

Logarithmic Sobolev Inequalities and Concentration of Measure for Convex Functions and Polynomial Chaoses

by

Radosław ADAMCZAK

Presented by Stanisław KWAPIEŃ

Summary. We prove logarithmic Sobolev inequalities and concentration results for convex functions and a class of product random vectors. The results are used to derive tail and moment inequalities for chaos variables (in the spirit of Talagrand and Arcones–Giné). We also show that the same proof may be used for chaoses generated by log-concave random variables, recovering results by Łochowski, and present an application to exponential integrability of Rademacher chaos.

1. Introduction. The paper is concerned with concentration properties of random vectors. We start with the following

DEFINITION 1. A real random variable ξ is said to have the *concentration property* of order $\alpha > 0$ with constants K, C if there exists $a \in \mathbb{R}$ such that for all $t \geq 0$,

$$(1) \quad \mathbb{P}(|\xi - a| \geq t) \leq C e^{-t^\alpha/K}.$$

It is easy to see that the concentration property implies that ξ has a finite moment and there exist constants C', K' , depending on α, C, K only, such that for all $t \geq 0$,

$$(2) \quad \mathbb{P}(|\xi - \mathbb{E}\xi| \geq t) \leq C' e^{-t^\alpha/K'}.$$

Moreover, by the Chebyshev inequality, the condition (2) is equivalent to the following moment estimates, valid for all $p \geq 1$:

$$(3) \quad \|\xi - \mathbb{E}\xi\|_p \leq K'' p^{1/\alpha}.$$

2000 *Mathematics Subject Classification*: Primary 60E15; Secondary 60B11.

Key words and phrases: log-Sobolev inequalities, concentration of measure, polynomial chaos.

More precisely, if (2) holds then so does (3) with K'' depending only on α, C', K' , whereas (3) implies (2) with C', K' depending only on K'', α .

In what follows we restrict our attention to random variables of the form $\xi = f(X)$, where X is a random vector in \mathbb{R}^n and the function f belongs to \mathcal{F} , a specified class of real, Borel measurable functions on \mathbb{R}^n (e.g. 1-Lipschitz functions or 1-Lipschitz (homogeneous) convex functions).

DEFINITION 2. We say that a random vector X in \mathbb{R}^n has the *concentration property* of order α with constants C, K with respect to a class \mathcal{F} of real, Borel measurable functions on \mathbb{R}^n if for every $f \in \mathcal{F}$ the random variable $f \circ X$ has the concentration property of order α with constants C, K .

The above definition seems justified as there are quite a few examples of pairs (X, \mathcal{F}) satisfying it. For instance, it is by now classical that if X is a standard Gaussian random vector in \mathbb{R}^n then it has the concentration property of order 2 with constants 1, 2 with respect to the class of 1-Lipschitz functions. Also random vectors in \mathbb{R}^n with independent uniformly bounded components have the concentration property of order 2 with constants independent of the dimension with \mathcal{F} being the class of 1-Lipschitz convex functions [19]. The latter example can be extended to arbitrary random vectors with bounded support, but the constants will then also depend on the mixing coefficients associated with the random vector [17].

We now briefly describe one of the most efficient tools for proving the concentration property (especially for product distributions), which has been developed over the past several years, namely the entropy method.

DEFINITION 3. Let ξ be a nonnegative random variable and $\Phi: \mathbb{R}_+ \rightarrow \mathbb{R}$ a convex function such that $\Phi'' > 0$ and $1/\Phi''$ is concave. Define the Φ -entropy of ξ by the formula

$$(4) \quad \text{Ent}_\Phi \xi = \mathbb{E}\Phi(\xi) - \Phi(\mathbb{E}\xi).$$

The most important examples are $\Phi(x) = x^2$ and $\Phi(x) = x \log x$. In these cases Ent_Φ becomes respectively the variance and the usual entropy of a random variable (which will be denoted simply by Ent). The notion is important from the concentration of measure point of view since we have

THEOREM 1 (Herbst argument, see [11], [12]). *Let $X \in \mathbb{R}^n$ be a random variable and \mathcal{F} a class of functions such that $\lambda f \in \mathcal{F}$ for all $f \in \mathcal{F}$ and $\lambda \geq 0$. Assume furthermore that for all $f \in \mathcal{F}$,*

$$(5) \quad \text{Ent} e^{f(X)} \leq C \mathbb{E} |\nabla f(X)|^2 e^{f(X)},$$

and the right-hand side is finite. Then for all $f \in \mathcal{F}$ with $|\nabla f| \leq 1$ and $t \geq 0$,

$$\mathbb{P}(f(X) \geq \mathbb{E}f(X) + t) \leq e^{-t^2/4C}.$$

A crucial property of Ent_Φ is the *tensorization*, which is described in the following

THEOREM 2 (see [4], [10]). *Consider a product probability space (Ω, μ) , where $\Omega = \bigotimes_{i=1}^n \Omega_i$ and $\mu = \bigotimes_{i=1}^n \mu_i$. Then for every nonnegative random variable ξ we have*

$$\text{Ent}_\Phi \xi \leq \sum_{i=1}^n \mathbb{E} \text{Ent}_{\Phi, \mu_i} \xi,$$

where $\text{Ent}_{\Phi, \mu_i} \xi$ denotes the value of the functional Ent_Φ at the function ξ , considered as a function on Ω_i , with the other coordinates fixed.

Thus if a random vector $X \in \mathbb{R}^n$ satisfies the inequality (5) for all $f \in \mathcal{F}$, then so does the random vector $X_1 \oplus \dots \oplus X_d \in (\mathbb{R}^n)^d$, where X_i are independent copies of X , for all functions $f: (\mathbb{R}^n)^d \rightarrow \mathbb{R}$ such that $f(x_1, \dots, x_{i-1}, \cdot, x_{i+1}, \dots, x_d) \in \mathcal{F}$ for all i and $x_1, \dots, x_n \in \mathbb{R}^n$, which can be used to obtain concentration inequalities. This method has led to concentration results for 1-Lipschitz functions of standard Gaussian vectors and 1-Lipschitz convex functions of uniformly bounded product distributions (see [12, Chapter 5]). In a slightly different setting it was also used to obtain concentration results for more general functions of independent random variables and also to some general moment inequalities for such functions [4]. We also mention that inequalities in the spirit of (5) with the left-hand side replaced by $\text{Var } f(X)$ (the so-called Poincaré inequalities) yield concentration property of order 1. There are also similar Łatała–Oleszkiewicz inequalities which imply concentration of order $\alpha \in (1, 2)$ (see [10]).

In this article we will present two results concerning concentration. First, in Section 2 we obtain some sufficient conditions for a real random variable to satisfy the logarithmic Sobolev inequality (5) for convex functions, which yields some subgaussian deviation inequalities. Then in Section 3 we will show that the concentration property of a random vector X with respect to the class of seminorms can be tensorized to obtain concentration inequalities for $X_1 \otimes \dots \otimes X_d$ (where X_i 's are independent copies of X), which gives some new and helps to recover known inequalities for polynomial chaoses. Finally, in the last section we present an application of these inequalities, by presenting a new proof of exponential integrability for Rademacher chaos process.

