Bounding suprema of canonical processes via convex hull *

Rafał Latała

Abstract

We discuss the method of bounding suprema of canonical processes based on the inclusion of their index set into a convex hull of a well-controlled set of points. While the upper bound is immediate, the reverse estimate was established to date only for a narrow class of regular stochastic processes. We show that for specific index sets, including arbitrary ellipsoids, regularity assumptions may be substantially weakened.

1 Formulation of the problem

Let $X = (X_1, \ldots, X_n)$ be a centered random vector with independent coordinates. To simplify the notation we will write

$$X_t = \langle t, X\rangle = \sum_i t_i X_i \quad \text{ for } t = (t_1, \ldots, t_n) \in \mathbb{R}^n.$$

Our aim is to estimate the expected value of the supremum of the process $(X_t)_{t\in T}$, i.e. the quantity

$$b_X(T) := \mathbb{E}\sup_{t\in T} X_t, \quad T \subset \mathbb{R}^n \text{ nonempty bounded}.$$

There is a long line of research devoted to bounding $b_X(T)$ via the chaining method (cf. the monograph [11]). However chaining methods do not work well for heavy-tailed random variables. In this paper we will investigate another approach based on the convex hull method.

First let us discuss an easy upper bound. Suppose that there exists $t_0, t_1, \ldots \in \mathbb{R}^n$ such that

$$T - t_0 \subset \overline{\mathrm{conv}}\{\pm t_i \colon\ i \geq 1\} \tag{1}$$

then for any $u > 0$,

$$\mathbb{E}\sup_{t\in T} X_t = \mathbb{E}\sup_{t\in T} X_{t-t_0} \leq \mathbb{E}\sup_{i\geq 1} |X_{t_i}| \leq u + \sum_{i\geq 1} \mathbb{E}|X_{t_i}| I_{\{|X_{t_i}|\geq u\}}.$$

*Supported by the National Science Centre, Poland grant 2015/18/A/ST1/00553

Indeed the equality above follows since $X_{t-t_0} = X_t - X_{t_0}$ and $\mathbb{E}X_{t_0} = 0$ and all inequalities are pretty obvious. To make the notation more compact let us define for nonempty countable sets $S \subset \mathbb{R}^n$

$$M_X(S) = \inf_{u>0} \Big[u + \sum_{t\in S} \mathbb{E}|X_t| I_{\{|X_t|\geq u\}}\Big], \quad \widetilde{M}_X(S) = \inf \Big\{m > 0\colon \sum_{t\in S} \mathbb{E}|X_t| I_{\{|X_t|\geq m\}} \leq m\Big\}.$$

It is easy to observe that

$$\widetilde{M}_X(S) \leq M_X(S) \leq 2\widetilde{M}_X(S). \tag{2}$$

To see the lower bound let us fix $u > 0$ and set $m = u + \sum_{t\in S} \mathbb{E}|X_t| I_{\{|X_t|\geq u\}}$ then

$$\sum_{t\in S} \mathbb{E}|X_t| I_{\{|X_t|\geq m\}} \leq \sum_{t\in S} \mathbb{E}|X_t| I_{\{|X_t|\geq u\}} \leq m,$$

so $\widetilde{M}_X(s) \leq m$. For the upper bound it is enough to observe that for $u > \widetilde{M}_X(S)$ we have $\sum_{t\in S} \mathbb{E}|X_t| I_{\{|X_t|\geq u\}} \leq u$.

We have thus shown that

$$b_X(T) \leq M_X(S) \leq 2\widetilde{M}_X(S) \quad \text{if} \quad T - t_0 \subset \overline{\mathrm{conv}}(S \cup -S). \tag{3}$$

Remark 1. The presented proof of (3) did not use independence of coordinates of X, the only required property is mean zero.

Main question. *When can we reverse bound* (3) *– what should be assumed about variables* X_i *(and the set* T*) in order that*

$$T - t_0 \subset \overline{\mathrm{conv}}(S \cup -S) \quad \textit{and} \quad M_X(S) \lesssim \mathbb{E} \sup_{t\in T} X_t \tag{4}$$

for some $t_0 \in \mathbb{R}^n$ *and nonempty countable set* $S \subset \mathbb{R}^n$*?*

Remark 2. It is not hard to show (see Section 3 below) that $M_X(S) \sim \mathbb{E}\max_i |X_{t_i}| = b_X(S \cup -S)$ if $S = \{t_1, \ldots, t_k\}$ and variables $(X_{t_i})_i$ are independent. Thus our main question asks whether the parameter $b_X(T)$ may be explained by enclosing a translation of T into the convex hull of points $\pm t_i$ for which variables X_{t_i} behave as though they are independent.

Remark 3. The main question is related to Talagrand conjectures about suprema of positive selector processes, c.f. [11, Section 13.1], i.e. the case when $T \subset \mathbb{R}^n_+$ and $\mathbb{P}(X_i \in \{0, 1\}) = 1$. Talagrand investigates possibility of enclosing T into a solid convex hull, which is bigger than the convex hull. On the other hand we think that in our question some regularity conditions on variables X_i is needed (such as $4 + \delta$ moment condition (10), which is clearly not satisfied for nontrivial classes of selector processes).

Remark 4. i) In the one dimensional case if $a = \inf T$, $b = \sup T$, then $T \subset [a, b] = \frac{a+b}{2} + \mathrm{conv}\{\frac{a-b}{2}, \frac{b-a}{2}\}$. Hence

$$b_{X_1}(T) = \mathbb{E}\max\{aX_1, bX_1\} = \frac{a+b}{2}\mathbb{E}X_1 + \mathbb{E}\Big|\frac{b-a}{2}X_1\Big| = \frac{b-a}{2}\mathbb{E}|X_1| \geq \widetilde{M}_{X_1}\Big(\Big\{\frac{b-a}{2}\Big\}\Big),$$

so this case is trivial. Thus in the sequel it is enough to consider $n \geq 2$.

ii) The set $V := \overline{\mathrm{conv}}(S \cup -S)$ is convex and origin-symmetric. Hence if $T = -T$ and $T - t_0 \subset V$ then $T + t_0 = -(-T - t_0) = -(T - t_0) \subset V$ and $T \subset \mathrm{conv}((T - t_0) \cup (T + t_0)) \subset V$. Thus for symmetric sets it is enough to consider only $t_0 = 0$.

iii) Observe that $b_X(\mathrm{conv}(T)) = b_X(T)$ and $T - t_0$ is a subset of a convex set if and only if $\mathrm{conv}(T) - t_0$ is a subset of this set. Moreover, if $T - T \subset V$ then $T - t_0 \subset V$ for any $t_0 \in V$ and $b_X(T - T) = b_X(T) + b_X(-T) = b_X(t) + b_{-X}(T)$. So if X is symmetric it is enough to consider symmetric convex sets T.

Notation. Letters c, C will denote absolute constants which value may differ at each occurence. For two nonnegative functions f and g we write $f \gtrsim g$ (or $g \lesssim f$) if $g \leq Cf$. Notation $f \sim g$ means that $f \gtrsim g$ and $g \gtrsim f$. We write $c(\alpha)$, $C(\alpha)$ for constants depending only on a parameter α and define accordingly relations $\gtrsim_\alpha$, $\lesssim_\alpha$, $\sim_\alpha$.

