A remark on the diameter of random sections of convex bodies

Shahar Mendelson\textsuperscript{1}

December 13, 2013

Abstract

We obtain a new upper estimate on the Euclidean diameter of the intersection of the kernel of a random matrix with iid rows with a given convex body. The proof is based on a small-ball argument rather than on concentration and thus the estimate holds for relatively general matrix ensembles.

1 Introduction

In this note we revisit the following problem.

Let $\mu$ be an isotropic measure on $\mathbb{R}^n$, and by ‘isotropic’ we mean a symmetric measure that satisfies
\[
\int_{\mathbb{R}^n} \langle x, t \rangle^2 d\mu(t) = \|x\|_2^2 \quad \text{for every } x \in \mathbb{R}^n.
\]

Given a random vector $X$ distributed according to $\mu$ and for $X_1, \ldots, X_k$ that are independent copies of $X$, let $\Gamma$ be the random matrix $k^{-1/2} \sum_{i=1}^k \langle X_i, \cdot \rangle e_i$.

Question 1.1 If $T \subset \mathbb{R}^n$ is a convex body (that is, a convex, centrally-symmetric set with a non-empty interior), what is the typical diameter of $T \cap \ker(\Gamma)$?

The origin of this problem was the study of the geometry of convex bodies, and in particular, Milman’s low-$M^*$ estimate [10] and subsequent estimates on the Gelfand widths of convex bodies, due to Pajor and Tomczak-Jaegermann [11, 12].

\textsuperscript{1}Department of Mathematics, Technion, I.I.T, Haifa, 32000, Israel. Email: shahar@tx.technion.ac.il. Partially supported by the Mathematical Sciences Institute – The Australian National University and by the Israel Science Foundation grant 900/10.
The focus of the original question had been the existence of a section of \( T \) of codimension \( k \) and of a small Euclidean diameter, and was established by estimating \( \mathbb{E} \text{diam}(T \cap E) \) from above, relative to the uniform measure on the Grassmann manifold \( G_{n-k,n} \).

In recent years, more emphasis has been put on other choices of measures on the Grassmann manifold, for example, using the distribution generated by kernels of matrices selected from some random ensemble – like \( \Gamma = k^{-1/2} \sum_{i=1}^k \langle X_i, \cdot \rangle e_i \) defined above.

The standard way of estimating \( \mathbb{E} \text{diam}(T \cap \ker(\Gamma)) \) for such matrix ensembles is based on the quadratic empirical processes indexed by linear forms associated with \( T \).

It is straightforward to show (see, for example, the discussion in [7]) that given \( r > 0 \), if

\[
\sup_{x \in T \cap rS^{n-1}} \left| \frac{1}{k} \sum_{i=1}^k \langle X_i, x \rangle^2 - \mathbb{E} \langle X, x \rangle^2 \right| \leq \frac{r^2}{2},
\]  

(1.1)

one has

\[
\frac{1}{2} \|x\|_{\ell_2^k}^2 \leq \|\Gamma x\|_{\ell_2^k}^2 \leq \frac{3}{2} \|x\|_{\ell_2^k}^2
\]

for every \( x \in T \) of \( \ell_2^k \) norm larger than \( r \). Hence, on the event given by (1.1), \( \text{diam}(T \cap \ker(\Gamma)) \leq r \).

Setting \( r_0(k, \delta) \) to be the smallest for which

\[
\text{Pr} \left( \sup_{x \in T \cap rS^{n-1}} \left| \frac{1}{k} \sum_{i=1}^k \langle X_i, x \rangle^2 - \mathbb{E} \langle X, x \rangle^2 \right| \leq \frac{r^2}{2} \right) \geq 1 - \delta,
\]

it follows that with probability at least \( 1 - \delta \),

\[
\text{diam}(T \cap \ker(\Gamma)) \leq r_0,
\]

and a similar argument may be used to control \( \mathbb{E} \text{diam}(T \cap \ker(\Gamma)) \).

Unfortunately, estimating the quadratic empirical process is a difficult task. In fact, one has a satisfactory estimate that holds for every convex body \( T \subset \mathbb{R}^n \) only for measures that are subgaussian or unconditional log-concave.

