Gaussian averages of interpolated bodies and applications to approximate reconstruction

Y. Gordon* A. E. Litvak S. Mendelson† A. Pajor

Abstract

We prove sharp bounds for the expectation of the supremum of the Gaussian process indexed by the intersection of B_p^n with ρB_q^n for $1 \leq p, q \leq \infty$ and $\rho > 0$, and by the intersection of $B_{p\infty}^n$ with ρB_2^n for $0 < p \leq 1$ and $\rho > 0$. We present an application of this result to a statistical problem known as the approximate reconstruction problem.

Keywords: Approximate reconstruction, Gaussian process, Gaussian averages, Gelfand widths, diameter of a section, interpolation, learning theory, "Low M^*-estimate".

1 Introduction

The motivation for the questions we study here came from problems in convex geometry and in nonparametric statistics (learning theory).

To formulate the main question we tackle, let $e_1, ..., e_n$ be the standard basis in $\mathbb{R}^n$ endowed with the canonical Euclidean structure, and set $\{g_i\}_1^n$ to be independent $\mathcal{N}(0,1)$ Gaussian random variables. Let $T \subset \mathbb{R}^n$ and consider the sets $T_\rho = T \cap \rho B_2^n$, where B_2^n is the unit Euclidean ball. Our aim is to bound $\ell_*(T_\rho) := \mathbb{E} \sup_{t \in T_\rho} \langle \sum_{i=1}^n g_i e_i, t \rangle$ as a function of ρ. Obtaining precise estimates for a general set T is virtually impossible, but as we show here, in some cases one can establish sharp bounds when T is B_p^n, the unit ball in ℓ_p^n, for $1 \leq p \leq \infty$ or $B_{p\infty}^n$, the unit ball in a weak-ℓ_p^n, for $0 < p \leq 1$. In fact, one can even obtain sharp bounds when the B_2^n is replaced by a B_q^n, the unit ball in ℓ_q^n.

Our main results are the following two theorems (see Section 5 for more precise formulations).

Theorem A *There exist absolute positive constants c, C, and $c_1 < 1$ for which the following holds. Let $\{g_i\}_{i \leq n}$ be independent $\mathcal{N}(0,1)$ Gaussian variables. Consider* $1 \leq$

*Partially supported by France-Israel exchange fund 2005389, and by the Fund for the Promotion of Research at the Technion

†Supported in part by an Australian Research Council Discovery grant DP0559465

$q_0 < q_1 \le \ln(2n)$*, let* r *be such that* $1/r = 1/q_0 - 1/q_1$*, set* $1 \le t \le c_1^{q_1/r} n^{1/r}$*, and put* $L = B_{p_0}^n \cap t B_{p_1}^n$*, where* $1/p_i + 1/q_i = 1$*,* $i = 0,1$*. Then*

$$c\ t^{r/q_0}\ \sqrt{q_0 + \ln(2n/t^r)} + t\sqrt{q_1} n^{1/q_1} \le \mathbb{E} \sup_{y \in L} \left\langle \sum_{i=1}^n g_i e_i,\, y \right\rangle$$

$$\le C\ t^{r/q_0}\ \sqrt{q_0 + \ln(2n/t^r)} + t\sqrt{q_1} n^{1/q_1}.$$

Theorem B *There are absolute positive constants* c *and* C *for which the following holds. Let* $\{g_i\}_{i \le n}$ *be independent, standard Gaussian variables. Set* $0 < p \le 1$ *and* $\gamma = 1/(1/p - 1/2)$*, let* $n^{-1/\gamma} < \rho < 1$ *and put* $K = B_{p\infty}^n \cap \rho B_2^n$*.*

(i) If $0 < p < 1$ *then*

$$c\rho^{2\frac{1-p}{2-p}}\ \sqrt{\ln\left(2n\rho^\gamma\right)} \le \mathbb{E} \sup_{y \in K} \left\langle \sum_{i=1}^n g_i e_i, y \right\rangle \le \frac{C}{1-p}\ \rho^{2\frac{1-p}{2-p}}\ \sqrt{\ln\left(2n\rho^\gamma\right)}.$$

