L_p moments of random vectors via majorizing measures

Olivier Guédon, Mark Rudelson *

Abstract

For a random vector X in $\mathbb{R}^n$, we obtain bounds on the size of a sample, for which the empirical p-th moments of linear functionals are close to the exact ones uniformly on a convex body $K \subset \mathbb{R}^n$. We prove an estimate for a general random vector and apply it to several problems arising in geometric functional analysis. In particular, we find a short Lewis type decomposition for any finite dimensional subspace of L_p. We also prove that for an isotropic log-concave random vector, we only need $\lfloor n^{p/2} \log n \rfloor$ sample points so that the empirical p-th moments of the linear functionals are almost isometrically the same as the exact ones. We obtain a concentration estimate for the empirical moments. The main ingredient of the proof is the construction of an appropriate majorizing measure to bound a certain Gaussian process.

1 Introduction

In many problems of geometric functional analysis it is necessary to approximate a given random vector by an empirical sample. More precisely, given a random vector $X \in \mathbb{R}^n$, we want to find the smallest number m such that the properties of X can be recovered from the empirical measure $1/m \sum_{j=1}^m \delta_{X_j}$, constructed with independent copies $X_1, \dots, X_m$ of the vector X. In particular, for $p \geq 2$ and for $y \in \mathbb{R}^n$, we want to approximate the moments $\mathbb{E}|\langle X, y\rangle|^p$ by the empirical averages $1/m \sum_{j=1}^m |\langle X_j, y\rangle|^p$ with high probability. Moreover, we require this approximation to be uniform over y belonging to some convex symmetric set in $\mathbb{R}^n$. A problem of this type was considered

*Research was supported in part by NSF grant DMS-024380.

in [5]. Formulated in analytic language, it asks about finding the smallest m and a set of points $x_1, \dots x_m \in X$ such that for any function f from an n-dimensional function space $F \subset L_1(X, \mu)$,

$$(1-\varepsilon)\|f\|_1 \le \frac{1}{m}\sum_{j=1}^{m} |f(x_j)| \le (1-\varepsilon)\|f\|_1.$$

Another example of such problems originates in Computer Science. The probabilistic algorithm for estimating the volume of an n-dimensional convex body, constructed by Kannan, Lovász, and Simonovits [13] required to bring the body to a nearly isotropic position as a preliminary step. To this end, one has to sample m random points $x_1, \dots, x_m$ in the body L so that the empirical isotropy tensor will be close to the exact one, namely

$$\left\| \frac{1}{m}\sum_{j=1}^{m} x_j \otimes x_j - \frac{1}{\text{vol}(L)} \int_L x \otimes x \, dx \right\| < \varepsilon. \tag{1}$$

This problem was attacked with different probabilistic techniques. The original estimate of [13] was significantly improved by Bourgain [4]. Using the decoupling method he proved that $m = C(\varepsilon) n \log^3 n$ vectors $x_1, \dots x_m$ uniformly distributed in the body L satisfy (1) with high probability. This estimate was farther improved to $(Cn/\varepsilon^2) \cdot \log^2(Cn/\varepsilon^2)$ in [25], [26]. The proof in [25] used majorizing measures, while the later proof in [26] was based on the non-commutative Khinchine inequality.

These problems were put into a general framework by Giannopoulos and Milman [9], who related them to the concentration properties of a random vector. Let $\alpha > 0$ and let ν be a probability measure on (X, Ω). For a function $f : X \to \mathbb{R}$ define the ψ_α-norm by

$$\|f\|_{\psi_\alpha} = \inf\{\lambda > 0 \mid \int_X \exp(|f|/\lambda)^\alpha d\nu \le 2\}.$$

Chebychev's inequality shows that the functions with bounded ψ_α-norm are strongly concentrated, namely $\nu\{x \mid |f(x)| > \lambda t\} \le C \exp(-t^\alpha)$. Let μ be a Borel measure in $\mathbb{R}^n$. It is called isotropic if

$$\int_{\mathbb{R}^n} x \otimes x \, d\mu(x) = \text{Id},$$

where Id is the identity operator in $\mathbb{R}^n$. Note that this normalization is consistent with the one used in [13, 25, 26]. The normalization used in

[19, 9] differs from it by the multiplicative coefficient L_μ^2, where L_μ is the *isotropic constant* of μ (see [19]).

The paper [9] considers isotropic measures which satisfy the ψ_α-condition for scalar products:

$$\|\langle \cdot, y\rangle\|_{\psi_\alpha} \le C$$

for all $y \in S^{n-1}$. Here and below $C, c, \dots$ denote absolute constants, whose value may change at each occurrence.

Note that by Borell's lemma, any log-concave measure in $\mathbb{R}^n$ satisfies the ψ_1-condition [20], [19]. Let $p \ge 1$ and let μ be an isotropic log-concave measure satisfying the ψ_α condition for scalar products with some $\alpha \in [1, 2]$. The central result of [9] provides an estimate for the minimal size of a set of independent random vectors $X_1, \dots, X_m$ distributed according to the measure μ such that the empirical p-moments satisfy the inequality

$$\Gamma_1(p) \le \left(\frac{1}{m} \sum_{j=1}^{m} |\langle X, y\rangle|^p \right)^{1/p} \le \Gamma_2(p), \quad \forall y \in S^{n-1}. \tag{2}$$

The ψ_α-condition implies that the $L_p(\mu)$ and $L_2(\mu)$-norms of the function $f_y(x) = \langle x, y\rangle$ are equivalent. Thus the inequality (2) means that the empirical p-moment of f_y is equivalent to the real p-moment up to a constant coefficient.

In the present paper we use a different approach to this problem based on the majorizing measure technique developed by Talagrand [28]. This approach lead to breakthrough results in various problems in probabilistic combinatorics and analysis (see [28] and references therein). In a similar context the majorizing measures were applied in [27] to select small almost orthogonal submatrices of an orthogonal matrix, and in [25] to prove the estimate (1) with small m.

To state the results we have to introduce some notation. Let $(\mathbb{R}^n, \langle \cdot, \cdot \rangle)$ be a Euclidean space, and let $|\cdot|_2$ be the associated Euclidean norm. For a symmetric convex body K in $\mathbb{R}^n$, we denote by $\|\cdot\|_K$ the norm, whose unit ball is K, and by $K^o = \{y \in \mathbb{R}^n \mid \forall x \in K, \langle x, y\rangle \le 1\}$ the polar of K. We assume that the body K has the modulus of convexity of power type $q \ge 2$ (see Section 2 for the definition). Classical examples of convex bodies satisfying this property are unit balls of finite dimensional subspaces of L_q [6] or of non-commutative L_q-spaces (like Schatten trace class matrices [29]). We denote by D the radius of the symmetric convex set K i.e. the smallest

D such that $K \subset DB_2^n$. For every $1 \leq q \leq +\infty$, we define q^* to be the conjugate of q, i.e. $1/q + 1/q^* = 1$.

Given a random vector X in $\mathbb{R}^n$, let $X_1, \ldots, X_m$ be m independent copies of X. Let $K \subset \mathbb{R}^n$ be a convex symmetric body. Denote by

$$V_p(K) = \sup_{y \in K} \left| \frac{1}{m} \sum_{j=1}^{m} |\langle X_j, y\rangle|^p \; - \; \mathbb{E}|\langle X, y\rangle|^p \right|$$

the maximal deviation of the empirical p-moment of X from the exact one. We would like to bound $V_p(K)$ under minimal assumptions on the body K and random vector X. This will allow us to choose the size of the sample m for which this deviation is small with high probability. Although the resulting statement is pretty technical, it is applicable to a wide range of problems arising in geometric functional analysis. We discuss some examples in Sections 3, 4.

To bound such random process, we must have some control of the random variable $\max_{1 \leq j \leq m} |X_j|_2$. To this end we introduce the parameter $\kappa_{p,m}(X)$, which plays a key role below

$$\kappa_{p,m}(X) = \left(\mathbb{E} \max_{1 \leq j \leq m} |X_j|_2^p \right)^{1/p}.$$

We prove the following estimate for $V_p(K)$.

Theorem 1 *Let $K \subset (\mathbb{R}^n, \langle \cdot, \cdot \rangle)$ be a symmetric convex body of radius D. Assume that K has modulus of convexity of power type q for some $q \geq 2$. Let $p \geq q$ and let q^* be the conjugate of q.*
Let X be a random vector in $\mathbb{R}^n$, and let $X_1, \ldots, X_m$ be independent copies of X. Assume that

$$C_{p,\lambda} \frac{(\log m)^{2/q^*}}{m} (D \cdot \kappa_{p,m}(X))^p \leq \delta^2 \cdot \sup_{y \in K} \mathbb{E}|\langle X, y\rangle|^p$$

for some $\delta < 1$. Then

$$\mathbb{E} V_p(K) \leq 2\delta \cdot \sup_{y \in K} \mathbb{E}|\langle X, y\rangle|^p.$$

The constant $C_{p,\lambda}$ in Theorem 1 depends on p and on the parameter λ in the definition of the modulus of convexity of power type q (see Section 2.1 for the definition).

Note that minimal assumptions on the vector X are enough to guarantee that $\mathbb{E}V_p(K)$ becomes small for large m. Indeed, assume that the variable $|X|_2$ possesses a finite moment of order $p+\varepsilon$ for some positive ε. Then

$$\kappa_{p,m}(X) \le \left(\mathbb{E}\sum_{j=1}^{m} |X_j|_2^{p+\varepsilon}\right)^{1/p+\varepsilon} \le m^{1/p+\varepsilon}(\mathbb{E}|X|_2^{p+\varepsilon})^{1/p+\varepsilon},$$

so the quantity

$$\frac{(\log m)^{2/q^*}}{m} \cdot \kappa_{p,m}^p(X)$$

tends to 0 when m goes to ∞. Moreover, in most cases, $\kappa_{p,m}(X)$ may be bounded by a simpler quantity:

$$\kappa_{p,m}(X) \le \left(\mathbb{E}\sum_{j=1}^{m} |X_j|_2^{p}\right)^{1/p} \le e\,(\mathbb{E}|X|^s)^{1/s} =: eM_s, \tag{3}$$

where $s = \max(p, \log m)$.