**Theorem 1.2** [6] There exist absolute constants \( c_1, c_2 \) and \( c_3 \) for which the following holds. Let \( \mu \) be an isotropic, \( L \)-subgaussian measure (and, in particular, for every \( x \in \mathbb{R}^n \), \( \|X, x\|_{\psi_2(\mu)} \leq L \|x\|_{\ell_2^k} \)).
Let \( T \subset \mathbb{R}^n \) and set \( d_T = \sup_{t \in T} \|t\|_2 \). For \( u \geq c_1 \), with probability at least

\[
1 - 2 \exp(-c_2 u^2 (\mathbb{E}\|G\|_{T^\circ})^2),
\]

\[
\sup_{x \in T} \left| \sum_{i=1}^k (\langle X_i, x \rangle^2 - \mathbb{E}|\langle X, x \rangle|^2) \right| \leq c_3 L^2 u^2 \left( (\mathbb{E}\|G\|_{T^\circ})^2 + u \sqrt{k d_T} \mathbb{E}\|G\|_{T^\circ} \right),
\]

where \( G = (g_1, ..., g_n) \) is the standard gaussian vector in \( \mathbb{R}^n \) and \( \mathbb{E}\|G\|_{T^\circ} = \mathbb{E}\sup_{t \in T}|\langle G, t \rangle| \).

A version of Theorem 1.2 has been established in [7] when \( T \subset S^{n-1} \) and with a weaker probability estimate.

Theorem 1.2 follows from a general bound on the quadratic empirical process that is based on a global complexity parameter of the indexing set [8], and that will not be defined here. Thanks to Talagrand’s Majorizing Measures Theorem (see the book [15] for a detailed survey on this topic), this complexity parameter is upper bounded by \( \sim \mathbb{E}\|Y\|_{T^\circ} \) in the subgaussian case, thus leading to Theorem 1.2. However, in other cases, controlling it is nontrivial.

One other case in which the global complexity may be upper bounded using a mean-width of \( T \), is when \( X \) is isotropic, unconditional and log-concave. Using the Bobkov-Nazarov Theorem [1], \( X \) is dominated by \( Y = (y_1, ..., y_n) \), a vector with independent, standard, exponential coordinates. One may show [8] that with high probability,

\[
\sup_{x \in T} \left| \sum_{i=1}^k (\langle X_i, x \rangle^2 - \mathbb{E}|\langle X, x \rangle|^2) \right| \lesssim (\mathbb{E}\|Y\|_{T^\circ})^2 + \sqrt{k d_T} \mathbb{E}\|Y\|_{T^\circ}. \tag{1.2}
\]

The proof of (1.2) is based on two additional observations. First, that when \( X \) is isotropic, unconditional and log-concave, the global complexity parameter of \( T \) may be bounded using a mixture of Talagrand’s \( \gamma_\alpha \) functionals, and second, that this mixture is equivalent to \( \mathbb{E}\|Y\|_{T^\circ} \) [14].

Additional bounds on the quadratic process are known for more general measures, but only for very specific choices of sets \( T \). The most important example is when \( T \) is the Euclidean ball, and the quadratic empirical process may be used to obtain a Bai-Yin type estimate on the largest and smallest singular values of \( \Gamma \) [8, 9].

At this point, it should be noted that (1.1) is a much stronger statement than what is actually needed to bound the diameter of \( T \cap \ker(\Gamma) \). Clearly, any sort of a positive lower bound on

\[
\inf_{x \in T \cap S^{n-1}} \|\Gamma x\|_{\ell_2^k} \tag{1.3}
\]
would suffice – rather than the ‘almost isometric’, two-sided bound that follows from bounds on the quadratic process.

Here, we will show that (1.3) holds for rather general matrix ensembles.

**Theorem 1.3** Let $X$ be an isotropic vector on $\mathbb{R}^n$ and assume that linear forms satisfy the following small-ball condition: that there is some $\lambda > 0$ for which

$$\Pr(|\langle x, X \rangle| \geq \lambda \|x\|_{\ell^2}) \geq 99/100$$

for every $x \in \mathbb{R}^n$.

Then, there exist a constant $c$ that depends only on $\lambda$, for which, with probability at least $3/4$,

$$\text{diam}(T \cap \ker(\Gamma)) \leq \frac{c}{\sqrt{k}} \cdot \max \left\{ \mathbb{E}\|G\|_{T^o}, \mathbb{E}\|k^{-1/2} \sum_{i=1}^k X_i\|_{T^o} \right\}.$$ 

Theorem 1.3 can be improved and extended in various ways.