Organization of the paper. In Section 2 we present another quantity $m_X(S)$, defined via L_p-norms of $(X_t)_{t\in S}$, and show that for regular variables X_i it is equivalent to $M_X(S)$. We also discuss there the relation of the convex hull method to the chaining functionals. In Section 3 we show that for $T = B_1^n$ the bound (3) may be reversed for arbitrary independent $X_1, \ldots, X_n$ and $S = \{e_1, \ldots, e_n\}$. Section 4 is devoted to the study of ellipsoids. First we show that for $T = B_2^n$ and symmetric p-stable random variables, $1 < p < 2$, one cannot reverse (3). Then we prove that under $4 + \delta$ moment condition our main question have the affirmative answer for $T = B_2^n$ and more general case of ellipsoids. We extend this result to the case of linear images of B_q^n-balls, $q \geq 2$ in Section 5. We conclude by discussing some open questions in the last section.

2 Regular growth of moments.

In this section we consider variables with regularly growing moments in a sense that

$$\|X_i\|_{2p} \leq \alpha \|X_i\|_p < \infty \quad \text{for } p \geq 1, \tag{5}$$

where $\|X\|_p = (\mathbb{E}|X|^p)^{1/p}$.

For such variables we will prove that there is alternate quantity equivalent to $M_X(S)$, namely

$$m_X(S) := \inf \sup_i \|X_{t_i}\|_{\log(e+i)}.$$

where the infimum runs over all numerations of $S = \{t_i\colon\ 1 \leq i \leq N\}$, $N \leq \infty$.

It is not hard to check (cf. Lemma 4.1 in [7]) that (5) yields

$$\|X_t\|_{2p} \leq C_0(\alpha)\|X_t\|_p \quad \text{for } p \geq 1 \tag{6}$$

and as a consequence we have for $p > 0$,

$$\mathbb{P}(|X_t| \geq e\|X_t\|_p) \leq e^{-p}, \quad \mathbb{P}(|X_t| \geq c_1(\alpha)\|X_t\|_p) \geq \min\{c_2(\alpha), e^{-p}\}, \tag{7}$$

where the first bound follows by Chebyshev's inequality and the second one by the Paley-Zygmund inequality.

Proposition 5. *Suppose that X_i are independent r.v's satisfying condition* (5). *Then* $M_X(S) \sim_\alpha m_X(S)$.

Proof. Let $S = \{t_i\colon\ 1 \leq i \leq N\}$ and $m := \sup_i \|X_{t_i}\|_{\log(e+i)}$. Then for $u > 1$,

$$\sum_{s\in S} \mathbb{P}(|X_s| \geq um) \leq \sum_{i=1}^N \mathbb{P}(|X_{t_i}| \geq u\|X_{t_i}\|_{\log(e+i)}) \leq \sum_{i=1}^N u^{-\log(e+i)}.$$

Therefore

$$\begin{aligned}\sum_{s\in S} \mathbb{E}|X_s| I_{\{|X_s|\geq e^2 m\}} &= \sum_{s\in S} \Big(e^2 m\mathbb{P}(|X_s| \geq e^2 m) + m\int_{e^2}^\infty \mathbb{P}(|X_s| \geq um)\mathrm{d}u\Big)\\ &\leq m\sum_{i=1}^N \Big(e^{2-2\log(e+i)} + \int_{e^2}^\infty u^{-\log(e+i)}\mathrm{d}u\Big)\\ &\leq m\sum_{i=1}^N \Big((e+i)^{-2}\Big(e^2 + \frac{1}{\log(e+i)-1}\Big)\Big) \leq 100m,\end{aligned}$$

which shows that $M_X(S) \leq 100 m_X(S)$ (this bound does not use neither regularity neither independence of X_i).

To establish the reverse inequality let us take any $m > 2M_X(S) \geq \widetilde{M}_X(S)$ and enumerate elements of S as $t_1, t_2, \ldots$ in such a way that that $i \to \mathbb{P}(|X_{t_i}| \geq m)$ is nonincreasing. By the definition of $\widetilde{M}_X(S)$ we have

$$\sum_{i=1}^N \mathbb{P}(|X_{t_i}| \geq m) \leq \frac{1}{m}\sum_{i=1}^N \mathbb{E}|X_{t_i}| I_{\{|X_{t_i}|\geq m\}} \leq 1.$$

In particular it means that $\mathbb{P}(|X_{t_i}| \geq m) \leq 1/i$. By (7) this yields that for $i > 1/c_2(\alpha)$ $\|X_{t_i}\|_{\log(i)} \leq m/c_1(\alpha)$. Since $\log(e+i)/\log(i) \leq 2$ for $i \geq 3$ we have $\|X_{t_i}\|_{\log(e+i)} \leq C(\alpha)m$ for large i. For $i \leq \max\{3, 1/c_2(\alpha)\}$ it is enough to observe that $\log(e+i) \leq 2^{k(\alpha)}$, so

$$\|X_{t_i}\|_{\log(e+i)} \leq C_0(\alpha)^{k(\alpha)}\mathbb{E}|X_{t_i}| \leq C_0(\alpha)^{k(\alpha)} M_X(S).$$

This shows that $\|X_{t_i}\|_{\log(e+i)} \lesssim_\alpha m$ for all i and therefore $m_X(S) \lesssim_\alpha M_X(S)$.

□

2.1 γ_X-functional

The famous Fernique-Talagrand theorem [3, 10] states that suprema of Gaussian processes may be estimated in geometrical terms by γ_2-functional. This result was extended in several directions. One of them is based on the so-called γ_X functional.

For a nonempty subset $T \subset \mathbb{R}^n$ we define

$$\gamma_X(T) := \inf \sup_{t\in T} \sum_{n=0}^{\infty} \Delta_{n,X}(A_n(t)),$$

where the infimum runs over all increasing sequences of partitions $(\mathcal{A}_n)_{n\geq 0}$ of T such that $\mathcal{A}_0 = \{T\}$ and $|\mathcal{A}_n| \leq N_n := 2^{2^n}$ for $n \geq 1$, $A_n(t)$ is the unique element of $\mathcal{A}_n$ which contains t and $\Delta_{n,X}(A)$ denotes the diameter of A with respect to the distance $d_n(s,t) := \|X_s - X_t\|_{2^n}$.

It is not hard to check that $b_X(T) \lesssim \gamma_X(T)$. The reverse bound was discussed in [6], where it was shown that it holds (with constants depending on β and λ) if

$$\|X_i\|_p \leq \beta \frac{p}{q} \|X_i\|_q \text{ and } \|X_i\|_{\lambda p} \geq 2\|X_i\|_p \text{ for all } i \text{ and } p \geq q \geq 2. \tag{8}$$

Moreover the condition $\|X_i\|_p \leq \beta\frac{p}{q}\|X_i\|_q$ is necessary in the i.i.d. case if the estimate $\gamma_X(T) \leq C b_X(T)$ holds with a constant independent on n and $T \subset \mathbb{R}^n$.

The next result may be easily deduced from the proof of [6, Corollary 2.7], but we provide its proof for the sake of completeness.

Proposition 6. *Let X_i be independent and satisfy condition* (5) *and let T be a nonempty subset of $\mathbb{R}^n$ such that $\gamma_X(T) < \infty$. Then there exists set $S \subset \mathbb{R}^n$ such that for any $t_0 \in T$, $T - t_0 \subset T - T \subset \overline{\mathrm{conv}}(S \cup -S)$ and $M_X(S) \lesssim m_X(S) \lesssim_\alpha \gamma_X(T)$.*

Proof. Wlog (since it is only a matter of rescaling) we may assume that $\mathbb{E}X_i^2 = 1$.

By the definition of $\gamma_X(T)$ we may find an increasing sequence of partitions $(\mathcal{A}_n)$ such that $\mathcal{A}_0 = \{T\}$, $|\mathcal{A}_j| \leq N_j$ for $j \geq 1$ and

$$\sup_{t\in T} \sum_{n=0}^{\infty} \Delta_{n,X}(A_n(t)) \leq 2\gamma_X(T). \tag{9}$$

For any $A \in \mathcal{A}_n$ let us choose a point $\pi_n(A) \in A$ and set $\pi_n(t) := \pi_n(A_n(t))$.