Theorem 1 improves the results of [9] in two ways. First, it contains an almost isometric approximation of the L_p-moments of the random vector by empirical samples (see Theorem 2 below). Second, the assumption on the norm of a random vector X used in Theorem 1 is weaker than the ψ_α-assumption on the scalar products, appearing in [9]. This allows to handle the situations, where the ψ_α-estimate does not hold (see e.g. approximate Lewis decompositions, discussed in Section 3).

While Theorem 1 combined with Chebychev's inequality provides a bound for $V_p(K)$, which holds with high probability, it is often useful to have this probability exponentially close to 1. Using a measure concentration result of Talagrand ([15] Theorem 6.21), we obtain such probability estimate in Theorem 4.

We apply Theorem 1 to isotropic log-concave random vectors. This class includes many naturally arising types of random vectors, in particular a vector uniformly distributed in an isotropic convex body (see Section 4 for exact definitions). The empirical moments of log-concave vectors have been extensively studied in the last years [13], [4], [26], [9], [8]. We will prove the following

Theorem 2 *For any $\varepsilon \in (0,1)$ and $p \geq 2$ there exists $n_0(\varepsilon, p)$ such that for any $n \geq n_0(\varepsilon, p)$, the following holds: let X be a log-concave isotropic random vector in $\mathbb{R}^n$, let $X_1, \ldots, X_m$ be independent copies of X, if*

$$m = \lfloor C_p \, \varepsilon^{-2} n^{p/2} \log n \rfloor$$

then for any $t > \varepsilon$, with probabiblity greater than $1 - C \exp(-\left(t/C'_p \varepsilon\right)^{1/p})$, for any $y \in \mathbb{R}^n$,

$$(1-t)\mathbb{E}|\langle X, y\rangle|^p \leq \frac{1}{m}\sum_{j=1}^{m} |\langle X_j, y\rangle|^p \leq (1+t)\mathbb{E}|\langle X, y\rangle|^p.$$

The constants C_p and C'_p are positive real numbers depending only on p.

Theorem 2 provides an almost isometric approximation of the exact moments, instead of the isomorphic estimates of [9], and achieves it with fewer sample vectors. In the case $p = 2$, it also improves the estimate of [26], and extends to the general setting the estimate obtained by Giannopoulos, Hartzoulaki and Tsolomitis [8] for a random vector uniformly distributed in a 1-unconditional isotropic convex body.

The rest of the paper is organized as follows. In Section 2 we formulate and prove the main results for abstract random vectors. The key step of the proof of Theorem 1 is the estimate of the Gaussian random process

$$Z_y = \sum_{j=1}^{m} g_j |\langle X_j, y\rangle|^p,$$

where g_j are independent standard Gaussian random variables $\mathcal{N}(0,1)$. To obtain such estimate we construct an appropriate majorizing measure and apply the Majorizing measure theorem of Talagrand [28]. In Sections 3 and 4, we provide applications of Theorem 1. Since we require only the existence of high order moments of the norm of X we can apply Theorem 1 to the measures supported by the contact points of a convex body, like in [24], [25], as well as to finding a short Lewis-type decomposition, as described in Section 3. In Section 4, we study in detail the case of log-concave random vectors X. In the last part of this paper, we extend the results obtained in [9] for a uniform distribution on a discrete cube to a general random vector X, which satisfies a ψ_2 estimate for the scalar products $\langle X, y\rangle$, $y \in \mathbb{R}^n$.

2 Maximal deviation of the empirical p-moment

2.1 Statement of the results

Let $K \subset \mathbb{R}^n$ be a convex symmetric body. The modulus of convexity of K is defined for any $\varepsilon \in (0, 2)$ by

$$\delta_K(\varepsilon) = \inf \left\{ 1 - \left\| \frac{x+y}{2} \right\|_K , \|x\|_K = 1, \|y\|_K = 1, \|x-y\|_K > \varepsilon \right\}.$$

We say that K has modulus of convexity of power type $q \geq 2$ if $\delta_K(\varepsilon) \geq c\varepsilon^q$ for every $\varepsilon \in (0, 2)$. It is known (see e.g., [23], Proposition 2.4 or [7]) that this property is equivalent to the fact that the inequality

$$\left\| \frac{x+y}{2} \right\|_K^q + \lambda^{-q} \left\| \frac{x-y}{2} \right\|_K^q \leq \frac{1}{2}(\|x\|_K^q + \|y\|_K^q).$$

holds for all $x, y \in \mathbb{R}^n$. Here $\lambda > 0$ is a constant depending only on c and q. Referring to this inequality below, we shall say that K has modulus of convexity of power type q with constant λ.

Our main result is the following theorem, which implies Theorem 1 from the Introduction.

Theorem 3 *Let $K \subset \mathbb{R}^n$ be a symmetric convex body of radius D. Assume that K has modulus of convexity of power type q with constant λ for some $q \geq 2$, and let q^* be the conjugate of q.*
Let X be a random vector in $\mathbb{R}^n$ and let $X_1, \dots, X_m$ be independent copies of X. For $p \geq q$ set

$$A = C^p \lambda^p \frac{(\log m)^{1/q^*}}{\sqrt{m}} (D\kappa_{p,m}(X))^{p/2} \text{ and } B = \sup_{y \in K} \mathbb{E}|\langle X, y \rangle|^p.$$

Then

$$\mathbb{E} \sup_{y \in K} \left| \frac{1}{m} \sum_{j=1}^{m} |\langle X_j, y \rangle|^p - \quad \mathbb{E}|\langle X, y \rangle|^p \right| \leq A^2 + A\sqrt{B}.$$

The assumption of Theorem 1 reads $A^2 \leq \delta^2 \cdot B$, hence $A^2 + A\sqrt{B} \leq 2\delta B$. Thus, Theorem 1 follows immediately from Theorem 3.

Remark. In fact we shall prove a slightly better inequality. Define

$$\kappa'_{p,m}(X,K) = \left(\mathbb{E} \max_{1\le j\le m} |X_j|_2^2 \max_{1\le j\le m} \|X_j\|_{K^o}^{p-2} \right)^{1/p},$$

then Theorem 3 holds, if the quantity $(D\kappa_{p,m}(X))^{p/2}$ is replaced by $D\kappa'_{p,m}(X,K)^{p/2}$. Since $K \subset DB_2^n$, it is clear that

$$\kappa'_{p,m}(X,K)^{p/2} \le D^{p/2-1} \kappa_{p,m}(X)^{p/2}.$$

The proof of this Theorem is based on the following lemma.

Lemma 1 *Let $K \subset \mathbb{R}^n$ be a symmetric convex body of radius D. Assume that K has modulus of convexity of power type q with constant λ for some $q \ge 2, \lambda > 0$. Let q^* be the conjugate of q.*
Then for every $p \ge q$, and every deterministic vectors $X_1, \ldots, X_m$ in $\mathbb{R}^n$,

$$\mathbb{E} \sup_{y\in K} \left| \sum_{j=1}^m \varepsilon_j |\langle X_j, y\rangle|^p \right| \le$$
$$C^p \, \lambda^p \, (\log m)^{1/q^*} \, D \max_{1\le j\le m} |X_j|_2 \sup_{y\in K} \left(\sum_{j=1}^m |\langle X_j, y\rangle|^{2(p-1)} \right)^{1/2}$$

where expectation is taken over the Bernoulli random variables $(\varepsilon_j)_{1\le j\le m}$.

The proof of the Lemma uses a specific construction of a majorizing measure. It will be presented in part 2.2.
Proof of Theorem 3. The proof is based on a standard symmetrization argument. We denote by $X'_1, \ldots, X'_m$ independent copies of $X_1, \ldots, X_m$. Let $(\varepsilon_j)_{j=1}^m$ be independent symmetric Bernoulli random variables, which are independent of all others. Then the expectation of

$$V_p(K) = \sup_{y\in K} \left| \frac{1}{m} \sum_{j=1}^m |\langle X_j, y\rangle|^p - \quad \mathbb{E}|\langle X, y\rangle|^p \right|$$

can be estimated as follows:

$$\begin{aligned}
m\,\mathbb{E}V_p(K) &= \mathbb{E}\sup_{y\in K}\left|\sum_{j=1}^m |\langle X_j, y\rangle|^p \;-\; m\,\mathbb{E}|\langle X, y\rangle|^p\right| \\
&= \mathbb{E}\sup_{y\in K}\left|\sum_{j=1}^m \big(|\langle X_j, y\rangle|^p - \mathbb{E}|\langle X'_j, y\rangle|^p\big)\right| \\
&\leq \mathbb{E}_X\mathbb{E}_{X'}\sup_{y\in K}\left|\sum_{j=1}^m \big(|\langle X_j, y\rangle|^p - |\langle X'_j, y\rangle|^p\big)\right| \\
&= \mathbb{E}_X\mathbb{E}_{X'}\mathbb{E}_\varepsilon\sup_{y\in K}\left|\sum_{j=1}^m \varepsilon_j\big(|\langle X_j, y\rangle|^p - |\langle X'_j, y\rangle|^p\big)\right| \\
&\leq 2\mathbb{E}_X\mathbb{E}_\varepsilon\sup_{y\in K}\left|\sum_{j=1}^m \varepsilon_j|\langle X_j, y\rangle|^p\right|.
\end{aligned}$$

Therefore, Lemma 1 implies

$$\mathbb{E}V_p(K) \leq C^p\,\lambda^p D\frac{(\log m)^{1/q^*}}{\sqrt{m}}\mathbb{E}_X\max_{1\leq j\leq m}|X_j|_2\sup_{y\in K}\left(\frac{1}{m}\sum_{j=1}^m|\langle X_j, y\rangle|^{2(p-1)}\right)^{1/2}.$$

Since $p \geq 2$, it is easy to see that

$$\begin{aligned}
\mathbb{E}_X\max_{1\leq j\leq m}|X_j|_2\sup_{y\in K}\left(\frac{1}{m}\sum_{j=1}^m|\langle X_j, y\rangle|^{2(p-1)}\right)^{1/2} &\leq \\
\mathbb{E}_X\max_{1\leq j\leq m}|X_j|_2\max_{1\leq j\leq m}\|X_j\|_{K^o}^{p/2-1}\sup_{y\in K}\left(\frac{1}{m}\sum_{j=1}^m|\langle X_j, y\rangle|^{p}\right)^{1/2} &\leq \\
\kappa'_{p,m}(X,K)^{p/2}\left(\mathbb{E}_X\sup_{y\in K}\frac{1}{m}\sum_{j=1}^m|\langle X_j, y\rangle|^{p}\right)^{1/2} &\leq \\
\kappa'_{p,m}(X,K)^{p/2}\left(\mathbb{E}V_p(K) + \sup_{y\in K}\mathbb{E}|\langle X, y\rangle|^p\right)^{1/2}&.
\end{aligned}$$

We get that $\mathbb{E}V_p(K) \leq A'(\mathbb{E}V_p(K)+B)^{1/2}$ where

$$A' = C^p\,\lambda^p D\frac{(\log m)^{1/q^*}}{\sqrt{m}}\kappa'_{p,m}(X,K)^{p/2} \text{ and } B = \sup_{y\in K}\mathbb{E}|\langle X, y\rangle|^p$$

which proves the announced result. □

We present now a deviation inequality for the positive random variable $V_p(K)$ under the assumption that $|X|_2$ satisfies some ψ_α estimate. Mendelson and Pajor [18] studied the same deviation inequality in the case $p = 2$ and $K = B_2^n$ using a symmetrization argument. Our approach is based on a concentration result of Talagrand (Theorem 6.21 in [15]).