First of all, the ‘correct’ upper estimate on the diameter should be based on a fixed point condition defined using the norms $\| \| (T^o \cap B^2)^o$ rather than the norm $\| \|_{T^o}$. Also, the constant probability estimate of $3/4$ may be improved significantly to $1 - 2 \exp(-ck)$ with a slightly more involved proof (see [4] for a similar argument). We will formulate, without proof, a more general version of Theorem 1.3 at the end of the note.

**Examples.**

1. If $X$ is an isotropic $L$-subgaussian vector, it is standard to verify that $k^{-1/2} \sum_{i=1}^k X_i$ is isotropic and $cL$-subgaussian for a suitable absolute constant $c$. Therefore,

$$\mathbb{E}\|k^{-1/2} \sum_{i=1}^k X_i\|_{T^o} \leq c_1 L \mathbb{E}\|G\|_{T^o},$$

and by Theorem 1.3 with probability at least $3/4$,

$$\text{diam}(T \cap \ker(\Gamma)) \leq c_1(\lambda, L) \frac{\mathbb{E}\|G\|_{T^o}}{\sqrt{k}}.$$ 

This coincides with the estimate from [7] (up to the ‘localization’ mentioned above) and with the classical result of [10] when $X$ is the standard gaussian vector.
2. If $X$ is an isotropic, unconditional, log-concave measure then so is $Z = k^{-1/2} \sum_{i=1}^{k} X_i$. By the Bobkov-Nazarov Theorem [1], both $Z$ and $G$ are strongly dominated by $Y$, the random vector with independent, standard exponential coordinates. Therefore, by Theorem 1.3, with probability at least $3/4$,

$$\text{diam}(T \cap \ker(\Gamma)) \leq c_2(\lambda) \frac{\mathbb{E}\|Y\|_{T^o}}{\sqrt{k}}.$$

3. Theorem 1.3 leads to a ‘heavy tails’ result in some cases. Since $X$ is symmetric, $\sum_{i=1}^{k} X_i$ has the same distribution as $\sum_{i=1}^{k} \varepsilon_i X_i$, where $(\varepsilon_i)_{i=1}^{k}$ are independent, symmetric $\{-1,1\}$-valued random variables that are independent of $(X_i)_{i=1}^{k}$. If $T^o$ has a Rademacher type 2 constant $R_2(T^o)$, then

$$\mathbb{E}\|k^{-1/2} \sum_{i=1}^{k} X_i\|_{T^o} = \mathbb{E}\|k^{-1/2} \sum_{i=1}^{k} \varepsilon_i X_i\|_{T^o} \leq R_2(T^o)(\mathbb{E}\|X\|_{T^o}^2)^{1/2},$$

and with probability at least $3/4$,

$$\text{diam}(T \cap \ker(\Gamma)) \leq \frac{c_3(\lambda)}{\sqrt{k}} \cdot \max \left\{ \mathbb{E}\|G\|_{T^o}, R_2(T^o)(\mathbb{E}\|X\|_{T^o}^2)^{1/2} \right\}.$$

For example, if $T = B_1^n$ and $X \in \beta B_\infty^n$ almost surely, then $\|X\|_{T^o} = \|X\|_{\ell_\infty} \leq \beta$, $R_2(\ell_\infty) \leq \sqrt{\log n}$ and $\mathbb{E}\|G\|_{\ell_\infty} \lesssim \sqrt{\log n}$. Therefore,

$$\text{diam}(B_1^n \cap \ker(\Gamma)) \leq c_3(\lambda) \beta \sqrt{\frac{\log n}{k}}.$$ 

2 Proof of Theorem 1.3

Lemma 2.1 Let $\zeta$ be a random variable that satisfies

$$\Pr(|\zeta| \geq \lambda \|\zeta\|_{L_2}) \geq 1 - \varepsilon$$

for constants $0 < \varepsilon < 1/12$ and $\lambda > 0$.