Let $M_n := \sum_{j=0}^{n} N_j$ for $n = 0, 1, \ldots$ (we put $N_0 := 1$). Then $\log(M_n+2) \leq 2^{n+1}$. Notice that there are $|\mathcal{A}_n| \leq N_n$ points of the form $\pi_n(t) - \pi_{n-1}(t)$, $t \in T$. So we may define s_k, $M_{n-1} \leq k < M_n$, $n = 1, 2, \ldots$ as some rearrangement (with repetition if $|\mathcal{A}_n| < N_n$) of points of the form $(\pi_n(t) - \pi_{n-1}(t))/\|X_{\pi_n(t)} - X_{\pi_{n-1}(t)}\|_{2^{n+1}}$, $t \in T$. Then $\|X_{s_k}\|_{\log(k+e)} \leq 1$ for all $k \geq 1$.

Observe that

$$\|t - \pi_n(t)\|_2 = \|X_t - X_{\pi_n(t)}\|_2 \leq \Delta_{2,X}(A_n(t)) \leq \Delta_{n,X}(A_n(t)) \to 0 \quad \text{ for } n \to \infty.$$

For any $s,t \in T$ we have $\pi_0(s) = \pi_0(t)$ and thus

$$s - t = \lim_{n\to\infty} (\pi_n(s) - \pi_n(t)) = \lim_{n\to\infty} \left(\sum_{k=1}^{n} (\pi_k(s) - \pi_{k-1}(s)) - \sum_{k=1}^{n} (\pi_k(t) - \pi_{k-1}(t)) \right).$$

This shows that

$$T - T \subset R\, \overline{\mathrm{conv}}\{\pm s_k \colon\ k \geq 1\},$$

where

$$\begin{aligned} R &:= 2\sup_{t\in T} \sum_{n=1}^{\infty} d_{n+1}(\pi_n(t), \pi_{n-1}(t)) \leq 2\sup_{t\in T} \sum_{n=1}^{\infty} \Delta_{n+1,X}(A_{n-1}(t)) \\ &\leq C(\alpha) \sup_{t\in T} \sum_{n=1}^{\infty} \Delta_{n-1,X}(A_{n-1}(t)) \leq 2C(\alpha)\gamma_X(T), \end{aligned}$$

where the second inequality follows by (6). Thus it is enough to define $S := \{Rs_k \colon\ k \geq 1\}$. $\square$

Remark 7. Proposition 6 together with the equivalence $b_X(T) \sim_{\alpha,\lambda} \gamma_X(T)$ shows that the main question has the affirmative answer for any bounded nonempty set T if symmetric random variables X_i satisfy moment bounds (5). We strongly believe that the condition $\|X_i\|_{\lambda p} \geq 2\|X_i\|_p$ is not necessary – equivalence of $b_X(T)$ and the convex hull bound was established in the case of symmetric Bernoulli r.v's ($\mathbb{P}(X_i = \pm 1) = 1/2$) in [1, Corollary 1.2]. However to treat the general case of r.v's satisfying only the condition $\|X_i\|_p \leq \beta \frac{p}{q} \|X_i\|_q$ one should most likely combine γ_X functional with a suitable decomposition of the process $(X_t)_{t\in T}$, as was done for Bernoulli processes.

3 Toy case: ℓ_1-Ball

Let us now consider a simple case of $T = B_1^n = \{t \in \mathbb{R}^n \colon\ \|t\|_1 \leq 1\}$. Let

$$u_0 := \inf\left\{ u > 0 \colon\ \mathbb{P}\big(\max_i |X_i| \geq u\big) \leq \frac{1}{2} \right\}.$$

Since

$$\mathbb{P}\big(\max_i |X_i| \geq u\big) \geq \frac{1}{2} \min\left\{1, \sum_i \mathbb{P}(|X_i| \geq u)\right\}$$

we get

$$\begin{aligned} \mathbb{E} \sup_{t\in B_1^n} X_t &= \mathbb{E} \max_{1\leq i\leq n} |X_i| = \int_0^\infty \mathbb{P}\big(\max_{1\leq i\leq n} |X_i| \geq u\big)\mathrm{d}u \geq \frac{1}{2}u_0 + \int_{u_0}^\infty \frac{1}{2} \sum_{i=1}^n \mathbb{P}(|X_i| \geq u)\mathrm{d}u \\ &= \frac{1}{2}u_0 + \frac{1}{2}\sum_{i=1}^n \int_{u_0}^\infty \mathbb{P}(|X_i| \geq u)\mathrm{d}u = \frac{1}{2}u_0 + \frac{1}{2}\sum_{i=1}^n \mathbb{E}(|X_i| - u_0)_+. \end{aligned}$$

Therefore

$$2u_0 + \sum_{i=1}^n \mathbb{E}|X_i| I_{\{|X_i| \geq 2u_0\}} \leq 2u_0 + 2\sum_{i=1}^n \mathbb{E}(|X_i| - u_0)_+ \leq 4\mathbb{E} \sup_{t \in B_1^n} X_t,$$

so that $M_X(\{e_i\colon\ i \leq n\}) \leq 4\mathbb{E}\sup_{t\in B_1^n} X_t$, where $(e_i)_{i\leq n}$ is the canonical basis of $\mathbb{R}^n$. Since $B_1^n \subset \mathrm{conv}\{\pm e_1, \ldots, \pm e_n\}$ we get the affirmative answer to the main question for $T = B_1^n$.

Proposition 8. *If $T = B_1^n$ then estimate (4) holds for arbitrary independent integrable r.v's $X_1, \ldots, X_n$ with $S = \{e_1, \ldots, e_n\}$ and $t_0 = 0$.*

4 Case II. Euclidean balls

Now we move to the case $T = B_2^n$. Then $\sup_{t\in T}\langle t, x\rangle = |x|$, where $|x| = \|x\|_2$ is the Euclidean norm of $x \in \mathbb{R}^n$.

4.1 Counterexample

In this subsection $X = (X_1, X_2, \ldots, X_n)$, where X_k have symmetric p-stable distribution with characteristic function $\varphi_{X_k}(t) = \exp(-|t|^p)$ and $p \in (1, 2)$. We will assume for convenience that n is even. Let G be a canonical n-dimensional Gaussian vector, independent of X. Then

$$\begin{aligned}\mathbb{E}|X| &= \mathbb{E}_X\mathbb{E}_G\sqrt{\frac{\pi}{2}}|\langle X, G\rangle| = \sqrt{\frac{\pi}{2}}\mathbb{E}_G\mathbb{E}_X|\langle X, G\rangle| = \sqrt{\frac{\pi}{2}}\mathbb{E}_G\|G\|_p\mathbb{E}|X_1| \\ &\sim_p \mathbb{E}\|G\|_p \sim (\mathbb{E}\|G\|_p^p)^{1/p} \sim n^{1/p}.\end{aligned}$$

Observe also that for $u > 0$, $\mathbb{P}(|X_1| \geq u) \sim_p \min\{1, u^{-p}\}$, so

$$\mathbb{E}|X_1| I_{\{|X_1|\geq u\}} \sim_p u\min\{1, u^{-p}\} + \int_u^\infty \min\{1, v^{-p}\}\mathrm{d}v \sim_p \min\{1, u^{1-p}\}, \quad u > 0$$

and

$$\mathbb{E}|X_t| I_{\{|X_t|\geq u\}} = \|t\|_p\mathbb{E}|X_1| I_{\{|X_1|\geq u/\|t\|_p\}} \sim_p \min\{\|t\|_p, u^{1-p}\|t\|_p^p\}, \quad u > 0,\ t \in \mathbb{R}^n.$$