Theorem 4 *With the same notation as in Theorem 3, let $V_p(K)$ be the random variable*

$$V_p(K) = \sup_{y\in K} \left| \frac{1}{m} \sum_{j=1}^{m} |\langle X_j, y\rangle|^p - \quad \mathbb{E}|\langle X, y\rangle|^p \right|.$$

Assume that $\||X|_2\|_{\psi_\alpha} < \infty$ for some $0 < \alpha \le p$. Then there exists a positive constant $c_{\alpha,p}$ depending only on α and p such that

$$\forall t > 0,\ \mathbb{P}(V_p(K) \ge t) \le 2\exp\left(-(t/Q)^{\alpha/p}\right)$$

where

$$Q = c_{\alpha,p}\left(\mathbb{E}V_p(K) + \frac{(\log m)^{p/\alpha}}{m} D^p \||X|_2\|_{\psi_\alpha}^p\right).$$

Remark. Observe that in the typical case, Q is of the order $\mathbb{E}V_p(K)$ for which we may use Theorem 3. By Lemma 2 (see below),

$$\kappa_{p,m}(X) \le C(p\log m)^{1/\alpha}\||X|_2\|_{\psi_\alpha}$$

therefore, using Theorem 3,

$$Q \le C_{\alpha,p}(2A_1^2 + A_1\sqrt{B})$$

where

$$A_1 = \lambda^p D^{p/2}\frac{(\log m)^{1/q^*+p/2\alpha}}{\sqrt{m}}\||X|_2\|_{\psi_\alpha} \text{ and } B = \sup_{y\in K}\mathbb{E}|\langle X, y\rangle|^p.$$

For the proof of this theorem, we need an elementary lemma.

Lemma 2 *Let $\delta > 0$ and let $Z_1, \dots, Z_m$ be independent copies of a random variable Z. Then*

$$\|\max_{j=1,\dots,m} |Z_j|\ \|_{\psi_\delta} \le C\log^{1/\delta} m \cdot \|Z\|_{\psi_\delta}.$$

Proof. Note that for any random variable Y the inequality $\|Y\|_{\psi_\delta} \leq A$ is equivalent to

$$\|Y\|_r \leq CAr^{1/\delta}$$

for all $r > 1$. Assume that $r < \log m$. Then

$$\begin{aligned} \| \max_{j=1\ldots m} |Z_j| \, \|_r &\leq \|(\sum_{j=1}^m |Z_j|^{\log m})^{1/\log m}\|_r \leq \left(\mathbb{E}(\sum_{j=1}^m |Z_j|^{\log m})^{r/\log m} \right)^{1/r} \\ &\leq \left(\sum_{j=1}^m \mathbb{E}|Z_j|^{\log m} \right)^{1/\log m} \leq C \log^{1/\delta} m \cdot \|Z\|_{\psi_\delta} . \end{aligned}$$

If $r > \log m$, then using $\max_{j=1,\ldots,m} a_j \leq (\sum_{j=1}^m a_j^r)^{1/r}$, we get

$$\| \max_{j=1,\ldots,m} |Z_j| \, \|_r \leq (\sum_{j=1}^m \mathbb{E}|Z_j|^r)^{1/r} \leq m^{1/r} \|Z\|_r \leq Cr^{1/\delta} \cdot \|Z\|_{\psi_\delta} .$$

These two inequalities imply the Lemma. □

Proof of Theorem 4. To any vector $x \in \mathbb{R}^n$ we associate the function f_x defined on K by

$$\begin{aligned} f_x : K &\to \mathbb{R} \\ y &\mapsto \tfrac{1}{m} \left(|\langle x, y \rangle|^p - \mathbb{E}|\langle X, y \rangle|^p \right) . \end{aligned}$$

Let f_X be the random vector of $L_\infty(K)$ associated to X. Now we apply Theorem 6.21 of Ledoux–Talagrand [15] to $\left\| \sum_{j=1}^m f_{X_j} \right\|$ where the X_j's are independent copies of X. By definition,

$$\| \sum_{j=1}^m f_{X_j} \|_{L_\infty(K)} = V_p(K),$$

and

$$\|f_X\|_{L_\infty(K)} \leq \frac{1}{m} \left(\sup_{y \in K} |\langle X, y \rangle|^p + \sup_{y \in K} \mathbb{E}|\langle X, y \rangle|^p \right) \leq \frac{D^p}{m} \left(|X|_2^p + \mathbb{E}|X|_2^p \right) .$$

Theorem 6.21 of Ledoux-Talagrand [15] states that if $\alpha/p \leq 1$, there exists a constant $c_{\alpha,p}$ depending only on α/p such that

$$\|V_p(K)\|_{\psi_{\alpha/p}} \leq c_{\alpha,p} \left(\mathbb{E}V_p(K) + \| \max_{1 \leq j \leq m} \|f_{X_j}\|_{L_\infty(K)} \|_{\psi_{\alpha/p}} \right) .$$

Moreover,

$$\| \max_{1\le j\le m} \|f_{X_j}\|_{L_\infty(K)} \|_{\psi_{\alpha/p}} \le \frac{2D^p}{m} \| \max_{1\le j\le m} |X_j|_2^p \|_{\psi_{\alpha/p}}$$
$$= \frac{2D^p}{m} \| \max_{1\le j\le m} |X_j|_2 \|_{\psi_\alpha}^p.$$

Lemma 2 implies

$$\| \max_{1\le j\le m} |X_j|_2 \|_{\psi_\alpha} \le C(\log m)^{1/\alpha} \| \, |X|_2\|_{\psi_\alpha}. \tag{4}$$

This proves that

$$\|V_p(K)\|_{\psi_{\alpha/p}} \le c_{\alpha,p} \left(\mathbb{E}V_p(K) + \frac{(\log m)^{p/\alpha}}{m} D^p \| \, |X|_2\|_{\psi_\alpha}^p \right).$$

The deviation inequality follows from the Chebychev inequality. □

2.2 Construction of majorizing measures

Let us recall the assumptions of Lemma 1. The ambient space is $\mathbb{R}^n$ equipped with a Euclidean structure and we denote by $|\cdot|_2$ the norm associated. The symmetric convex body K has a modulus of convexity of power type $q \ge 2$ with a constant λ, which means that

$$\forall x, y \in \mathbb{R}^n, \left\| \frac{x+y}{2} \right\|_K^q + \lambda^{-q} \left\| \frac{x-y}{2} \right\|_K^q \le \frac{1}{2}(\|x\|_K^q + \|y\|_K^q). \tag{5}$$

and satisfies also the inclusion $K \subset DB_2^n$, which means that

$$\forall x \in \mathbb{R}^n, |x|_2 \le D\|x\|_K.$$

Let $p \ge q \ge 2$, and $X_1, \dots, X_m$ be m fixed vectors in $\mathbb{R}^n$. We define the random process V_y for all $y \in \mathbb{R}^n$ by

$$V_y = \sum_{j=1}^m \varepsilon_j |\langle X_j, y\rangle|^p,$$

where ε_j are independent symmetric Bernoulli random variables. It is well known that this process satisfies a sub-Gaussian tail estimate: $\forall y, \overline{y} \in \mathbb{R}^n$, $\forall t > 0$,

$$P(|V_y - V_{\overline{y}}| \ge t) \le 2\exp\left(-\frac{ct^2}{\tilde{d}^2(y, \overline{y})} \right)$$

where

$$\tilde{d}^2(y,\overline{y}) = \sum_{j=1}^{m} \left(|\langle X_j, y\rangle|^p - |\langle X_j, \overline{y}\rangle|^p\right)^2 .$$

Instead of working with this function which is not a metric, it will be preferable to consider the following quasi-metric

$$d^2(y,\overline{y}) = \sum_{j=1}^{m} |\langle X_j, y-\overline{y}\rangle|^2 \left(|\langle X_j, y\rangle|^{2(p-1)} + |\langle X_j, \overline{y}\rangle|^{2(p-1)}\right) .$$

The following propositions state inequalities that we will need to prove Lemma 1. Proposition 1 gives some information concerning the geometry of the balls associated to the metric d and Proposition 2 explains relation between metric d, new Euclidean norm and the following norm defined by

$$\|x\|_\infty = \max_{1\le j\le m} |\langle X_j, x\rangle|.$$

We denote by $\mathcal{B}_\rho(x)$ the ball of center x with radius ρ for the quasi-metric d.

Proposition 1 *For all $y, \overline{y} \in K$*

$$\tilde{d}(y,\overline{y}) \le p\, d(y,\overline{y}), \tag{6}$$

$$d(y,\overline{y}) \le \sqrt{2}\|y-\overline{y}\|_\infty \sup_{y\in K} \left(\sum_{j=1}^{m} |\langle X_j, y\rangle|^{2(p-1)}\right)^{1/2}, \tag{7}$$

$$\|y-\overline{y}\|_\infty \le D \max_{1\le j\le m} |X_j|_2 \; \|y-\overline{y}\|_K. \tag{8}$$

Moreover, the quasi-metric d satisfies the generalized triangle inequality, and for any point x, the ball $\mathcal{B}_\rho(x)$ is a convex set: for all $u_1,\dots,u_N \in \mathbb{R}^n$ and all $x, y, z \in \mathbb{R}^n$,

$$d(u_1,u_N) \le 2p \sum_{i=1}^{N-1} d(u_i,u_{i+1}) \text{ and } d^2(x, \frac{y+z}{2}) \le \frac{1}{2}\big(d^2(x,y) + d^2(x,z)\big). \tag{9}$$

To prove it, we will need the following basic inequalities on real numbers.

