a_{ij}^2)^{1/2}$ and

$$\left\| (a_{ij}) \right\|_{\{1\}\{2\}} := \sup \left\{ \sum_{ij} a_{ij} x_i y_j : \|x\|_2 \leq 1, \|y\|_2 \leq 1 \right\}.$$
The upper part of the estimate (2) was obtained in [6]; the lower one is much easier (cf. [7]). One of the reasons why the case where \( d = 2 \) turned out to be relatively simple is that every square matrix is orthogonally equivalent to the diagonal matrix.

For \( d \geq 3 \), Borell [4] and Arcones and Giné [3] showed that

\[
\|S\|_p \sim \sum_{k=1}^d p^{k/2} \mathbb{E} \sup \left\{ \sum_i a_i \prod_{l=1}^{k} x_{i_l}^{(k)} \prod_{l=k+1}^{d} g_{i_l}^{(l)} : \|x^{(l)}\|_2 \leq 1, 1 \leq l \leq k \right\}.
\]

The above formula gives the precise dependence on \( p \), but unfortunately involves suprema of empirical processes that are, in general, not easy to estimate. (For generalizations of the above formula to the non-Gaussian case, cf. [1] and [10].) In this paper, we present bounds on moments and tails that involve only deterministic quantities.

The paper is organized as follows. In the next section, we present notation and definitions that will be used in the rest of the paper and formulate main results. In Section 3 we obtain bounds on entropy numbers for distances on products of Euclidean balls. This will provide a crucial tool to estimate suprema of certain Gaussian processes that naturally appear in the study of Gaussian chaos. Finally, in the last section, we present proofs of main results.

2. Notation and main results. We use the letter \( C \) to denote universal positive constants that may change from occurrence to occurrence and \( C(d) \) to denote positive constants depending only on \( d \). \( C(d) \) may also differ at each occurrence.

Let \( I = (a_{i,j})_{1 \leq i \leq n, 1 \leq j \leq n} \) be a finite multi-indexed matrix of order \( d \). If \( i \in \{1, \ldots, n\}^d \) and \( I \subset \{1, \ldots, d\} \), then we define \( i_I := (i_j)_{j \in I} \). For disjoint nonempty subsets \( I_1, \ldots, I_k \) of \( \{1, \ldots, d\} \), we put

\[
\|A\|_{I_1, \ldots, I_k} := \sup \left\{ \sum_i a_{i_{I_1} \ldots x_{I_k}} : \sum_i x_{i_{I_1} \ldots x_{I_k}}^2 \leq 1, \ldots, \sum_i x_{i_{I_1} \ldots x_{I_k}}^2 \leq 1 \right\}.
\]

Thus, for example,

\[
\|A\|_{\{1, \ldots, d\}} = \left( \sum_i a_{i}^2 \right)^{1/2}
\]

and

\[
\|(a_{i,j,k})\|_{\{1\}\{2,3\}} = \sup \left\{ \left( \sum_{jk} \left( \sum_i a_{i,j,k} x_i \right)^2 \right)^{1/2} : \sum_i x_i^2 \leq 1 \right\}
\]
By $S(k,d)$, we denote a set of all partitions of $\{1,\ldots,d\}$ into $k$ nonempty disjoint sets $I_1,\ldots,I_k$. For $p \geq 1$, we put

$$m_p(A) := \sum_{k=1}^{d} p^{k/2} \sum_{(I_1,\ldots,I_k) \in S(k,d)} \|A\|_{I_1,\ldots,I_k}.$$ 

Our main result is the following:

**Theorem 1.** For any multi-indexed finite matrix $A = (a_{i_1 \ldots i_d})_{1 \leq i_1,\ldots,i_d \leq n}$ and $p \geq 2$, we have

$$\frac{1}{C(d)} m_p(A) \leq \left\langle \left\langle \sum_i a_i \prod_{j=1}^{d} g_{i,j}^{(j)} \right\rangle \right\rangle_p \leq C(d) m_p(A).$$

(4)

Theorem 1 may be easily translated into the following two-sided estimate for tails:

**Corollary 1.** For any $t > 0$ and $d \geq 2$, we have

$$\frac{1}{C(d)} \exp\left[ -C(d) \min_{1 \leq k \leq d} \left( \min_{(I_1,\ldots,I_k) \in S(k,d)} \left( \frac{t}{\|A\|_{I_1,\ldots,I_k}} \right)^{2/k} \right) \right] \leq P\left( \left\| \sum_i a_i \prod_{j=1}^{d} g_{i,j}^{(j)} \right\| \geq t \right) \leq C(d) \exp\left[ -\frac{1}{C(d)} \min_{1 \leq k \leq d} \left( \min_{(I_1,\ldots,I_k) \in S(k,d)} \left( \frac{t}{\|A\|_{I_1,\ldots,I_k}} \right)^{2/k} \right) \right].$$

In view of (3), it is clear that the proof of (4) should be based on the estimation of norms of some random Gaussian matrices. The next theorem is, in our opinion, of independent interest and has recently been applied in [2] to obtain moment estimates for canonical $U$-statistics.

**Theorem 2.** For any $d \geq 2$ and any finite matrix $A$, we have for $p \geq 2$,

$$E\left\| \sum_{i_d} a_{i_d} g_i \right\|_{\{1\} \ldots \{d-1\}} \leq C(d) p^{(1-d)/2} m_p(A).$$

(5)
Remark. We suspect that a stronger estimate may actually hold, namely
\[
\mathbb{E} \left\| \left( \sum_{i,d} a_{i} G_{i,d} \right) \right\|_{\{1\} \ldots \{d-1\}} \\
\leq C(d) \inf_{\rho \geq 1} (s_{d-1}(A) + p^{1/2} A\|_{\{1\} \ldots \{d\}} + p^{(2-d)/2} A\|_{\{1, \ldots, d\}}) \\
\leq C(d) s_{d-1}(A) + \|A\|^{(d-2)/(d-1)}_{\{1\} \ldots \{d\}} \|A\|^{1/(d-1)}_{\{1, \ldots, d\}},
\]
where
\[
s_{d-1}(A) := \sum_{1 \leq j \leq d-1} \|A\|_{\{j, d\}, \{\{1\} \ldots \{d-1\}, \{1 \leq l \leq d-1, l \neq j\}}.
\]
However, we are not able to show this result for \(d > 3\).