Hence

$$\sum_{t\in S}\|t\|_p^p \lesssim_p u^p \quad \text{ for } u > \widetilde{M}_X(S).$$

Suppose that $B_2^n \subset \overline{\mathrm{conv}}(S \cup -S)$ and $M_X(S) \sim \widetilde{M}_X(S) < \infty$. We may then enumerate elements of S as $(t_k)_{k=1}^N$, $N \leq \infty$ in such a way that $(\|t_k\|_p)_{k=1}^N$ is nonincreasing. Obviously

$N \geq n$ (otherwise $\mathrm{conv}(S \cup -S)$ would have empty interior). Take $u > \widetilde{M}_X(S)$ and set $E := \mathrm{span}(\{t_k\colon\ k \leq n/2\})$. Then $\|t_k\|_p^p \leq C_p u^p/n$ for $k > n/2$. Thus

$$B_2^n \subset \overline{\mathrm{conv}}(S \cup -S) \subset E + \overline{\mathrm{conv}}(\{\pm t_k\colon\ k > n/2\}) \subset E + \Big(\frac{C_p}{n}\Big)^{1/p} u B_p^n.$$

Let $F = E^{\perp}$ and P_F denotes the ortogonal projection of $\mathbb{R}^n$ onto the space F. Then $\mathrm{dim}F = \mathrm{dim}E = n/2$ and

$$B_2^n \cap F = P_F(B_2^n) \subset \Big(\frac{C_p}{n}\Big)^{1/p} u P_F(B_p^n).$$

In particular

$$n^{-1/2} \sim \mathrm{vol}_{n/2}^{2/n}(B_2^n \cap F) \leq \Big(\frac{C_p}{n}\Big)^{1/p} u \mathrm{vol}_{n/2}^{2/n}(P_F(B_p^n)).$$

By the Rogers-Shephard inequality [8] and inclusion $B_2^n \subset n^{1/p-1/2} B_p^n$ we have

$$\mathrm{vol}_{n/2}(P_F(B_p^n)) \leq \binom{n}{n/2} \frac{\mathrm{vol}_n(B_p^n)}{\mathrm{vol}_{n/2}(B_p^n \cap E)} \leq 2^n \frac{\mathrm{vol}_n(B_p^n)}{\mathrm{vol}_{n/2}(n^{1/2-1/p} B_2^n \cap E)} \leq (C n^{-1/p})^{n/2}.$$

This shows that $u \gtrsim_p n^{2/p-1/2}$. Thus $M_X(S) \gtrsim_p n^{2/p-1/2} \gg n^{1/p} \sim_p b_X(B_2^n)$ and our question has a negative answer in this case.

4.2 $4+\delta$ moment condition

In this part we establish positive answer to the main question in the case $T = B_2^n$ under the following $4+\delta$ moment condition

$$\exists_{r \in (4,8], \lambda < \infty}\ (\mathbb{E} X_i^r)^{1/r} \leq \lambda (\mathbb{E} X_i^2)^{1/2} < \infty \quad i = 1, \ldots, n. \tag{10}$$

The restriction $r \leq 8$ is just for convenience. The following easy consequence of (10) will be helpful in the sequel.

Lemma 9. *Suppose that $X_1, \ldots, X_n$ are independent mean zero r.v's satisfying condition* (10)*. Then for any $1 \leq p \leq r$,*

$$\Big\| \sum_{i=1}^n u_i X_i \Big\|_p \sim_\lambda \Big\| \sum_{i=1}^n u_i X_i \Big\|_2 = \Big(\sum_{i=1}^n u_i^2 \mathbb{E} X_i^2 \Big)^{1/2} \tag{11}$$

and

$$\Big\| \sum_{1 \leq i < j \leq n} u_{ij} X_i X_j \Big\|_p \sim_\lambda \Big\| \sum_{1 \leq i < j \leq n} u_{ij} X_i X_j \Big\|_2 = \Big(\sum_{1 \leq i < j \leq n} u_{ij}^2 \mathbb{E} X_i^2 \mathbb{E} X_j^2 \Big)^{1/2}. \tag{12}$$

Proof. Since it is only a matter of scaling wlog we may and will assume that $\mathbb{E}X_i^2 = 1$ for all i.

Rosenthal's inequality [9] gives for $2 \le p \le r$ (recall that $r \in (4,8]$, so constants below do not depend on r)

$$\Big\|\sum_{i=1}^n u_i X_i\Big\|_p \sim \Big(\sum_i \mathbb{E}|u_iX_i|^2\Big)^{1/2} + \Big(\sum_i \mathbb{E}|u_iX_i|^p\Big)^{1/p} \sim_\lambda \Big(\sum_i u_i^2\Big)^{1/2} + \Big(\sum_i |u_i|^p\Big)^{1/p}$$
$$\sim \Big(\sum_i u_i^2\Big)^{1/2}.$$

To estimate $\|S\|_p$ for $1 \le p \le 2$ and $S = \sum_{i=1}^n u_iX_i$ it is enough to note that $\|S\|_1 \le \|S\|_p \le \|S\|_2$ and $\|S\|_2 \le \|S\|_4^{1/3}\|S\|_1^{2/3} \sim_\lambda \|S\|_2^{1/3}\|S\|_1^{2/3}$, so $\|S\|_p \sim \|S\|_2$.

To prove the last part of the assertion we will use the hypercontractive method. Observe that for a real number u there exists $\theta \in [0,1]$ such that

$$\begin{aligned}(1+u)^r &\le \Big(1 + ru + \frac{r(r-1)}{2}(1+\theta u)^{r-2}u^2\Big)I_{\{|u|<1\}} + (2|u|)^r I_{\{|u|\ge 1\}}\\ &\le 1 + ru + r^2 2^{r-3}u^2 + 2^r|u|^r.\end{aligned}$$

Hence (note that $\lambda \ge 1$, $\mathbb{E}X_i = 0$, $\mathbb{E}X_i^2 = 1$ and $\mathbb{E}|X_i|^r \le \lambda^r$)

$$\mathbb{E}\Big(1 + \frac{1}{32\lambda}uX_i\Big)^r \le 1 + r^2 2^{r-3}\frac{u^2}{1024} + 2^{-4r}|u|^r \le 1 + \frac{ru^2}{4} + \frac{|u|^r}{2} \le 1 + \max\Big\{\frac{r}{2}u^2, |u|^r\Big\}.$$

Since
$$(\mathbb{E}(1+uX_i)^2)^{r/2} = (1+u^2)^{r/2} \ge 1 + \max\Big\{\frac{r}{2}u^2, |u|^r\Big\}$$

we get $\|1 + \frac{1}{32\lambda}uX_i\|_r \le \|1 + uX_i\|_2$ for any $u \in \mathbb{R}$ and the hypercontractivity method (cf. [5, Theorem 6.5.2]) yields (12) for $p = r$. The case $1 \le p \le r$ may be obtained in the same way as in the proof of (11). □

Observe that (10) implies that $\mathrm{Var}(X_i^2) \le (\lambda^4 - 1)(\mathbb{E}X_i^2)^2$, so $\mathrm{Var}(|X|^2) \le \sum_i(\lambda^4 - 1)(\mathbb{E}X_i^2)^2 \le (\lambda^4 - 1)(\mathbb{E}|X|^2)^2$. This yields that $\mathbb{E}|X|^4 \le \lambda^4(\mathbb{E}|X|^2)^2$ and $(\mathbb{E}|X|^2)^{1/2} \le \lambda^2\mathbb{E}|X|$.