Lemma 3 *For every $x, y \in \mathbb{R}^+$ and $p \ge 2$, we have*

$$|x^p - y^p| \le p|x-y|\sqrt{x^{2p-2} + y^{2p-2}} \tag{10}$$

Moreover, if $f(s,t) = |s-t|\sqrt{|s|^{2p-2}+|t|^{2p-2}}$ *then for all* $r_1, \ldots, r_N \in \mathbb{R}$

$$f(r_1, r_N) \leq 2p \sum_{i=1}^{N-1} f(r_i, r_{i+1})$$

and for all $r,\ s,\ t \in \mathbb{R}$*,*

$$f(r,(s+t)/2)^2 \leq \big(f(r,s)^2 + f(r,t)^2\big)/2$$

Proof. The first inequality is straightforward. To prove the second one, consider two cases. When $r_1 r_N \geq 0$, since

$$|r_1 - r_N|\sqrt{|r_1|^{2p-2}+|r_N|^{2p-2}} \leq \sqrt{2}\Big||r_1|^p - |r_N|^p\Big|,$$

the conclusion follows from the triangle inequality and inequality (10). When $r_1 r_N \leq 0$, we can assume without loss of generality that $r_1 \geq 0$ and $r_N \leq 0$. Then

$$\begin{aligned} f(r_1, r_N) &= (r_1 + |r_N|)\sqrt{r_1^{2p-2}+|r_N|^{2p-2}} \leq (r_1 + |r_N|)(r_1^{p-1} + |r_N|^{p-1}) \\ &\leq 2(r_1^p + |r_N|^p). \end{aligned}$$

Let $m < N$ be a number such that $r_m \geq 0$ and $r_{m+1} \leq 0$. Then

$$r_1^p + |r_N|^p \leq \sum_{i=1}^{m-1} \Big||r_i|^p - |r_{i+1}|^p\Big| + r_m^p + |r_{m+1}|^p + \sum_{i=m+1}^{N-1} \Big||r_i|^p - |r_{i+1}|^p\Big|.$$

Combining the previous inequalities with (10), we get

$$\begin{aligned} f(r_1, r_N) &\leq 2\sum_{i=1}^{m-1} \Big||r_i|^p - |r_{i+1}|^p\Big| + 2(r_m^p + |r_{m+1}|^p) + 2\sum_{i=m+1}^{N-1} \Big||r_i|^p - |r_{i+1}|^p\Big| \\ &\leq 2p\sum_{i=1}^{m-1} f(r_i, r_{i+1}) + 2(r_m - r_{m+1})(r_m^{p-1} + |r_{m+1}|^{p-1}) \\ &+ 2p\sum_{i=m+1}^{N-1} f(r_i, r_{i+1}) \\ &\leq 2p\sum_{i=1}^{m-1} f(r_i, r_{i+1}) + 2\sqrt{2} f(r_m, r_{m+1}) + 2p\sum_{i=m+1}^{N-1} f(r_i, r_{i+1}) \end{aligned}$$

which proves the announced result. The last inequality follows from the fact that for $p \geq 2$, the function $v \mapsto (1-v)^2(1+v^{2p-2})$ is convex on $\mathbb{R}$, which can be checked by computing the second derivative. $\square$

Proof of Proposition 1. Inequalities (6) and (9) clearly follow from the three inequalities proved in Lemma 3. Inequalities (7) and (8) follow from simple observations about d and the fact that $K \subset DB_2^n$. $\square$

Proposition 2 *Let* $M = \sup_{y\in K} \sum_{j=1}^{m} |\langle X_j, y\rangle|^{2(p-1)}$. *For a fixed* $u \in K$, *we define the Euclidean norm* $|\cdot|_{\mathcal{E}_u}$ *associated to* u *by*

$$|z|_{\mathcal{E}_u}^2 = \sum_{\ell=1}^{m} |\langle X_\ell, z\rangle|^2 |\langle X_\ell, u\rangle|^{2(p-1)}, \ \forall z \in \mathbb{R}^n.$$

Then the following inequality holds for all z, $\overline{z} \in \mathbb{R}^n$:

$$d^2(z,\overline{z}) \leq 2\cdot 4^{p-1}\big(|z-\overline{z}|_{\mathcal{E}_u}^2 + M\ \|z-\overline{z}\|_\infty^2 (\|z-u\|_K^{2p-2} + \|\overline{z}-u\|_K^{2p-2})\big).$$

Proof. By homogeneity of the statement, we can assume that

$$M = \sup_{y\in K} \sum_{j=1}^{m} |\langle X_j, y\rangle|^{2(p-1)} = 1.$$

For any $z \in \mathbb{R}^n$, let $L_z = \{\ell \in \{1,\dots,m\} \,\big|\, |\langle X_\ell, z\rangle| \geq 2|\langle X_\ell, u\rangle|\}$. Then by convexity of the function $t \mapsto t^{2p-2}$, we have

$$\begin{aligned}\sum_{\ell\in L_z} |\langle X_\ell, z\rangle|^{2(p-1)} &\leq 2^{2p-3}\sum_{\ell\in L_z} |\langle X_\ell, z-u\rangle|^{2(p-1)} + 2^{2p-3}\sum_{\ell\in L_z} |\langle X_\ell, u\rangle|^{2(p-1)} \\ &\leq 2^{2p-3}\sum_{\ell\in L_z} |\langle X_\ell, z-u\rangle|^{2(p-1)} + \frac{1}{2}\sum_{\ell\in L_z} |\langle X_\ell, z\rangle|^{2(p-1)},\end{aligned}$$

which proves (since $M = 1$) that for any $z \in \mathbb{R}^n$,

$$\sum_{\ell\in L_z} |\langle X_\ell, z\rangle|^{2(p-1)} \leq 4^{p-1}\|z-u\|_K^{2p-2}.$$

Hence, for any $z, \overline{z} \in \mathbb{R}^n$,

$$\begin{aligned}\sum_{\ell\in L_z} |\langle X_\ell, z-\overline{z}\rangle|^2 |\langle X_\ell, z\rangle|^{2(p-1)} &\leq \|z-\overline{z}\|_\infty^2 \sum_{\ell\in L_z} |\langle X_\ell, z\rangle|^{2(p-1)} \\ &\leq 4^{p-1}\|z-u\|_K^{2p-2}\,\|z-\overline{z}\|_\infty^2.\end{aligned}$$

For any $l \notin L_z$ we have $|\langle X_l, z\rangle| \le 2|\langle X_l, u\rangle|$, so

$$\sum_{\ell\notin L_z} |\langle X_\ell, z-\overline{z}\rangle|^2 |\langle X_\ell, z\rangle|^{2(p-1)} \le 4^{p-1} \sum_{\ell=1}^m |\langle X_\ell, z-\overline{z}\rangle|^2 |\langle X_\ell, u\rangle|^{2(p-1)}.$$

The same inequalities hold if we exchange the roles of z and $\overline{z}$. To compute $d^2(z_i, z_j)$, we split the sum in four parts and apply the inequalities above:

$$\begin{aligned}
&d^2(z,\overline{z}) = \sum_{\ell=1}^m |\langle X_\ell, z-\overline{z}\rangle|^2 |\langle X_\ell, z\rangle|^{2(p-1)} + |\langle X_\ell, z-\overline{z}\rangle|^2 |\langle X_\ell, \overline{z}\rangle|^{2(p-1)} \\
&= \sum_{\ell\in L_z} |\langle X_\ell, z-\overline{z}\rangle|^2 |\langle X_\ell, z\rangle|^{2(p-1)} + \sum_{\ell\notin L_z} |\langle X_\ell, z-\overline{z}\rangle|^2 |\langle X_\ell, z\rangle|^{2(p-1)} \\
&+ \sum_{\ell\in L_{\overline{z}}} |\langle X_\ell, z-\overline{z}\rangle|^2 |\langle X_\ell, \overline{z}\rangle|^{2(p-1)} + \sum_{\ell\notin L_{\overline{z}}} |\langle X_\ell, z-\overline{z}\rangle|^2 |\langle X_\ell, \overline{z}\rangle|^{2(p-1)} \\
&\le 2\cdot 4^{p-1} \left(|z-\overline{z}|^2_{\mathcal{E}_u} + \|z-\overline{z}\|^2_\infty (\|z-u\|_K^{2p-2} + \|z-u\|_K^{2p-2}) \right). \qquad \square
\end{aligned}$$

Proof of Lemma 1. By inequality (6), we may treat V_y as a sub-Gaussian process with the quasi-metric $p \cdot d$. By homogeneity of the statement, we can assume that

$$\sup_{y\in K} \sum_{j=1}^m |\langle X_j, y\rangle|^{2(p-1)} = 1.$$

Denote $Q = \max_{1\le j\le m} |X_j|_2$. We want to show that

$$\mathbb{E} \sup_{y\in K} |V_y| \le C^p \, \lambda^p \; Q \; (\log m)^{1/q^*} \; D. \tag{11}$$

By Proposition 1, the diameter of the set K with respect to the metric d is bounded by $2\sqrt{2}QD$. Let r be a fixed number chosen such that $r = cp^2$ for a large universal constant c and k_0 be the largest integer such that $r^{-k_0} \ge 2\sqrt{2}QD$.

The proof of inequality (11) is based on the majorizing measure theory of Talagrand [28]. The following theorem is a combination of Proposition 2.3, Theorem 4.1 and Proposition 4.5 of [28]. Note that assuming that $r \ge 2$, one can set $K(r) = C$ in Proposition 2.3, and $K(2, 1, r) = C$ in Proposition 4.5.