3. Entropy estimates and Gaussian processes. By \(\gamma_{n,t}\), we will denote the distribution of \(tG_n\), where \(G_n = (g_1, \ldots, g_n)\) is a canonical \(n\)-dimensional Gaussian vector. We also put \(G^{(i)}_n := (g^{(i)}_1, \ldots, g^{(i)}_n)\) for i.i.d. copies of \(G_n\).

If \(\rho\) is a metric on a set \(T\), \(N(T, \rho, t)\) is the minimal number of closed balls of radius \(t\) that are necessary to cover \(T\). The closed unit Euclidean ball in \(\mathbb{R}^n\) is denoted by \(B^n_2\).

Lemma 1. For any norms \(\alpha_1, \alpha_2\) on \(\mathbb{R}^n\), \(y \in S \subseteq B^n_2\) and \(t > 0\),
\[
\gamma_{n,t}(x : \alpha_1(x - y) \leq 4t \mathbb{E} \alpha_1(G_n), \alpha_2(x) \leq 4t \mathbb{E} \alpha_2(G_n) + \alpha_2(y)) \geq \frac{1}{2} e^{-t^2/2}.
\]

Proof. Let
\[
K := \{x \in \mathbb{R}^n : \alpha_1(x) \leq 4t \mathbb{E} \alpha_1(G_n), \alpha_2(x) \leq 4t \mathbb{E} \alpha_2(G_n)\}.
\]
By Chebyshev’s inequality,
\[
1 - \gamma_{n,t}(K) \leq \mathbb{P}(\alpha_1(tG_n) > 4\mathbb{E} \alpha_1(tG_n)) + \mathbb{P}(\alpha_2(tG_n) > 4\mathbb{E} \alpha_2(tG_n)) \leq 1/2.
\]
By the symmetry of \(K\), we obtain for any \(y \in B^n_2\),
\[
\gamma_{n,t}(y + K) = e^{-|y|^2/(2t^2)} \int_{K} e^{(y,x)/t^2} d\gamma_{n,t}(x) \\
= e^{-|y|^2/(2t^2)} \int_{K} \frac{1}{2} (e^{(y,x)/t^2} + e^{-(y,x)/t^2}) d\gamma_{n,t}(x) \\
\geq \exp(-t^2/2) \gamma_{n,t}(K) \geq \frac{1}{2} \exp(-t^2/2).
\]
Finally, note that if \(x \in y + K\), then \(\alpha_1(x - y) \leq 4t \mathbb{E} \alpha_1(G_n)\) and \(\alpha_2(x) \leq \alpha_2(x - y) + \alpha_2(y) \leq 4t \mathbb{E} \alpha_2(G_n) + \alpha_2(y)\). □
Let $\alpha$ be a norm on $\mathbb{R}^{n_1 \cdots n_d}$ and the distance $\rho_\alpha$ on $\mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_d}$ be defined by

$$\rho_\alpha(x, y) := \alpha \left( \bigotimes_{i=1}^{d} x^i - \bigotimes_{i=1}^{d} y^i \right)$$

for $x = (x^1, \ldots, x^d), y = (y^1, \ldots, y^d)$, where $\bigotimes_{i=1}^{d} x^i := (\prod_{k=1}^{d} x^i_k)_{i_1 \leq n_1, \ldots, i_d \leq n_d}$. For $t > 0$, $T \subset \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_d}$, we put

$$W_d^T(\alpha, t) := \sum_{k=1}^{d} t^k \sum_{I \subset \{1, \ldots, d\}} W_I^T(\alpha),$$

where

$$W_I^T(\alpha) := \sup_{x \in T} \mathbb{E}_\alpha \left( \prod_{k \notin I} x^{i_k} \prod_{k \in I} g_i^{(k)} \right).$$

To simplify the notation, we will write $W_d$ and $W_I$ instead of $W_d^T(\alpha, t)$ and $W_I^T(\alpha)$ if $T = B^{n_1}_2 \times \cdots \times B^{n_d}_2$.

**Lemma 2.** For any $t > 0$ and $x \in B^{n_1}_2 \times \cdots \times B^{n_d}_2$, we have

$$\gamma_{n_1 + \cdots + n_d, t}(B_\alpha(x, W_d^x(\alpha, 4t))) \geq 2^{-d} \exp(-dt^{-2}/2).$$

**Proof.** We will proceed by induction on $d$. For $d = 1$, we have

$$B_\alpha(x, W_d^x(\alpha, 4t)) = \{ y \in \mathbb{R}^{n_1} : \alpha(x - y) \leq 4t \mathbb{E}_\alpha(G_{n_1}) \}$$

and (6) then follows by Lemma 1.

Now, suppose that (6) holds for $d - 1$. We will show that it is also satisfied for $d$.

