Rev. Mat. Iberoamericana **24** (2008), no. 3, 1075–1095

Majorizing measures and proportional subsets of bounded orthonormal systems

Olivier Guédon, Shahar Mendelson, Alain Pajor and Nicole Tomczak-Jaegermann

Abstract

In this article we prove that for any orthonormal system $(\varphi_j)_{j=1}^n \subset L_2$ that is bounded in L_∞, and any $1 < k < n$, there exists a subset I of cardinality greater than $n - k$ such that on $\mathrm{span}\{\varphi_i\}_{i \in I}$, the L_1 norm and the L_2 norm are equivalent up to a factor $\mu(\log \mu)^{5/2}$, where $\mu = \sqrt{n/k}\sqrt{\log k}$. The proof is based on a new estimate of the supremum of an empirical process on the unit ball of a Banach space with a good modulus of convexity, via the use of majorizing measures.

1. Introduction

We study some natural empirical processes determined by uniformly convex Banach spaces with modulus of convexity of power type 2. Results of this kind were extensively studied in a Hilbertian setting, and became an important tool for investigations, for example, of the behaviour of various random sets of vectors (as in [14, 15, 7, 16, 2, 6]). We then apply these results to address a problem of selecting a subset of a bounded orthonormal system (for example, a set of characters), in the spirit of a result of Bourgain (see [20]) and of Talagrand [20], that also has applications to a sparse reconstruction ([2, 16, 6]). A particular case of this result, formulated in Theorem 3, is a so called Kashin's splitting of a set of $2k$ orthonormal vectors in L_2 that is bounded in L_∞.

2000 Mathematics Subject Classification: 46B07, 46B09, 42A05, 42A61.
Keywords: Empirical process, majorizing measure, orthonormal system.

Theorem. *There exist two positive constants c and C such that for any even integer n and any orthonormal system $(\varphi_j)_{j=1}^n$ in L_2 with $\|\varphi_j\|_{L_\infty} \leq L$ for $1 \leq j \leq n$, we can find a subset $I \subset \{1, \ldots, n\}$ with $n/2 - c\sqrt{n} \leq |I| \leq n/2 + c\sqrt{n}$ such that for every $a = (a_i) \in \mathbb{C}^n$,*

$$\Big\| \sum_{i \in I} a_i \varphi_i \Big\|_{L_2} \leq C\ L \sqrt{\log n} (\log \log n)^{5/2} \Big\| \sum_{i \in I} a_i \varphi_i \Big\|_{L_1}$$

and

$$\Big\| \sum_{i \notin I} a_i \varphi_i \Big\|_{L_2} \leq C\ L \sqrt{\log n} (\log \log n)^{5/2} \Big\| \sum_{i \notin I} a_i \varphi_i \Big\|_{L_1}.$$

This result strengthens Theorem 2.4 of [6] and is almost optimal. We have been told by J. Bourgain that the term $\sqrt{\log n}$ is necessary. We would like to thank L. Rodriguez-Piazza for showing us the details of the proof of this optimality in the case of the Walsh system and for allowing us to present the argument at the end of this paper. We thank also the referee who remarks that this construction allows to show that this bound is almost optimal for the trigonometric system as well. The technical proof of Theorem 1 about empirical processes will be presented in the first part. It is based on a construction of majorizing measures developped in [14] and [7] and the main new ideas we use are some observations on packing and covering numbers.

2. Maximal deviation of the empirical moment

We begin this section with some definitions and notation. If E is a normed space we denote by E^* the dual space to E and the dual norm is denoted by $\| \ \|_*$. The modulus of convexity of E is defined for any $\varepsilon \in (0, 2)$ by

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

We say that E has modulus of convexity of power type 2 if there is a constant c such that $\delta_E(\varepsilon) \geq c\varepsilon^2$ for every $\varepsilon \in (0, 2)$. It is well-known (see e.g., [11], Proposition 2.4) that this property is equivalent to the fact that the inequality

$$\left\| \frac{x+y}{2} \right\|^2 + \lambda^{-2} \left\| \frac{x-y}{2} \right\|^2 \leq \frac{1}{2}(\|x\|^2 + \|y\|^2) \tag{2.1}$$

holds for all $x, y \in E$ (where $\lambda > 0$ is a constant depending only on c). If (2.1) is satisfied then we say that E has modulus of convexity of power type 2 with constant λ (in such a case it is clear that $\delta_E(\varepsilon) \geq \varepsilon^2/2\lambda^2$).

The notions of type (and cotype) of a Banach space were studied extensively during the 70's (see, for example, [10] and the survey [9]). A Banach space E has type p if there is a constant C such that for every $N \in \mathbb{N}$ and every $x_1, \ldots, x_N$

$$\mathbb{E}\Big\| \sum_{i=1}^{N} g_i x_i \Big\| \leq C \Big(\sum_{i=1}^{N} \|x_i\|^p \Big)^{1/p} \tag{2.2}$$

where $g_1, \ldots, g_N$ are standard independent gaussian variables (that is $g_i \sim \mathcal{N}(0,1)$). The smallest constant C for which (2.2) holds is called the type p constant of E and is denoted by $T_p(E)$.

Moreover, it is well known that if E has modulus of convexity of power type 2 then the dual space E^* has also modulus of smoothness of power type 2, and therefore, E^* has type 2 (see, for example, [8, Theorem 1.e.16]).

2.1. Results on empirical processes

Our first theorem generalizes a result of Rudelson [15], proved in a Hilbertian setting, to the case of a Banach space with modulus of convexity of power type 2. In fact, in Theorem 1 we solve a question left open in [7] by removing the condition on the distance of E to an Euclidean space of the same dimension.

Theorem 1. *There exists an absolute constant C for which the following holds. Let E be a Banach space with modulus of convexity of power type 2 with constant λ. Then, for every vectors $X_1, \ldots, X_m$ in E^*,*

$$\begin{aligned} \mathbb{E} \sup_{y \in B_E} \Big| \sum_{j=1}^{m} \varepsilon_j |\langle X_j, y\rangle|^2 \Big| &\leq \\ &\leq C\,\lambda^4\, T_2(E^*) \sqrt{\log m} \max_{1\leq j\leq m} \|X_j\|_* \sup_{y\in B_E} \Big(\sum_{j=1}^{m} |\langle X_j, y\rangle|^2 \Big)^{1/2}, \end{aligned} \tag{2.3}$$

where the expectation is taken over the i.i.d. Bernoulli random variables $(\varepsilon_j)_{1\leq j\leq m}$ and $T_2(E^*)$ is the type 2 constant of E^*.

The proof of Theorem 1 uses the same construction of a majorizing measure used in [14] and in [7], and this construction was inspired by the work of Talagrand in [18]. The improvement in Theorem 1 compared with [7] comes from entropy estimates that will be presented in Section 2.2. We will only sketch the construction of the majorizing measure in Section 2.3 and explain how the argument from [7] may be adapted using the new estimates.

As it has been proved in [15], one can apply a symmetrization argument due to Giné and Zinn [5] combined with Theorem 1 and deduce the following result.