The next fact is pretty standard, we prove it for completeness.

Lemma 10. *For any k there exists $T \subset B_2^k$ with $|T| \le 5^k$ such that $B_2^k \subset 2\mathrm{conv}(T)$.*

Proof. Let T be the maximal $\frac{1}{2}$-separated set in B_2^k, the standard volumetric argument shows that $|T| \le 5^k$. We have $B_2^k \subset T + \frac{1}{2}B_2^k \subset \mathrm{conv}(T) + \frac{1}{2}B_2^k$, so $B_2^k \subset 2\mathrm{conv}(T)$. □

The next lemma comes from [4].

Lemma 11. *For any $1 \le k \le n$ there exists $T \subset B_2^n$ with $|T| \le \frac{2n}{k}5^k$ such that $B_2^n \subset 2\sqrt{\frac{2n}{k}}\mathrm{conv}(T)$.*

Proof. Let $l = \lceil n/k \rceil \le 2n/k$ and $\mathbb{R}^n = F_1 \oplus \cdots \oplus F_l$ be an orthogonal decomposition of $\mathbb{R}^n$ into spaces of dimension at most k. By Lemma 10 we can find $T_i \subset B_2(F_i) := B_2^n \cap F_i$ such that $B_2(F_i) \subset 2\mathrm{conv}(T_i)$ and $|T_i| \le 5^k$. Let $T := \bigcup_{i \le l} T_i$. Then $T \subset B_2^n$ and $|T| \le l5^k \le \frac{2n}{k}5^k$.

Fix now $x \in B_2^n$ and x_i denotes its orthogonal projection on F_i. Observe that

$$\sum_{i \le l} \|x_i\| \le \sqrt{l}\Big(\sum_{i \le l} \|x_i\|^2\Big)^{1/2} \le \sqrt{l}.$$

Therefore

$$x \subset \sqrt{l}\mathrm{conv}\Big\{0, \frac{x_1}{\|x_1\|}, \ldots, \frac{x_l}{\|x_l\|}\Big\} \subset \sqrt{l}\mathrm{conv}\Big(\bigcup_{i \le l} B_2(F_i)\Big) \subset 2\sqrt{l}\mathrm{conv}(T).$$

□

Lemma 12. *Let Y be a vector uniformly distributed over S^{n-1}. Then*

$$\mathbb{E}|\langle Y, t\rangle| I_{\{|\langle Y,t\rangle| \ge u\}} \le \min\Big\{\frac{|t|}{\sqrt{n}}, \frac{2(|t|^2 + nu^2)}{nu} e^{-nu^2/(2|t|^2)}\Big\} \quad t \in \mathbb{R}^n,\ u > 0.$$

Proof. Observe that $\langle Y, t\rangle$ is distributed as $|t|Y_1$. Hence

$$\mathbb{E}|\langle Y, t\rangle| I_{\{|\langle Y,t\rangle| \ge u\}} = |t|\mathbb{E}|Y_1| I_{\{|Y_1| \ge u/|t|\}}.$$

We have $\mathbb{E}|Y_1| \le (\mathbb{E}|Y_1|^2)^{1/2} = n^{-1/2}$. Moreover $\mathbb{P}(Y_1 \ge v) \le \exp(-nv^2/2)$ for $v \ge 0$ (cf. [12]). Therefore

$$\begin{aligned}\mathbb{E}|Y_1| I_{\{|Y_1| \ge u\}} &\le u\mathbb{P}(|Y_1| \ge u) + \int_u^\infty \mathbb{P}(|Y_1| \ge v)\mathrm{d}v \le 2ue^{-nu^2/2} + 2\int_u^\infty e^{-nv^2/2}\mathrm{d}v \\ &\le 2ue^{-nu^2/2} + 2\int_u^\infty \frac{nv}{nu} e^{-nv^2/2}\mathrm{d}v = \frac{2(1 + nu^2)}{nu} e^{-nu^2/2}.\end{aligned}$$

□

Now we are able to show that (4) holds for $T = B_2^n$ under $4 + \delta$ moment condition.

Proposition 13. *Let $X_1, \ldots, X_n$ be independent centered r.v's with variance 1 satisfying condition* (10)*. Then there exists $S \subset \mathbb{R}^n$ such that $|S| \le 10n^2$, $B_2^n \subset \mathrm{conv}(S)$ and*

$$M_X(S) \lesssim_{r,\lambda} \sqrt{n} \sim_\lambda \mathbb{E}|X| = b_X(B_2^n).$$

Proof. By the Rosenthal inequality [9] we have (recall that $r \in (4,8]$),

$$\big\| |X|^2 - n \big\|_{r/2} = \Big\| \sum_{i=1}^n (X_i^2 - 1) \Big\|_{r/2} \lesssim \Big(\sum_{i=1}^n \mathrm{Var}(X_i^2) \Big)^{1/2} + \Big(\sum_{i=1}^n \mathbb{E}|X_i^2 - 1|^{r/2} \Big)^{2/r}$$
$$\lesssim_\lambda n^{1/2} + n^{2/r} \le 2n^{1/2}.$$

Therefore

$$\mathbb{E}|X| I_{\{|X| \ge \sqrt{2n}\}} \le \mathbb{E}\sqrt{2(|X|^2 - n)} I_{\{|X| \ge \sqrt{2n}\}} \le \sqrt{2} n^{1/2 - r/2} \mathbb{E}(|X|^2 - n)^{r/2} \le C(\lambda) n^{1/2 - r/4}. \tag{13}$$

By Lemma 11 (applied with $k = c(r)\log n$) there exists $t_1, \dots, t_N$ such that $B_2^n \subset \mathrm{conv}\{t_1, \dots, t_N\}$, $N \le 10 n^{1/2 + r/8}$ and $|t_i| \le C(r)\sqrt{n/\log n}$, $1 \le i \le N$. Let U be the random rotation (uniformly distributed on $O(n)$) then Ut_i is distributed as $|t_i|Y$, where Y has uniform distribution on S^{n-1}. Thus by Lemma 12,

$$\begin{aligned}
\mathbb{E}_U \mathbb{E}_X |\langle X, Ut_i\rangle| I_{\{|\langle X, Ut_i\rangle| \ge u\}} &= \mathbb{E}_X \mathbb{E}_Y |\langle Y, |t_i| X\rangle| I_{\{|\langle Y, |t_i| X\rangle| \ge u\}} \\
&\le \mathbb{E} \min\Big\{ \frac{|t_i||X|}{\sqrt{n}}, \frac{2(|t_i|^2|X|^2 + nu^2)}{nu} e^{-nu^2/(2|t_i|^2|X|^2)} \Big\} \\
&\le \frac{|t_i|}{\sqrt{n}} \mathbb{E}|X| I_{\{|X| \ge \sqrt{2n}\}} + \frac{4|t_i|^2 + 2u^2}{u} e^{-u^2/(4|t_i|^2)}.
\end{aligned}$$

Recall that $|t_i| \lesssim_r \sqrt{n/\log n}$ so for sufficiently large $C(r)$ we get by (13),

$$\mathbb{E}_U \mathbb{E}_X |\langle X, Ut_i\rangle| I_{\{|\langle X, Ut_i\rangle| \ge C(r)\sqrt{n}\}} \le C(\lambda) n^{-r/4} |t_i| + n^{-2} \le C(r,\lambda) n^{1/2 - r/4}.$$

As a consequence there exists $U \in O(n)$ such that

$$\sum_{i=1}^N \mathbb{E}_X |\langle X, Ut_i\rangle| I_{\{|\langle X, Ut_i\rangle| \ge C(r)\sqrt{n}\}} \le N C(r,\lambda) n^{1/2 - r/4} \le 10 C(r,\lambda) n^{1 - r/8}. \tag{14}$$

Thus if we put $S := \{Ut_1, \dots, Ut_N\}$ we will have $\mathrm{conv}(S) = U\mathrm{conv}\{t_1, \dots, t_N\} \supset B_2^n$ and $M_X(S) \le C'(r,\lambda)\sqrt{n}$.