(ii) If $p = 1$ *then*

$$c\left(\ln\left(2n\rho^2\right)\right)^{3/2} \le \mathbb{E} \sup_{y \in K} \left\langle \sum_{i=1}^n g_i e_i,\, y \right\rangle \le C\left(\ln\left(2n\rho^2\right)\right)^{3/2}.$$

The notion in learning theory that motivated this study is *localization*. Since we do not want to present a detailed discussion concerning learning theory, let us present one concrete problem in which the question we study is essential.

Let $T \subset \mathbb{R}^n$ be a given set, which we assume to be convex and symmetric. A point $t_0 \in T$ is selected, and the goal of the learner is to approximate it with respect to the Euclidean norm (denoted below by $\|\cdot\|_2$). The data one is given to accomplish this task is a set of random linear measurements $(\langle X_i, t_0 \rangle)_{i=1}^k$, where $X_1, ..., X_k$ are independent random variables, distributed according to a probability measure μ on $\mathbb{R}^n$. For every such data set one produces $\hat{t} \in \mathbb{R}^n$ according to some rule, and the hope is to show that with high probability (with respect to the product measure μ^k), $\|t_0 - \hat{t}\|_2$ is small.

The measure μ plays an important role here, and the idea is that it should be as general as possible, specifically, it should not depend on the particular choice of the set T.

This problem, called the *approximate reconstruction problem*, and problems of a similar flavor including the new direction of sparse approximation theory called *compressed sensing* have been studied by various authors in the last few years (see, e.g. [CDS, CT1, CT2, D, DE, DET, RV]). In all these results the main focus was on the case where μ is the standard Gaussian measure on $\mathbb{R}^n$ and T is the set of sparse vectors $\{x \in \mathbb{R}^n \mid |\text{supp } x| \le s\}$ for some s, or $T = B_1^n$, or $T = B_{p\infty}^n$. In [MePT] a more general problem was solved – for an arbitrary convex, centrally symmetric set T and isotropic,

L-subgaussian measures. Recall that μ is isotropic if for every $t \in \mathbb{R}^n$, $\mathbb{E}\langle X, t\rangle^2 = \|t\|_2^2$, and is L-subgaussian if for every $t \in \mathbb{R}^n$,

$$\Pr\left(|\langle X, t\rangle| \geq uL\|t\|_2\right) \leq 2\exp(-u^2)$$

for every $u \geq 1$.

It turns out (see Section 7 for more details) that the key parameter that governs the degree of approximation, $r_k^*(\theta, T)$, is given by

$$r_k^*(\theta, T) := \inf\left\{\rho > 0 \;\middle|\; \rho \geq 2\ell_*(T_\rho)/\theta\sqrt{k}\right\}, \tag{1}$$

where $\ell_*(T_\rho)$ was defined above and $\theta = c/L^2$ for some absolute constant c. More precisely, one can show that if one selects $\hat{t} \in T$ for which $\langle X_i, \hat{t}\rangle = \langle X_i, t_0\rangle$ for every $1 \leq i \leq k$, then with high probability, $\|\hat{t} - t_0\|_2 \leq c_1 r_k^*(\theta, T)$, where c_1 is an absolute constant.

Note that $r_k^*(\theta, T)$ is governed by the quantity we are interested in – the expectation of the supremum of the Gaussian process indexed by an intersection body. We show the details in Section 7.