2. Logarithmic Sobolev inequalities and concentration of measure for convex functions

DEFINITION 4. For $m > 0$ and $\sigma \geq 0$ let $\mathcal{M}(m, \sigma^2)$ denote the class of probability distributions μ on \mathbb{R} for which

$$v^+(A) \leq \sigma^2 \mu(A)$$

for all sets A of the form $A = [x, \infty)$, $x \geq m$, and

$$v^-(A) \leq \sigma^2 \mu(A)$$

for all sets A of the form $A = (-\infty, -x]$, $x \geq m$, where v^+ is the measure on $[m, \infty)$ with density $g(x) = x\mu([x, \infty))$ and v^- is the measure on $(-\infty, -m]$ with density $g(x) = -x\mu((-\infty, x])$.

PROPOSITION 1. *Let μ be a probability distribution on \mathbb{R} . Then the following conditions are equivalent:*

- (i) $\mu \in \mathcal{M}(m, \sigma^2)$ for some m, σ ,
- (ii) we have

$$\begin{aligned} \mu([x + C/x, \infty)) &\leq \alpha \mu([x, \infty)), \\ \mu((-\infty, -x - C/x]) &\leq \alpha \mu((-\infty, -x]), \end{aligned}$$

for some $C > 0$, $\alpha < 1$ and all $x \geq m$,

where the constants in (ii) depend only on the constants in (i) and vice versa. For instance if (i) holds, we can take $C = 2\sigma^2$ and $\alpha = 1/2$.

Proof. Assume (i) holds. Then for $x \geq m$ we have

$$\begin{aligned} \sigma^2 \mu([x, \infty)) &\geq \int_x^{x+2\sigma^2/x} y \mu([y, \infty)) dy \geq x \frac{2\sigma^2}{x} \mu([x + 2\sigma^2/x, \infty)) \\ &= 2\sigma^2 \mu([x + 2\sigma^2/x, \infty)), \end{aligned}$$

which clearly implies the first inequality of (ii). The second inequality follows similarly.

Suppose now that (ii) is satisfied and for $x \geq m$ define the sequence $a_0 = x$, $a_{n+1} = a_n + C/a_n$. Then it is easy to see that $a_n \rightarrow \infty$ and therefore

$$\begin{aligned} \int_x^\infty t \mu([t, \infty)) dt &\leq \sum_{n=0}^\infty a_{n+1} \mu([a_n, \infty)) (a_{n+1} - a_n) \leq K_1 \sum_{n=0}^\infty \alpha^n \mu([a_0, \infty)) \\ &\leq K_2 \mu([x, \infty)). \end{aligned}$$

We can proceed analogously to obtain the condition on the left tail. ■

REMARK. It is also worth noting that for a real random variable X , the condition $\mathcal{L}(X) \in \mathcal{M}(m, \sigma^2)$ is equivalent to $\mathbb{E}X^2 \mathbf{1}_{\{X \geq t\}} \leq (t^2 + 2\sigma^2) \mathbb{P}(X \geq t)$ and $\mathbb{E}X^2 \mathbf{1}_{\{X \leq -t\}} \leq (t^2 + 2\sigma^2) \mathbb{P}(X \leq -t)$ for all $t \geq m$.

EXAMPLE. Of course all measures with bounded support belong to $\mathcal{M}(m, 0)$ for some m . Other examples of measures from $\mathcal{M}(m, \sigma^2)$ are absolutely continuous distributions μ satisfying the inequalities

$$\frac{d}{dt} \log \mu([t, \infty)) \leq -\frac{t}{\sigma^2}, \quad \frac{d}{dt} \log \mu((-\infty, -t]) \leq -\frac{t}{\sigma^2}$$

for $t \geq m$. In particular, if μ has density of the form $e^{-V(x)}$ with $V'(x) \geq x/\sigma^2$ and $V'(-x) \leq -x/\sigma^2$ then $\mu \in \mathcal{M}(1, \sigma^2)$.

Now we are ready to state the main result of this section.

THEOREM 3. *Let X_1, \dots, X_n be independent random variables such that $\mathcal{L}(X_i) \in \mathcal{M}(m, \sigma^2)$ and let $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}$ be a smooth convex function. Then*

$$\text{Ent } e^{\varphi(X_1, \dots, X_n)} \leq C(m, \sigma^2) \mathbb{E} e^{\varphi(X_1, \dots, X_n)} |\nabla \varphi(X_1, \dots, X_n)|^2.$$

Hence for every 1-Lipschitz convex function $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}$ and all $t \geq 0$,

$$\mathbb{P}(\varphi(X_1, \dots, X_n) \geq \mathbb{E}\varphi(X_1, \dots, X_n) + t) \leq e^{-t^2/4C(m, \sigma^2)}.$$

Before we proceed to the proof of Theorem 3, we will need a few lemmas.

LEMMA 1. *Let $\mu \in \mathcal{M}(m, \sigma^2)$. Then for all functions $f: \mathbb{R} \rightarrow \mathbb{R}_+$ which are nonincreasing for $x \leq x_0$ and nondecreasing for $x \geq x_0$, we have*

$$\int_{\tilde{m}}^{\infty} f(x) x \mu([x, \infty)) dx \leq 2\sigma^2 \int_{\mathbb{R}} f(x) d\mu(x),$$

where $\tilde{m} = m \vee (\sqrt{2}\sigma) + 2\sigma^2/(m \vee (\sqrt{2}\sigma))$.

Proof. First notice that by standard approximation arguments the inequalities of Definition 4 are also satisfied for sets $A = (x, \infty)$, $x \geq m$. We have

$$\begin{aligned} (6) \quad \int_{\tilde{m}}^{\infty} f(x) x \mu([x, \infty)) dx &= \int_{\tilde{m}}^{\infty} \int_0^{\infty} \mathbf{1}_{\{s \leq f(x)\}} x \mu([x, \infty)) ds dx \\ &= \int_0^{\infty} v^+(\{x \geq \tilde{m}: f(x) \geq s\}) ds. \end{aligned}$$

The set $A = \{x \geq \tilde{m}: f(x) \geq s\}$ is either a half-line contained in $[m, \infty)$ or a disjoint union of such a half-line and an interval I with left end \tilde{m} . In the former case we have $v^+(A) \leq \sigma^2 \mu(A) \leq \sigma^2 \mu(\{x \in \mathbb{R}: f(x) \geq s\})$.