If $\zeta_1, \ldots, \zeta_k$ are independent copies of $\zeta$, then with probability at least $1 - 2^{-6k}$ there is a subset $J \subset \{1, \ldots, k\}$ of cardinality at least $(1 - 6\varepsilon)k$, and for every $j \in J$,

$$|\zeta_j| \geq \lambda \|\zeta\|_{L_2}.$$
Proof. It suffices to show that no more than $6\varepsilon k$ of the $|\zeta_i|$'s are smaller than $\lambda \|\zeta\|_{L^2}$. By a binomial estimate, if $6\varepsilon k \leq k/2$,

$$
Pr(\exists J \subset \{1, \ldots, k\}, |J| = 6\varepsilon k, |\zeta_j| \leq \lambda \|\zeta\|_{L^2} \text{ if } j \in J) \\
\leq \binom{k}{6\varepsilon k} Pr(6\varepsilon k (|\zeta| \leq \lambda \|\zeta\|_{L^2}) \leq \left(\frac{\varepsilon}{6\varepsilon}\right)^{6\varepsilon k} \cdot \varepsilon^{6\varepsilon k} \leq 2^{-6\varepsilon k}.
$$

\[\square\]

Let $\{\zeta_i : 1 \leq i \leq N\}$ be a collection of random variables, and for every $i$ let $Z_i \in \mathbb{R}^k$ be a random vector with independent coordinates, distributed according to the random variable $\zeta_i$. Denote by $Z_i(j)$ the $j$-th coordinate of $Z_i$.

**Corollary 2.2** If each $\zeta_i$ satisfies the small-ball condition (2.1) and $N \leq 2^{3\varepsilon k}$, then with probability at least $1 - 2^{-3\varepsilon k}$, for every $1 \leq i \leq N$ there is a subset $J_i \subset \{1, \ldots, k\}$ of cardinality at least $(1 - 6\varepsilon)k$, and

$$
|Z_i(j)| \geq \lambda \|\zeta_i\|_{L^2} \text{ for every } j \in J_i.
$$

**Proof of Theorem 1.3.** Let $\varepsilon = 1/600$ and observe that by the small ball assumption and since $X$ is isotropic,

$$
Pr(\langle X, x \rangle \geq \lambda \|x\|_{\ell^2}) \geq 99/100 \geq (1 - \varepsilon)
$$

for every $x \in \mathbb{R}^n$.

Fix $r > 0$ to be named later and set $T_r = T \cap rS^{n-1}$. Let

$$
\rho = c \frac{\mathbb{E}\|G\|_{T^\circ r}}{\varepsilon k}
$$

for a suitable absolute constant $c$ and set $V_r \subset T_r$ to be a maximal $\rho$-separated subset of $T_r$ with respect to the $\ell^2_2$ norm. Sudakov’s inequality (see, e.g. [13, 5]) shows that for the right choice of $c$, $|V_r| \leq 2^{3\varepsilon k}$.

Let

$$
\zeta^i = \langle X, v_i \rangle, \; v_i \in V_r, \; 1 \leq i \leq 2^{3\varepsilon k},
$$

and set $Z_i = (\zeta^i_{j=1})^k$, a vector whose coordinates are independent copies of $\zeta^i$.

Applying Corollary 2.2 to the set $\{Z_i : 1 \leq i \leq 2^{3\varepsilon k}\}$, it follows that with probability at least $1 - 2^{-3\varepsilon k}$, for every $v \in V_r$ there is a subset $J_v \subset \{1, \ldots, k\}$, $|J_v| \geq (1 - 6\varepsilon)k = 99k/100$, and for every $j \in J_v$,

$$
\langle X_j, v \rangle \geq \lambda \|v\|_{\ell^2_v} = \lambda r,
$$

(2.2)
and the last equality holds because $V_r \subset rS^{n-1}$.

For every $x \in T_r$, let $\pi(x)$ be the nearest point to $x$ in $V_r$ with respect to the $\ell_2^n$ norm. Therefore,

$$E|\langle X, x - \pi(x) \rangle| \leq \|x - \pi(x)\|_{L^2(\mu)} = \|x - \pi(x)\|_{\ell_2^n} \leq \rho.$$ 

By the Giné-Zinn symmetrization inequality [3], the contraction inequality for Bernoulli processes (see, e.g., [5]), and since $x - \pi(x) \in 2T \cap \rho B_2^n$ for every $x \in T_r$,

$$E \sup_{x \in T_r} \frac{1}{k} \sum_{i=1}^{k} |\langle X_i, x - \pi(x) \rangle| \leq \rho + E \sup_{x \in T_r} \left| \frac{1}{k} \sum_{i=1}^{k} |\langle X_i, x - \pi(x) \rangle| - E|\langle X, x - \pi(x) \rangle| \right| \leq \rho.$$ 

$$\leq \rho + \frac{2}{k} E \sup_{x \in T_r} \left| \sum_{i=1}^{k} \varepsilon_i |\langle X_i, x - \pi(x) \rangle| \right| \leq \rho + \frac{2}{k} E \sup_{x \in T_r} \sum_{i=1}^{k} \varepsilon_i |\langle X_i, x - \pi(x) \rangle| \leq \rho + \frac{2}{k} \sum_{i=1}^{k} \|X_i\|_{(2T \cap \rho B_2^n)^c}.$$ 

$$\leq \rho + \frac{2}{k} \sum_{i=1}^{k} \|X_i\|_{(2T \cap \rho B_2^n)^c} \leq \rho + \frac{4}{k} \sum_{i=1}^{k} \|X_i\|_{(T \cap \rho B_2^n)^c}.$$ 