The geometric applications are related to Dvoretzky type results and estimates on diameters of sections of convex bodies. Recall the following variant of so-called "Low M^*-estimate", which was first proved in [Mi1, Mi2], then improved in [PT1, PT2]. The version we use here is from [Go1]. Given convex centrally-symmetric body T in $\mathbb{R}^n$ and $1 \leq k \leq n$ if

$$k > \left(\frac{\ell_*(T_\rho)}{\omega_k \rho}\right)^2,$$

where $1 - 1/(4\sqrt{k}) < \omega_k := \sqrt{\frac{2}{k}}\Gamma\left(\frac{k+1}{2}\right)/\Gamma\left(\frac{k}{2}\right) < 1$, then a "random" k-codimensional subspace E of $\mathbb{R}^n$ satisfies

$$T \cap E \subset \rho B_2^n.$$

In other words, if we control $\ell_*(T_\rho)$ then we control the diameter of k-codimensional section of T for an appropriate k. Thus our main results, Theorems A and B, have immediate consequences for diameters of sections. We provide precise estimates for some cases in Section 6.

2 Preliminaries and Notation

Let $\|\cdot\|_2$ and $\langle\cdot,\cdot\rangle$ denote a fixed (canonical) Euclidean norm and inner product on $\mathbb{R}^n$. The canonical basis of $\mathbb{R}^n$ is denoted by $e_1, \ldots, e_n$. For $1 \leq p \leq \infty$ set $\|\cdot\|_p$ to be the ℓ_p^n-norm, i.e.

$$\|x\|_p = \left(\sum_i |x_i|^p\right)^{1/p} \quad \text{for} \quad p < \infty, \qquad \|x\|_\infty = \sup_i |x_i|,$$

and let B_p^n be their unit balls.

Given a sequence $\{a_i\}_{i=1}^n$, let $\{a_i^*\}_{i=1}^n$ be the non-increasing rearrangement of $\{|a_i|\}_{i=1}^n$. We will also need the definition of the weak-ℓ_p^n-norm, $\|\cdot\|_{p\infty}$ for $0<p<\infty$, given by

$$\|x\|_{p\infty} = \sup_{1\le k\le n} k^{1/p}\ x_k^*$$

with the unit ball $B_{p\infty}^n = \left\{x \in \mathbb{R}^n \ \mid \ x_k^* \le k^{-1/p} \text{ for every } \ k \le n\right\}$.

The convex hull of a set A is denoted by conv A.

Let $K \subset \mathbb{R}^n$ be a centrally symmetric (with respect to the origin) compact convex set. As usual, the Minkowski functional of K is denoted by $\|\cdot\|_K$ and defined by

$$\|x\|_K = \inf\{\lambda > 0 \ \mid \ x \in \lambda K\}.$$

The polar of K is

$$K^\circ = \{x \mid \langle x,y\rangle \le 1 \ \text{ for every } \ y \in K\}.$$

Note that K is the unit ball of the normed space $X = (\mathbb{R}^n, \|\cdot\|_K)$ and that $X^* = (\mathbb{R}^n, \|\cdot\|_{K^\circ})$. Moreover, given symmetric convex bodies K and L we have

$$(K\cap L)^\circ = \text{conv }(K^\circ \cup L^\circ) \qquad \text{and} \qquad (\text{conv }(K\cup L))^\circ = K^\circ \cap L^\circ.$$

In particular,

$$\left(\text{conv}\left(B_{q_0}^n \bigcup \frac{1}{t} B_{q_1}^n\right)\right)^\circ = B_{p_0}^n \cap t B_{p_1}^n,$$

where $1/p_i + 1/q_i = 1$ and $t > 0$.

Throughout this note we denote by $\{g_i\}$ and $\{g_{i,j}\}$ collections of independent $\mathcal{N}(0,1)$ Gaussian random variables.

Given two functions F and G we write $F \sim G$ if there are absolute positive constants c, C such that $cF \le G \le CF$.

Finally, all absolute constants are positive and denoted by c or C. Their actual values may change from line to line.

3 Norm estimates on Gaussian vectors

In this section we recall some well known results and develop some new ones regarding the expectations of Gaussian variables. We deal with a sequence of n independent $\mathcal{N}(0,1)$ Gaussian random variables, $g_1,\dots,g_n$, and compute expectations of some functionals of the rearranged sequence $g_1^*,\dots,g_n^*$.