Let us first observe that

$$\alpha \left( \bigotimes_{i=1}^{d} x^i - \bigotimes_{i=1}^{d} y^i \right) \leq \alpha^1(x^d - y^d) + \alpha_y \left( \bigotimes_{i=1}^{d-1} x^i - \bigotimes_{i=1}^{d-1} y^i \right),$$

where $\alpha^1$ and $\alpha_y$ are norms on $\mathbb{R}^{n_d}$ and $\mathbb{R}^{n_1 \cdots n_{d-1}}$, respectively, defined by

$$\alpha^1(z) := \alpha \left( \bigotimes_{i=1}^{d-1} x^i \otimes z \right) \quad \text{and} \quad \alpha_y(z) := \alpha(z \otimes y).$$

Then, obviously,

$$\mathbb{E}\alpha^1(G_{n_d}) = W_{d-1}^{[x]}(\alpha).$$
Moreover, if we put \( \pi(x) = (x^1, \ldots, x^{d-1}) \) and define a norm \( \alpha^2_t \) on \( \mathbb{R}^{n_d} \) by the formula
\[
\alpha^2_t(y) := W^\pi_{d-1}(\alpha_y, t),
\]
then
\[
(9) \quad tE\alpha^2_t(G_{n_d}) + \alpha^2_t(x^d) = \sum_{I \subset \{1, \ldots, d\}, I \neq \emptyset, \{d\}} t^{|I|}W^I_f(\alpha, t).
\]
Observe also that by the induction assumption, we have for any \( \gamma \)
\[
(10) \quad \gamma_{n_1 + \cdots + n_{d-1}, t}(\gamma) \geq 2^{1-d} \exp(-d \cdot t^{-2}/2).
\]
Finally, let
\[
A(x) := \left\{ y \in \mathbb{R}^{n_1 + \cdots + n_d} : \alpha(x^d - y^d) \leq 4tE\alpha^1(\alpha^1), \right. \\
\left. \alpha^2_t(y^d) \leq 4tE\alpha^2_t(G_{n_d}) + \alpha^2_t(x^d), \right. \\
\left. \alpha^2_t\left(\bigotimes_{i=1}^{d-1} x^i - \bigotimes_{i=1}^{d-1} y^i\right) \leq \alpha^2_t(y^d) \right\}.
\]
By (7)–(9), we get \( A(x) \subset B_\alpha(x, W_d^I(\alpha, 4t)) \) and, therefore, by (10), Lemma 1 and Fubini’s theorem, we get
\[
\gamma_{n_1 + \cdots + n_{d-1}, t}(B_\alpha(x, W_d^I(\alpha, 4t))) \geq \gamma_{n_1 + \cdots + n_d, t}(A(x)) \geq 2^{-d} \exp(-dt^{-2}/2). \tag*{□}
\]

**Corollary 2.** For any \( T \subset B_2^{n_1} \times \cdots \times B_2^{n_d} \) and \( t \in (0, 1] \),
\[
N(T, \rho_\alpha, W_d^T(\alpha, t)) \leq \exp(Cdt^{-2}).
\]
In particular,
\[
N(B_2^{n_1} \times \cdots \times B_2^{n_d}, \rho_\alpha, W_d(\alpha, t)) \leq \exp(Cdt^{-2}).
\]

**Proof.** Obviously, \( W_d^T(\alpha, t) \geq \sup_{x \in T} W_d^\pi(x)(\alpha, t) \). Therefore, by Lemma 2, we have for any \( x \in T \),
\[
(11) \quad \gamma_{n_1 + \cdots + n_d, t}(B_\alpha(x, W_d^T(\alpha, 4t))) \geq 2^{-d} \exp(-dt^{-2}/2).
\]
Suppose that there exist \( x_1, \ldots, x_N \in T \) such that \( \rho_\alpha(x_i, x_j) > W_d^T(\alpha, t/2) \) for \( i \neq j \). Then the sets \( B_\alpha(x_i, W_d^T(\alpha, t/2)) \) are disjoint, so by (11), we obtain \( N \leq 2^d \exp(32dt^{-2}) \). Hence,
\[
N(T, \rho_\alpha, W_d^T(\alpha, t)) \leq 2^d \exp(32dt^{-2}) \leq \exp(33dt^{-2}).
\]
To finish this section, let us recall standard estimates for Gaussian processes.

**Lemma 3.** Let \((X_t)_{t \in T}\) be a centered Gaussian process and \(T = \bigcup_{l=1}^{m} T_l\).

Then

\[
E \sup_{t \in T} X_t \leq \max_l E \sup_{t \in T_l} X_t + C \sqrt{\log m} \sup_{t, s \in T} (E(X_t - X_s)^2)^{1/2}.
\]

**Proof.** Obviously, \(E \sup_{t \in T} X_t = E \max_l \sup_{t \in T_l} (X_t - X_{t_0})\) for any \(t_0 \in T\). The lemma follows by integration by parts and the classical estimate (cf. [8], Theorem 7.1)

\[
P\left(\sup_{t \in T_l} (X_t - X_{t_0}) \geq E \sup_{t \in T_l} X_t + u \sup_{t \in T_l} (E(X_t - X_t_0)^2)^{1/2}\right) \leq \exp(-u^2/2)
\]

for \(u > 0\). □

**Lemma 4.** Let \((X_t)_{t \in T}\) be a centered Gaussian process. Then for any \(p \geq 2\),

\[
\frac{1}{C} \left(\|\sup_{t \in T} X_t\|_1 + \sqrt{p}\sigma\right) \leq \|\sup_{t \in T} X_t\|_p \leq \|\sup_{t \in T} X_t\|_1 + C \sqrt{p}\sigma,
\]

where \(\sigma := \sup_{t \in T}(EX_t^2)^{1/2}\).

**Proof.** The lower bound follows from the easy estimates

\[
\|\sup_{t \in T} X_t\|_p \geq \|\sup_{t \in T} \max(X_t, 0)\|_p \geq \sup_{t \in T} \|\max(X_t, 0)\|_p \geq \sup_{t \in T} \|X_t\|_p/2
\]

and the upper one by the concentration of suprema of Gaussian processes (cf. [8], Theorem 7.1) and integration by parts. □

**4. Proofs.** Let us start with some additional notation. For a matrix \(A = (a_{i})_{1 \leq i_1, \ldots, i_d \leq n}\) of order \(d \geq 2\), we set

\[
s_{d-1}(A) := \sum_{1 \leq j \leq d-1} \|A\|_{(j,d),\{\{t\}: 1 \leq t \leq d-1, t \neq j}\}
\]

and for \(1 \leq k \leq d-2\),

\[
s_k(A) := \sum_{(I_1, \ldots, I_k) \in S(k,d)} \|A\|_{I_1,\ldots,I_k}.
\]
On $\mathbb{R}^{(d-1)n} = (\mathbb{R}^n)^{d-1}$, we introduce the distance $\rho_A$ by the formula

$$
\rho_A(x, y) := \left( \sum_{i_d} \left( \sum_{i_1, \ldots, i_{d-1}} a_i \left( \prod_{k=1}^{d-1} x_{i_k}^{i_k} - \prod_{k=1}^{d-1} y_{i_k}^{i_k} \right) \right)^2 \right)^{1/2},
$$

where $x = (x^1, \ldots, x^{d-1})$ and $y = (y^1, \ldots, y^{d-1})$. We have

$$
\rho_A(x, y) = (E(X - X)^2)^{1/2},
$$

where $X := \sum_{i_1, \ldots, i_d} a_i \prod_{k=1}^{d-1} x_{i_k}^{i_k} g_{i_d}$.