Theorem 2. *There exists an absolute constant C for which the following holds. Let E be a Banach space with modulus of convexity of power type 2 with constant λ. Let X be a random vector in E and set $X_1, \dots, X_m$ to be independent copies of X. If*

$$A = C\lambda^4 T_2(E^*)\sqrt{\frac{\log m}{m}}\left(\mathbb{E}\max_{1\le j\le m}\|X_j\|_*^2\right)^{1/2} \quad \text{and } \sigma^2 = \sup_{y\in B_E}\mathbb{E}|\langle X, y\rangle|^2$$

then

$$\mathbb{E}\sup_{y\in B_E}\left|\frac{1}{m}\sum_{j=1}^m |\langle X_j, y\rangle|^2 - \mathbb{E}|\langle X, y\rangle|^2\right| \le A^2 + \sigma A.$$

We omit further details and refer the reader to one of the articles [15, 7, 16, 6].

2.2. Covering and packing numbers

The new ingredient of our proof, in comparison with [7], are estimates on packing and covering numbers which we shall now discuss.

Definition 1. *Let T and B be symmetric convex bodies in a Banach space E. Define $N(T, B)$ to be the minimal number of translates of B needed to cover T, i.e.*

$$N(T, B) = \inf\left\{N, \exists\{x_1, \dots, x_N\} \subset T \text{ such that } T \subset \cup_{j=1}^N x_j + B\right\}.$$

We denote by $M(T, B)$ the maximal number of disjoint translates of B by elements of T, i.e.

$$M(T, B) = \sup\left\{N, \exists\{x_1, \dots, x_N\} \subset T \text{ such that } x_j - x_k \notin B \text{ for } j \ne k\right\}.$$

It is well known that these two quantities are related by

$$N(T, B) \le M(T, B) \le N(T, B/2). \tag{2.4}$$

Note that if E is a normed space, B_E is its unit ball and $B = \varepsilon B_E$, then $M(T, \varepsilon B_E)$ is the cardinality of a maximal ε-separated subset of T with respect to the norm in E. Also, if $S : X \to Y$ is an operator, we sometimes denote $M(SB_X, \varepsilon B_Y)$ by $M(S : X \to Y, \varepsilon)$, and use a similar notation for the covering numbers N.

The following is a combination of results that appeared in [1] and in [3] (see Lemma 3.3 in [18]). We repeat its proof since we need precise estimates on the dependence on the modulus of convexity of the space E.

Lemma 1. *Let E be a Banach space of modulus of convexity of power type 2 with constant λ and denote by $T_2(E^*)$ the type 2 constant of E^*.*

Let $X_1,\dots,X_m$ be vectors in E^* such that for every $1\le j\le m$, $\|X_j\|_*\le L$, and define for every $y\in E$,

$$\|y\|_{\infty,m}=\max_{1\le \ell\le m}\big|\langle X_\ell,y\rangle\big|.$$

If $\varepsilon>0$ and $x_1,\dots,x_N\in B_E$ are ε-separated with respect to $\|\cdot\|_{\infty,m}$ then

$$\varepsilon\sqrt{\log N}\le C\lambda^2\ T_2(E^*)\ L\sqrt{\log m}$$

where C is an absolute constant.

Proof. Let $(e_i)_{i=1}^m$ be the standard basis in ℓ_1^m and define $S:\ell_1^m\to E^*$ by $Se_\ell=X_\ell$ for every $\ell=1,\dots,m$. Hence, for every $\ell\ne\ell'$, $|S^*(x_\ell-x_{\ell'})|_\infty\ge\varepsilon$ where $|\cdot|_\infty$ denotes the norm in ℓ_∞^m. Thus,

$$N\le M\big(S^*:E\to\ell_\infty^m,\varepsilon\big).$$

Since $\|S\|=\max_{1\le\ell\le N}\|X_\ell\|_*\le L$ and E^* has type 2, Proposition 1 in [3] yields that

$$\varepsilon\sqrt{\log N(S:\ell_1^m\to E^*,\varepsilon)}\le C\,L\,T_2(E^*)\sqrt{\log m}$$

where C is an absolute constant. By (2.4), a similar estimate holds for $M(S:\ell_1^m\to E^*,\varepsilon)$, where C is replaced by a new constant, also denoted by C. Define the function

$$f(\varepsilon)=\varepsilon\sqrt{\log M(S:\ell_1^m\to E^*,\varepsilon)}$$

and observe that f is bounded by $C\,L\,T_2(E^*)\sqrt{\log m}$.

Since E has modulus of convexity of power type 2 with constant λ, Proposition 2 in [1] states that for every $\theta\ge\varepsilon/5$,

$$M\big(S^*:E\to\ell_\infty^m,\varepsilon\big)\le M\big(S^*:E\to\ell_\infty^m,\theta\big)M\big(S:\ell_1^m\to E^*,c\varepsilon\delta_E(\varepsilon/5\theta)\big)$$

and recall that $\delta_E(\varepsilon/\theta)\ge\varepsilon^2/8\lambda^2\theta^2$. Therefore, if

$$h(\varepsilon)=\varepsilon\sqrt{\log M(S^*:E\to\ell_\infty^m,\varepsilon)},$$

then the previous inequality implies that

$$h(\varepsilon)\le\frac{\varepsilon}{\theta}\,h(\theta)+c'\,\left(\frac{\lambda^2\theta^2}{\varepsilon^2}\right)f\left(\frac{c\varepsilon^3}{\theta^2\lambda^2}\right)$$

where c,c' are absolute constants.

Choosing $\theta = 2\varepsilon$ and taking the supremum in ε, it follows that h is bounded by

$$C' \lambda^2 L T_2(E^*) \sqrt{\log m}$$

proving the announced result. ■

The next lemma is a simple application of Sudakov's inequality [17].

Lemma 2. *There exists an absolute constant C for which the following holds. Let E be a Banach space such that E^* has type 2. Assume that $X_1, \ldots, X_m$ are vectors in E^* with $\sup_{1 \le j \le m} \|X_j\|_* \le L$. Let $\alpha_1, \ldots, \alpha_m$ such that $\sum_{\ell=1}^m \alpha_\ell^2 \le 1$ and for every $y \in E$ set*

$$|y|_{\mathcal{E}}^2 = \sum_{\ell=1}^m |\langle X_\ell, y\rangle|^2 \alpha_\ell^2.$$