The first two lemmas are known and are derived by direct calculations (see, e.g. Example 10 in [GoLSW2] for Lemma 3.2). They show that the expectation of g_k^* behaves quite regularly as a function of k and n, but the behavior is different for "large" and "small" k.

Lemma 3.1 *Let $1 \le k \le n/2$ and set $\{g_i\}_{i=1}^n$ to be independent $\mathcal{N}(0,1)$ Gaussian random variables. Then*

$$\mathbb{E}\, g_k^* \sim \sqrt{\ln \frac{n}{k}}.$$

In particular,

$$\mathbb{E} \sum_{i=1}^{k} g_i^* \sim \; k\sqrt{\ln \frac{n}{k}}.$$

Lemma 3.2 *Let $n/2 \le k \le n$ and set $\{g_i\}_{i=1}^n$ to be independent $\mathcal{N}(0,1)$ Gaussian random variables. Then*

$$\sqrt{\frac{\pi}{2}}\; \frac{n+1-k}{n+1} \le \mathbb{E}\; g_k^* \le \sqrt{2\pi}\; \frac{n+1-k}{n+1}.$$

We will also require the following Lemma, which is a specific application of Example 16 in [GoLSW1].

Lemma 3.3 *Let $1 \le q \le \ln(2n)$ and $1 \le k \le n/2$. Then*

$$\left(\mathbb{E} \sum_{i=1}^{k} (g_i^*)^q \right)^{1/q} \sim k^{1/q} \sqrt{q + \ln \frac{n}{k}}.$$

We now turn to two corollaries of Lemma 3.1 and Lemma 3.3 which will be used below.

Corollary 3.4 *Let $1 \le q \le \ln(2n)$ and $1 \le k \le n$. If $\{g_i\}_{i=1}^n$ are independent $\mathcal{N}(0,1)$ Gaussian random variables then*

$$\mathbb{E} \left(\sum_{i=1}^{k} (g_i^*)^q \right)^{1/q} \sim k^{1/q}\; \sqrt{q + \ln \frac{2n}{k}}.$$

Proof: Without loss of generality, assume that $k \le n/2$. The upper bound follows immediately from Lemma 3.3 and a comparison between the first and the q-th moments.

To obtain the lower bound, note that by Lemma 3.1 for every $m \le k$,

$$\mathbb{E} \left(\sum_{i=1}^{k} (g_i^*)^q \right)^{1/q} \ge \mathbb{E} \left(\sum_{i=1}^{m} (g_i^*)^q \right)^{1/q}$$

$$\ge m^{-1+1/q}\; \mathbb{E} \sum_{i=1}^{m} g_i^* \ge c\; m^{1/q}\; \sqrt{\ln \frac{2n}{m}},$$

where $c > 0$ is an absolute constant. Choosing $m = [1 + k/e^q]$ we obtain the desired result. □

Remark 3.5 *Corollary 3.4 can be used to show that for* $1 \le q \le \ln(2n)$

$$\mathbb{E}\left(\sum_{i=1}^{n} |g_i|^q\right)^{1/q} \sim n^{1/q}\ \sqrt{q}.$$

Of course, this estimate is well known and can be obtained using direct calculations. Note also that if $q \ge \ln(2k)$ *then* $g_1^* \sim \left(\sum_{i=1}^{k} (g_i^*)^q\right)^{1/q}$. *Hence, for* $q \ge \ln(2k)$ *we have*

$$\mathbb{E}\left(\sum_{i=1}^{k} (g_i^*)^q\right)^{1/q} \sim \sqrt{\ln(2n)}.$$

Corollary 3.6 *There is an absolute positive constant* $c_1 < 1$ *for which the following holds. If* $1 \le q \le \ln(2n)$ *then for every* $k \le c_1^q n$,

$$\mathbb{E}\left(\sum_{i=k+1}^{n} (g_i^*)^q\right)^{1/q} \sim \sqrt{q}\ n^{1/q},$$

where $\{g_i\}_{i=1}^n$ *are independent* $\mathcal{N}(0,1)$ *Gaussian random variables.*