Now consider the latter case. Denote the right end of I by t . Let $a = m \vee (\sqrt{2}\sigma)$. As $t \geq \tilde{m} = a + 2\sigma^2/a \geq a$, we obtain

$$\begin{aligned} v^+(A) &\leq v^+([\tilde{m}, \infty)) \leq \sigma^2 \mu([a, \infty)) \leq 2\sigma^2 \mu([a, a + 2\sigma^2/a)) \\ &\leq 2\sigma^2 \mu([a, t)) \leq 2\sigma^2 \mu(\{x \leq t: f(x) \geq s\}), \end{aligned}$$

where the second inequality follows from the assumption $\mu \in \mathcal{M}(m, \sigma^2)$, the third from Proposition 1 and the last one from the observation that $x_0 \geq t$ and thus f is nonincreasing for $x \leq t$. Now we can write

$$\begin{aligned} 2\sigma^2 \int_{\mathbb{R}} f(x) d\mu(x) &= 2\sigma^2 \int_{\mathbb{R}} \int_0^\infty \mathbf{1}_{\{s \leq f(x)\}} ds d\mu(x) \\ &= 2\sigma^2 \int_0^\infty \mu(\{x \in \mathbb{R} : f(x) \geq s\}) ds, \end{aligned}$$

which together with (6) allows us to complete the proof. ■

LEMMA 2. *If X is a random variable such that $\mathcal{L}(X) \in \mathcal{M}(m, \sigma^2)$ then $\mathbb{P}(|X| \geq t) \leq C_1(m, \sigma^2)e^{-t^2/C_2(m, \sigma^2)}$ for all $t \geq 0$.*

Proof. Obviously it is sufficient to prove the inequality for $t \geq 4m$. Define $g(x) = \int_x^\infty y\mathbb{P}(X \geq y) dy$. Then for $x \geq m$,

$$xg(x) \leq x\sigma^2\mathbb{P}(X \geq x),$$

and thus

$$\int_z^\infty xg(x) dx \leq \sigma^2g(z), \quad z \geq m.$$

Let $f(z) = \int_z^\infty xg(x) dx$. Since the function $x \mapsto xg(x)$ is continuous, we can rewrite the above inequality as

$$f'(z) \leq -\frac{1}{\sigma^2}zf(z), \quad z \geq m,$$

which gives $f(z) \leq C \exp(-z^2/2\sigma^2)$ with C depending only on m and σ^2 . Now, as g is nonincreasing, for $z \geq m$ we have

$$g(2z) \leq \frac{1}{z^2}f(z) \leq \frac{C}{z^2}e^{-z^2/2\sigma^2}$$

and similarly

$$\mathbb{P}(X \geq 4x) \leq \frac{g(2x)}{4x^2} \leq \frac{C}{4x^4}e^{-x^2/2\sigma^2}$$

for $x \geq m$.

The lower tail can be dealt with analogously. ■

LEMMA 3. *Let $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ be a smooth convex Lipschitz function and X a random variable with $\mathcal{L}(X) \in \mathcal{M}(m, \sigma^2)$. Then there exists a constant $C(m, \sigma^2)$ such that*

$$(7) \quad \int_0^\infty \int_0^\infty \varphi'(x)\varphi'(y)e^{\varphi(y)}\mathbb{P}(X \leq x \wedge y)\mathbb{P}(X \geq x \vee y) dx dy \leq C(m, \sigma^2)\mathbb{E}\varphi'(X)^2e^{\varphi(X)}.$$

Proof. Let us first notice that the left-hand side of (7) is equal to

$$\int_0^\infty \int_0^y \varphi'(x)\varphi'(y)(e^{\varphi(x)} + e^{\varphi(y)})\mathbb{P}(X \leq x)\mathbb{P}(X \geq y) dx dy.$$

Since φ is convex, there exists a point x_0 (possibly 0 or infinity) at which φ attains its minimum on $[0, \infty]$. Moreover φ is nonincreasing on $(0, x_0)$ and nondecreasing on (x_0, ∞) . Therefore for $x, y \in (0, x_0)$ with $x \leq y$ one has

$$\varphi'(x)\varphi'(y)(e^{\varphi(x)} + e^{\varphi(y)}) \leq 2\varphi'(x)^2e^{\varphi(x)}.$$

Thus for \tilde{m} being the constant defined in Lemma 1 we have

$$\begin{aligned} (8) \quad & \int_0^{x_0 \wedge \tilde{m}} \int_0^y \varphi'(x)\varphi'(y)(e^{\varphi(x)} + e^{\varphi(y)})\mathbb{P}(X \leq x)\mathbb{P}(X \geq y) \, dx \, dy \\ & \leq 2 \int_0^{x_0 \wedge \tilde{m}} \varphi'(x)^2e^{\varphi(x)}\mathbb{P}(X \leq x) \int_x^{x_0 \wedge \tilde{m}} \mathbb{P}(X \geq y) \, dy \, dx \\ & \leq 2\tilde{m}\mathbb{E} \int_0^{x_0 \wedge \tilde{m}} \varphi'(x)^2e^{\varphi(x)}\mathbf{1}_{\{X \leq x\}} \, dx \leq 2\tilde{m}^2\mathbb{E}\varphi'(X)^2e^{\varphi(X)}, \end{aligned}$$

where the last inequality follows from the fact that if $x_0 > 0$, then $\varphi'(x)^2e^{\varphi(x)} \leq \varphi'(y)^2e^{\varphi(y)}$ for $y \leq x \leq x_0$.

On the other hand, for $x_0 < x < y$ we have $\varphi'(x)\varphi'(y)(e^{\varphi(x)} + e^{\varphi(y)}) \leq 2\varphi'(y)^2e^{\varphi(y)}$. Obviously this is also the case if $x < x_0 < y$, so

$$\begin{aligned} (9) \quad & \int_{x_0 \vee \tilde{m}}^\infty \int_0^y \varphi'(x)\varphi'(y)(e^{\varphi(x)} + e^{\varphi(y)})\mathbb{P}(X \leq x)\mathbb{P}(X \geq y) \, dx \, dy \\ & \leq 2 \int_{x_0 \vee \tilde{m}}^\infty \varphi'(y)^2e^{\varphi(y)}y\mathbb{P}(X \geq y) \, dy \leq 4\sigma^2\mathbb{E}\varphi'(X)^2e^{\varphi(X)}, \end{aligned}$$

by Lemma 1, since $\mathcal{L}(X) \in \mathcal{M}(m, \sigma^2)$.

So it remains to estimate the integral over the interval $(x_0 \wedge \tilde{m}, x_0 \vee \tilde{m})$. Let us consider two cases:

(i) $x_0 < \tilde{m}$. Then

$$\begin{aligned} (10) \quad & \int_{x_0}^{\tilde{m}} \int_0^y \varphi'(x)\varphi'(y)(e^{\varphi(x)} + e^{\varphi(y)})\mathbb{P}(X \leq x)\mathbb{P}(X \geq y) \, dx \, dy \\ & \leq 2 \int_{x_0}^{\tilde{m}} \int_{x_0}^y \varphi'(y)^2e^{\varphi(y)}\mathbb{P}(X \geq y) \, dx \, dy \\ & \leq 2\tilde{m}\mathbb{E} \int_{x_0}^{\tilde{m}} \varphi'(y)^2e^{\varphi(y)}\mathbf{1}_{\{X \geq y\}} \, dy \leq 2\tilde{m}^2\mathbb{E}\varphi'(X)^2e^{\varphi(X)}. \end{aligned}$$