For $T \subset \mathbb{R}^{(d-1)n}$, we put

$$
\Delta_A(T) := \sup\{\rho_A(x, y) : x, y \in T\}.
$$

Let us note that in particular, we have

$$
\Delta_A((B_n^2)^{d-1}) \leq 2 \sup\{\rho_A(x, 0) : x \in (B_n^2)^{d-1}\} = 2 \|A\|_{\{1\} \ldots \{d\}}.
$$

(13)

For a set $T \subset \mathbb{R}^{(d-1)n}$ and $I \subset \{1, \ldots, d-1\}$, we put

$$
W^T_I(A) := \sup_{x \in T} \left( \sum_{i \in I \setminus \{d\}} \left( \sum_{i_1, \ldots, i_{d-1}} a_i \prod_{k \in \{1, \ldots, d-1\} \setminus I} x_{i_k}^{i_k} \right)^2 \right)^{1/2}
$$

and for $1 \leq k \leq d-1$,

$$
W^T_k(A) := \sum_{I \subset \{1, \ldots, d-1\}, \#I = k} W^T_I(A).
$$

The next lemma shows how the results of the previous section may be adapted to the case of the particular metric $\rho_A$.

**Lemma 5.** For any $0 < t \leq 1$ and $T \subset (B_2^n)^{d-1}$, we have

$$
N\left( T, \rho_A, \sum_{k=1}^{d-1} t^k W^T_k(A) \right) \leq \exp(\mathcal{C}dt^{-2}).
$$

(15)

**In particular,**

$$
N\left( T, \rho_A, t W^T_1(A) + \sum_{k=2}^{d-1} t^k s_{d-k}(A) \right) \leq \exp(\mathcal{C}dt^{-2}).
$$

(16)

**Proof.** Note that $\rho_A = \rho_\alpha$, where for $z \in \mathbb{R}^{d-1}$,

$$
\alpha(z) := \left( \sum_{i_d} \left( \sum_{i_1, \ldots, i_{d-1}} a_i \prod_{k=1}^{d-1} x_{i_k}^{i_k} \right)^2 \right)^{1/2}.
$$
We have for any $\mathbf{x} \in (\mathbb{R}^n)^{d-1}$ and $I \subset \{1, \ldots, d-1\}$,
\[
\mathbb{E}^\alpha \left( \prod_{k \in \{1, \ldots, d-1\} \setminus I} x_{i_k}^k \prod_{k \in I} g_{i_k}^{(k)} \right) \leq \left( \mathbb{E}^{\alpha^2} \left( \prod_{k \in \{1, \ldots, d-1\} \setminus I} x_{i_k}^k \prod_{k \in I} g_{i_k}^{(k)} \right) \right)^{1/2} = \left( \sum_{i \in I \cup \{d\}} \left( \sum_{i \in \{1, \ldots, d-1\} \setminus I} a_i \prod_{k \in \{1, \ldots, d-1\} \setminus I} x_{i_k}^k \right)^2 \right)^{1/2}.
\]
Hence,
\[
W_f^T(\alpha) \leq W_f^T(A)
\]
and (15) immediately follows by Corollary 2.

Inequality (15) implies (16), since
\[
W_f^T(A) \leq W_f^{(B_2^n)^{d-1}}(A) = \|A\|_{I \cup \{d\}, \{l\} : l \in \{1, \ldots, d-1\} \setminus I}.
\]

We are now ready to present a stronger version of Theorem 2. To formulate it, let us define for $T \subset (\mathbb{R}^n)^{d-1}$,
\[
F_A(T) := \mathbb{E} \sup_{\mathbf{x} \in T} \left( \sum_{i} a_i \prod_{k=1}^{d-1} x_{i_k}^k g_{i_d} \right).
\]

**Theorem 3.** For any $T \subset (B_2^n)^{d-1}$ and $p \geq 1$,
\[
F_A(T) \leq C(d) \left( \sqrt{p} \Delta_A(T) + \sum_{k=1}^{d-1} p^{(1-k)/2} s_{d-k}(A) \right). \tag{17}
\]

Let us observe that $\mathbb{E} \|(\sum_i a_i g_{i_d})\|_{\{1\} \ldots \{d-1\}} = F_A((B_2^n)^{d-1})$ and, therefore, Theorem 3 implies Theorem 2 since by (14), $\Delta_A((B_2^n)^{d-1}) \leq 2\|A\|_{\{1\} \ldots \{d\}}$.

We will prove (17) by induction on $d$, but first we will show several consequences of the theorem. In the next three lemmas, we shall assume that Theorem 3 (and thus also Theorem 2) holds for all matrices of order smaller than $d$.