Theorem [28] *Let $r \ge 2$. Let $\phi_k : K \to \mathbb{R}^+$ for $k \ge k_0$ be a family of maps satisfying the following assumption: there exists $A > 0$ such that for any point $x \in K$, for any $k \ge k_0$ and any $N \in \mathbb{N}$*

$$(H)\begin{cases} \textit{for any points } x_1,\ldots,x_N \in \mathcal{B}_{r^{-k}}(x) \textit{ with } d(x_i,x_j) \geq r^{-k-1}, i \neq j \\ \textit{we have } \max_{i=1,\ldots,N} \phi_{k+2}(x_i) \geq \phi_k(x) + \dfrac{1}{A}\, r^{-k}\sqrt{\log N}. \end{cases}$$

Then for any fixed $y_0 \in K$,

$$\mathbb{E}\sup_{y\in K} |V_y - V_{y_0}| \leq c\ A \cdot \sup_{k\geq k_0, x\in K} \phi_k(x).$$

To obtain the conclusion of Lemma 1, set $y_0 = 0$.
To complete the proof , we have to define the functionals $\phi_k : K \to \mathbb{R}^+$. Let k_1 be the smallest integer such that $r^{-k_1} \leq QD/\sqrt{n}$. For $k \geq k_1 + 1$, set

$$\phi_k(x) = 1 + \frac{1}{2\log r} + \frac{\sqrt{n}}{Q\,D\,(\log m)^{1/q^*}} \sum_{l=k_1}^{k} r^{-l}\sqrt{\log(1+4QDr^l)}.$$

Note that in this range of k the functionals ϕ_k do not depend on x. We shall show that with this choice of ϕ_k, the condition (H) follows from the classical volumetric estimate of the covering numbers.

For $k_0 \leq k \leq k_1$, the functionals ϕ_k are defined by

$$\phi_k(x) = \min\{\|y\|_K^q, y \in \mathcal{B}_{4pr^{-k}}(x)\} + \frac{k-k_0}{\log m}.$$

Since $q \geq 2$ then $1 \leq q^* \leq 2$ and $(\log m)^{1/q^*} \geq \sqrt{\log m}$. It is easy to see using definitions of k_0 and k_1 that

$$\sup_{x\in K, k\geq k_0} \phi_k(x) \leq c.$$

We shall prove that our functionals satisfy condition (H) for

$$A = (C\lambda)^p\ Q\ D\ (\log m)^{1/q^*}$$

where C is a large numerical constant. That will conclude the proof of Lemma 1 with a new constant C. □

Proof of condition (H)**.** Let $N \in \mathbb{N}$, $x \in K$, $x_1,\ldots,x_N \in \mathcal{B}_{r^{-k}}(x)$ with $d(x_i,x_j) \geq r^{-k-1}$. We have to prove that

$$\max_{i=1,\ldots,N} \phi_{k+2}(x_i) - \phi_k(x) \geq \frac{r^{-k}\sqrt{\log N}}{(C\lambda)^p\ Q\ D\ (\log m)^{1/q^*}}.$$

For $k \geq k_1 - 1$, we always have

$$\phi_{k+2}(x_i) - \phi_k(x) \geq \frac{\sqrt{n \log(1 + 4QDr^{k+2})}}{QD(\log m)^{1/q^*}} \, r^{-k-2}.$$

Since the points $x_1, \dots, x_N$ are well separated in the metric d, they are also well separated in the norm $\|\cdot\|_K$. Indeed, by (7) and (8), we have

$$\|x_i - x_j\|_K \geq r^{-k-1}/QD\sqrt{2}.$$

By the classical volumetric estimate, the maximal cardinality of a t-net in a convex symmetric body $K \subset \mathbb{R}^n$ with respect to $\|\cdot\|_K$ does not exceed $(1 + 2/t)^n$. Therefore,

$$\sqrt{\log N} \leq \sqrt{n \log(1 + 2\sqrt{2}QDr^{k+1})},$$

which proves the desired inequality.

The case $k_0 \leq k \leq k_1 - 2$ is much more difficult. Our proof uses estimates of the covering numbers, in particular, the dual Sudakov inequality [22]. Recall that the covering number $N(W, \|\cdot\|_X, t)$ is the minimal cardinality of $\|\cdot\|_X$-balls of radius t needed to cover the W.

For $j = 1, \dots, N$ denote by $z_j \in K$ the points which satisfy $\|z_j\|_K^q = \min\{\|y\|_K^q, y \in \mathcal{B}_{4pr^{-k-2}}(x_j)\}$. Denote by $u \in K$ a point such that $\|u\|_K^q = \min\{\|y\|_K^q, y \in \mathcal{B}_{4pr^{-k}}(x)\}$. Set

$$\theta = \max_j \|z_j\|_K^q - \|u\|_K^q.$$

Then we have $\max_j \phi_{k+2}(x_j) - \phi_k(x) = \theta + \frac{2}{\log m}$. We shall prove that

$$\theta + \frac{2}{\log m} \geq r^{-k}\sqrt{\log N}/A. \tag{12}$$

Since $d(x_i, x_j) \geq r^{-k-1}$, $z_l \in \mathcal{B}_{4pr^{-k-2}}(x_l)$, and d satisfies a generalized triangle inequality, the points $(z_j)_{1 \leq j \leq N}$ remain well separated. Indeed,

$$r^{-k-1} \leq d(x_i, x_j) \leq 2p(d(x_i, z_i) + d(z_i, z_j) + d(z_j, x_j)) \leq 2pd(z_i, z_j) + 16p^2r^{-k-2}$$

and since $r = cp^2$, we have

$$d(z_i, z_j) \geq r^{-k-1}/cp$$

for all $i \neq j$. Recall that $r = cp^2$. Using again the generalized triangle inequality, we get that

$$d(x, z_j) \leq 2p(d(x, x_j) + d(x_j, z_j)) \leq 2p(r^{-k} + 4pr^{-k-2}) \leq 4pr^{-k}.$$

It means that $z_j \in \mathcal{B}_{4pr^{-k}}(x)$, $u \in \mathcal{B}_{4pr^{-k}}(x)$, and the convexity of the balls for the quasi-metric d proved in Proposition 1 implies $(u + z_j)/2 \in \mathcal{B}_{4pr^{-k}}(x)$. Since K has modulus of convexity of power type q, inequality (5) holds. By the definition of u, we get that for all $j = 1, \ldots, N$

$$\lambda^{-q} \left\| \frac{z_j - u}{2} \right\|_K^q \leq \frac{1}{2} (\|z_j\|_K^q + \|u\|_K^q) - \left\| \frac{z_j + u}{2} \right\|_K^q \leq \frac{\|z_j\|_K^q - \|u\|_K^q}{2} \leq \frac{\theta}{2}.$$

This proves that $\forall j = 1, \ldots, N, \|z_j - u\|_K \leq 2\lambda\theta^{1/q}$. Let $\delta > 0$. Consider the set

$$U = u + 2\lambda\theta^{1/q}K$$

which contains all the z_j's and let S be the maximal number of points in U that are 2δ separated in $\|\cdot\|_\infty$. Then U is covered by S subsets of diameter smaller than 2δ in $\|\cdot\|_\infty$ metric, and so $S \leq N(U, \|\cdot\|_\infty, 2\delta)$. Set

$$\delta = \tilde{c}^p \lambda^{1-p} r^{-k} \theta^{1/q-1}$$

where the constant $\tilde{c}$ will be chosen later. Since $U = u + 2\lambda\theta^{1/q}K$ and $K \subset DB_2^n$,the dual Sudakov inequality [22] implies

$$\sqrt{\log S} \leq \sqrt{\log N(B_2^n, \|\cdot\|_\infty, \delta/D\lambda\theta^{1/q})} \leq c\ D\ \lambda\ \theta^{1/q}\ \mathbb{E}\|G\|_\infty / \delta.$$

Here G denotes a standard Gaussian vector in $\mathbb{R}^n$. It is well known that

$$\mathbb{E}\|G\|_\infty = \mathbb{E} \max_{j=1,\ldots,m} |\langle X_j, G\rangle| \leq c\ Q\ \sqrt{\log m}.$$

We consider now two cases.

First, assume that $S \geq \sqrt{N}$. Then by previous estimate and the definition of δ, we get

$$\sqrt{\log N} \leq c\ Q\ \lambda\ D\ \sqrt{\log m}\ \theta^{1/q} / \delta \leq \theta\ \tilde{c}^p\ r^k\ Q\ \lambda^p\ D\ \sqrt{\log m}$$

which easily proves (12) (since $q^* \leq 2$).

The second case is when $S \leq \sqrt{N}$. Since U is covered by S balls of diameter smaller than 2δ in $\|\cdot\|_\infty$, there exists a subset J of $\{1, \ldots, N\}$ with $\#J \geq \sqrt{N}$ such that

$$\forall i, j \in J, \|z_i - z_j\|_\infty \leq 2\delta.$$

By Proposition 2 applied to the Euclidean norm defined by

$$|y|_{\mathcal{E}_u}^2 = \sum_{\ell=1}^{m} |\langle X_\ell, y\rangle|^2 |\langle X_\ell, u\rangle|^{2(p-1)},$$

we get that

$$d^2(z_i, z_j) \leq 2 \cdot 4^{p-1} \left(|z_i - z_j|_{\mathcal{E}_u}^2 + 4^p \lambda^{2p-2} \theta^{(2p-2)/q} \delta^2 \right).$$

Since $\theta \leq 1$ and $q \leq p$, the definition of δ implies

$$4^{2p} \lambda^{2p-2} \theta^{(2p-2)/q} \delta^2 \leq (4\tilde{c})^{2p} r^{-2k} \theta^{2(p/q-1)} \leq (4\tilde{c})^{2p} r^{-2k}.$$

Recall that $d(z_i, z_j) \geq r^{-k-1}/cp$ and $r = cp^2$. Hence,

$$r^{-2k}/cp^6 \leq d(z_i, z_j)^2 \leq 2 \cdot 4^{p-1} |z_i - z_j|_{\mathcal{E}_u}^2 + 2(4\tilde{c})^{2p} r^{-2k}.$$

Choosing $\tilde{c}$ small enough, we get that for all $i, j \in J$,

$$|z_i - z_j|_{\mathcal{E}_u} \geq r^{-k-1} c^p.$$

Since $K \subset DB_2^n$, we have the following estimate for the covering numbers:

$$\begin{aligned} \#J \leq N(U, |\cdot|_{\mathcal{E}_u}, c^p r^{-k-1}) &= N(K, |\cdot|_{\mathcal{E}_u}, c^p r^{-k-1}/2\lambda\theta^{1/q}) \\ &\leq N(B_2^n, |\cdot|_{\mathcal{E}_u}, c^p r^{-k-1}/2\lambda\theta^{1/q} D). \end{aligned}$$