Hence, by the choice of $\rho$ and the trivial inclusion $T \cap \rho B_2^n \subset T$,

$$E \sup_{x \in T_r} \frac{1}{k} \sum_{i=1}^{k} |\langle X_i, x - \pi(x) \rangle| \leq c \frac{E\|G\|_{T^c}}{\sqrt{k}} + \frac{4}{k} \sum_{i=1}^{k} \|X_i\|_{T^c}.$$ (2.3)

Set $A = E \sup_{x \in T_r} k^{-1} \sum_{i=1}^{k} |\langle X_i, x - \pi(x) \rangle|$ and let

$$W_r = \left\{ (\langle X_i, x - \pi(x) \rangle)_{i=1}^{k} : x \in T_r \right\}.$$ 

Note that for $0 < \delta < 1$, with probability at least $1 - \delta$,

$$W_r \subset k(A/\delta)B_1^k.$$ 

On that event, if $(w_i)_{i=1}^{k} \in W_r$ and $(w_i^*)_{i=1}^{k}$ is a non-increasing rearrangement of $(|w_i|)_{i=1}^{k}$,

$$w_{k/100}^* \leq \frac{\|w\|_{l_1^k}}{k/100} \leq \frac{100}{\delta} A.$$ 

Thus, for every $x \in T_r$ there is a subset $J_x' \subset \{1, \ldots, k\}$ of cardinality at least $99k/100$, and for every $j \in J_x'$,
\[ |\langle X_i, x - \pi(x) \rangle| \leq \frac{100}{\delta} A. \quad (2.4) \]

Fix \( X_1, \ldots, X_k \) in the intersection of the two events defined in (2.2) and (2.4). For every \( x \in T_r \), set \( I_x = J_x \cap J_{\pi(x)} \). Observe that \(|I_x| \geq 98k/100 \) and that for every \( i \in I_x \),

\[ |\langle X_i, x \rangle| \geq |\langle X_i, \pi(x) \rangle| - |\langle X_i, x - \pi(x) \rangle| \geq \lambda r - \frac{100}{\delta} A \]

\[ \geq \lambda r - \frac{c_1}{\delta \sqrt{k}} \left( \mathbb{E} \|G\|_{T^0} + \mathbb{E} \left\| \frac{1}{\sqrt{k}} \sum_{i=1}^{k} X_i \right\|_{T^0} \right). \]

Therefore, if

\[ r \geq \frac{c_2(\lambda, \delta)}{\sqrt{k}} \cdot \max \left\{ \mathbb{E} \|G\|_{T^0}, \mathbb{E} \left\| \frac{1}{\sqrt{k}} \sum_{i=1}^{k} X_i \right\|_{T^0} \right\}, \]

then with probability at least \( 1 - \delta - 2^{-3c_1k} = 1 - \delta - 2^{-k/200} \), for each \( x \in T_r \), \(|\langle X_i, x \rangle| \geq (\lambda/2) \|x\|_{\ell_2^k}\) on at least \( 98k/100 \) coordinates; Thus,

\[ \inf_{x \in T_r} \|\Gamma x\|_{\ell_2^k} \geq \lambda \|x\|_{\ell_2^k}. \quad (2.5) \]

Finally, using the convexity of \( T \) and since the condition in (2.5) is positive-homogeneous, (2.5) holds for any \( x \in T \) with \( \|x\|_{\ell_2^k} \geq r \), as claimed.

\section{Concluding Comments}

The proof of Theorem 1.3 has two components. The first is based on a small-ball estimate for linear functionals and does not require additional information on their tails. Thus, this part holds even for heavy-tailed ensembles.

The more restrictive condition is on the random vector \( k^{-1/2} \sum_{i=1}^{k} X_i \). Still, it is far easier to handle the norm \( \| \sum_{i=1}^{k} X_i \|_{T^0} \) than the supremum of the quadratic empirical process indexed by \( T \).