For $\mathbf{x}, \mathbf{y} \in (\mathbb{R}^n)^{d-1}$, we set
\[
\tilde{\alpha}_A(\mathbf{x}) = \sum_{1 \leq j \neq k \leq d-1} \left\| \sum_{i_j} a_i x_{i_j}^k \right\|_{\{k,d\} \{\{l\} : 1 \leq l \leq d-1, l \neq k,j\}},
\]
\[
\tilde{\rho}_A(\mathbf{x}, \mathbf{y}) := \tilde{\alpha}_A(\mathbf{x} - \mathbf{y}) = \sum_{1 \leq j \neq k \leq d-1} \left\| \sum_{i_j} a_i (x_{i_j}^k - y_{i_j}^k) \right\|_{\{k,d\} \{\{l\} : 1 \leq l \leq d-1, l \neq k,j\}}
\]
and for $T \subset (\mathbb{R}^n)^{d-1}$,
\[
\tilde{\alpha}_A(T) := \sup \{ \tilde{\alpha}_A(\mathbf{x}) : \mathbf{x} \in T \}.
\]
LEMMA 6. For any $p \geq 1$ and $l \geq 0$,
\[ N \left( (B^n_2)^{d-1}, \tilde{\rho}, 2^{l-1} \sum_{k=1}^{d-1} p^{(1-k)/2} s_{d-k}(A) \right) \leq \exp(C(d)2^l p). \]

PROOF. Note that $\tilde{\alpha}_A$ is a norm on $(\mathbb{R}^n)^{d-1} = \mathbb{R}^{(d-1)n}$ and that
\[ E\tilde{\alpha}(G_{(d-1)n}) = \sum_{1 \leq j \neq k \leq d-1} E \left\| \sum_{ij} a_{ij} g_{ij} \right\|_{\{k,d\} \{\{l\} : 1 \leq l \leq d-1, l \neq k,j \}}. \]
Let us fix $1 \leq j \neq k \leq d-1$ and observe that
\[ \left\| \sum_{ij} a_{ij} g_{ij} \right\|_{\{k,d\} \{\{l\} : 1 \leq l \leq d-1, l \neq k,j \}} = \left\| \sum_{l=1}^{d-1} b_{i_1, ..., i_{d-1}, j} g_{i_{d-1}} \right\|_{\{1, ..., d-2\}} \]
for an appropriately chosen matrix $B = (b_{i_1, ..., i_{d-1}})$ (we treat a pair of indices $k, d$ as a single index and renumerate indices in such a way that $j$ would become $d - 1$). Moreover, for any $1 \leq l \leq d-1$,
\[ \sum_{(I_1, ..., I_l) \in S(l,d-1)} \|B\|_{I_1, ..., I_l} = \sum_{(I_1, ..., I_l) \in S(l,d) \{k,d\} \in I_1} \|A\|_{I_1, ..., I_l} \leq s_l(A). \]
Thus, by (5) (applied to the matrix $B$ of order $d - 1$),
\[ E \left\| \sum_{ij} a_{ij} g_{ij} \right\|_{\{k,d\} \{\{l\} : 1 \leq l \leq d-1, l \neq k,j \}} = E \left\| \sum_{l=1}^{d-1} b_{i_1, ..., i_{d-1}} g_{i_{d-1}} \right\|_{\{1, ..., d-2\}} \leq C(d) \sum_{l=1}^{d-1} p^{(2-d+l)/2} \sum_{(I_1, ..., I_l) \in S(l,d-1)} \|B\|_{I_1, ..., I_l} \leq C(d) \sum_{s=1}^{d-1} p^{(2-s)/2} s_{d-s}(A). \]
Hence, by Corollary 2 (with $d = 1$), we have for $t \in (0, 1]$,
\[ N \left( (B^n_2)^{d-1}, \tilde{\rho}, C(d) t \sum_{k=1}^{d-1} p^{(2-k)/2} s_{d-k}(A) \right) \leq \exp(Ct^{-2}) \]
and it suffices to make the substitution $t = (C(d)2^l \sqrt{p})^{-1}$. □
Lemma 7. Suppose that $d \geq 3$, $y \in (B_2^n)^{d-1}$ and $T \subset (B_2^n)^{d-1}$. Then for any $p \geq 1$ and $l \geq 0$, we can find a decomposition
\[ T = \bigcup_{j=1}^{N} T_j, \quad N \leq \exp(C(d)2^l p) \]
such that for each $j \leq N$,
\[ F_\Delta(y + T_j) \leq F_\Delta(T_j) \tag{18} \]
and
\[ \Delta_\Delta(T_j) \leq 2^{-l} p^{-1/2} \Delta_\Delta(T) + 2^{-2l} \sum_{k=2}^{d-1} p^{-k/2} s_{d-k}(A). \tag{19} \]

Proof. For $I \not\subset \{1, \ldots, d - 1\}$, $x, \tilde{x} \in (\mathbb{R}^n)^{d-1}$ and $S \subset (\mathbb{R}^n)^{d-1}$, let us define
\[ \rho_A^y, I(x, \tilde{x}) := \left( \sum_{i_1, \ldots, i_{d-1}} \prod_{i \in I} y_{i_k}^k \left( \prod_{j \leq d-1, j \not\in I} x_{i_j}^j \prod_{j \leq d-1, j \not\in I} \tilde{x}_{i_j}^j \right) \right)^{1/2}, \]
\[ \Delta_A^y, I(S) := \sup \{ \rho_A^y, I(x, \tilde{x}) : x, \tilde{x} \in S \} \]
and
\[ F_A^y, I(S) := \mathbb{E} \sup_{x \in S} \sum_{i} a_i \prod_{k \in I} y_{i_k}^k \prod_{j \leq d-1, j \not\in I} x_{i_j}^j, \]
Note that if $I \neq \emptyset$, then (17) applied to the matrix
\[ A(y, I) := \left( \sum_{i} a_i \prod_{k \in I} y_{i_k}^k \right) \]
of order $d - \#I < d$ gives for any $S \subset (B_2^n)^{d-1}$ and $q \geq 1$,
\[ F_A^y, I(S) \leq C(d - \#I) \left( q^{1/2} \Delta_A^y, I(S) + \sum_{k=1}^{d - \#I - 1} q^{(1-k)/2} s_{d-k}(A(y, I)) \right). \]
But, $s_{d-\#I-k}(A(y, I)) \leq s_{d-k}(A)$ for $k \geq 2$ and $s_{d-\#I-1}(A(y, I)) \leq \Delta_A(y)$, hence,
\[ F_A^y, I(S) \leq C(d) \left( q^{1/2} \Delta_A^y, I(S) + \Delta_A(y) + \sum_{k=2}^{d-1} q^{(1-k)/2} s_{d-k}(A) \right). \]
(20)
Since $E \sum a_i \prod_{k \leq d-1} y_{ik}^k g_{ik} = 0$, we get
\begin{equation}
F_A(y + S) - F_A(S) \leq \sum_{\varnothing \neq I \subseteq \{1, \ldots, d-1\}} F^y_A(S).
\end{equation}
Thus, we may apply $2^{d-1} - d$ times (16) with $t = 2^{-l} p^{-1/2}$ and find a decomposition $T = \bigcup_{j=1}^N T_j$, $N \leq \exp(C(d)2^{2l}p)$, such that for each $j$ and $I \subseteq \{1, \ldots, d-1\}$, $0 \leq \#I \leq d - 3$,
\begin{equation}
\Delta^n_A(T_j) \leq 2^{-l} p^{-1/2} \bar{\alpha}_A(T) + 2^{-2l} \sum_{k=2}^{d-1} p^{-k/2} s_{d-k}(A).
\end{equation}
Moreover, if $I \subseteq \{1, \ldots, d-1\}$ with $\#I = d - 2$, then $A(y, I)$ is a matrix of order 2 and for $S \subseteq (B_2^n)^{d-1}$,
\begin{equation}
F^y_A(S) \leq \|A(y, I)\|_{HS} \leq \bar{\alpha}_A(y).
\end{equation}
Estimate (22) reduces to (19) for $I = \varnothing$ and (18) follows by (20) with $q = 2^{2l}p$ and (21)--(23).