(ii) $x_0 > \tilde{m}$. We can obviously assume that $\tilde{m} \geq 1$. Then as before

$$\begin{aligned}
 (11) \quad & \int_{\tilde{m}}^{x_0} \int_0^y \varphi'(x)\varphi'(y)(e^{\varphi(x)} + e^{\varphi(y)})\mathbb{P}(X \leq x)\mathbb{P}(X \geq y) \, dx \, dy \\
 & \leq 2 \int_{\tilde{m}}^{x_0} \int_0^y \varphi'(x)^2 e^{\varphi(x)} \mathbb{P}(X \leq x)\mathbb{P}(X \geq y) \, dx \, dy \\
 & = 2 \int_0^{x_0} \int_{x \vee \tilde{m}}^{x_0} \varphi'(x)^2 e^{\varphi(x)} \mathbb{P}(X \leq x)\mathbb{P}(X \geq y) \, dy \, dx \\
 & \leq 2 \int_0^{x_0} \varphi'(x)^2 e^{\varphi(x)} \mathbb{P}(X \leq x) \int_{\tilde{m} \vee x}^{\infty} y \mathbb{P}(X \geq y) \, dy \, dx \\
 & \leq 2\sigma^2 \int_0^{x_0} \varphi'(x)^2 e^{\varphi(x)} \mathbb{P}(X \leq x)\mathbb{P}(X \geq x) \, dx \\
 & \leq 2\sigma^2 \int_0^{\tilde{m}} \varphi'(x)^2 e^{\varphi(x)} \mathbb{P}(X \leq x) \, dx + 2\sigma^2 \int_{\tilde{m}}^{x_0} \varphi'(x)^2 e^{\varphi(x)} x \mathbb{P}(X \geq x) \, dx \\
 & \leq 2\sigma^2(\tilde{m} + 2\sigma^2)\mathbb{E}\varphi'(X)^2 e^{\varphi(X)}.
 \end{aligned}$$

Bringing together (8)–(11) completes the proof. ■

LEMMA 4. Let $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ be a smooth convex Lipschitz function, nonincreasing on $(-\infty, 0)$, and X be a random variable with $\mathcal{L}(X) \in \mathcal{M}(m, \sigma^2)$. Then

$$\begin{aligned}
 & \int_{\{(x,y) \in \mathbb{R}^2 : xy \leq 0\}} \varphi'(x)\varphi'(y)e^{\varphi(y)}\mathbb{P}(X \leq x \wedge y)\mathbb{P}(X \geq x \vee y) \, dx \, dy \\
 & \leq C(m, \sigma^2)\mathbb{E}\varphi'(X)^2 e^{\varphi(X)}.
 \end{aligned}$$

Proof. Assume without loss of generality that $m \geq 1$ and let \tilde{m} be the constant defined in Lemma 1. For $x < 0 < y$ we have either $\varphi'(x)\varphi'(y) \leq 0$, or $\varphi'(x) \leq \varphi'(y) < 0$ and $\varphi(x) \geq \varphi(y)$, so

$$\begin{aligned}
 & \int_{-\infty}^0 \int_0^{\infty} \varphi'(x)\varphi'(y)e^{\varphi(y)}\mathbb{P}(X \leq x)\mathbb{P}(X \geq y) \, dy \, dx \\
 & \leq \int_{-\infty}^0 \int_0^{\infty} \varphi'(x)^2 e^{\varphi(x)} \mathbb{P}(X \leq x)\mathbb{P}(X \geq y) \, dy \, dx \\
 & \leq C(m, \sigma^2) \int_{-\infty}^0 \varphi'(x)^2 e^{\varphi(x)} \mathbb{P}(X \leq x) \, dx,
 \end{aligned}$$

where the last inequality follows from the fact that by Lemma 2,

$$\int_0^\infty \mathbb{P}(X \geq y) dy = \mathbb{E}X_+ \leq C(m, \sigma^2).$$

Also

$$\begin{aligned} \int_0^\infty \int_{-\infty}^0 \varphi'(x)\varphi'(y)e^{\varphi(y)}\mathbb{P}(X \leq y)\mathbb{P}(X \geq x) dy dx \\ \leq \int_{-\infty}^0 \int_0^\infty \varphi'(y)^2 e^{\varphi(y)}\mathbb{P}(X \leq y)\mathbb{P}(X \geq x) dx dy \\ \leq C(m, \sigma^2) \int_{-\infty}^0 \varphi'(y)^2 e^{\varphi(y)}\mathbb{P}(X \leq y) dy. \end{aligned}$$

Now

$$\begin{aligned} \int_{-\infty}^{-\tilde{m}} \varphi'(x)^2 e^{\varphi(x)}\mathbb{P}(X \leq x) dx &\leq \int_{-\infty}^{-\tilde{m}} \varphi'(x)^2 e^{\varphi(x)}(-x)\mathbb{P}(X \leq x) dx \\ &\leq 2\sigma^2 \mathbb{E}\varphi'(X)^2 e^{\varphi(X)} \end{aligned}$$

by Lemma 1, as $\mathcal{L}(X) \in \mathcal{M}(m, \sigma^2)$. Moreover,

$$\begin{aligned} \int_{-\tilde{m}}^0 \varphi'(x)^2 e^{\varphi(x)}\mathbb{P}(X \leq x) dx &= \mathbb{E} \int_{-\tilde{m}}^0 \varphi'(x)^2 e^{\varphi(x)} \mathbf{1}_{\{X \leq x\}} dx \\ &\leq \tilde{m} \mathbb{E}\varphi'(X)^2 e^{\varphi(X)}. \blacksquare \end{aligned}$$

Proof of Theorem 3. We will follow Ledoux’s approach for bounded variables. Due to the tensorization property of entropy (Theorem 2), it is enough to prove the theorem for $n = 1$. Also, by the standard approximation argument, we can restrict our attention to convex Lipschitz functions only. Let now Y be an independent copy of X . By Jensen’s inequality we have

$$\begin{aligned} \text{Ent } e^{\varphi(X)} &= \mathbb{E}\varphi(X)e^{\varphi(X)} - \mathbb{E}e^{\varphi(X)} \log \mathbb{E}e^{\varphi(X)} \\ &\leq \frac{1}{2} \mathbb{E}(\varphi(X) - \varphi(Y))(e^{\varphi(X)} - e^{\varphi(Y)}) \\ &= \mathbb{E}(\varphi(X) - \varphi(Y))(e^{\varphi(X)} - e^{\varphi(Y)})\mathbf{1}_{\{X \leq Y\}} \\ &= \mathbb{E} \int_{\mathbb{R}} \int_{\mathbb{R}} \varphi'(x)\varphi'(y)e^{\varphi(y)}\mathbf{1}_{\{X \leq x \leq Y\}}\mathbf{1}_{\{X \leq y \leq Y\}} dx dy \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} \varphi'(x)\varphi'(y)e^{\varphi(y)}\mathbb{P}(X \leq x \wedge y)\mathbb{P}(X \geq x \vee y) dx dy. \end{aligned}$$

Since $\mathcal{L}(-X) \in \mathcal{M}(m, \sigma^2)$, we can assume that the minimal value of φ is attained at some point of the right half-axis (possibly at ∞). Splitting now

the double integral into four integrals depending on the signs of x and y and using Lemmas 3 and 4 we obtain the desired inequality. Note that we can use Lemma 3 to handle the integration over $(-\infty, 0)^2$ again by change of variables and the fact that $\mathcal{L}(-X) \in \mathcal{M}(m, \sigma^2)$. The tail inequality follows from the entropy estimates by Theorem 1. ■

REMARK. One would obviously like to characterize all real random variables X such that the random vectors (X_1, \dots, X_n) (where X_i 's are independent copies of X) have the concentration property of order 2 for 1-Lipschitz convex functions with constants independent of the dimension n . Each such variable must of course have the concentration property of order 2 itself. This, however, is not sufficient, as concentration with respect to convex functions implies hypercontractivity (see [5]), which is equivalent to some regularity of the tail. In particular, it follows that $\mathbb{E}X^2 \mathbf{1}_{\{X \geq t\}} \leq Ct^2 \mathbb{P}(X \geq t)$ for t large. This condition is weaker than $\mathcal{L}(X) \in \mathcal{M}(m, \sigma^2)$ for some m, σ but hypercontractivity is also weaker than the concentration property of order 2, uniformly over the dimension n .