If $\varepsilon > 0$ and $x_1, \ldots, x_N \in B_E$ are ε-separated with respect to $|\cdot|_{\mathcal{E}}$, then

$$\varepsilon \sqrt{\log N} \le C T_2(E^*) L.$$

Proof. Let $\mathcal{E}_1$ be the ellipsoid in $\mathbb{R}^m$ consisting of all $y \in \mathbb{R}^m$ such that $|y|_{\mathcal{E}_1} := \left(\sum_{\ell=1}^m \alpha_\ell^2 \langle y, e_\ell\rangle^2\right)^{1/2} \le 1$. Set $H = (\mathbb{R}^m, |\cdot|_{\mathcal{E}_1})$ and define the operator $S : H^* \to E^*$ by $Se_i = X_i$, where $(e_i)_{i=1}^m$ is the standard basis in $\mathbb{R}^m$. Note that for every x_i, x_j,

$$\left|S^*(x_i - x_j)\right|_{\mathcal{E}_1}^2 = \sum_{\ell=1}^m \alpha_\ell^2 \langle S^*(x_i - x_j), e_\ell\rangle^2 = \left|x_i - x_j\right|_{\mathcal{E}}^2,$$

and thus the points $\{S^*(x_1), \ldots, S^*(x_N)\}$ are ε-separated in $|\cdot|_{\mathcal{E}_1}$ and belong to $S^*(B_E)$. By (2.4), $N \le N(S^*(B_E), (\varepsilon/2)\mathcal{E}_1)$. On the other hand, $\mathcal{E}_1 = T^{-1}(B_2^m)$ where $T : \mathbb{R}^m \to \mathbb{R}^m$ is the diagonal operator $Te_\ell = \alpha_\ell e_\ell$. Applying Sudakov's inequality [17],

$$\text{(2.5)} \quad \varepsilon \sqrt{\log N} \le \varepsilon \sqrt{\log N(S^*(B_E), (\varepsilon/2)T^{-1}(B_2^m))} \le C \mathbb{E} \sup_{z \in T(S^*(B_E))} \langle G, z\rangle$$

where G is a canonical Gaussian vector in ℓ_2^m and C is an absolute constant. Moreover

$$\begin{aligned} \mathbb{E} \sup_{z \in T(S^*(B_E))} \langle G, z\rangle &= \mathbb{E} \sup_{y \in B_E} \langle TG, S^* y\rangle \\ = \mathbb{E} \Big\| \sum_{\ell=1}^m \alpha_\ell g_\ell X_\ell \Big\|_* &\le T_2(E^*) \Big(\sum_{\ell=1}^m \alpha_\ell^2 \|X_\ell\|_*^2 \Big)^{1/2} \le T_2(E^*) L \end{aligned}$$

where we have used the type 2 inequality for E^* and the fact that for every $1 \le \ell \le m$, $\|X_\ell\|_* \le L$. The result now follows from (2.5). ■

2.3. Construction of a majorizing measure

The construction that we present here is the same one that was presented in [14] and in [7]. Let $X_1, \dots, X_m$ be m fixed vectors in E^* and define the random process $\{V_y : y \in B_E\}$ by

$$V_y = \sum_{j=1}^{m} \varepsilon_j \big| \langle X_j, y \rangle \big|^2,$$

where ε_j are independent symmetric Bernoulli random variables.

Our aim is to show that when E has modulus of convexity of power type 2,

$$\mathbb{E} \sup_{y \in B_E} |V_y| \leq$$

$$\leq C\, \lambda^4\, T_2(E^*) \max_{1 \leq j \leq m} \|X_j\|_* \sqrt{\log m} \left(\sup_{y \in B_E} \sum_{i=1}^{m} |\langle X_j, y \rangle|^2 \right)^{1/2} \tag{2.6}$$

for a suitable absolute constant C.

It is known that the process $\{V_y : y \in B_E\}$ satisfies a sub-Gaussian tail estimate, namely, that for every $y, \overline{y} \in E$ and any $t > 0$,

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

where

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

and c is an absolute constant.

It will be preferable to consider the following quasi-metric

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

and the quasi-norm $\| \cdot \|_{\infty,m}$ endowed on E by

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

The proof of inequality (2.6) is based on the majorizing measure theory of Talagrand [19]. The following theorem is a combination of Proposition 2.3, Theorem 4.1 and Proposition 4.4 of [19].

Theorem [19] *Assume that the process $\{V_y : y \in B_E\}$ is subgaussian with respect to a metric d. Let $r \geq 2$ and $k_0 \in \mathbb{Z}$ be the largest integer such that r^{-k_0} is greater than the radius of B_E with respect to the metric d. For every $k \geq k_0$ let $\phi_k : B_E \to \mathbb{R}^+$ be a family of maps satisfying the following assumption: there exists $A > 0$ such that for any point $x \in B_E$, any $k \geq k_0$ and any $N \in \mathbb{N}$*

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

Then

$$\mathbb{E} \sup_{y \in B_E} |V_y - V_0| \leq c\, A \cdot \sup_{k \geq k_0, x \in K} \phi_k(x).$$

The construction requires certain properties of the quasi metric d and the quasi norm $\| \ \|_{\infty,m}$. We refer to Propositions 1 and 2 in [7] for precise properties of these metrics and list the ones we require in the following lemma.

Lemma 3. *For every $y, \overline{y} \in \mathbb{R}^n$ and every $u \in B_E$,*

$$\tilde{d}(y, \overline{y}) \leq 2\, d(y, \overline{y}), \tag{2.7}$$

$$d(y, \overline{y}) \leq \sqrt{2} \|y - \overline{y}\|_{\infty,m} \sqrt{M}, \tag{2.8}$$

$$\|y - \overline{y}\|_{\infty,m} \leq \max_{1 \leq j \leq m} \|X_j\|_* \, \|y - \overline{y}\|, \tag{2.9}$$

$$d^2(z, \overline{z}) \leq 8 \left(|z - \overline{z}|^2_{\mathcal{E}_u} + M\, \|z - \overline{z}\|^2_{\infty,m} (\|z - u\|^2 + \|\overline{z} - u\|^2) \right), \tag{2.10}$$

where

$$|x|^2_{\mathcal{E}_u} = \sum_{i=1}^m \langle X_i, x \rangle^2 \langle X_i, u \rangle^2 \ \textit{and}\ M = \sup_{y \in B_E} \sum_{j=1}^m |\langle X_j, y \rangle|^2.$$

Moreover, for every $x \in E$ and $\rho > 0$, the ball (with respect to the quasi-metric d) centered in x and with radius ρ, denoted by $\mathcal{B}_x(\rho)$, is convex.

Note that by combining (2.8) and (2.9) it follows that for every $\rho > 0$ and every $x \in B_E$, $\inf_{y \in \mathcal{B}_\rho(x)} \|y\|$ is attained.

Proof of Theorem 1. Since there is only a finite number of points $X_1, \dots, X_m$ then by passing to a quotient of E we can assume that E is a finite dimensional space. We will denote its dimension by n and obviously one can assume that $m \geq n$. Also, recall that if E is a Banach space with modulus of convexity of power type 2 with constant λ then every quotient of E satisfies that property with a constant smaller than λ (see, e.g., [8]).

By the homogeneity of the statement we can assume that

$$\sup_{y\in B_E}\sum_{j=1}^{m}|\langle X_j,y\rangle|^2=1 \tag{2.11}$$

and by inequality (2.7), V_y is a sub-Gaussian process with the quasi-metric $2d$.

Therefore, if we denote

$$L=\max_{1\le j\le m}\|X_j\|_*$$

our aim is to show that

$$\mathbb{E}\sup_{y\in B_E}|V_y|\le C\,\lambda^4\,T_2(E^*)\,L\,\sqrt{\log m}$$

for an absolute constant C.