\textbf{Lemma 8.} Suppose that $S$ is a finite subset of $(B_2^n)^{d-1}$, with $\#S \geq 2$, such that $S - S \subseteq (B_2^n)^{d-1}$. Then there exist finite sets $S_i \subseteq (B_2^n)^{d-1}$ and $y_i \in S$, $i = 1, \ldots, N$, such that:
\begin{itemize}
  \item [(i)] $2 \leq N \leq \exp(C(d)2^{2l}p)$,
  \item [(ii)] $S = \bigcup_{i=1}^N (y_i + S_i)$, $S_i - S_i \subseteq S - S$, $\#S_i \leq \#S - 1$,
  \item [(iii)] $\Delta_A(S_i) \leq 2^{-2l} \sum_{k=1}^{d-1} p^{-k/2} s_{d-k}(A)$,
  \item [(iv)] $\bar{\alpha}_A(S_i) \leq 2^{-l} \sum_{k=1}^{d-1} p^{(1-k)/2} s_{d-k}(A)$
\end{itemize}
and
\begin{itemize}
  \item [(v)] $F_A(y_i + S_i) \leq F_A(S_i) + C(d)\tilde{\alpha}_A(S) + 2^{-l} \sum_{k=1}^{d-1} p^{(1-k)/2} s_{d-k}(A)$.
\end{itemize}
\textbf{Proof.} By Lemma 6, we get
\begin{equation}
S = \bigcup_{i=1}^{N_1} (y_i + T_i), \quad N_1 \leq \exp(C(d)2^{2l}p),
\end{equation}
$y_i \in S$, $0 \in T_i$ and
\begin{equation}
\bar{\alpha}_A(T_i) \leq 2^{-l} \sum_{k=1}^{d-1} p^{(1-k)/2} s_{d-k}(A).
\end{equation}
Note that $T_i \subset S - y_i \subset S - S \subset (B^2_n)^{d-1}$. Hence, by Lemma 7 (with $l + 1$ instead of $l$), we get

$$T_i = \bigcup_{j=1}^{N_2} T_{i,j}, \quad N_2 \leq \exp(C(d)2^d p),$$

where

$$F_A(y_i + T_{i,j}) \leq F_A(T_{i,j}) + C(d)\left(\bar{\alpha}_A(y_i) + \bar{\alpha}_A(T_i) + 2^{-l-1} \sum_{k=2}^{d-1} p^{(1-k)/2} s_{d-k}(A)\right)$$

$$\leq F_A(T_{i,j}) + C(d)\left(\bar{\alpha}_A(S) + 2^{-l} \sum_{k=2}^{d-1} p^{(1-k)/2} s_{d-k}(A)\right)$$

and

$$\Delta_A(T_{i,j}) \leq 2^{-l-1} p^{-1/2} \bar{\alpha}_A(T_i) + 2^{-2l-2} \sum_{k=2}^{d-1} p^{-k/2} s_{d-k}(A)$$

$$\leq 2^{-2l} \sum_{k=1}^{d-1} p^{-k/2} s_{d-k}(A).$$

Therefore,

$$S = \bigcup_{i,j}(y_i + T_{i,j}).$$

We have $N = N_1 N_2 \leq \exp(C(d)2^{d/p})$. Moreover, we may obviously assume that $N \geq 2$, and by making the sets $T_{i,j}$ disjoint, we may assume that $\#T_{i,j} \leq \#S - 1$. Obviously, $T_{i,j} - T_{i,j} \subset S - S$ and $\bar{\alpha}_A(T_{i,j}) \leq \bar{\alpha}_A(T_i) \leq 2^{-l} \sum_{k=1}^{d-1} p^{(1-k)/2} s_{d-k}(A)$. □

**Proof of Theorem 3.** We proceed by induction on $d$. For $d = 2$ and $A = (a_{ij})$, we have

$$F_A(T) \leq F_A(B^2_n) = \mathbf{E}\left(\sum_i \left(\sum_j a_{ij}g_j\right)^2\right)^{1/2} \leq \|A\|_{HS} = s_1(A).$$