Proof: First observe that the upper estimate is simple. Indeed,

$$\mathbb{E}\left(\sum_{i=k+1}^{n} (g_i^*)^q\right)^{1/q} \le \mathbb{E}\left(\sum_{i=1}^{n} (g_i^*)^q\right)^{1/q} = \mathbb{E}\left(\sum_{i=1}^{n} |g_i|^q\right)^{1/q} \sim n^{1/q}\ \sqrt{q}$$

by Remark 3.5.

Now let us prove the lower estimate. Using Remark 3.5 again, we obtain that there exists an absolute positive constant c_2 such that

$$\mathbb{E}\left(\sum_{i=1}^{n} |g_i|^q\right)^{1/q} \ge 2c_2 n^{1/q}\ \sqrt{q}.$$

Therefore, applying Corollary 3.4,

$$\begin{aligned}\mathbb{E}\left(\sum_{i=k+1}^{n} (g_i^*)^q\right)^{\frac{1}{q}} &\ge \mathbb{E}\left(\sum_{i=1}^{n} |g_i|^q\right)^{\frac{1}{q}} - \mathbb{E}\left(\sum_{i=1}^{k} (g_i^*)^q\right)^{\frac{1}{q}} \\ &\ge 2c_2 n^{\frac{1}{q}}\ \sqrt{q} - C\ k^{\frac{1}{q}}\ \sqrt{q + \ln\frac{2n}{k}},\end{aligned}$$

where C is an absolute constant. Since the function $f(x) = x^{2/q}(q + \ln(2n/x))$ is increasing on $[0, n]$, it is evident that if $k \le c_1^q n$ for some $0 < c_1 < 1$ then

$$\begin{aligned}\mathbb{E}\left(\sum_{i=k+1}^{n} (g_i^*)^q\right)^{\frac{1}{q}} &\ge 2c_2 n^{\frac{1}{q}} \sqrt{q} - C\ c_1\ n^{\frac{1}{q}}\ \sqrt{q \ln(2e/c_1)} \\ &= (2c_2 - c_1 C \ln(2e/c_1))\, n^{\frac{1}{q}} \sqrt{q}.\end{aligned}$$

The desired result is evident by choosing $0 < c_1 < 1$ such that $c_1 C \ln(2e/c_1) \le c_2$. □

4 Interpolation results

We begin this section with the following two known interpolation results. We present the proof of the second one for the sake of completeness. The proof of the first one can be obtained in a similar way (see [H]).

Lemma 4.1 *There exists an absolute constant $c > 0$ for which the following holds. Let $1 \le q_0 < q_1 < \infty$, set r to satisfy $1/r = 1/q_0 - 1/q_1$ and put $1 \le t \le n^{1/r}$. If $K = \mathrm{conv}\left(B^n_{q_0} \cup \frac{1}{t} B^n_{q_1}\right)$ then for every $x \in \mathbb{R}^n$,*

$$c\left(\left(\sum_{i\le t^r}(x_i^*)^{q_0}\right)^{1/q_0} + t\left(\sum_{i>t^r}(x_i^*)^{q_1}\right)^{1/q_1}\right) \le$$

$$\le \|x\|_K \le \left(\sum_{i\le t^r}(x_i^*)^{q_0}\right)^{1/q_0} + t\left(\sum_{i>t^r}(x_i^*)^{q_1}\right)^{1/q_1}.$$