By inequality (2.8), the diameter of B_E with respect to the metric d is bounded by $2\sqrt{2}L$. Let r be a fixed number chosen large enough, set k_0 to be the largest integer such that $r^{-k_0}\ge 2\sqrt{2}L$ and put k_1 to be the smallest integer such that $r^{-k_1}\le L/\sqrt{n}$, where n is the dimension of E. We shall use the same definition of the functionals $\phi_k : B_E\to\mathbb{R}^+$ as in [14] and [7], namely:

$$\begin{cases}\forall k\ge k_1+1, \phi_k(x)=1+\dfrac{1}{2\log r}+\dfrac{\sqrt{n}}{L\sqrt{\log m}}\displaystyle\sum_{\ell=k_1}^{k}r^{-\ell}\sqrt{\log(1+4Lr^{\ell})}\\ \forall k_0\le k\le k_1, \phi_k(x)=\min\left\{\|y\|^2, y\in\mathcal{B}_{8r^{-k}}(x)\right\}+\dfrac{k-k_0}{\log m}.\end{cases}$$

It is easy to verify using definitions of k_0 and k_1 that

$$\sup_{x\in B_E,k\ge k_0}\phi_k(x)\le c,$$

where c is an absolute constant.

It remains to prove that our functionals satisfy condition (H) for

$$A=C\,\lambda^4\,T_2(E^*)\,L\,\sqrt{\log m},$$

and that will conclude the proof of Theorem 1. ■

Proof of condition (H)**.** Fix integers N and k, let $x\in B_E$ and $x_1,\ldots,x_N\in\mathcal{B}_{r^{-k}}(x)$ for which $d(x_i,x_j)\ge r^{-k-1}$.

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

$$\phi_{k+2}(x_i)-\phi_k(x)\ge\frac{\sqrt{n\log(1+4Lr^{k+2})}}{L\sqrt{\log m}}\,r^{-k-2}.$$

Since the points $x_1, \ldots, x_N$ are well separated with respect to the metric d, then by (2.8) and (2.9),

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

By a classical volumetric estimate (see, for example [12]), $N(B_E, tB_E) \leq (1+2/t)^n$ where n is the dimension of E. Therefore,

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

which proves the desired inequality.

The case $k_0 \leq k \leq k_1 - 2$ is more delicate and the main ingredients in this part are the entropy estimates proved in part 2.2.

For $j = 1, \ldots, N$ denote by $z_j \in B_E$ points at which $\min\{\|y\|^2, y \in \mathcal{B}_{8r^{-k-2}}(x_j)\}$ is attained and set $u \in B_E$ to be a point at which $\min\{\|y\|^2, y \in \mathcal{B}_{8r^{-k}}(x)\}$ is attained. Put

$$\theta = \max_j \|z_j\|^2 - \|u\|^2,$$

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

$$\theta + \frac{2}{\log m} \geq r^{-k}\sqrt{\log N}/A. \tag{2.12}$$

Following [7], it is evident from the properties of $\mathcal{B}_\rho$ (see Proposition 1 and 2 in [7] and Lemma 3) that for any $i \neq j$, $d(z_i, z_j) \geq cr^{-k-1}$ and that $d(x, z_j) \leq 8r^{-k}$. It implies that $(z_j + u)/2 \in \mathcal{B}_{8r^{-k}}(x)$ and by the definition of u, $\|u\| \leq \|z_j + u\|/2$. Since B_E has modulus of convexity of power type 2, then for all $j = 1, \ldots, N$ $\|z_j - u\| \leq \lambda\sqrt{2\theta}$. Thus, the set

$$U = u + \lambda\sqrt{2\theta}B_E$$

contains all the z_j's.

Fix an absolute constant $\tilde{c}$ to be named later, set

$$\delta = \tilde{c}\lambda^{-1}r^{-k}\theta^{-1/2}$$

and let S be the maximal number of points in U that are δ separated in $\|\cdot\|_{\infty,m}$. By Lemma 1,

$$\delta\sqrt{\log S} \leq C\,\lambda^2\,T_2(E^*)\,L\sqrt{\log m}\;\lambda\;\sqrt{\theta}$$

where C is an absolute constant. We consider now two cases.

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

$$\sqrt{\log N} \leq C\, T_2(E^*)\, L\sqrt{\log m}\ \lambda^3\ \sqrt{\theta}/\,\delta \leq C\,\theta\ \ r^k\,\lambda^4\,T_2(E^*)L\,\sqrt{\log m}$$

which proves (2.12).

The second case is when $S \leq \sqrt{N}$. Since S is the maximal number of points in U that are δ separated with respect to $\|\cdot\|_{\infty,m}$, U is covered by S balls of diameter smaller than δ in $\|\cdot\|_{\infty,m}$. Thus, there exists a subset J of $\{1,\ldots,N\}$ with cardinality $|J| \geq \sqrt{N}$ such that

$$\forall i,j \in J, \quad \|z_i - z_j\|_{\infty,m} \leq \delta.$$

Recall that for any $y \in E$,

$$|y|^2_{\mathcal{E}_u} = \sum_{\ell=1}^{m} |\langle X_\ell, y\rangle|^2 |\langle X_\ell, u\rangle|^2 = \sum_{\ell=1}^{m} |\langle X_\ell, y\rangle|^2 \alpha_\ell^2.$$

It is evident from (2.11) that $\sum \alpha_\ell^2 \leq 1$ and from (2.10) that for every $z, \overline{z}, u \in E$,

$$d^2(z,\overline{z}) \leq 8\big(|z-\overline{z}|^2_{\mathcal{E}_u} + \|z-\overline{z}\|^2_{\infty,m}(\|z-u\|^2 + \|\overline{z}-u\|^2)\big).$$

Since $d(z_i,z_j) \geq cr^{-k-1}$,

$$\|z_i - u\| \leq 2\lambda\sqrt{\theta} \qquad \text{and} \qquad \|z_i - z_j\|_{\infty,m} \leq \delta$$

for any $i,j \in J$, we can define $\tilde{c}$ small enough such that for all $i \neq j \in J$,

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

Thus, there are $|J|$ points in $u + \lambda\sqrt{2\theta}B_E$ that are cr^{k-1}-separated for $|\cdot|_{\mathcal{E}_u}$. Using Lemma 2,

$$\sqrt{\log |J|} \leq CT_2(E^*)L\ r^k\ \lambda\sqrt{\theta},$$

and a simple application of the arithmetic-geometric means inequality proves that

$$\sqrt{\log N} \leq CT_2(E^*)L\ r^k\ \lambda\sqrt{\log m}\left(\theta + \frac{2}{\log m}\right),$$

completing the proof of (2.12) (because $\lambda \geq 1$). ■

3. Selecting an arbitrary proportion of a bounded orthonormal system

In this section, we will prove the

Theorem 3. *There exist two positive constants c, C such that for any orthonormal system $(\varphi_j)_{j=1}^n$ in L_2 with $\|\varphi_j\|_{L_\infty} \le L$ for $1 \le j \le n$, the following holds.*

1) *For any $1 < k < n$ there exists a subset $I \subset \{1, \dots, n\}$ with $|I| \ge n-k$ such that for every $a = (a_i) \in \mathbb{C}^n$,*

$$\Big\| \sum_{i \in I} a_i \varphi_i \Big\|_{L_2} \le C\ \mu\ (\log \mu)^{5/2} \Big\| \sum_{i \in I} a_i \varphi_i \Big\|_{L_1} \tag{3.1}$$

where $\mu = L\sqrt{n/k}\,\sqrt{\log k}$.