Suppose that $d \geq 3$ and that (17) holds for matrices of order smaller than $d$. Let us put $\Delta_0 := \Delta_A(T)$, $\Delta_0 := \bar{\alpha}_A((B^2_n)^{d-1}) \leq C(d) s_{d-1}(A)$ and

$$\Delta_l := 2^{2-2l} \sum_{k=1}^{d-1} p^{-k/2} s_{d-k}(A),$$

$$\bar{\Delta}_l := 2^{1-l} \sum_{k=1}^{d-1} p^{(1-k)/2} s_{d-k}(A) \quad \text{for } l \geq 1.$$
Suppose first that $T \subset \frac{1}{2}(B^n_2)^{d-1}$ and define
\[ c_T(r, l) := \sup \{ F_A(S) : S \subset (B^n_2)^{d-1}, S - S \subset T - T, \#S \leq r, \Delta_A(S) \leq \Delta_l, \tilde{\alpha}_A(S) \leq \tilde{\Delta}_l \}. \]

Note that any subset $S \subset T$ satisfies $\Delta_A(S) \leq \Delta_0$ and $\tilde{\alpha}_A(S) \leq \tilde{\Delta}_0$, therefore,
\[ c_T(r, 0) \geq \sup \{ F_A(S) : S \subset T, \#S \leq r \}. \]

Obviously, $c_T(1, l) = 0$. We will now show that for $r \geq 2$,
\[ c_T(r, l) \leq c_T(r - 1, l + 1) + C(d) \left( \tilde{\Delta}_l + 2^l \sqrt{p} \Delta_l + 2^{-l} \sum_{k=1}^{d-1} p^{(1-k)/2} s_{d-k}(A) \right). \]

Indeed, let us take $S \subset (B^n_2)^{d-1}$ as in the definition of $c_T(r, l)$. Then by Lemma 8, we may find a decomposition $S = \bigcup_{i=1}^\infty (y_i + S_i)$ satisfying (i)–(v). Hence, by Lemma 3 and (13), we have
\[ F_A(S) \leq C \sqrt{\log N} \Delta_A(S) + \max_i F_A(y_i + S_i) \leq C(d) \left( \tilde{\alpha}_A(S) + 2^l \sqrt{p} \Delta_l + 2^{-l} \sum_{k=1}^{d-1} p^{(1-k)/2} s_{d-k}(A) \right) + \max_i F_A(S_i). \]

We have $\Delta_A(S_i) \leq \Delta_{l+1}$, $\tilde{\alpha}_A(S_i) \leq \tilde{\Delta}_{l+1}$, $S_i - S_i \subset S - S \subset T - T$ and $\#S_i \leq \#S - 1 \leq r - 1$, thus $\max_i F_A(S_i) \leq c_T(r - 1, l + 1)$ and (26) yields (25).

By (25), we immediately obtain
\[ c_T(r, 0) \leq c_T(1, r - 1) + C(d) \sum_{l=0}^\infty \left( \tilde{\Delta}_l + 2^l \sqrt{p} \Delta_l + 2^{-l} \sum_{k=1}^{d-1} p^{(1-k)/2} s_{d-k}(A) \right) \]
\[ \leq C(d) \left( \sqrt{p} \Delta_A(T) + \sum_{k=1}^{d-1} p^{(1-k)/2} s_{d-k}(A) \right). \]

For $T \subset \frac{1}{2}(B^n_2)^{d-1}$, we have by (24),
\[ F_A(T) = \sup \{ F_A(S) : S \subset T, \#S < \infty \} \leq \sup_r c_T(r, 0) \leq C(d) \left( \sqrt{p} \Delta_A(T) + \sum_{k=1}^{d-1} p^{(1-k)/2} s_{d-k}(A) \right). \]
Finally, if $T \subset (B_2^n)^{d-1}$, then $\frac{1}{2}T \subset \frac{1}{2}(B_2^n)^{d-1}$ and $\Delta_A(\frac{1}{2}T) = 2^{1-d}\Delta_A(T)$, hence,

$$F_A(T) = 2^{d-1}F_A(\frac{1}{2}T) \leq C(d)\left(\sqrt{p}\Delta_A(T) + \sum_{k=1}^{d-1} p^{(1-k)/2}s_{d-k}(A)\right).$$

**Proof of Theorem 1.** First, we prove by induction on $d$ the estimate from below. For $d = 1$, it follows by (1). Suppose that the lower estimate holds for matrices of order smaller than $d$. Then by the induction assumption, we have for any matrix $B = (b_l)_{i_1,\ldots,i_{d-1}}$,

$$\left\|\sum_{i} b_i \prod_{j=1}^{d-1} g_{i_j}^{(j)}\right\|_p \geq C(d-1)^{-1}\sqrt{p}\left(\sum_{i} b_i^2\right)^{1/2} \geq C(d)^{-1}\left\|\sum_{i} b_i g_{i_1,\ldots,i_{d-1}}\right\|_p,$$

where $(g_{i_1,\ldots,i_{d-1}})$ is a sequence of i.i.d. $\mathcal{N}(0,1)$ r.v.’s independent of $(g_{i_j}^{(d)})$. Therefore, by (2),

$$\left\|\sum_{i} a_i \prod_{j=1}^{d} g_{i_j}^{(j)}\right\|_p \geq C(d)^{-1}\left\|\sum_{i} a_i g_{i_1,\ldots,i_{d-1},i_d}^{(d)}\right\|_p \geq C(d)^{-1}\sqrt{p}\|A\|_{1,\ldots,d}.$$

Let $(I_1,\ldots,I_k) \in S(k,d)$ with $k \geq 2$ and $\sum_{i} a_{I_i} (x_{I_i}^{(l)})^2 \leq 1$ for $l = 1,\ldots,k$. Then by the induction assumption applied twice [first conditionally on $(g_{i_j}^{(j)})_{j \in I_1}$], we have

$$\left\|\sum_{i} a_i \prod_{j=1}^{d} g_{i_j}^{(j)}\right\|_p \geq C(d-\#I_1)^{-1}p^{(k-1)/2}\left\|\sum_{i} a_i \prod_{j \in I_1} g_{i_j}^{(j)} \prod_{l=2}^{k} x_{I_l}^{(l)}\right\|_p \geq (C(d-\#I_1)C(\#I_1))^{-1}p^{k/2}\sum_{i} a_i \prod_{l=1}^{k} x_{I_l}^{(l)}$$

and, hence,

$$\left\|\sum_{i} a_i \prod_{j=1}^{d} g_{i_j}^{(j)}\right\|_p \geq C(d)^{-1}p^{k/2}\|A\|_{I_1,\ldots,I_k}.$$

Inequalities (27) and (28) imply the lower part of estimate (4).