□

4.3 Ellipsoids

We now extend the bounds from the previous subsection to the case of ellipsoids, i.e. sets of the form

$$\mathcal{E} := \Big\{ t \in \mathbb{R}^n \colon \ \sum_{i=1}^n \frac{\langle t, u_i\rangle^2}{a_i^2} \le 1 \Big\}, \tag{15}$$

where $u_1, \dots, u_n$ is an orthonormal system in $\mathbb{R}^n$ and $a_1, \dots, a_n > 0$.

Observe that

$$\sup_{t\in\mathcal{E}}\langle t,x\rangle=\sqrt{\sum_{i=1}^n a_i^2\langle x,u_i\rangle^2}.$$

To treat this case we will need the following Lemma.

Lemma 14. *Let $X=(X_1,\dots,X_n)$, where X_i are independent mean zero and variance one r.v's satisfying $4+\delta$ condition* (10).
i) For any $a_1,\dots,a_n\ge 0$ and any o.n. vectors $u_1,\dots,u_n$,

$$\mathbb{E}\Big(\sum_{k=1}^n a_k^2\langle X,u_k\rangle^2\Big)^{1/2}\sim_\lambda\Big(\mathbb{E}\sum_{k=1}^n a_k^2\langle X,u_k\rangle^2\Big)^{1/2}=\Big(\sum_{k=1}^n a_k^2\Big)^{1/2}.$$

ii) For any $n\times n$ matrix B,

$$\Big(\mathbb{E}(|BX|^2-\|B\|_{\mathrm{HS}}^2)^{r/2}\Big)^{2/r}\le C(\lambda)\|B^TB\|_{\mathrm{HS}}^{1/2}.$$

In particular for any linear supspace $E\subset\mathbb{R}^n$ od dimension $k\in\{1,\dots,n\}$,

$$\Big(\mathbb{E}(|P_EX|^2-k)^{r/2}\Big)^{2/r}\le C(\lambda)k^{1/2}.$$

Proof. Part i) follows from Lemma 9.

To show part ii) let $B=(b_{ij})_{i,j=1}^n$, $e_1,e_2,\dots,e_n$ be the canonical basis of $\mathbb{R}^n$ and let

$$\sigma_{i,j}:=\sum_{l=1}^n b_{l,i}b_{l,j}=\langle Be_i,Be_j\rangle,\quad 1\le i,j\le n.$$

Then

$$\begin{aligned}\big\||BX|^2-\|B\|_{\mathrm{HS}}^2\big\|_{r/2}&=\Big\|\sum_{i=1}^n(X_i^2-1)\sigma_{i,i}+\sum_{1\le i\ne j\le n}X_iX_j\sigma_{i,j}\Big\|_{r/2}\\&\le\Big\|\sum_{i=1}^n(X_i^2-1)\sigma_{i,i}\Big\|_{r/2}+\Big\|\sum_{1\le i\ne j\le n}X_iX_j\sigma_{i,j}\Big\|_{r/2}.\end{aligned}$$

Applying Rosenthal's inequality we get

$$\begin{aligned}\Big\|\sum_{i=1}^n(X_i^2-1)\sigma_{i,i}\Big\|_{r/2}&\lesssim\Big(\sum_{i=1}^n\mathrm{Var}(X_i^2)\sigma_{i,i}^2\Big)^{1/2}+\Big(\sum_{i=1}^n\mathbb{E}(X_i^2-1)^{r/2}\sigma_{i,i}^{r/2}\Big)^{2/r}\\&\lesssim_\lambda\Big(\sum_{i=1}^n\sigma_{i,i}^2\Big)^{1/2}+\Big(\sum_{i=1}^n\sigma_{i,i}^{r/2}\Big)^{2/r}\le 2\Big(\sum_{i=1}^n\sigma_{i,i}^2\Big)^{1/2}.\end{aligned}$$

Hypercontractive method (as in the proof of Lemma 9) yields

$$\Big\|\sum_{i\neq j} X_iX_j\sigma_{i,j}\Big\|_{r/2} \lesssim_\lambda \Big\|\sum_{i\neq j} X_iX_j\sigma_{i,j}\Big\|_2 = \Big(\sum_{i\neq j}\sigma_{i,j}^2\Big)^{1/2}.$$

Finally

$$\Big(\sum_{i=1}^n \sigma_{i,i}^2\Big)^{1/2} + \Big(\sum_{i\neq j}\sigma_{i,j}^2\Big)^{1/2} \leq 2\Big(\sum_{i,j}\sigma_{i,j}^2\Big)^{1/2} = 2\|B^TB\|_{\mathrm{HS}}.$$

□

Now we state and prove the main result of this section.

Theorem 15. *Let $X_1,\ldots,X_n$ be independent centered r.v's satisfying the condition* (10) *and let T be an ellipsoid in $\mathbb{R}^n$. Then there exists $S\subset\mathbb{R}^n$ such that $|S|\leq 10n^2$, $T\subset$ conv(S) and*

$$M_X(S) \lesssim_{r,\lambda} b_X(T).$$

Proof. Since it is only a matter of scaling we may and will assume that $\mathbb{E}X_i^2=1$ for all i. Let $T=\mathcal{E}$ be an ellipsoid of the form (15). Then the first part of Lemma 14 yields

$$\mathbb{E}\sup_{t\in\mathcal{E}} X_t = \mathbb{E}\Big(\sum_{k=1}^n a_k^2\langle X,u_k\rangle^2\Big)^{1/2} \sim_\lambda \sqrt{\sum_{k=1}^n a_k^2}.$$

By homogenity we may assume that $\sum_{k=1}^n a_k^2 = 1$.

Define

$$I_k := \{i\colon\ 2^{-k-1} < a_i \leq 2^{-k}\},\ \ n_k := |I_k|,\ \ J := \{k\in\mathbb{Z}\colon\ I_k\neq\emptyset\},\ \ E_k := \mathrm{span}\{u_i\colon i\in I_k\}.$$

Then

$$1 \leq \sum_{k\in J} n_k 2^{-2k} < 4. \tag{16}$$

In particular J is a subset of nonnegative integers.