We would also like to point out that all Borel probability measures μ on the real line, which satisfy the logarithmic Sobolev inequality for all smooth (not necessarily log-convex) functions, belong to $\mathcal{M}(m, \sigma^2)$ for some m, σ . Thus $\bigcup_{m, \sigma} \mathcal{M}(m, \sigma^2)$ is strictly larger than the class of all measures satisfying the logarithmic Sobolev inequality for all smooth functions. More precisely, we have the following

PROPOSITION 2. *Let μ be a Borel probability measure on \mathbb{R} for which there exists $C < \infty$ such that for all smooth functions,*

$$\text{Ent } f(X)^2 \leq C \mathbb{E}|f'(X)|^2,$$

where X is a random variable with the law μ . Then there exist constants $m, \sigma < \infty$ such that $\mu \in \mathcal{M}(m, \sigma^2)$.

Proof. From the Bobkov–Götze criterion (see [2]) it follows that if n is the density of the absolutely continuous part of μ and M is a median of μ , then for some constant K and all $x \geq M$,

$$\mu([x, \infty)) \log \frac{1}{\mu([x, \infty))} \int_M^x \frac{1}{n(t)} dt < K.$$

Thus (since μ has the concentration property of order 2) from Hölder's inequality we get (using the above inequality for $x + 1/x$ instead of x)

$$\begin{aligned} \frac{D}{K} x^2 \mu([x + 1/x, \infty)) &\leq \frac{1}{\int_M^{x+1/x} (1/n(t)) dt} \leq \frac{\mu([x, x + 1/x])}{\left(\int_x^{x+1/x} 1 dt\right)^2} \\ &= x^2 \mu([x, x + 1/x]) \end{aligned}$$

for some $D > 0$ and x large enough, which implies $\mu([x + 1/x, \infty)) \leq \alpha\mu([x, \infty))$ with $\alpha = (K/D)/(K/D + 1)$. Since a similar condition on the left tail can be proven analogously, the claim follows by Proposition 1. ■

3. Concentration for seminorms on the tensor product and random chaoses

3.1. A tensorization inequality for seminorms

THEOREM 4. *Consider a random vector $X \in \mathbb{R}^n$ for which there exists a constant K such that for every seminorm $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}_+$ we have $\mathbb{E}\varphi(X) < \infty$ and for every $p \geq 1$,*

$$\|\varphi(X) - \mathbb{E}\varphi(X)\|_p \leq K\sqrt{p} \sup_{|\alpha|=1} \varphi(\alpha).$$

Then if X_1, \dots, X_d are independent copies of X and $\psi: \bigotimes_{i=1}^d \mathbb{R}^n \rightarrow \mathbb{R}_+$ a seminorm, we have

$$\begin{aligned} & \left\| \psi\left(\bigotimes_{i=1}^d X_i\right) - \mathbb{E}\psi\left(\bigotimes_{i=1}^d X_i\right) \right\|_p \\ & \leq K_d \sum_{I \subseteq \{1, \dots, d\}, I \neq \emptyset} p^{\#I/2} \mathbb{E} \sup_{|\alpha_k| \leq 1, k \in I} \psi\left(\bigotimes_{i=1}^d X_{i, I, (\alpha_k)_{k \in I}}\right), \end{aligned}$$

where

$$X_{i, I, (\alpha_k)_{k \in I}} = \begin{cases} X_i & \text{if } i \notin I, \\ \alpha_i & \text{if } i \in I, \end{cases}$$

K_d is a constant depending only on K and d , and $|\alpha|$ denotes the Euclidean norm of the vector α .

Proof. We use induction on d . For $d = 1$, the statement of the theorem is just its hypothesis. Assume that the statement is true for fewer than $d \geq 2$ copies of X . Using conditionally the induction assumption for $d_1 = 1$ and the function

$$\varphi_1(x) = \psi(X_1 \otimes \dots \otimes X_{d-1} \otimes x),$$

we obtain

$$\mathbb{E}_{X_d} \left| \psi\left(\bigotimes_{i=1}^d X_i\right) - \mathbb{E}_{X_d} \psi\left(\bigotimes_{i=1}^d X_i\right) \right|^p \leq K^p p^{p/2} \sup_{|\alpha| \leq 1} \psi\left(\bigotimes_{i=1}^{d-1} X_i \otimes \alpha\right)^p.$$

Now notice that the function $\varphi_2(x) = \sup_{|\alpha| \leq 1} \psi(x \otimes \alpha)$ is a seminorm on $\bigotimes_{i=1}^{d-1} \mathbb{R}^d$ and thus we can apply the induction assumption and the triangle inequality in L^p , which together with the Fubini Theorem gives

$$\begin{aligned}
 (12) \quad & \mathbb{E} \left| \psi \left(\bigotimes_{i=1}^d X_i \right) - \mathbb{E}_{X_d} \psi \left(\bigotimes_{i=1}^d X_i \right) \right|^p \\
 & \leq \tilde{K}_{d-1}^p \sum_{I \subseteq \{1, \dots, d\}, d \in I} p^{p\#I/2} \left(\mathbb{E} \sup_{|\alpha_k| \leq 1, k \in I} \psi \left(\bigotimes_{i=1}^d X_{i,I,(\alpha_k)_{k \in I}} \right) \right)^p,
 \end{aligned}$$

where \tilde{K}_{d-1} depends only on K_{d-1} and d .