Moreover, if $q_1 = \infty$, then, denoting $q = q_0 \in [1, \infty)$,

$$\|x\|_K \ \sim \ \left(\sum_{i\le t^q}(x_i^*)^{q}\right)^{1/q}.$$

Lemma 4.2 *There exists an absolute constant $c > 0$ for which the following holds. Let $0 < p \le 1$, set $\gamma = 1/(1/p - 1/2)$ and put $n^{-1/\gamma} < \rho < 1$. If $K = B^n_{p\infty} \cap \rho B^n_2$ then*

$$c\left(\rho\ |||x||| + \sum_{i>m} i^{-1/p} x_i^*\right) \le \sup_{y\in K}\langle x, y\rangle \le \rho\ |||x||| + \sum_{i>m} i^{-1/p} x_i^*, \tag{2}$$

where $m = [1/\rho^\gamma]$ and

$$|||x||| = \left(\sum_{i\le m}(x_i^*)^2\right)^{1/2}.$$

Proof: Fix $x \in \mathbb{R}^n$, $x \ne 0$ and without loss of generality assume that $x_1 \ge x_2 \ge \ldots \ge x_n \ge 0$. Recall that $y \in K$ if and only if $y \in \rho B^n_2$ and $y_i^* \le i^{-1/p}$ for every $i \le n$. Applying Hardy-Littlewood inequality for rearrangements we obtain for every $y \in K$

$$\langle x, y\rangle \le \sum_{i=1}^n |x_i y_i| \le \sum_{i=1}^n x_i^* y_i^* = \sum_{i\le m} x_i y_i^* + \sum_{i>m} x_i y_i^* \le \rho\ |||x||| + \sum_{i>m} i^{-1/p} x_i,$$

which shows the right hand side inequality in (2).

To prove the left hand side of (2), first consider $y \in \mathbb{R}^n$ defined by $y_i = \rho x_i/|||x|||$ for $i \le m$ and $y_i = 0$ for $i > m$. Clearly, $y \in \rho B^n_2$. Note that for every $i \le m$

$$x_i \le \left(\frac{1}{i}\ \sum_{j\le i} x_j^2\right)^{1/2} \le |||x|||/\sqrt{i}.$$

Thus,

$$y_i \le \frac{\rho}{\sqrt{i}} \le \frac{1}{m^{1/p-1/2}\sqrt{i}} \le \frac{1}{i^{1/p}},$$

implying that $y \in B_{p\infty}$, and hence $y \in K$. Therefore

$$\sup_{z\in K} \langle x, z\rangle \ge \langle x, y\rangle = \sum_{i\le m} \rho x_i^2/|||x||| = \rho|||x|||.$$

Now take $y = \sum_{i>m} i^{-1/p}\, e_i$. It is evident that $y \in B_{p\infty}$ and, since $(m+1)^{-1/\gamma} \le \rho$,

$$\sum_{i>m} y_i^2 = \sum_{i>m} i^{-2/p} \le (m+1)^{-2/p} + \int_{m+1}^{\infty} x^{-2/p}\, dx \le \frac{2}{2-p}\, \rho^2.$$

Thus $y \in \sqrt{2}\rho B_2^n$, which implies that $y \in \sqrt{2}K$. Therefore

$$\sqrt{2}\, \sup_{z\in K} \langle x, z\rangle \ge \langle x, y\rangle \ge \sum_{i>m} \frac{x_i}{i^{1/p}},$$

and we obtain that

$$\sup_{z\in K} \langle x, z\rangle \ge \max\left\{\rho|||x|||,\ \frac{1}{\sqrt{2}}\sum_{i>m}\frac{x_i}{i^{1/p}}\right\},$$

which completes the proof. □

Remark 4.3 *Note that if $\rho \le n^{-1/\gamma}$ then $K = \rho B_2^n$. Also note that if $p < 1$ then*

$$\sum_{i>m} \frac{x_i^*}{i^{1/p}} \le \frac{\sqrt{2}}{1-p}\, \rho\, |||x|||.$$

Indeed, since $x_i^* \le |||x|||/\sqrt{m}$ for every $i \ge m$ and $m \le 1/\rho^\gamma \le m+1$, then

$$\begin{aligned}\sum_{i>m} \frac{x_i^*}{i^{1/p}} &\le \frac{|||x|||}{\sqrt{m}} \sum_{i>m} i^{-1/p} \le |||x|||\, \sqrt{\frac{2}{m+1}} \left(\frac{1}{(m+1)^{1/p}} + \int_{m+1}^{\infty} x^{-1/p}\, dx\right) \\ &\le \frac{\sqrt{2}}{1-p}\, \rho\, |||x|||.\end{aligned}$$

5 Gaussian averages of interpolated bodies

Now we are ready to formulate our main results.