We claim that for any positive sequence $(c_k)_{k\in J}$ such that $\sum_k c_k^{-2}\leq 1$,

$$\mathcal{E}\subset \mathrm{conv}\Big(\bigcup_{k\in J} c_k2^{-k}B_2^{I_k}\Big), \quad \text{where} \quad B_2^{I_k} := B_2^n\cap E_k.$$

Indeed, let $P_kx := \sum_{i\in I_k}\langle x,u_i\rangle u_i$ be the projection of x onto E_k, then

$$x = \sum_{k\in J} c_k^{-1}2^k|P_kx|c_k2^{-k}\frac{P_kx}{|P_kx|}$$

and for $x \in \mathcal{E}$,

$$\sum_{k\in J} c_k^{-1} 2^k |P_k x| \le \sqrt{\sum_{k\in J} c_k^{-2}}\sqrt{\sum_{k\in J} 2^{2k}|P_k x|^2} \le \sqrt{\sum_{k\in J}\sum_{i\in I_k} \frac{\langle x, u_i\rangle^2}{a_i^2}} \le 1.$$

Let us for a moment fix $k \in J$. By Lemma 11 (applied with $k = c(r)\log n_k$) there exists $t_1, \dots, t_{N_k} \in E_k$ such that $B_2^{I_k} \subset \mathrm{conv}\{t_1, \dots, t_{N_k}\}$, $N_k \le 10 n_k^{1/2+r/8}$ and $|t_i| \le C(r)\sqrt{n_k/\log(n_k)}$. Let U be the random rotation of E_k (uniformly distributed on $O(E_k)$) then Ut_i is distributed as $|t_i|Y$, where Y has uniform distribution on $S^{I_k} := S^{n-1} \cap E_k$. Thus by Lemma 12,

$$\begin{aligned}
&\mathbb{E}_U \mathbb{E}_X |\langle X, Ut_i\rangle| I_{\{|\langle X, Ut_i\rangle| \ge u\}} \\
&\qquad = \mathbb{E}_X \mathbb{E}_Y |\langle Y, |t_i| P_{E_k} X\rangle| I_{\{|\langle Y, |t_i| P_{E_k} X\rangle| \ge u\}} \\
&\qquad \le \mathbb{E} \min\Big\{ \frac{|t_i||P_{E_k}X|}{\sqrt{n_k}}, \frac{2(|t_i|^2|P_{E_k}X|^2 + n_k u^2)}{n_k u} e^{-n_k u^2/(2|t_i|^2|P_{E_k}X|^2)} \Big\} \\
&\qquad \le \frac{|t_i|}{\sqrt{n_k}} \mathbb{E}|P_{E_k}X| I_{\{|P_{E_k}X| \ge \sqrt{2n_k}\}} + \frac{4|t_i|^2 + 2u^2}{u} e^{-u^2/(4|t_i|^2)}.
\end{aligned}$$

We have

$$\begin{aligned}
\mathbb{E}|P_{E_k}X| I_{\{|P_{E_k}X| \ge \sqrt{2n_k}\}} &\le \sqrt{2}\mathbb{E}(|P_{E_k}X|^2 - n_k)^{1/2} I_{\{|P_{E_k}X| \ge \sqrt{2n_k}\}} \\
&\le \sqrt{2} n_k^{1/2-r/2} \mathbb{E}(|P_{E_k}X|^2 - n_k)^{r/2} \le C(\lambda) n_k^{1/2-r/4},
\end{aligned}$$

where the last bound follows by Lemma 14. Recall that $|t_i| \lesssim_r \sqrt{n_k/\log n_k}$, thus for sufficiently large $C(r)$ we get

$$\mathbb{E}_U \mathbb{E}_X |\langle X, Ut_i\rangle| I_{\{|\langle X, Ut_i\rangle| \ge C(r)\sqrt{n}\}} \le C(\lambda) n_k^{-r/4}|t_i| + n_k^{-2} \le C(r,\lambda) n_k^{1/2-r/4}.$$

As a consequence there exists $U \in O(E_k)$ such that

$$\sum_{i=1}^{N_k} \mathbb{E}_X |\langle X, Ut_i\rangle| I_{\{|\langle X, Ut_i\rangle| \ge C(r)\sqrt{n_k}\}} \le N_k C(r,\lambda) n_k^{1/2-r/4} \le 10 C(r,\lambda) n_k^{1-r/8}.$$

Define $S_k = \{t_{k,1}, \dots, t_{k,N_k}\} := \{Ut_1, \dots, Ut_{N_k}\}$. Then $\mathrm{conv}(S_k) = U\mathrm{conv}\{t_1, \dots, t_{N_k}\} \supset B_2^{I_k}$, $N_k \le 10 n_k^{1/2+r/8} \le 10 n_k^2$ and

$$\sum_{i=1}^{N_k} \mathbb{E}_X |\langle X, t_{k,i}\rangle| I_{\{|\langle X, t_{k,i}\rangle| \ge C(r)\sqrt{n_k}\}} \le C(r,\lambda) n_k^{1-r/8}.$$

Set $c_k := 2^{k+2}(2^k+n_k)^{-1/2}$. By (16) we get $\sum_{k\in J} c_k^{-2} \le 1$, so

$$\mathcal{E} \subset \operatorname{conv}\Big(\bigcup_{k\in J} c_k 2^{-k} B_2^{I_k}\Big) \subset \operatorname{conv}(\{c_k 2^{-k} t_{k,i}\colon\ k\in J, i\le N_k\}) := \operatorname{conv}(S).$$

We have

$$|S| = \sum_{k\in J} N_k \le \sum_{k\in J} 10 n_k^2 \le 10\Big(\sum_{k\in J} n_k\Big)^2 = 10n^2.$$

Moreover,

$$\begin{aligned}\sum_{s\in S}\mathbb{E}|\langle s,X\rangle| I_{\{|\langle s,X\rangle|\ge 4C(r)\}} &= \sum_{k\in J} 2^{-k}c_k \sum_{i=1}^{N_k}\mathbb{E}|\langle t_{k,i},X\rangle| I_{\{2^{-k}c_k|\langle t_{k,i},X\rangle|\ge 4C(r)\}}\\ &\le \sum_{k\in J} 4(2^k+n_k)^{-1/2}\sum_{i=1}^{N_k}\mathbb{E}|\langle t_{k,i},X\rangle| I_{\{|\langle t_{k,i},X\rangle|\ge C(r)\sqrt{n_k}\}}\\ &\le \sum_{k\in J} 4(2^k+n_k)^{-1/2} C(r,\lambda) n_k^{1-r/8}\\ &\le 4C(r,\lambda)\sum_{k\in J}(2^k+n_k)^{1/2-r/8} \le 4C(r,\lambda)\sum_{k\ge 0} 2^{k(1/2-r/8)}\\ &\le C'(r,\lambda),\end{aligned}$$

which shows that $M_X(S) \sim \widetilde{M}_X(S) \lesssim_{\lambda,r} 1 \sim b_X(\mathcal{E})$.

□

5 Case III. ℓ_q^n-balls, $2 < q \le \infty$

It turns out that results of the previous sections may be easily applied to get estimates in the case when $T = B_q^n$ is the unit ball in ℓ_q^n and $q \in (2,\infty]$. In the whole section by q' we will denote the Hölder dual of q, i.e. $q' = \frac{q}{q-1}$, $2 \le q < \infty$ and $q' = 1$ for $q = \infty$.

Proposition 16. *Let $X_1,\ldots,X_n$ be independent centered r.v's with variance 1 satisfying condition* (10)*. Then there exists $S \subset \mathbb{R}^n$ such that $|S| \le 10n^2$, $B_q^n \subset \operatorname{conv}(S)$ and*

$$M_X(S) \lesssim_{r,\lambda} n^{1/q'} \sim_\lambda b_X(B_q^n).$$

Proof. Since $q' \in (1,2]$, condition (10) yields $\|X_i\|_{q'} \sim_\lambda \|X_i\|_{2q'} \sim_\lambda \|X_i\|_2 = 1$ and hence $(\mathbb{E}\|X\|_{q'}^{2q'})^{1/(2q')} \sim_\lambda (\mathbb{E}\|X\|_{q'}^{q'})^{1/q'}$. Therefore

$$b_X(B_q^n) = \mathbb{E}\sup_{t\in B_q^n}\langle t,X\rangle = \mathbb{E}\|X\|_{q'} \sim_\lambda \big(\mathbb{E}\|X\|_{q'}^{q'}\big)^{1/q'} \sim_\lambda n^{1/q'}.$$

Hölder's inequality implies $B_q^n \subset n^{1/2-1/q}B_2^n = n^{1/q'-1/2}B_2^n$ and the assertion easily follows from Proposition 13. □

Now let us consider the case of linear transformation of ℓ_q^n-ball, i.e. $T = AB_q^n$. Next simple lemma shows how to estimate $b_X(T)$.