Now we would like to estimate $\mathbb{E} |\mathbb{E}_{X_d} \psi(\bigotimes_{i=1}^d X_i) - \mathbb{E} \psi(\bigotimes_{i=1}^d X_i)|^p$. To this end consider $\varphi_3: \bigotimes_{i=1}^{d-1} \mathbb{R}^n \rightarrow \mathbb{R}_+$ defined as $\varphi_3(x) = \mathbb{E} \psi(x \otimes X_d)$. It is easy to see that φ_3 is a seminorm and thus, by the induction assumption,

$$\begin{aligned}
 (13) \quad & \mathbb{E} \left| \varphi_3 \left(\bigotimes_{i=1}^{d-1} X_i \right) - \mathbb{E} \varphi_3 \left(\bigotimes_{i=1}^{d-1} X_i \right) \right|^p \\
 & \leq \tilde{K}_{d-1}^p \sum_{I \subseteq \{1, \dots, d-1\}, I \neq \emptyset} p^{p\#I/2} \left(\mathbb{E} \sup_{|\alpha_k| \leq 1, k \in I} \varphi_3 \left(\bigotimes_{i=1}^{d-1} X_{i,I,(\alpha_k)_{k \in I}} \right) \right)^p.
 \end{aligned}$$

Now it is enough to note that for each $I \subseteq \{1, \dots, d-1\}$,

$$\mathbb{E} \sup_{|\alpha_k| \leq 1, k \in I} \varphi_3 \left(\bigotimes_{i=1}^{d-1} X_{i,I,(\alpha_k)_{k \in I}} \right) \leq \mathbb{E} \sup_{|\alpha_k| \leq 1, k \in I} \psi \left(\bigotimes_{i=1}^d X_{i,I,(\alpha_k)_{k \in I}} \right),$$

which together with (12) and (13) completes the proof. ■

Notice that by Theorem 3 for all product random vectors $X \in \mathbb{R}^n$ with 1-dimensional marginals in $\mathcal{M}(m, \sigma^2)$ and all seminorms $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}_+$ we have $\|(\varphi(X) - \mathbb{E} \varphi(X))_+\|_p \leq K \sqrt{p} \sup_{|\alpha| \leq 1} \varphi(\alpha)$, with K depending only on m and σ^2 . Thus the same proof, with formal changes only, gives

THEOREM 5. *Let $X_1, \dots, X_n \in \mathbb{R}^n$ be independent random vectors with independent components and all 1-dimensional marginals in $\mathcal{M}(m, \sigma^2)$. Then for every seminorm $\psi: \bigotimes_{i=1}^d \mathbb{R}^n \rightarrow \mathbb{R}_+$ we have*

$$\begin{aligned}
 & \left\| \left(\psi \left(\bigotimes_{i=1}^d X_i \right) - \mathbb{E} \psi \left(\bigotimes_{i=1}^d X_i \right) \right)_+ \right\|_p \\
 & \leq K_d \sum_{I \subseteq \{1, \dots, d\}, I \neq \emptyset} p^{\#I/2} \mathbb{E} \sup_{|\alpha_k| \leq 1, k \in I} \psi \left(\bigotimes_{i=1}^d X_{i,I,(\alpha_k)_{k \in I}} \right),
 \end{aligned}$$

where K_d depends only on m, σ^2 and d .

By the Chebyshev inequality we can obtain from the above theorems a corollary concerning the tail behaviour of $\psi(\bigotimes_{i=1}^d X_i)$. We give it only for Theorem 4; for Theorem 5 it is analogous but deals with the upper tail only.

COROLLARY 1. Under the assumption of Theorem 4 there exist constants K_d , depending only on K and d , such that for all $t \geq 1$,

$$\begin{aligned} & \mathbb{P}\left(\left|\psi\left(\bigotimes_{i=1}^d X_i\right) - \mathbb{E}\psi\left(\bigotimes_{i=1}^d X_i\right)\right|\right) \\ & \geq K_d \sum_{I \subseteq \{1, \dots, d\}, I \neq \emptyset} t^{\#I/2} \mathbb{E} \sup_{|\alpha_k| \leq 1, k \in I} \psi\left(\bigotimes_{i=1}^d X_{i, I, (\alpha_k)_{k \in I}}\right) \leq e^{-t}. \end{aligned}$$

3.2. Chaos random variables. The above theorems can be rewritten in terms of decoupled polynomial chaoses. Below we present such a version as a corollary (actually equivalent to Theorem 4), so that the reader could compare it with existing results on chaos random variables.

Let $X^{(1)}, \dots, X^{(d)}$ be independent copies of $X = (X_1, \dots, X_n) \in \mathbb{R}^n$ and consider a *homogeneous decoupled chaos* of order d , i.e. a random variable of the form

$$(14) \quad Z = \sup_{t \in \mathcal{T}} \left| \sum_{i_1, \dots, i_d=1}^n t_{i_1 \dots i_d} X_{i_1}^{(1)} \dots X_{i_d}^{(d)} \right|,$$

where \mathcal{T} is a countable, bounded set of functions $t: \{1, \dots, n\}^d \rightarrow \mathbb{R}$.

Let us introduce

DEFINITION 5. For $I \subseteq \{1, \dots, d\}$ let

$$\|\mathcal{T}\|_I = \mathbb{E} \sup_{t \in \mathcal{T}} \sup_{|\alpha^{(k)}| \leq 1, k \in I} \left| \sum_{i_1, \dots, i_d=1}^n t_{i_1 \dots i_d} \prod_{k \in I} \alpha_{i_k}^{(k)} \prod_{k \notin I} X_{i_k}^{(k)} \right|,$$

where the second supremum is taken over all $(\alpha^{(k)})_{k \in I} \in (\mathbb{R}^n)^{\#I}$, and $|\alpha^{(k)}|$ stands for the Euclidean norm of the vector $\alpha^{(k)}$.

COROLLARY 2. Let $X = (X_1, \dots, X_n)$ have the concentration property with constants C, K with respect to 1-Lipschitz seminorms on \mathbb{R}^n . Then for any integer $d \geq 1$, there exists a constant K_d , depending only on d and C, K , such that for any homogeneous chaos Z (as defined in (14)) and any $p \geq 1$,

$$(15) \quad \|Z - \mathbb{E}Z\|_p \leq K_d \sum_{I \subseteq \{1, \dots, d\}, I \neq \emptyset} p^{\#I/2} \|\mathcal{T}\|_I.$$

COROLLARY 3. There exist constants K_d such that for all $t \geq 1$,

$$\mathbb{P}\left(|Z - \mathbb{E}Z| \geq K_d \sum_{I \subseteq \{1, \dots, d\}, I \neq \emptyset} t^{\#I/2} \|\mathcal{T}\|_I\right) \leq e^{-t}.$$

REMARK. Usually one is interested in the “undecoupled” chaos, i.e. a random variable of the form

$$Z = \sup_{t \in \mathcal{T}} \left| \sum_{i_1, \dots, i_d=1}^n t_{i_1 \dots i_d} X_{i_1} \cdots X_{i_d} \right|,$$

where X_1, \dots, X_n are independent random variables and for all $t \in \mathcal{T}$ the number $t_{i_1 \dots i_d}$ is invariant under permutations of coordinates and nonzero only if the coordinates are pairwise distinct. The analogues of the above theorems for such chaoses generated by Gaussian variables were obtained by Borell [3] and Arcones and Giné [1]. The Rademacher case was considered by Talagrand [18] (chaos of order 2) and Boucheron, Bousquet, Lugosi and Massart [4] (chaos of arbitrary order d). It is easy to see that each decoupled chaos can be represented as an undecoupled one, but the aforementioned results do not recover the lower tail inequalities (except for the case $d = 2$). Also the methods are quite different and do not allow treating both Gaussian and Rademacher variables in a unified way.

3.2.1. Chaoses generated by symmetric random variables with log-concave tails. Now we would like to point out that the proof of Theorem 4 can be actually used in a slightly different setting, namely for chaoses generated by independent random variables with logarithmically concave tails. Such variables have been investigated by Latała [7] and Łochowski [14].