2) *Moreover, if n is an even integer, there exists a subset $I \subset \{1, \dots, n\}$ with $n/2 - c\sqrt{n} \le |I| \le n/2 + c\sqrt{n}$ such that for every $a = (a_i) \in \mathbb{C}^n$,*

$$\Big\| \sum_{i \in I} a_i \varphi_i \Big\|_{L_2} \le C\ L\sqrt{\log n}(\log\log n)^{5/2} \Big\| \sum_{i \in I} a_i \varphi_i \Big\|_{L_1}$$

and

$$\Big\| \sum_{i \notin I} a_i \varphi_i \Big\|_{L_2} \le C\ L\sqrt{\log n}(\log\log n)^{5/2} \Big\| \sum_{i \notin I} a_i \varphi_i \Big\|_{L_1}.$$

Remark. Theorem 3 is satisfactory when k is an arbitrary proportion of n. It strengthens Theorem A from [6] and extends the result of Talagrand [20] which is applicable only when $|I| = n - k$ is a sufficiently small proportion of n. Here we get an arbitrary proportion, with the distance $\sqrt{\log n}(\log\log n)^{5/2}$, which is the right power of $\log n$ as we will see at the end of the paper.

Define $\rho = \rho_{k,n}$, the restricted Kolmogorov k-width of the system as the smallest number ρ such that there exists a subset $I \subset \{1, \dots, n\}$ with $|I| \ge n - k$ satisfying

$$\Big\| \sum_{i \in I} a_i \varphi_i \Big\|_{L_2} \le \rho \Big\| \sum_{i \in I} a_i \varphi_i \Big\|_{L_1} \tag{3.2}$$

for every $a = (a_i) \in \mathbb{C}^n$ (see [6] section 3). It was proved in [2] that $\rho = O(\sqrt{n/k}\log^3 n)$. This result was improved to $\rho = O(\sqrt{n/k}\sqrt{\log n}\ \log^{3/2} k)$ in [16].

The proof of Theorem 3 is based on a random method and follows the argument given in [6] for proving Theorem 2.1. However, instead of working with the space L_1, which is not uniformly convex, we will approximate it by an L_p space for p "close" to 1 and use the full strength of the estimate given in Theorem 2.

Proposition 1. *There exist two positive constants c, C such that for any orthonormal system $(\varphi_j)_{j=1}^n$ in L_2 with $\|\varphi_j\|_{L_\infty} \le L$ for $1 \le j \le n$, the following holds.*

1) *For any $p \in (1,2)$ and any $1 < k < n$ there exists a subset $I \subset \{1,\dots,n\}$ with $|I| \ge n-k$ such that for every $a = (a_i) \in \mathbb{C}^n$,*

$$\Big\| \sum_{i\in I} a_i\varphi_i \Big\|_{L_2} \le \frac{C}{(p-1)^{5/2}}\, L\, \sqrt{n/k}\,\sqrt{\log k}\, \Big\| \sum_{i\in I} a_i\varphi_i \Big\|_{L_p}. \tag{3.3}$$

2) *Moreover, if n is an even integer, there exists a subset $I \subset \{1,\dots,n\}$ with $n/2 - c\sqrt{n} \le |I| \le n/2 + c\sqrt{n}$ such that for every $a = (a_i) \in \mathbb{C}^n$,*

$$\Big\| \sum_{i\in I} a_i\varphi_i \Big\|_{L_2} \le \frac{C}{(p-1)^{5/2}}\, L\, \sqrt{n/k}\,\sqrt{\log k}\, \Big\| \sum_{i\in I} a_i\varphi_i \Big\|_{L_p}$$

and

$$\Big\| \sum_{i\notin I} a_i\varphi_i \Big\|_{L_2} \le \frac{C}{(p-1)^{5/2}}\, L\, \sqrt{n/k}\,\sqrt{\log k}\, \Big\| \sum_{i\notin I} a_i\varphi_i \Big\|_{L_p}.$$

Proof of Proposition 1. Let X be the random vector taking the value φ_i with probability $1/n$ and denote by E the complex vectorial space spanned by $\{\varphi_1,\dots,\varphi_n\}$. Let $1 < k < n$ and let $X_1,\dots,X_k$ be independent copies of X. We define an operator $\Gamma : E \to \ell_2^k$ by

$$\forall y \in E, \quad \Gamma y = \sum_{i=1}^k \langle X_i, y\rangle e_i$$

where $(e_1,\dots,e_k)$ denotes the canonical basis of ℓ_2^k. Since $(\varphi_j)_{j=1}^n$ is an orthonormal system of L_2, the basic properties of Γ are:

$$\begin{cases} (i)\ \mathbb{E}\,\|\Gamma y\|_{\ell_2^k}^2 = \dfrac{k}{n}\,\|y\|_{L_2}^2, \\ (ii)\ \ker\Gamma = \operatorname{span}\big\{(\varphi_j)_{j=1}^n \setminus (X_i)_{i=1}^k\big\}. \end{cases}$$

We shall first prove that for any $\delta \in (0,1)$, with probability greater than $1-\delta$, the set $(\varphi_i)_{i\in I} = \{(\varphi_j)_{j=1}^n \setminus (X_i)_{i=1}^k\}$ satisfies

$$\forall (a_i)_{i\in I} \in \mathbb{C}^{|I|}, \Big\| \sum_{i\in I} a_i\varphi_i \Big\|_{L_2} \le \frac{C}{\delta\,(p-1)^{5/2}}\, L\, \sqrt{n/k}\,\sqrt{\log k} \Big\| \sum_{i\in I} a_i\varphi_i \Big\|_{L_p}$$

for a universal constant C.

Let $S_{L_2} = \{y : \|y\|_{L_2} = 1\}$ be the unit sphere in L_2 and observe that for any star-shaped subset $T \subset L_2$ the following holds: if $\rho > 0$ satisfies

$$\sup_{y \in T \cap \rho S_{L_2}} \left| \sum_{j=1}^{k} \langle X_j, y \rangle^2 \; - \; \frac{k}{n}\, \rho^2 \right| \le \frac{k\rho^2}{3n}, \tag{3.4}$$

then

$$\operatorname{diam}(\ker \Gamma \cap T) \le \rho. \tag{3.5}$$

Indeed, condition (3.4) implies that for all $y \in T \cap \rho S_{L_2}$,

$$\frac{2k\rho^2}{3n} \le \|\Gamma y\|_{\ell_2^k}^2 \le \frac{4k\rho^2}{3n}. \tag{3.6}$$

The homogeneity of (3.6) and the fact that T is star-shaped imply that if the lower bound in (3.6) holds for all $y \in T \cap \rho S_{L_2}$, then the same lower bound also holds for all $y \in T$ with $\|y\|_{L_2} \ge \rho$. This in turn shows that if $y \in \ker \Gamma \cap T$ then $\|y\|_{L_2} \le \rho$, as required in (3.5).