Theorem 5.1 *There are absolute positive constants c, C, and $c_1 < 1$ for which the following holds. Let $\{g_i\}_{i\le n}$ be independent $\mathcal{N}(0,1)$ Gaussian variables. Consider $1 \le q_0 < q_1 \le \infty$, let r be such that $1/r = 1/q_0 - 1/q_1$, set $1 \le t \le n^{1/r}$ and put $K = \mathrm{conv}\left(B_{q_0}^n \cup \frac{1}{t} B_{q_1}^n\right)$, $L = K^\circ = B_{p_0}^n \cap t B_{p_1}^n$, where $1/p_i + 1/q_i = 1$, $i = 0, 1$.*

(i) If $q_0 \geq \ln(2n)$ then

$$\mathbb{E} \sup_{y \in L} \left\langle \sum_{i=1}^{n} g_i e_i,\, y \right\rangle = \mathbb{E} \left\| \sum_{i=1}^{n} g_i e_i \right\|_K \sim \sqrt{\ln(2n)}.$$

(ii) If $q_0 < \ln(2n) \leq q_1$ then

$$\mathbb{E} \sup_{y \in L} \left\langle \sum_{i=1}^{n} g_i e_i,\, y \right\rangle = \mathbb{E} \left\| \sum_{i=1}^{n} g_i e_i \right\|_K \sim t\, \sqrt{q_0 + \ln(2n/t^{q_0})}.$$

(iii) If $q_1 < \ln(2n)$ and $t > c_1^{q_1/r} n^{1/r}$ then

$$c\, \sqrt{q_0}\, n^{1/q_0} \leq \mathbb{E} \sup_{y \in L} \left\langle \sum_{i=1}^{n} g_i e_i,\, y \right\rangle = \mathbb{E} \left\| \sum_{i=1}^{n} g_i e_i \right\|_K \leq C\, c_1^{-q_1/r}\, \sqrt{q_0}\, n^{1/q_0}.$$

(iv) If $q_1 < \ln(2n)$ and $t \leq c_1^{q_1/r} n^{1/r}$ then

$$\mathbb{E} \sup_{y \in L} \left\langle \sum_{i=1}^{n} g_i e_i,\, y \right\rangle = \mathbb{E} \left\| \sum_{i=1}^{n} g_i e_i \right\|_K \sim t \sqrt{q_1} n^{1/q_1}.$$

Proof:
(i) In this case $e^{-1} B_\infty^n \subset B_{q_0}^n \subset B_\infty^n$ and thus the same is true for K. The estimate is known for the unit cube (see e.g. Lemma 4.14 of [Pi], or just use Lemma 3.1), from which the claim follows.
(ii) Here, $e^{-1} B_\infty^n \subset B_{q_1}^n \subset B_\infty^n$. Therefore, setting $T = \text{conv}\left(B_{q_0}^n \cup \frac{1}{t} B_\infty^n\right)$ and applying Lemma 4.1, we obtain

$$\|x\|_K \sim \|x\|_T \sim \left(\sum_{i \leq t^{q_0}} (x_i^*)^{q_0} \right)^{1/q_0}.$$