DEFINITION 6. Let $\mathcal{N} = (X_i^{(k)})_{k \leq d, i \leq n}$ be a matrix of independent symmetric random variables with logarithmically concave tails, i.e. random variables such that the functions

$$\mathcal{N}_i^{(k)}(t) = -\log \mathbb{P}(|X_i^{(k)}| \geq t), \quad t \geq 0,$$

are convex. Furthermore assume (as a matter of normalization) that

$$\inf\{t: \mathcal{N}_i^{(k)}(t) \geq 1\} = 1,$$

and define modified functions $\tilde{\mathcal{N}}_i^{(k)}$ by the formula

$$\tilde{\mathcal{N}}_i^{(k)}(t) = \begin{cases} t^2 & \text{for } |t| \leq 1. \\ \mathcal{N}_i^{(k)}(|t|) & \text{for } |t| \geq 1. \end{cases}$$

Let now \mathcal{T} be a countable set of functions $t: \{1, \dots, n\}^d \rightarrow \mathbb{R}$ and Z a random variable defined by (14). Moment estimates for Z will be expressed in terms of the following quantities:

DEFINITION 7. For $I \subseteq \{1, \dots, d\}$ and $p \geq 1$ define

$$\|\mathcal{T}\|_{\mathcal{N}, I, p} = \mathbb{E} \sup_{t \in \mathcal{T}} \sup_{\alpha^{(k)} \in \mathcal{A}_{k,p}, k \in I} \left| \sum_{i_1, \dots, i_d=1}^n t_{i_1 \dots i_d} \prod_{k \in I} \alpha_{i_k}^{(k)} \prod_{k \notin I} X_{i_k}^{(k)} \right|,$$

where

$$\mathcal{A}_{k,p} = \left\{ \alpha \in \mathbb{R}^n : \sum_{i=1}^n \tilde{\mathcal{N}}_i^{(k)}(\alpha_i) \leq p \right\}.$$

The following result was proved for $d = 1$ by Latała [7] and for arbitrary d by Łochowski [14].

THEOREM 6 (Latała, Łochowski). *There exist constants K_d (depending only on d) such that for all $p \geq 1$,*

$$\frac{1}{K_d} \sum_{I \subseteq \{1, \dots, d\}} \|\mathcal{T}\|_{\mathcal{N}, I, p} \leq \|Z\|_p \leq K_d \sum_{I \subseteq \{1, \dots, d\}} \|\mathcal{T}\|_{\mathcal{N}, I, p}.$$

PROPOSITION 3. *The conclusion of Theorem 6 for $d = 1$ implies that conclusion for arbitrary d .*

Proof. When we rewrite the inequalities of Theorem 6 in the language of seminorms, the proof becomes analogous to the proof of Theorem 4. One has simply to notice that the factors $p^{\#I/2}$ do not appear, as the dependence on p is incorporated in the sets $\mathcal{A}_{k,p}$ which replace the unit Euclidean ball in the supremum.

As for the lower estimate, the proof is even simpler but we present it here for the sake of completeness (written in the “chaos language”). Obviously $\|Z\|_p \geq \|Z\|_1 = \|\mathcal{T}\|_{\mathcal{N}, \emptyset, p}$. Moreover for any nonempty set $I \subseteq \{1, \dots, d\}$, say with $r \in I$, we have by the induction hypothesis

$$\begin{aligned} \|Z\|_p &\geq \frac{1}{K_1} \left\| \sup_{t \in \mathcal{T}} \sup_{\alpha^{(r)} \in \mathcal{A}_{r,p}} \left| \sum_{i_1, \dots, i_d=1}^n t_{i_1 \dots i_d} \prod_{k \neq r} X_{i_k}^{(k)} \alpha_{i_r}^{(r)} \right| \right\|_p \\ &\geq \frac{1}{K_1 K_{d-1}} \|\mathcal{T}\|_{\mathcal{N}, I, p}. \quad \blacksquare \end{aligned}$$

Using the Chebyshev inequality and the Paley–Zygmund inequality together with the hypercontractive properties of chaoses (see [15]) we obtain

COROLLARY 4. *There exist constants K_d such that for all $t \geq 1$,*

$$\begin{aligned} \mathbb{P}\left(Z \geq K_d \sum_{I \subseteq \{1, \dots, d\}} \|\mathcal{T}\|_{\mathcal{N}, I, t}\right) &\leq e^{-t}, \\ \mathbb{P}\left(Z \geq \frac{1}{K_d} \sum_{I \subseteq \{1, \dots, d\}} \|\mathcal{T}\|_{\mathcal{N}, I, t}\right) &\geq e^{-t} \wedge c. \end{aligned}$$

REMARK. The bound of Theorem 6 is also valid in the case of undecoupled chaoses due to the decoupling results by de la Peña and Montgomery-Smith [16], which say that tails and moments of the decoupled and undecoupled chaos with the same coefficients are equivalent.

REMARK. All estimates presented so far are expressed in terms of expected values of empirical processes, which themselves are in general troublesome and difficult to estimate. One would obviously want to obtain moment estimates in terms of deterministic quantities at least in the real-valued case (i.e. when \mathcal{T} is a singleton). This has been done by Latała in [7] for $d = 2$ and log-concave random variables, and recently in [8] for arbitrary d and Gaussian chaoses.

4. An application. Finally, we would like to argue that estimates in the spirit of Section 3, although nondeterministic, may be of some use. We will demonstrate it by presenting a sketch of a new (at least to the author’s best knowledge) proof of exponential integrability of a generalized Rademacher chaos process for general d , which we believe is simpler than the preceding ones. The general result and the proof for $d \leq 2$ may be found in the monograph by Ledoux and Talagrand [13].

Let us first recall the general setting. We deal with a Banach space B for which there exists a countable set D of linear functionals from the unit ball of B' such that for each $x \in B$ we have $\|x\| = \sup_{f \in D} |f(x)|$. A B -valued random variable X is a *homogeneous Rademacher chaos* of order d if there is a sequence $(x_{i_1 \dots i_d})_{i_1, \dots, i_d \in \mathbb{N}} \in B$ ($x_{i_1 \dots i_d}$ invariant under permutation of coordinates and nonzero only if i_1, \dots, i_d are pairwise distinct) such that for every $f \in D$ the multiple series $\sum_{i_1, \dots, i_d} f(x_{i_1 \dots i_d}) \varepsilon_{i_1} \cdots \varepsilon_{i_d}$ converges almost surely and $(\sum_{i_1, \dots, i_d} f(x_{i_1 \dots i_d}) \varepsilon_{i_1} \cdots \varepsilon_{i_d})_{f \in D}$ has the same distribution as $(f(X))_{f \in D}$.