Since $\ker \Gamma = \operatorname{span}\{(\varphi_j)_{j=1}^n \setminus (X_i)_{i=1}^k\}$, we get for $T = B_{L_p} \cap E$ that whenever ρ satisfies (3.4), then

$$\left\| \sum_{i \in I} a_i \varphi_i \right\|_{L_2} \le \rho \left\| \sum_{i \in I} a_i \varphi_i \right\|_{L_p}$$

for all scalars a_i, proving that (3.3) is satisfied with the constant ρ.

In order to find ρ that satisfies (3.4) with positive probability we use Theorem 2. Denote by E_p the complex Banach space spanned by $\varphi_1, \ldots, \varphi_n$ endowed with the norm defined by

$$\|y\| = \left(\frac{\|y\|_{L_p}^2 + \rho^{-2}\|y\|_{L_2}^2}{2} \right)^{1/2}.$$

It is clear that $(B_{L_p} \cap E) \cap \rho B_{L_2} \subset B_{E_p} \subset \sqrt{2}(B_{L_p} \cap E) \cap \rho B_{L_2}$ and that the following properties are satisfied:

$$\begin{cases} E_p \text{ is a Banach space with modulus of convexity of power type 2} \\ \text{with constant } \lambda^{-2} = p(p-1)/8, \\ E_p^* \text{ is a Banach space of type 2 and } T_2(E_p^*) \le C\sqrt{q} = C\sqrt{p/(p-1)}. \end{cases}$$

Indeed, the first property follows from Clarkson inequality [4], that for any $f, g \in L_p$,

$$\left\| \frac{f+g}{2} \right\|_{L_p}^2 + \frac{p(p-1)}{8} \left\| \frac{f-g}{2} \right\|_{L_p}^2 \le \frac{1}{2}\big(\|f\|_{L_p}^2 + \|g\|_{L_p}^2 \big).$$

The second property is evident because for any $q \ge 2$, L_q has type 2 with constant $C\sqrt{q}$.

By property (i) of Γ, we get, taking $T = B_{L_p} \cap E$ and applying Theorem 2 to E_p,

$$\begin{aligned}
\mathbb{E} \sup_{y \in T \cap \rho S_{L_2}} \left| \sum_{j=1}^{k} \langle X_j, y\rangle^2 - \frac{k}{n}\rho^2 \right| &= \mathbb{E} \sup_{y \in T \cap \rho S_{L_2}} \left| \sum_{j=1}^{k} (\langle X_j, y\rangle^2 - \mathbb{E}\langle X_j, y\rangle^2) \right| \\
&\le \mathbb{E} \sup_{y \in T \cap \rho B_{L_2}} \left| \sum_{j=1}^{k} (\langle X_j, y\rangle^2 - \mathbb{E}\langle X_j, y\rangle^2) \right| \\
&\le \mathbb{E} \sup_{y \in B_{E_p}} \left| \sum_{j=1}^{k} (\langle X_j, y\rangle^2 - \mathbb{E}\langle X_j, y\rangle^2) \right| \\
&\le k(A^2 + A\,\sigma)
\end{aligned}$$

where $\sigma = \sup_{y \in B_{E_p}} \|y\|_{L_2}/\sqrt{n} \le \sqrt{2}\rho/\sqrt{n}$ and

$$\begin{aligned}
A &= C\,\lambda^4\, T_2(E_p^*) \sqrt{\frac{\log k}{k}} \big(\mathbb{E} \max_{j \in J} \|\varphi_j\|^2_{E_p^*}\big)^{1/2} \\
&\le C\,(p-1)^{-5/2}\ L\ \sqrt{\frac{\log k}{k}}
\end{aligned}$$

since for every $1 \le j \le n$, $\|\varphi_j\|_{E_p^*} \le \|\varphi_j\|_{L_\infty} \le L$. By Chebychev inequality, for any $\delta \in (0,1)$, we conclude that with probability greater than $1-\delta$, there exists $X_1, \ldots, X_k$ such that for any positive ρ,

$$\sup_{y \in T \cap \rho S_{L_2}} \left| \|\Gamma y\|^2_{\ell_2^k} \ - \ \frac{k}{n}\rho^2 \right| \le \frac{k}{\delta}\big(A^2 \ + \ A\,\sigma\big).$$

To conclude, we choose a constant c large enough such that, for

$$\rho = c\sqrt{\frac{n}{k}}\ \sqrt{\log k}\ \frac{L}{\delta\,(p-1)^{5/2}},$$

the inequality (3.4) is satisfied with probability greater than $1-\delta$.

The cardinality of the set $(\varphi_i)_{i \in I} = \{(\varphi_j)_{j=1}^n \setminus (X_i)_{i=1}^k\}$ is greater than $n-k$ and the first part of the proposition is proven choosing $\delta = 1/2$.

To prove the second part of the proposition, we make two more observations. First, if $\rho > 0$ satisfies (3.4) then we have proved that (3.6) holds true. The upper estimate of this inequality implies that for all $y \in T \cap \rho S_{L_2}$,

$$\sum_{i \in I} \langle \varphi_i, y\rangle^2 \ge \sum_{i=1}^{n} \langle \varphi_i, y\rangle^2 - \sum_{j=1}^{k} \langle X_j, y\rangle^2 \ge \rho^2 \left(1 - \frac{4k}{3n}\right).$$

Since T is star shaped, we conclude as before that this inequality holds for all $y \in T$ for which $\|y\|_{L_2} \geq \rho$. If $k < 3n/4$ then we have proved that if $y \in T$ and $\langle \varphi_i, y \rangle = 0$ for all $i \in I$ (i.e. $y \in T \cap (\ker \Gamma)^{\perp}$), then $\|y\|_{L_2} \leq \rho$. But $(\ker \Gamma)^{\perp} = \operatorname{span}\{(\varphi_i)_{i \notin I}\}$ and we conclude that if $\rho > 0$ satisfies (3.4) with $k < 3n/4$ and $T = B_{L_p} \cap E$ then for any $(a_i)_{i=1}^n \in \mathbb{C}^n$,

$$\Big\| \sum_{i \in I} a_i \varphi_i \Big\|_{L_2} \leq \rho \Big\| \sum_{i \in I} a_i \varphi_i \Big\|_{L_p}$$

and

$$\Big\| \sum_{i \notin I} a_i \varphi_i \Big\|_{L_2} \leq \rho \Big\| \sum_{i \notin I} a_i \varphi_i \Big\|_{L_p}.$$

Secondly, it is not difficult to prove with a combinatorial argument (see Lemma 1.3 in [6]) that if $k = [\lambda n]$ with $\lambda = \log 2 < 3/4$ then with probability greater than $3/4$,

$$n/2 - c\sqrt{n} \leq |I| = n - |\{X_1, \ldots, X_k\}| \leq n/2 + c\sqrt{n}, \tag{3.7}$$

for some absolute constant $c > 0$. Choosing $\delta = 1/2$, we get that with a positive probability, both inequalities (3.4) and (3.7) are satisfied. This concludes the proof of the second point of the Proposition 1. ■