Recall that G denotes a standard Gaussian vector in $\mathbb{R}^n$. By the dual Sudakov inequality [22], we have

$$\begin{aligned} \sqrt{\log N(B_2^n, |\cdot|_{\mathcal{E}_u}, \frac{c^p r^{-k-1}}{2\lambda D \theta^{1/q}})} &\leq C^p \; r^{k+1} \; \theta^{1/q} \; \lambda \; D \; \mathbb{E}|G|_{\mathcal{E}_u} \\ &\leq C^p \; r^{k+1} \; \theta^{1/q} \; \lambda \; D \left(\mathbb{E}|G|_{\mathcal{E}_u}^2 \right)^{1/2}. \end{aligned}$$

Since for all $y \in K$, $\sum_{j=1}^m |\langle X_j, y\rangle|^{2(p-1)} \leq 1$, we obtain

$$\mathbb{E}|G|_{\mathcal{E}_u}^2 = \sum_{\ell=1}^{m} |X_\ell|_2^2 |\langle X_\ell, u\rangle|^{2(p-1)} \leq Q^2.$$

Since $\#J \geq \sqrt{N}$, we have $\sqrt{\log N} \leq C^p r^{k+1} \lambda D Q \theta^{1/q}$ with a universal constant C. Moreover, by Young's inequality

$$\theta^{1/q} \leq (\log m)^{1/q^*} \left(\theta/q + 1/(q^* \log m)\right)$$

and since $q^* \leq 2 \leq q$ and $\lambda \geq 1$, we get

$$\sqrt{\log N} \leq (C\lambda)^p \; r^{k+1} \; D \; Q \; (\log m)^{1/q^*} \left(\theta + \frac{2}{\log m}\right).$$

This completes the proof of (12) and the proof of condition (H) for the functionals ϕ_k. □

3 Approximate Lewis decomposition

It is well known that if E is an n-dimensional subspace of L_p, then E is $(1+\varepsilon)$-isomorphic to an n-dimensional subspace of ℓ_p^N with N depending on n, p and ε. Lewis [16] proved that any linear subspace E of ℓ_p^N possesses a special decomposition of the identity. More precisely, there exists a Euclidean structure on E with the scalar product $\langle \cdot, \cdot \rangle$, vectors $y_1, \ldots, y_N \in E$ and scalars $c_1, \ldots, c_N > 0$ such that

$$\begin{cases} \forall i, \langle y_i, y_i \rangle = 1, \\ \|x\|_E = \left(\displaystyle\sum_{i=1}^N c_i |\langle x, y_i \rangle|^p \right)^{1/p}, \forall x \in E, \\ \mathrm{Id}_E = \displaystyle\sum_{i=1}^N c_i y_i \otimes y_i. \end{cases}$$

Denote by $(H, |\cdot|_H)$ the linear space E equipped with this Euclidean structure. Recall that p^* denotes the conjugate of p. In the following Theorem, we prove that both spaces E and H can be $(1+\varepsilon)$-embedded in ℓ_p^m and ℓ_2^m respectively via the same linear operator $T : \mathbb{R}^N \to \mathbb{R}^m$, whenever m is of the order of $\varepsilon^{-2} n^{p/2} \log^{2/p^*}(n/\varepsilon^{4/p})$. This extends a classical result of Bourgain, Lindenstrauss and Milman [5] (and [15] for a better dependance on ε) and some results in [24] concerning the number of contact points of a convex body needed to approximate the identity decomposition.

Theorem 5 *Let E be an n-dimensional subspace of L_p for some $p \geq 2$. Then for every $\varepsilon > 0$ there exists a Euclidean structure $H = (E, \langle \cdot, \cdot \rangle)$ on E and m points $x_1, \dots, x_m$ in E with*

$$m \leq \frac{C^p}{\varepsilon^2}\ n^{p/2}\ \log^{2/p^*}\left(\frac{n}{\varepsilon^{4/p}}\right) \leq \frac{C^p}{\varepsilon^2}\ n^{p/2} \log^2\left(\frac{n}{\varepsilon^{4/p}}\right)$$

such that $\forall j, |x_j|_H = 1$ and for all $y \in E$,

$$\begin{cases} (1-\varepsilon)\|y\|_E \leq \left(\dfrac{n}{m}\displaystyle\sum_{j=1}^m |\langle y, x_j\rangle|^p\right)^{1/p} \leq (1+\varepsilon)\|y\|_E \\ (1-\varepsilon)|y|_H \leq \left(\dfrac{n}{m}\displaystyle\sum_{j=1}^m |\langle y, x_j\rangle|^2\right)^{1/2} \leq (1+\varepsilon)|y|_H. \end{cases}$$

Proof. Let X be the random vector taking values y_i with probability c_i/n. Then for all $y \in E$,

$$\mathbb{E}|\langle X, y\rangle|^p = \|y\|_E^p/n \quad \text{and} \quad \mathbb{E}|\langle X, y\rangle|^2 = |y|_H^2/n \quad \text{and} \quad |X|_H = 1.$$

We will apply Theorem 1 twice: first time for the unit ball of E, and then for the unit ball of H.

Since E is a subspace of L_p, by Clarkson's inequality [6], B_E has modulus of convexity of power type p with constant $\lambda = 1$. From Lewis decomposition, we get $\|y\|_E \leq |y|_H \leq n^{\frac{1}{2}-\frac{1}{p}}\|y\|_E$ which means that for $D = n^{\frac{1}{2}-\frac{1}{p}}$,

$$B_H \subset B_E \subset DB_H.$$

Let $X_1, \dots, X_m$ be independent copies of X, then

$$\sup_{y \in B_E} \mathbb{E}|\langle X, y\rangle|^p = 1/n \quad \text{and} \quad \kappa_{p,m}(X) = \left(\mathbb{E} \max_{1\leq j\leq m} |X_j|_H^p\right)^{1/p} = 1.$$

Applying Theorem 1 with $\delta = \varepsilon$, we get that if $m \geq C^p n^{p/2}(\log m)^{2/p^*}/\varepsilon^2$, then

$$\mathbb{E} \sup_{y \in B_E}\left|\frac{n}{m}\sum_{j=1}^m |\langle X_j, y\rangle|^p - \|y\|_E^p\right| \leq \varepsilon.$$

Now, we apply Theorem 1 for $K = B_H$ which clearly has modulus of convexity of power type 2 (i.e. satisfies inequality (5) for $q = 2$). In that case, $D = 1$, and

$$\sup_{|y|_H \leq 1} \mathbb{E}|\langle X, y\rangle|^2 = 1/n \quad \text{and} \quad \kappa_{2,m}(X) = \left(\mathbb{E} \max_{1\leq j\leq m} |X_j|_2^2\right)^{1/2} = 1.$$

Applying Theorem 1 for $q = p = 2$ and $\delta = \varepsilon$, we get that if $m \geq C^2 n \log m/\varepsilon^2$,

$$\mathbb{E} \sup_{|y|_H \leq 1} \left| \frac{n}{m} \sum_{j=1}^{m} |\langle X_j, y\rangle|^2 - |y|_H^2 \right| \leq \varepsilon.$$

Choosing the smallest integer m such that, for a new constant $\tilde{C}$,

$$m \geq \frac{\tilde{C}^p}{\varepsilon^2} \; n^{p/2} \; (\log n/\varepsilon^{4/p})^{2/p^*}$$

we get by Chebychev's inequality that there exist m vectors $x_1, \ldots, x_m$ of Euclidean norm 1 such that for all $y \in E$,

$$\left| \frac{n}{m} \sum_{j=1}^{m} |\langle x_j, y\rangle|^p - \|y\|_E^p \right| \leq \varepsilon \|y\|_E^p$$

and

$$\left| \frac{n}{m} \sum_{j=1}^{m} |\langle x_j, y\rangle|^2 - |y|_H^2 \right| \leq \varepsilon |y|_H^2$$

which gives the desired result. □

4 Isotropic log-concave vectors in $\mathbb{R}^n$

We investigate the case of X being an isotropic log-concave vector in $\mathbb{R}^n$ (or also a vector uniformly distributed in an isotropic convex body). Let us recall some definitions and classical facts about log-concave measures. A probability measure μ on $\mathbb{R}^n$ is said to be log-concave if for every compact sets A, B, and every $\lambda \in [0,1]$, $\mu(\lambda A + (1-\lambda)B) \geq \mu(A)^\lambda \mu(B)^{1-\lambda}$. There is always a Euclidean structure $\langle \cdot, \cdot \rangle$ on $\mathbb{R}^n$ for which this measure is isotropic, i.e. for every $y \in \mathbb{R}^n$,

$$\mathbb{E}\langle X, y\rangle^2 = \int_{\mathbb{R}^n} \langle x, y\rangle^2 d\mu(x) = |y|_2^2.$$

A particular case of a log-concave probability measure is the normalized uniform (Lebesgue) measure on a convex body. Borell's inequality [3] (see also [20, 19]) implies that the linear functionals $x \mapsto \langle x, y\rangle$ satisfy Khintchine type inequalities with respect to log-concave probability measures. Namely, if $p \geq 2$, then for every $y \in \mathbb{R}^n$,

$$\left(\mathbb{E}\langle X, y\rangle^2\right)^{1/2} \leq (\mathbb{E}|\langle X, y\rangle|^p)^{1/p} \leq Cp\left(\mathbb{E}\langle X, y\rangle^2\right)^{1/2}, \tag{13}$$

or in other words

$$\|\langle\cdot, y\rangle\|_{\psi_1} \leq C\left(\mathbb{E}\langle X, y\rangle^2\right)^{1/2}.$$

We have stated in (3) that it is easy to deduce some information about the parameter $\kappa_{p,m}(X)$ from the behavior of the moment M_s of order $s = \max(p, \log m)$ of the Euclidean norm of the random vector X. These moments were studied for a random vector uniformly distributed in an isotropic 1-unconditional convex body in [2], and for a vector uniformly distributed in the unit ball of a Schatten trace class in [12], where it was proved that when $s \leq c\sqrt{n}$, M_s is of the same order as M_2 (up to constant not depending on s). Very recently, Paouris [21] proved that the same statement is valid for any log-concave isotropic random vector in $\mathbb{R}^n$. We state precisely his result.