THEOREM 7. *Let X be a homogeneous Rademacher chaos of order d . Then*

$$\mathbb{E}e^{\alpha\|X\|^{2/d}} < \infty \quad \text{for all } \alpha \in \mathbb{R}.$$

Before we proceed to the proof, we need

LEMMA 5. *There exist constants L_d (depending only on d) with the property that for every $\alpha > 0$ there exists a constant $\varepsilon(\alpha)$ such that for every homogeneous Rademacher chaos X of order d and every M satisfying $\mathbb{P}(\|X\| > M) < \varepsilon(\alpha)$ we have*

$$\mathbb{P}(\|X\| > L_d M t^{d/2}) < e^{-\alpha t} \quad \text{for all } t \geq 1.$$

Proof. Rademacher variables are log-concave, so we can apply Corollary 4 for finite sums. It is easy to see that

$$(16) \quad \mathcal{A}_{k,p} = \mathcal{A}_p = \left\{ (\alpha_i) : \sum \alpha_i^2 \leq p, |\alpha_i| \leq 1 \right\}.$$

Due to convergence, using standard arguments one can extend the tail estimates to the general Rademacher chaos X . Therefore, set

$$\phi(t) = \sum_{I \subseteq \{1, \dots, d\}} \|\mathcal{T}\|_{\mathcal{N}, I, t},$$

where $\mathcal{T} = \{(f(x_{i_1 \dots i_d})) : f \in D\}$. From (16) it follows that

$$(17) \quad \phi(xt) \leq t^{d/2} \phi(x)$$

for all $x \geq 0$ and $t \geq 1$. As ϕ is increasing, by the last inequality it is continuous. If $\phi(x) \leq K_d M$ for every x (K_d being the constant from Corollary 4), we have, for $t \geq 1$,

$$\mathbb{P}(\|X\| \geq MK_d^2 t^{d/2}) \leq \inf_x \mathbb{P}(\|X\| \geq K_d \phi(x)) \leq \inf_x e^{-x} = 0.$$

Otherwise $K_d M = \phi(x)$ for some x . Thus

$$\varepsilon \geq \mathbb{P}(\|X\| \geq M) = \mathbb{P}(\|X\| \geq \phi(x)/K_d) \geq c \wedge e^{-x},$$

which for ε small enough yields $x \geq -\log \varepsilon$. Moreover (17) gives $K_d M t^{d/2} \geq \phi(tx)$ for $t \geq 1$ and thus for $\varepsilon < e^{-\alpha}$,

$$\mathbb{P}(\|X\| \geq K_d^2 M t^{d/2}) \leq \mathbb{P}(\|X\| \geq K_d \phi(tx)) \leq e^{-tx} \leq e^{t \log \varepsilon} \leq e^{-\alpha t}. \blacksquare$$

Proof of Theorem 7 (sketch). We will proceed by induction on d . Let $S_n = \sum_{i_1, \dots, i_d \geq n} x_{i_1 \dots i_d} \varepsilon_{i_1} \dots \varepsilon_{i_d}$. Then $\|S_n\|$ is a reversed submartingale with $\mathbb{E}\|S_n\| \leq \mathbb{E}\|X\|$ and thus it converges to some random variable, which by the zero-one law must be almost surely constant. Thus there exists M such that for all $\varepsilon > 0$ we have $\mathbb{P}(\|S_n\| \geq M) < \varepsilon$ for n large enough. By Lemma 5, for every $\beta > 0$ there is n such that for all $t \geq 1$,

$$\mathbb{P}(\|S_n\| \geq L_d M t^{d/2}) < e^{-\beta t}$$

or equivalently

$$\mathbb{P}(\alpha \|S_n\|^{2/d} \geq t) \leq \exp\left(-\frac{\beta t}{\alpha (L_d M)^{2/d}}\right)$$

for $t \geq (L_d M)^{2/d} \alpha$, which clearly implies $\mathbb{E}e^{\alpha \|S_n\|^{2/d}} < \infty$ for $\alpha < \beta / (L_d M)^{2/d}$. Since $X - S_n$ is a finite sum of chaoses of orders lower than d (when $d > 1$) or a bounded random variable (for $d = 1$), its integrability properties allow us to use Hölder’s inequality and obtain $\mathbb{E}e^{\alpha \|X\|^{2/d}} < \infty$ also for $\alpha < \beta / (L_d M)^{d/2}$. This allows us to finish the proof as β can be chosen arbitrarily large. \blacksquare

Acknowledgements. The author would like to express his gratitude to Prof. Rafał Łatała for introducing him to the subject and all the inspiring conversations.

References

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

- [2] S. G. Bobkov and F. Götze, *Exponential integrability and transportation cost related to logarithmic Sobolev inequalities*, J. Funct. Anal. 163 (1999), 1–28.
- [3] C. Borell, *On a Taylor series of a Wiener polynomial*, in: Seminar Notes on Multiple Stochastic Integration, Polynomial Chaos and Their Integration, Case Western Reserve Univ., Cleveland, 1984.
- [4] S. Boucheron, O. Bousquet, G. Lugosi and P. Massart, *Moment inequalities for functions of independent random variables*, Ann. Probab. 33 (2005), 514–560.
- [5] P. Hitczenko, S. Kwapiień, W. V. Li, G. Schechtman, T. Schlumprecht and J. Zinn, *Hypercontractivity and comparison of moments of iterated maxima and minima of independent random variables*, Electronic J. Probab. 3 (1998).
- [6] R. Latała, *Tail and moment estimates for sums of independent random vectors with logarithmically concave tails*, Studia Math. 118 (1996), 301–304.
- [7] —, *Tail and moment estimates for some types of chaos*, Studia Math. 135 (1999), 39–53.
- [8] —, *Estimation of moments and tails of Gaussian chaoses*, preprint, 2005, <http://www.arxiv.org/abs/math.PR/0505313>.
- [9] R. Latała and R. Łochowski, *Moment and tail estimates for multidimensional chaos generated by positive random variables with logarithmically concave tails*, Progr. Probab. 56 (2003), 77–92.
- [10] R. Latała and K. Oleszkiewicz, *Between Sobolev and Poincaré*, in: Geometric Aspects of Functional Analysis, Israel Seminar (GAFA) 1996–2000, Lecture Notes in Math. 1745, Springer, Berlin, 2000, 147–168.
- [11] M. Ledoux, *On Talagrand’s deviation inequalities for product measures*, ESAIM: Probab. Statist. 1 (1996), 63–87.
- [12] —, *The Concentration of Measure Phenomenon*, Math. Surveys Monogr. 89, Amer. Math. Soc., 2001.
- [13] M. Ledoux and M. Talagrand, *Probability in Banach Spaces*, Springer, New York, 1991.
- [14] R. Łochowski, *Moment and tail estimates for multidimensional chaoses generated by symmetric random variables with logarithmically concave tails*, preprint.
- [15] V. H. de la Peña and E. Giné, *Decoupling. From Dependence to Independence*, Springer, 1999.
- [16] V. H. de la Peña and S. Montgomery-Smith, *Bounds on the tail probability of U -statistics and quadratic forms*, Bull. Amer. Math. Soc. 31 (1994), 223–227.
- [17] P. M. Samson, *Concentration of measure inequalities for Markov chains and Φ -mixing processes*, Ann. Probab. 28 (2000), 416–461.
- [18] M. Talagrand, *New concentration inequalities in product spaces*, Invent. Math. 126 (1996), 505–563.
- [19] —, *A new look at independence*, Ann. Probab. 24 (1996), 1–34.

Radosław Adamczak
Institute of Mathematics
Polish Academy of Sciences
Śniadeckich 8, P.O. Box 21
00-956 Warszawa 10, Poland
E-mail: R.Adamczak@impan.gov.pl

Received January 10, 2005;
received in final form July 20, 2005

(7429)