Now (again by induction on $d$), we prove the estimate from above. For $d \leq 2$, the estimate follows by (1) and (2). Suppose that $d \geq 3$ and the estimate
holds for chaoses of order smaller than \( d \). By the induction assumption, we have

\[
\left\| \sum_{i} a_i \prod_{j=1}^{d} g_{i,j}^{(j)} \right\|_{p}^{\frac{1}{p}} \leq C(d-1) \sum_{k=1}^{d-1} p^{k/2} \sum_{(I_1, \ldots, I_k) \in S(k,d-1)} \left( E \left\| \sum_{i} a_i g_{i,k}^{(d)} \right\|_{I_1, \ldots, I_k}^{p} \right)^{1/p}.
\] (29)

However, for \((I_1, \ldots, I_k) \in S(k,d-1)\), we have by (12),

\[
\left( E \left\| \sum_{i} a_i g_{i,k}^{(d)} \right\|_{I_1, \ldots, I_k}^{p} \right)^{1/p} \leq C \sqrt{\|A\|_{I_1, \ldots, I_k}}^{d} + E \left\| \sum_{i} a_i g_{i,k}^{(d)} \right\|_{I_1, \ldots, I_k}.
\] (30)

Theorem 2 gives

\[
E \left\| \sum_{i} a_i g_{i,k}^{(d)} \right\|_{I_1, \ldots, I_k} \leq C(k+1)p^{-k/2} m_p(A).
\] (31)

Inequalities (29)–(31) then yield the upper estimate in (4). \( \square \)

**Proof of Corollary 1.** Let \( S := \sum_{i} a_i \prod_{j=1}^{d} g_{i,j}^{(j)} \). By Chebyshev’s inequality and (4), one gets for \( p \geq 2 \),

\[
P(|S| \geq eC(d)m_p(A)) \leq P(|S| \geq e\|S\|_p) \leq e^{-p}.
\] (32)

Since \( \|S\|_2 \leq C_1(d)\|S\|_p \) (cf. [9], Section 3.2, or use (4) and \( m_{2p}(A) \leq 2^{d/2}m_p(A) \)), we get by the Paley–Zygmund inequality for \( q \geq 2 \),

\[
P(S \geq 2^{-1}\|S\|_q) = P(|S|^q \geq 2^{-q}E|S|^q) \geq (1 - 2^{-q})^2 \frac{(E|S|^q)^2}{E|S|^{2q}} \geq (2C_1(d))^{-2q}.
\]

Fix \( p > 0 \) and take \( q := p/(2\ln(2C_1(d))) \). Then \( q \geq 2 \) for \( p \geq p_0(d) \) and \( \|S\|_q \geq C(d)^{-1}m_q(A) \geq C(d)^{-1}(\max(2\ln(2C_1(d)),1))^{-d/2}m_p(A) = 2C_2(d)^{-d}m_p(A) \). Thus, for \( p \geq p_0(d) \),

\[
P(|S| \geq C_2(d)^{-d}m_p(A)) \geq P(S \geq 2^{-1}\|S\|_q) \geq (2C_1(d))^{-2q} = e^{-p}
\]

and, therefore, for any \( p > 0 \),

\[
P(|S| \geq C_2(d)^{-d}m_p(A)) \geq \min(e^{-p_0(d)}), e^{-p}).
\] (33)
Finally, note that if $m_p(A) = s \geq m_2(A)$, then $p$ is comparable (with constants depending only on $d$) with

$$\min\{\frac{s}{\|A\|_{I_1,\ldots,I_k}}^{2/k} : 1 \leq k \leq d, (I_1,\ldots,I_k) \in S(k,d)\}$$

and, therefore, Corollary 1 follows by (32) and (33). □

Acknowledgment. The author would like to thank the referee for a careful reading of the manuscript and valuable remarks.

REFERENCES

[1] Adamczak, R. (2005). Logarithmic Sobolev inequalities and concentration of measure for convex functions and polynomial chaos. Bull. Polish Acad. Sci. Math. 53 221–238. MR2163396

[2] Adamczak, R. (2005). Moment inequalities for $U$-statistics. Ann. Probab. 34 2288–2314.

[3] Arcones, M. and Giné, E. (1993). On decoupling, series expansions, and tail behaviour of chaos processes. J. Theoret. Probab. 6 101–122. MR1201060

[4] Borell, C. (1984). On the Taylor series of a Wiener polynomial. In Seminar Notes on Multiple Stochastic Integration, Polynomial Chaos and Their Integration. Case Western Reserve Univ., Cleveland.

[5] de la Peña, V. H. and Montgomery-Smith, S. (1994). Bounds for the tail probabilities of $U$-statistics and quadratic forms. Bull. Amer. Math. Soc. 31 223–227. MR1261237

[6] Hanson, D. L. and Wright, F. T. (1971). A bound on tail probabilities for quadratic forms in independent random variables. Ann. Math. Statist. 42 1079–1083. MR0279864

[7] Latala, R. (1999). Tail and moment estimates for some types of chaos. Studia Math. 135 39–53. MR1686370

[8] Ledoux, M. (2001). The Concentration of Measure Phenomenon. Amer. Math. Soc., Providence, RI. MR1849347

[9] Ledoux, M. and Talagrand, M. (1991). Probability in Banach Spaces: Isoperimetry and Processes. Springer, New York. MR1102015

[10] Lochowski, R. (2006). Moment and tail estimates for multidimensional chaos generated by symmetric random variables with logarithmically concave tails. Banach Center Publ. 72 161–176.