Theorem [21] *There exist constants $c, C > 0$ such that for any* log-*concave isotropic random vector X in $\mathbb{R}^n$, for any $p \leq c\sqrt{n}$,*

$$(\mathbb{E}|X|_2^p)^{1/p} \leq C\left(\mathbb{E}|X|_2^2\right)^{1/2}.$$

From this sharp estimate, we will deduce the following

Lemma 4 *Let X be an isotropic log-concave random vector in $\mathbb{R}^n$ and let $(X_j)_{1\leq j\leq m}$ be independent copies of X. If $m \leq e^{c\sqrt{n}}$, then for any $p \geq 2$*

$$\kappa_{p,m}(X) = \left(\mathbb{E}\max_{1\leq j\leq m}|X_j|_2^p\right)^{1/p} \leq \begin{cases} C\sqrt{n} \text{ if } p \leq \log m \\ C\ p\ \sqrt{n} \text{ if } p \geq \log m \end{cases}$$

Proof. Since X is isotropic, and for every $y \in \mathbb{R}^n$, $\mathbb{E}\langle X, y\rangle^2 = |y|_2^2$, we get $\mathbb{E}|X|_2^2 = n$. By Borell's inequality [3], $\forall q \geq 2$, $(\mathbb{E}|X|_2^q)^{1/q} \leq Cq\sqrt{n}$. Therefore if $p \geq \log m$,

$$\left(\mathbb{E}\max_{1\leq j\leq m}|X_j|_2^p\right)^{1/p} \leq \left(\mathbb{E}\sum_{1\leq j\leq m}|X_j|_2^p\right)^{1/p} \leq Cpm^{1/p}\sqrt{n} \leq Cp\sqrt{n}.$$

If $p \leq \log m$, by (3)

$$\left(\mathbb{E}\max_{1\leq j\leq m}|X_j|_2^p\right)^{1/p} \leq e\left(\mathbb{E}|X|_2^{\log m}\right)^{1/\log m}.$$

Since $m \le e^{c\sqrt{n}}$, $\log m \le c\sqrt{n}$, the Theorem of Paouris implies

$$\left(\mathbb{E}|X|_2^{\log m}\right)^{1/\log m} \le C\sqrt{n},$$

which concludes the proof of the Lemma. □

Corollary 1 *Let X be an isotropic log-concave random vector in $\mathbb{R}^n$, and let $(X_j)_{1\le j\le m}$ be independent copies of X. Then for every $m \le e^{c\sqrt{n}}$*

$$\| \max_{1\le j\le m} |X_j|_2 \|_{\psi_1} \le C\sqrt{n}.$$

Proof. By Lemma 4, we know that

$$\forall r \ge 2, \left(\mathbb{E} \max_{1\le j\le m} |X_j|_2^r\right)^{1/r} \le Cr\sqrt{n}$$

which proves the claimed estimate for the ψ_1-norm. □

Remark. Recall that for a random isotropic log-concave vector, Borell's inequality implies that

$$\| \, |X|_2 \, \|_{\psi_1} \le C\sqrt{n}.$$

Therefore, a direct application of Lemma 2 is not enough to obtain the desired estimate.

We are now able to give a proof of Theorem 2. It is based on the estimates of $\kappa_{p,m}(X)$ proved above.

Proof of Theorem 2. Let $\varepsilon \in (0,1)$ and $p \ge 2$ and set $n_0(\varepsilon, p) = c_p + \varepsilon^{-4/p}$ where c_p depends only on p. For any $n \ge n_0(\varepsilon, p)$, for any log-concave isotropic random vector X in $\mathbb{R}^n$, set

$$V_p = \sup_{y\in B_2^n} \left| \frac{1}{m} \sum_{j=1}^m |\langle X_j, y\rangle|^p \; - \; \mathbb{E}|\langle X, y\rangle|^p \right|$$

where $X_1, \ldots, X_m$ are independent copies of X. Assume that $p \le \log m$ and $m \le e^{c\sqrt{n}}$ then by Lemma 4, we know that

$$\kappa_{p,m}(X)^p \le c_1^p n^{p/2}.$$

We shall use Theorem 1 with $K = B_2^n$ which is uniformly convex of power type 2 with constant 1 and for which $D = 1$. By (13),

$$1 \le \sup_{y\in B_2^n} \mathbb{E}|\langle X, y\rangle|^p \le p^p,$$

therefore Theorem 1 implies that for every $\delta \in (0,1)$, satisfying $C^p n^{p/2}(\log m) \le \delta^2 m$, we have

$$\mathbb{E} \sup_{y \in B_2^n} \left| \frac{1}{m} \sum_{j=1}^{m} |\langle X_j, y\rangle|^p \ - \ \mathbb{E}|\langle X, y\rangle|^p \right| \le 2\delta p^p.$$

By taking δ such that $2\delta p^p = \varepsilon$, we deduce that

$$\text{if } m \ge C'_p\ \varepsilon^{-2} n^{p/2} \log(n\varepsilon^{-4/p}) \text{ then } \mathbb{E}V_p \le \varepsilon.$$

Since $n \ge n_0(\varepsilon, p)$, it is easy to see that if

$$m = \lfloor C_p\ \varepsilon^{-2} n^{p/2} \log n \rfloor,$$

then $m \ge C'_p\ \varepsilon^{-2} n^{p/2} \log(n\varepsilon^{-4/p})$, $m \le e^{c\sqrt{n}}$ and $p \le \log m$ which allows us to use the estimate $\mathbb{E}V_p \le \varepsilon$.

To get a deviation inequality for V_p, we will apply a result similar to Theorem 4. We know by Corollary 1 that

$$\| \max_{1 \le j \le m} |X|_2 \, \|_{\psi_1} \le C\sqrt{n}.$$

Following the proof of Theorem 4 and replacing inequality (4) by the previous estimate, we easily see that

$$\|V_p\|_{\psi_{1/p}} \le C_p \left(\mathbb{E}V_p + \frac{2n^{p/2}}{m} \right).$$

Since $\lfloor C_p\ \varepsilon^{-2} n^{p/2}\ \log n \rfloor = m$ then

$$\mathbb{E}V_p \le \varepsilon \text{ and } 2n^{p/2}/m \le \varepsilon$$

and we deduce from the Chebychev inequality that for any $t > 0$,

$$\mathbb{P}(V_p \ge t) \le C \exp(-\left(t/C'_p\varepsilon\right)^{1/p}).$$

Therefore, for any $t \ge \varepsilon$, with probability greater than $1 - C\exp(-\left(t/C'_p\varepsilon\right)^{1/p})$, $V_p \le t$ which means that

$$\forall y \in \mathbb{R}^n, \left| \frac{1}{m} \sum_{j=1}^{m} |\langle X_j, y\rangle|^p \ - \ \mathbb{E}|\langle X, y\rangle|^p \right| \le t|y|_2^p.$$

Since $|y|_2 = (\mathbb{E}\langle X, y\rangle^2)^{1/2} \le (\mathbb{E}|\langle X, y\rangle|^p)^{1/p}$, we get the claimed result of Theorem 2. $\square$

Remark. Since by Borell's inequality (13), for any $y \in \mathbb{R}^n$,

$$|y|_2 = \left(\mathbb{E}\langle X, y\rangle^2\right)^{1/2} \le (\mathbb{E}|\langle X, y\rangle|^p)^{1/p} \le Cp\left(\mathbb{E}\langle X, y\rangle^2\right)^{1/2} = Cp|y|_2,$$

it is clear that Theorem 2 improves the results of Giannopoulos and Milman [9].

5 When the linear functionals associated to the random vector X satisfy a ψ_2 condition

Let start this section considering the case when X is a Gaussian vector in $\mathbb{R}^n$. Let $X_j,\ j = 1, \dots, m$, be independent copies of X. For $t \in \mathbb{R}^m$ denote by $X_{t,y}$ the Gaussian random variable

$$X_{t,y} = \sum_{j=1}^{m} t_j\langle X_j, y\rangle.$$

Observe that if p^* denotes the conjugate of p, then

$$\sup_{t\in B_{p^*}^m} X_{t,y} = \left(\sum_{j=1}^{m} |\langle X_j, y\rangle|^p\right)^{1/p}.$$

Let Z and Y be Gaussian vectors in $\mathbb{R}^m$ and $\mathbb{R}^n$ respectively. Using Gordon's inequalities [10], it is easy to show that whenever $\mathbb{E}|Z|_p \ge \varepsilon^{-1}\mathbb{E}|Y|_2$ (i.e. for a universal constant c, $m \ge c^p p^{p/2}\varepsilon^{-p}n^{p/2}$)

$$\begin{aligned}\mathbb{E}|Z|_p - \mathbb{E}|Y|_2 \le \mathbb{E}\inf_{y\in S^{n-1}}\left(\sum_{j=1}^{m} |\langle X_j, y\rangle|^p\right)^{1/p} &\le \\ \le \mathbb{E}\sup_{y\in S^{n-1}}\left(\sum_{j=1}^{m} |\langle X_j, y\rangle|^p\right)^{1/p} &\le \mathbb{E}|Z|_p + \mathbb{E}|Y|_2,\end{aligned}$$

where $(\mathbb{E}|Z|_p + \mathbb{E}|Y|_2)/(\mathbb{E}|Z|_p - \mathbb{E}|Y|_2) \le (1+\varepsilon)/(1-\varepsilon)$. It is therefore possible to get (with high probability with respect to the dimension n, see [11]) a family of m random vectors $X_1, \dots, X_m$ such that for every $y \in \mathbb{R}^n$,

$$A\ |y|_2 \le \left(\frac{1}{m}\sum_{j=1}^{m} |\langle X_j, y\rangle|^p\right)^{1/p} \le A\ \frac{1+\varepsilon}{1-\varepsilon}\ |y|_2.$$

This argument significantly improves the bound on m in Theorem 2 for Gaussian random vectors.

In this part we will be interested in isomorphic moment estimates (instead of almost isometric as in Theorem 2). We will be able to extend the estimate for the Gaussian random vector to random vector X satisfying the ψ_2 condition for linear functionals $y \mapsto \langle X, y\rangle$ with the same dependance on m.