The estimate in Theorem 1.3 can be improved using what is, by now, a standard argument. First, observe that all the inequalities leading to (2.3) hold in probability and not just in expectation (see, for example, [16, 2]). Keeping the ‘localization’ level \( r \), one can define two fixed points:

\[ \rho_k(\delta, Q_1) = \inf \left\{ \rho : P_{\rho} \left( \left\| k^{-1/2} \sum_{i=1}^{k} X_i \|_{(T \cap \rho B_2^k)^0} \geq Q_1 \rho \sqrt{k} \right) \leq \delta \right\}, \]

8
and
\[ r_k(Q_2) = \inf \{ r : \mathbb{E}\|G\|(T \cap rS^{n-1})^c \leq Q_2 r \sqrt{k} \} . \]

It is straightforward to verify that there are constants \( Q_1 \) and \( Q_2 \) that depend only on \( \lambda \), for which, with probability at least \( 1 - \delta - 2^{-k/200} \), if
\[ \|x\|_{\ell_2^n} \gtrsim \max\{\rho_k(\delta, Q_1), r_k(Q_2)\}, \]
then
\[ \|\Gamma x\|_{\ell_2^k} \gtrsim \lambda \|x\|_{\ell_2^n}. \]

Thus, on the same event,
\[ \text{diam}(T \cap \ker(\Gamma)) \lesssim \max\{\rho_k(\delta, Q_1), r_k(Q_2)\}. \]

Finally, it is possible to use a slightly more involved, empirical processes based method, that leads to an exponential probability estimate of \( 1 - 2 \exp(-ck) \) in Theorem 1.3. A result of a similar flavour, concerning the smallest singular value of a random matrix with iid rows may be found in [4].

Since the goal in this note was to present the idea of using a simple small-ball argument, rather than pursuing an optimal result, we have opted to present this proof.

References

[1] S.G. Bobkov, F.L. Nazarov, On convex bodies and log-concave probability measures with unconditional basis, Geometric Aspects of Functional Analysis, Lecture Notes in Mathematics 1807, 53-69, 2003.

[2] R. M. Dudley, *Uniform Central Limit Theorems*, Cambridge Studies in Advanced Mathematics 63, Cambridge University Press, 1999.

[3] E. Giné and J. Zinn, Some limit theorems for empirical processes, Annals of Probability 12(4), 929–989, 1984.

[4] V. Koltchinskii, S. Mendelson, bounding the smallest singular value of a random matrix without concentration, preprint.

[5] M. Ledoux, M. Talagrand, *Probability in Banach spaces. Isoperimetry and processes*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3), vol. 23. Springer-Verlag, Berlin, 1991.
[6] S. Mendelson, On the geometry of subgaussian coordinate projections, preprint.

[7] S. Mendelson, A. Pajor and N. Tomczak-Jaegermann, Reconstruction and subgaussian operators, Geometric and Functional Analysis, 17, 1248–1282, 2007.

[8] Mendelson, S. and Paouris, G.: On generic chaining and the smallest singular values of random matrices with heavy tails, Journal of Functional Analysis, 262(9), 3775–3811 (2012)

[9] S. Mendelson, G. Paouris, On the singular values of random matrices, Journal of the European Mathematical Society, to appear.

[10] V. Milman, Random subspaces of proportional dimension of finite dimensional normed spaces: approach through the isoperimetric inequality. Lecture Notes in Mathematics, 1166, 106–115, 1985.

[11] A. Pajor, N. Tomczak-Jaegermann, Subspaces of small codimension of finite-dimensional Banach spaces, Proceedings of the American Mathematical Society 97(4), 637-642, 1986.

[12] A. Pajor, N. Tomczak-Jaegermann, Nombres de Gelfand et sections euclidiennes de grande dimension. (French) [Gelfand numbers and high-dimensional Euclidean sections] Séminaire d'Analyse Fonctionelle 1984/1985, Publ. Math. Univ. Paris VII, 26, Univ. Paris VII, Paris, 37–47, 1986.

[13] G. Pisier, The volume of convex bodies and Banach space geometry, Cambridge University Press, 1989.

[14] M. Talagrand, The supremum of some canonical processes, American Journal of Mathematics 116, 283–325, 1994.

[15] M. Talagrand, The generic chaining, Springer, 2005.

[16] A.W. Van der Vaart, J.A. Wellner, Weak convergence and empirical processes, Springer Verlag, 1996.