Proof of Theorem 3. For any $p \in (1, 2)$, Hölder inequality states that for $\theta = (2-p)/p$,

$$\|f\|_{L_p} \leq \|f\|_{L_1}^{\theta} \|f\|_{L_2}^{1-\theta}.$$

Let $\mu = L\sqrt{n/k}\sqrt{\log k}$ and choose $p = 1 + 1/\log \mu$. Using Proposition 1, there is a subset I of cardinality greater than $n - k$ for which

$$\Big\| \sum_{i \in I} a_i \varphi_i \Big\|_{L_2} \leq C_p\, \mu \Big\| \sum_{i \in I} a_i \varphi_i \Big\|_{L_p}$$

where $C_p = C/(p-1)^{5/2}$. By the choice of p and Hölder inequality,

$$\Big\| \sum_{i \in I} a_i \varphi_i \Big\|_{L_2} \leq \mu\, C_p^{p/(2-p)} \mu^{2(p-1)/(2-p)} \Big\| \sum_{i \in I} a_i \varphi_i \Big\|_{L_1} \leq C\, \mu\, h(\mu) \Big\| \sum_{i \in I} a_i \varphi_i \Big\|_{L_1}$$

where $h(\mu) = (\log \mu)^{5/2}$ and C is an absolute constant.

The same argument works for the second part of the Theorem 3. ■

Appendix

As we mentioned in the Introduction, from a result of J. Bourgain Theorem 3 is "almost" optimal. We would like to thank L. Rodriguez-Piazza for showing us the details of the proof of this optimality in the case of the Walsh system. We consider the Walsh system on $L_2[0,1]$ and take the n first functions with $n = 2^N$. In that case, it is well known that $(\varphi_1, \dots, \varphi_n)$ form a commutative multiplicative group that we denote by $(G, \cdot)$. The main result is based on constructing in any subset of a group a translate of a subgroup of big cardinality.

Lemma 4. *Let $(G, \cdot)$ be the multiplicative group generated by the 2^N first Walsh functions. For any $c \in (0,1)$ and for any subset $\Lambda \subset G$ with $|\Lambda| \geq c\,2^N$, we can find $b \in G$ and a subgroup Γ of G such that, whenever* $\log(1/c) \geq 1/\,2^{N/2}$,

$$\begin{cases} 1.\quad |\Gamma| = 2^p \geq N\log 2 \big/ \left(3\log(1/c)\right) \\ 2.\quad b\cdot\Gamma \subset \Lambda. \end{cases}$$

Assuming this result, we are able to prove the almost optimality of Theorem 3 in the case of the Walsh system $(\varphi_1, \dots, \varphi_n)$ with $n = 2^N$. Let I be a subset of cardinality $n - k$ with $k \geq \sqrt{n}$. Taking $c = 1 - k/n$, we have $\log(1/c) \geq k/n \geq 1/\sqrt{n}$ and Lemma 4 states that there exists $b \in (\phi_i)_{i=1}^n$ and a subgroup Γ of G such that $b \cdot \Gamma \subset (\phi_i)_{i\in I}$ and $|\Gamma| \geq (n\log n)/(20k)$. However, on any subgroup Γ of G, the L_1 norm and the L_2 norm can not be compared with a better estimate than $\sqrt{|\Gamma|}$ since

$$\Big\| \sum_{\gamma\in\Gamma} \gamma \Big\|_{L_2} = \sqrt{|\Gamma|}\, \Big\| \sum_{\gamma\in\Gamma} \gamma \Big\|_{L_1}.$$

Therefore, inequality (3.1) can not be satisfied with a better constant than $\sqrt{n\log n/20k}$. In particular when k is proportional to n, the factor $\sqrt{\log n}$ is necessary in (3.1).

Proof of Lemma 4. We will prove the result in the case of the abelian additive group $G = (\{0,1\}^N, +)$. Note that this lands to a minor change of notation. Let $n = 2^N$ be the cardinality of this group.

Let $\gamma_1 \in G \setminus \{0\}$ be such that $|(\gamma_1 + \Lambda) \cap \Lambda|$ is maximal and define $\Lambda_1 = (\gamma_1 + \Lambda) \cap \Gamma$. Applying convolution, it is evident that

$$\sum_{g\in G} |(g + \Lambda) \cap \Lambda| = |\Lambda|^2$$

and thus

$$|\Lambda_1| \geq \frac{1}{n-1} \sum_{g \in G \setminus \{0\}} |(g+\Lambda) \cap \Lambda| = \frac{|\Lambda|^2 - |\Lambda|}{n-1} \geq c^2 n \frac{1-1/cn}{1-1/n} \geq c^2 n \left(1 - \frac{1}{cn}\right).$$

For notational convenience, we define $\gamma_0 = 0$ and $\Lambda_0 = \Lambda$. We iterate this procedure to construct a family of points γ_j and sets Λ_j such that

$$\begin{cases} (i) & \gamma_j \notin \mathrm{gr}\{\gamma_1, \dots, \gamma_{j-1}\}, \\ (ii) & \Lambda_j = (\gamma_j + \Lambda_{j-1}) \cap \Lambda_{j-1} \text{ has maximal cardinality,} \\ (iii) & |\Lambda_j| \geq c^{2^j} n \left(1 - \frac{2^{j-1}}{n} \left(1/c + \dots + 1/c^{2^{(j-1)}}\right)\right), \end{cases}$$

where $\mathrm{gr}\{\gamma_1, \dots, \gamma_{j-1}\}$ is the group generated by $\{\gamma_1, \dots, \gamma_{j-1}\}$. Once γ_{j-1} and Λ_{j-1} are constructed, we again use the fact that

$$\sum_{g \in G} |(g + \Lambda_{j-1}) \cap \Lambda_{j-1}| = |\Lambda_{j-1}|^2$$

and deduce that

$$\sum_{g \in G \setminus \mathrm{gr}\{\gamma_1, \dots, \gamma_{j-1}\}} |(g + \Lambda_{j-1}) \cap \Lambda_{j-1}| \geq |\Lambda_{j-1}|^2 - 2^{j-1} |\Lambda_{j-1}|.$$

Therefore, there exists $\gamma_j \notin \mathrm{gr}\{\gamma_1, \dots, \gamma_{j-1}\}$ and $\Lambda_j = (\gamma_j + \Lambda_{j-1}) \cap \Lambda_{j-1}$ of maximal cardinality such that

$$|\Lambda_j| \geq |\Lambda_{j-1}| \frac{|\Lambda_{j-1}| - 2^{j-1}}{n - 2^{j-1}}$$

and (iii) follows from a straightforward computation.