Lemma 17. *Let $X = (X_1, \ldots, X_n)$, where X_i are independent mean zero and variance one r.v's satisfying $4+\delta$ condition (10). Then for any $n \times n$ matrix A and $2 \le q \le \infty$ we have*

$$b_X(AB_q^n) = b_{A^TX}(B_q^n) \sim_\lambda \Big(\sum_{i=1}^n |Ae_i|^{q'}\Big)^{1/q'}.$$

Proof. Observe that

$$\sup_{t \in AB_q^n} \langle X, t\rangle = \sup_{t \in B_q^n} \langle A^T X, t\rangle = \Big(\sum_{i=1}^n |\langle A^T X, e_i\rangle|^{q'}\Big)^{1/q'} = \Big(\sum_{i=1}^n |\langle X, Ae_i\rangle|^{q'}\Big)^{1/q'}.$$

Condition (10) (see Lemma 9) implies that

$$\|\langle X, Ae_i\rangle\|_{2q'} \sim_\lambda \|\langle X, Ae_i\rangle\|_{q'} \sim_\lambda \|\langle X, Ae_i\rangle\|_2 = |Ae_i|.$$

Hence $\|\sup_{t\in AB_q^n}\langle X,t\rangle\|_{2q'} \sim_\lambda \|\sup_{t\in AB_q^n}\langle X,t\rangle\|_{q'}$ and

$$b_X(AB_q^n) = \Big\|\sup_{t\in AB_q^n}\langle X,t\rangle\Big\|_1 \sim_\lambda \Big\|\sup_{t\in AB_q^n}\langle X,t\rangle\Big\|_{q'} \sim_\lambda \Big(\sum_{i=1}^n |Ae_i|^{q'}\Big)^{1/q'}.$$

□

As in the proof of Proposition 16 we may include linear image of B_q^n into ellipsoid with the comparable b_X-bound and deduce from Theorem 15 the following more general result.

Theorem 18. *Let $X_1, \ldots, X_n$ be independent centered r.v's satisfying condition (10) and let $T = AB_q^n$ for some $2 \le q \le \infty$ and an $n \times n$ matrix A. Then there exists $S \subset \mathbb{R}^n$ such that $|S| \le 10n^2$, $T \subset \mathrm{conv}(S)$ and*

$$M_X(S) \lesssim_{r,\lambda} b_X(T).$$

Proof. Since it is only a matter of scaling we may and will assume that $\mathbb{E}X_i^2 = 1$ for all i. By Lemma 17 it is enough to show that

$$M_X(S) \lesssim_{r,\lambda} \Big(\sum_{i=1}^n |Ae_i|^{q'}\Big)^{1/q'}.$$

By homogenity we may assume that $\sum_{i=1}^n |Ae_i|^{q'} = 1$. Case $q = 2$ was treated in Theorem 15, so we may assume that $q > 2$, i.e. $q' < 2$. Moreover, we may assume that $Ae_i \neq 0$ for all i.

Let $\lambda_i := |Ae_i|^{1-q'/2}$. Observe that if $t \in B_q^n$ then by Hölder's inequality

$$\begin{aligned}\sum_{i=1}^n |\lambda_i t_i|^2 &\le \Big(\sum_{i=1}^n |t_i|^q\Big)^{2/q}\Big(\sum_{i=1}^n |\lambda_i|^{2q/(q-2)}\Big)^{(q-2)/q}\\ &= \Big(\sum_{i=1}^n |t_i|^q\Big)^{2/q}\Big(\sum_{i=1}^n |Ae_i|^{q'}\Big)^{(q-2)/q} \le 1.\end{aligned}$$

This shows that $D^{-1}B_q^n \subset B_2^n$, where $D := \mathrm{diag}(d_1, \ldots, d_n)$ and $d_i := |Ae_i|^{q'/2-1}$. Hence $AB_q^n \subset ADB_2^n$ and

$$b_X(ADB_2^n) \sim_\lambda \Big(\sum_{i=1}^n |ADe_i|^2\Big)^{1/2} = \Big(\sum_{i=1}^n |Ae_i|^{q'}\Big)^{1/2} = 1.$$

We get the assertion applying Theorem 15 for the ellipsoid ADB_2^n.

□

6 Concluding remarks and open questions

We have shown that the main question has the affirmative answer in the case T is an ellipsoid (or more general linear image of ℓ_q^n-ball, $2 \le q \le n$) if X_i are independent mean zero r.v's satisfying the $4+\delta$ moment condition (10). The following questions are up to our best knowledge open.

- Does (4) holds for $T = B_q^n$, $1 < q < 2$ and X_i satisfying $4+\delta$ moment condition?
- John's theorem states that for any convex symmetric set T in $\mathbb{R}^n$ there exists an ellipsoid $\mathcal{E}$ such that $\mathcal{E} \subset T \subset \sqrt{n}\mathcal{E}$. Hence Theorem 15 implies that under $4+\delta$ condition (10) one may find finite set S such that $T \subset \mathrm{conv}(S \cup -S)$ and $M_X(S) \le C(r,\lambda)\sqrt{n}b_X(T)$. We do not whether one may improve upon $\sqrt{n}$ factor for general sets T.
- Are there heavy-tailed random variables X_i such that (4) holds for arbitrary set T (for heavy-tailed r.v's approach via chaining functionals described in Subsection 2.1 fails to work)?
- Let X_i be heavy-tailed symmetric Weibull r.v's (i.e. symmetric variables with tails $\exp(-t^r)$, $0 < r < 1$). Bogucki [2] was able to obtain two-sided bounds for $b_X(T)$ with the use of random permutations (which may be eliminated if T is permutationally invariant). We do not know if the convex hull method works in this case.

References

[1] W. Bednorz, R. Latała, *On the boundedness of Bernoulli processes*, Ann. of Math. (2) **180** (2014), 1167–1203.

[2] R. Bogucki, *Suprema of canonical Weibull processes*, Statist. Probab. Lett. **107** (2015).

[3] X. Fernique, *Regularité des trajectoires des fonctions aléatoires gaussiennes*, École d'Été de Probabilités de Saint-Flour, IV-1974, Lecture Notes in Mathematics 480, 1–96, Springer, Berlin, 1975.

[4] M. Kochol, *Constructive approximation of a ball by polytopes*, Math. Slovaca **44** (1994), 99–105.

[5] S. Kwapień, W. A. Woyczyński, *Random Series and Stochastic Integrals: Single and Multiple*, Birkhauser, Boston, 1992.

[6] R. Latała, T. Tkocz, *A note on suprema of canonical processes based on random variables with regular moments*, Electron. J. Probab. **20** (2015), no. 36, 1–17.

[7] R. Latała, M. Strzelecka, *Comparison of weak and strong moments for vectors with independent coordinates*, Mathematika **64** (2018), 211–229.

[8] C. A. Rogers, G. C. Shephard, *Convex bodies associated with a given convex body*, J. London Math. Soc. **33** (1958), 270–281.

[9] H. P. Rosenthal *On the subspaces of L^p ($p > 2$) spanned by sequences of independent random variables*, Israel J. Math. **8** (1970), 273–303.

[10] M. Talagrand, *Regularity of Gaussian processes*, Acta Math. **159** (1987), 99–149.

[11] M. Talagrand, *Upper and lower bounds for stochastic processes*, 2nd edition, Springer, Cham, 2021.

[12] T. Tkocz, *An upper bound for spherical caps*, Amer. Math. Monthly **119** (2012), 606–607.