Recall that a random variable Z satisfies the ψ_2 condition if and only if for any $\lambda \in \mathbb{R}$

$$\mathbb{E}\exp(\lambda Z) \leq 2\exp(c\lambda^2 \cdot \|Z\|_2^2).$$

We prove the following

Theorem 6 *Let X be an isotropic random vector in $\mathbb{R}^n$ such that all functionals $y \mapsto \langle X, y\rangle$ satisfy the ψ_2 condition. Let $X_1, \ldots, X_m$ be independent copies of X. Then for every $p \geq 2$ and every $m \geq n^{p/2}$*

$$\mathbb{E}\sup_{y\in B_2^n}\left(\frac{1}{m}\sum_{j=1}^{m}|\langle X_j, y\rangle|^p\right)^{1/p} \leq c\sqrt{p}.$$

Note that the results of Part 3 of [9] follow immediately from Theorem 6, since the random vector with independent ± 1 coordinates satisfies the ψ_2 condition for scalar products.

Proof. Since X is isotropic,

$$\|\langle X, y\rangle\|_{\psi_2} \leq c\,\|\langle X, y\rangle\|_2 = c|y|.$$

Hence, for any $\lambda \in \mathbb{R}$

$$\mathbb{E}\exp\lambda\langle X, y\rangle \leq 2e^{c\lambda^2|y|_2^2}.$$

Writing

$$\Delta = X_{t,y} - X_{t',y'} = \sum_{j=1}^{m}\left((t_j - t_j')\langle X_j, y\rangle + t_j'\langle X_j, y - y'\rangle\right),$$

it is easy to find a new constant $c \geq 1$ such that for every $t, t' \in B_{p^*}^m$, $y, y' \in B_2^n$ and every $\lambda \in \mathbb{R}^+$,

$$\mathbb{E}\exp(\lambda\Delta) \leq 2e^{c\lambda^2(|t-t'|_2^2+|y-y'|_2^2)}.$$

This means that $\|\Delta\|_{\psi_2} \le c(|t-t'|_2^2 + |y-y'|_2^2)^{1/2}$, and so $X_{t,y}$ is a sub-Gaussian random process with respect to the distance

$$d\big((t,y);(t',y')\big) = \big(|t-t'|_2^2 + |y-y'|_2^2\big)^{1/2}.$$

Let $G_{t,y} = \langle Z, t\rangle + \langle Y, y\rangle$, where $Z \in \mathbb{R}^m$ and $Y \in \mathbb{R}^n$ are two independent Gaussian vectors. Then

$$\big(\mathbb{E}|G_{t,y} - G_{t',y'}|^2\big)^{1/2} = d\big((t,y);(t',y')\big)$$

The natural metric for the random process $X_{t,y}$ is bounded by the metric of the process $G_{t,y}$. The Majorizing Measure theorem of Talagrand [28] implies that

$$\mathbb{E} \sup_{(t,y)\in V} X_{t,y} \le C \sup_{(t,y)\in V} G_{t,y}$$

for any compact set $V \subset \mathbb{R}^m \times \mathbb{R}^n$. Therefore,

$$\begin{aligned}
&\mathbb{E} \sup_{y\in B_2^n} \left(\frac{1}{m}\sum_{j=1}^m |\langle X_j, y\rangle|^p\right)^{1/p} = \frac{1}{m^{1/p}} \mathbb{E} \sup_{t\in B_{p^*}^m} \sup_{y\in B_2^n} \sum_{j=1}^m t_j \langle X_j, y\rangle \\
&\le \frac{C}{m^{1/p}} \mathbb{E} \sup_{t\in B_{p^*}^m} \sup_{y\in B_2^n} G_{t,y} = \frac{C}{m^{1/p}} \big(\mathbb{E}|Z|_p + \mathbb{E}|Y|_2\big) \\
&\le C\big(\sqrt{p} + \frac{\sqrt{n}}{m^{1/p}}\big).
\end{aligned}$$

This proves that if $m \ge n^{p/2}$, then

$$\mathbb{E} \sup_{y\in B_2^n} \left(\frac{1}{m}\sum_{j=1}^m |\langle X_j, y\rangle|^p\right)^{1/p} \le c\sqrt{p},$$

as claimed. □

Remark. Let X be an isotropic random vector in $\mathbb{R}^n$ satisfying the ψ_2 estimate for the scalar products. It is not difficult to see, using Corollary 2.7 in [9], that if $m \ge Cn$, then with probability greater than $3/4$

$$c_2 \, |y|_2 \le \left(\frac{1}{m}\sum_{j=1}^m |\langle X_j, y\rangle|^p\right)^{1/p}$$

for every $y \in \mathbb{R}^n$. Therefore, using Theorem 6, it is easy to deduce that if $m \geq n^{p/2}$, then with probability greater than 1/2

$$\forall y \in \mathbb{R}^n \quad c_2\ |y|_2 \leq \left(\frac{1}{m} \sum_{j=1}^{m} |\langle X_j, y \rangle|^p \right)^{1/p} \leq c_1 \sqrt{p}\, |y|_2$$

with universal constants $c_1, c_2 \geq 1$. This generalizes results of [9] and gives an isomorphic version of the result of Klartag and Mendelson [14] valid for every $p \geq 2$.

Acknowledgement. A part of this work was done when the first author was visiting University of Missouri (Columbia). We wish to thank this institution for its hospitality.

References

[1] S. Alesker, ψ_2-estimate for the Euclidean norm on a convex body in isotropic position, *Geometric aspects of functional analysis (Israel, 1992–1994)*, 1–4, Birkhäuser, Basel, 1995.

[2] S. Bobkov, S. Nazarov, On convex bodies and log-concave probability measures with unconditional basis, *Geometric aspects of functional analysis*, 53–69, Lecture Notes in Math., 1807, Springer, Berlin, 2003.

[3] C. Borell, Complements of Lyapunov's inequality, *Math. Ann.* **205** (1973), 323–331.

[4] J. Bourgain, Random points in isotropic convex sets, *Convex geometric analysis (Berkeley, CA, 1996)*, 53–58, Cambridge Univ. Press, Cambridge, 1999

[5] J. Bourgain, J. Lindenstrauss and V. Milman, Approximation of zonoids by zonotopes, *Acta Math.* **162** (1989), no. 1-2, 73–141.

[6] J. A. Clarkson, Uniformly convex spaces, *Trans. Amer. Math. Soc.* **40** (1936), no. 3, 396–414

[7] T. Figiel, An example of infinite dimensional reflexive Banach space non-isomorphic to its Cartesian square, *Studia Math.* **42** (1972), 295–306.

[8] A. A. Giannopoulos, M. Hartzoulaki and A. Tsolomitis, Random points in isotropic unconditional convex bodies, *J. Lond. Math. Soc.* 72 (2005), 779-798.

[9] A. A. Giannopoulos and V. D. Milman, Concentration property on probability spaces, *Adv. Math.*, **156** (2000), no. 1, 77–106.

[10] Y. Gordon, Gaussian processes and almost spherical sections of convex bodies, *Ann. Probab.* **16** (1988), no. 1, 180–188.

[11] Y. Gordon, On Milman's inequality and random subspaces which escape through a mesh in R^n *Geometric aspects of functional analysis (1986/87)*, 84–106, Lecture Notes in Math., 1317, Springer, Berlin, 1988

[12] O. Guédon, G. Paouris, Concentration of mass on the Schatten classes, to appear in *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques.*

[13] R. Kannan, L. Lovász, M. Simonovits, Random walks and an $O^*(n^5)$ volume algorithm for convex bodies, *Random Structures Algorithms* **11** (1997), no. 1, 1–50.

[14] B. Klartag, S. Mendelson, Empirical processes and random projections, to appear in *J. Funct. Anal.*

[15] M. Ledoux and M. Talagrand, *Probability in Banach spaces*, Springer, Berlin, 1991.

[16] D. R. Lewis, Finite dimensional subspaces of L_p, *Studia Math.* **63** (1978), no. 2, 207–212.

[17] L. Lovász and M. Simonovits, Random walks in a convex body and an improved volume algorithm, *Random Structures Algorithms* **4** (1993), no. 4, 359–412.

[18] S. Mendelson and A. Pajor, On singular values of matrices with independent rows, to appear in *Bernoulli.*

[19] V. D. Milman and A. Pajor, Isotropic position and inertia ellipsoids and zonoids of the unit ball of a normed n-dimensional space. Geometric aspects of functional analysis (1987–88), 64–104, Lecture Notes in Math., 1376, Springer, Berlin, 1989.

[20] V. D. Milman and G. Schechtman, *Asymptotic theory of finite-dimensional normed spaces*, Springer, Berlin, 1986

[21] G. Paouris, Concentration of mass on symmetric convex bodies, to appear in *Geom. Funct. Anal.*

[22] A. Pajor and N. Tomczak-Jaegermann, Subspaces of small codimension of finite-dimensional Banach spaces, *Proc. Amer. Math. Soc.* **97** (1986), no. 4, 637–642.

[23] G. Pisier, Martingales with values in uniformly convex spaces, *Israel J. Math.* **20** (1975), no. 3-4, 326–350.

[24] M. Rudelson, Contact points of convex bodies, *Israel J. Math.* **101** (1997), 93–124.

[25] M. Rudelson, Random vectors in isotropipc position, MSRI preprint.

[26] M. Rudelson, Random vectors in isotropipc position, *J. Funct. Anal.* **164** (1999), no. 1, 60–72.

[27] M. Rudelson, Almost orthogonal submatrices of an orthogonal matrix, *Israel J. of Math.* **111** (1999), 143-155.

[28] M. Talagrand, Majorizing measures: the generic chaining, *Ann. Probab.* **24** (1996), no. 3, 1049–1103.

[29] N. Tomczak-Jaegermann, The moduli of smoothness and convexity and the Rademacher averages of trace classes $S_p (1 \leq p < \infty)$ *Studia Math.* **50** (1974), 163–182.

Olivier Guédon:
Université Paris 6, Institut de Mathématiques de Jussieu, Projet Analyse Fonctionnelle, 4, place Jussieu, 75005 Paris, France;
e-mail: guedon@ccr.jussieu.fr

Mark Rudelson:
Department of Mathematics, University of Missouri, Columbia, MO 65211, USA;
e-mail: rudelson@math.missouri.edu