This construction can be continued as long as

$$1 - \frac{2^{p-1}}{n} \left(1/c + \ldots + 1/c^{2^{(p-1)}}\right) > 0.$$

This means that if p is an integer such that

$$n > p 2^{(p-1)} / c^{2^{(p-1)}} \tag{3.8}$$

then $\Lambda_p \neq \emptyset$. Let $b \in \Lambda_p$, we shall prove that

$$b + \mathrm{gr}\{\gamma_1, \dots, \gamma_p\} \subset \Lambda.$$

Indeed let $x \in b + \mathrm{gr}\{\gamma_1, \dots, \gamma_p\}$ and define $(\varepsilon_1, \dots, \varepsilon_p) \in \{0,1\}^p$ such that $x = b + \sum \varepsilon_i \gamma_i$. Since $b \in \Lambda_p = \Lambda_{p-1} \cap (\gamma_p + \Lambda_{p-1})$, there exists $\lambda_{p-1} \in \Lambda_{p-1}$ such that $b = \varepsilon_p \gamma_p + \lambda_{p-1}$. If one repeats the same argument for λ_{p-1} instead of b, then at the last step, we get $\lambda_1 \in \Lambda_1 = \Lambda \cap (\gamma_1 + \Lambda)$ and again, there exists $\lambda \in \Lambda$ such that $\lambda_1 = \varepsilon_1 \gamma_1 + \lambda$. Summarizing, we have found $\lambda \in \Lambda$ such that

$$b = \lambda + \sum_{i=1}^{p} \varepsilon_i \gamma_i.$$

Therefore

$$x = b + \sum_{i=1}^{p} \varepsilon_i \gamma_i = \lambda \in \Lambda.$$

To conclude, we have to notice that if p is chosen to be the integer such that $2^p \geq \dfrac{N \log 2}{3 \log(1/c)} > 2^{p-1}$ then

$$p 2^{(p-1)} \left(\frac{1}{c}\right)^{2^{(p-1)}} < \frac{N^2 \log 2}{3 \log(1/c)} \left(\frac{1}{c}\right)^{N \log 2 / 3 \log(1/c)} \leq N^2 2^{N/2 + N/3} \leq 2^N$$

using the fact that $\log(1/c) \geq 1/\, 2^{N/2}$. The inequality (3.8) is satisfied and Lemma 4 holds true. ∎

Remark. We would like to thank the referee for pointing out to us that the bound of Theorem 3 is almost optimal for the trigonometric system as well. Indeed, let $n = 2^N$ and let $\phi_j(t) = \exp(2i\pi j t)$ for $j = 0, \dots, n-1$. Define a one-to-one mapping $F : \{0,1\}^N \to \{0, \dots, n-1\}$ by

$$F(\varepsilon_1, \dots, \varepsilon_N) = \sum_{j=1}^{N} \varepsilon_j 2^{j-1}.$$

Let $I \subset \{0, \dots, n-1\}$ be a subset of cardinality $n - k$. By Lemma 4, the set $\Lambda = F^{-1}(I)$ contains $b \cdot \Gamma$ where $\Gamma = \mathrm{gr}(\gamma_1, \dots, \gamma_p)$. This means that for any set $J \subset \{1, \dots, p\}$, $F(b) + \sum_{j \in J} F(\gamma_j) \in I$. Therefore, the function

$$\psi(t) = \phi_{F(b)} \prod_{j=1}^{p} \left(1 + \phi_{F(\gamma_j)}\right)$$

is contained in $\mathrm{span}(\phi_j, j \in J)$. Obviously,

$$\|\psi\|_{L_2} = \sqrt{|\Gamma|} \|\psi\|_{L_1}.$$

References

[1] Bourgain, J., Pajor, A., Szarek, S. J. and Tomczak-Jaegermann, N.: On the duality problem for entropy numbers of operators. In *Geometric Aspects of Functional Analysis (1987–88)*, 50–63. Lecture Notes in Math. **1376**. Springer, Berlin, 1989.

[2] Candes, E. and Tao, T.: Near-optimal signal recovery from random projections: universal encoding strategies? *IEEE Trans. Inf. Theory* **52** (2006), no. 12, 5406–5425.

[3] Carl, B.: Inequalities of Bernstein-Jackson-type and the degree of compactness of operators in Banach spaces. *Ann. Inst. Fourier* (*Grenoble*) **35** (1985), no. 3, 79–118.

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

[5] Giné, E. and Zinn, J.: Some limit theorems for empirical processes. *Ann. Probab.* **12** (1984), no. 4, 929–998.

[6] Guédon, O., Mendelson, S., Pajor, A. and Tomczak-Jaegermann, N.: Subspaces and orthogonal decompositions generated by bounded orthogonal systems. *Positivity* **11** (2007), no. 2, 269–283.

[7] Guédon, O. and Rudelson, M.: L_p-moments of random vectors via majorizing measures. *Adv. Math.* **208** (2007), no. 2, 798–823.

[8] Lindenstrauss, J. and Tzafriri, L.: *Classical Banach spaces. II. Function spaces.* Ergebnisse der Mathematik und ihrer Grenzgebiete [Results in Mathematics and Related Areas] **97**. Springer-Verlag, Berlin-New York, 1979.

[9] Maurey, B.: Type, cotype and K-convexity. In *Handbook of the geometry of Banach spaces, Vol. 2*, 1299–1332. North-Holland, Amsterdam, 2003.

[10] Maurey, B. and Pisier, G.: Séries de variables aléatoires vectorielles indépendantes et propriétés géométriques des espaces de Banach. *Studia Math.* **58** (1976), no. 1, 45–90.

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

[12] Pisier, G.: *The volume of convex bodies and Banach space geometry.* Cambridge Tracts in Mathematics **94**. Cambridge University Press, Cambridge, 1989.

[13] Rodriguez-Piazza, L.: Personal communication.

[14] Rudelson, M.: Random vectors in the isotropic position. *MSRI preprint* (1996).

[15] Rudelson, M.: Random vectors in the isotropic position. *J. Funct. Anal.* **164** (1999), no. 1, 60–72.

[16] Rudelson, M. and Vershynin, R.: On sparse reconstruction from Fourier and Gaussian measurements. *Comm. Pure Appl. Math.* **61** (2008), no. 8, 1025–1045.

[17] SUDAKOV, V. N.: Gaussian random processes and measures of solide angles in Hilbert spaces. *Soviet. Math. Dokl.* **12** (1971), 412–415.

[18] TALAGRAND, M.: Sections of smooth convex bodies via majorizing measures. *Acta Math.* **175** (1995), no. 2, 273–300.

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

[20] TALAGRAND, M.: Selecting a proportion of characters. *Israel J. Math.* **108** (1998), 173–191.

Recibido: 7 de noviembre de 2007

Olivier Guédon
Université Pierre et Marie Curie, Paris 6
Institut de Mathématiques de Jussieu
4 place Jussieu, 75005 Paris, France
guedon@math.jussieu.fr

Shahar Mendelson
CMA, The Australian National University
Canberra, ACT 0200, Australia
and
Department of Mathematics, Technion
Haifa, 32000, Israel
shahar.mendelson@anu.edu.au

Alain Pajor
Université Paris-Est
Équipe d'Analyse et Mathématiques Appliquées
5, boulevard Descartes, Champs sur Marne
77454 Marne-la-Vallée, Cedex 2, France
Alain.Pajor@univ-mlv.fr

Nicole Tomczak-Jaegermann
Dept. of Math. and Stat. Sciences
University of Alberta
Edmonton, Alberta, Canada, T6G 2G1
nicole.tomczak@ualberta.ca

The second author is partially supported by an Australian Research Council Discovery grant and an Israel Sciences foundation grant 666/06. The fourth author holds the Canada Research Chair in Geometric Analysis.