By Corollary 3.4,

$$\mathbb{E} \left(\sum_{i \leq t^{q_0}} (g_i^*)^{q_0} \right)^{1/q_0} \sim t \sqrt{q_0 + \ln \frac{2n}{t^{q_0}}},$$

from which the desired result follows.
(iii) Since $B_{q_1}^n \subset n^{1/r} B_{q_0}^n$ then $B_{q_0}^n \subset K \subset c_1^{-q_1/r} B_{q_0}^n$ and the estimate is known (see Remark 3.5).
(iv) First we observe that by Lemma 4.1, Corollary 3.4, and Corollary 3.6 one has

$$\mathbb{E} \left\| \sum_{i=1}^{n} g_i e_i \right\|_K \sim t^{r/q_0} \sqrt{q_0 + \ln(2n/t^r)} + t \sqrt{q_1} n^{1/q_1} = t \left(t^{r/q_1} \sqrt{q_0 + \ln(2n/t^r)} + \sqrt{q_1} n^{1/q_1} \right).$$

Maximizing the function $f(s) = s\sqrt{q_0 + \ln(2n/s^{q_1})}$, it is not hard to see that

$$t^{r/q_1}\ \sqrt{q_0 + \ln(2n/t^r)} \le 3\sqrt{q_1} n^{1/q_1},$$

which implies the desired result. □

Theorem 5.2 *There are absolute positive constants c and C for which the following holds. Let $\{g_i\}_{i \le n}$ be independent, standard Gaussian variables. Set $0 < p \le 1$ and $\gamma = 1/(1/p - 1/2)$, let $n^{-1/\gamma} < \rho < 1$ and put $K = B_{p\infty} \cap \rho B_2^n$.*

(i) If $0 < p < 1$ then

$$c\rho^{2\frac{1-p}{2-p}}\ \sqrt{\ln(2n\rho^\gamma)} \le \mathbb{E} \sup_{y \in K} \left\langle \sum_{i=1}^n g_i e_i, y \right\rangle \le \frac{C}{1-p}\ \rho^{2\frac{1-p}{2-p}}\ \sqrt{\ln(2n\rho^\gamma)}.$$

(ii) If $p = 1$ then

$$\mathbb{E} \sup_{y \in K} \left\langle \sum_{i=1}^n g_i e_i,\ y \right\rangle \sim \ \left(\ln\left(2n\rho^2\right)\right)^{3/2}.$$

Proof: As in Lemma 4.2 denote

$$|||x||| = \left(\sum_{i \le m} (x_i^*)^2\right)^{1/2},$$

where $m = [1/\rho^\gamma]$. By Corollary 3.4

$$\mathbb{E}\ |||\sum_{i=1}^n g_i e_i||| \sim \sqrt{m\ \ln\frac{2n}{m}}.$$

Applying Lemma 4.2 and Remark 4.3, there are absolute constants c and C such that for $p < 1$

$$c\ \rho\ \sqrt{m\ \ln\frac{2n}{m}} \le \mathbb{E} \sup_{y \in K} \left\langle \sum_{i=1}^n g_i e_i,\ y \right\rangle \le \frac{C\rho}{1-p}\ \sqrt{m\ \ln\frac{2n}{m}},$$

which proves (i).

Now, let $p = 1$. Then $m = [1/\rho^2]$ and thus, by Corollary 3.4,

$$\rho\,\mathbb{E}\ |||\sum_{i=1}^n g_i e_i||| \sim \sqrt{\ln\frac{2n}{m}}.$$

Using the assertion of Lemma 4.2, it suffices to bound $\sum_{i>m} g_i^*/i$. To that end, note that there are absolute positive constants C_1, C_2 and C_3 for which the following holds.

(a) By Lemma 3.2, for every $m > n/2$,

$$\mathbb{E} \sum_{i>m} \frac{g_i^*}{i} \sim \sum_{i>m} \frac{1}{i} \, \frac{n-i+1}{n} \sim \frac{(n-m)^2}{n^2},$$

and thus

$$\mathbb{E} \sum_{i>m} \frac{g_i^*}{i} \leq C_1.$$

(b) By Lemma 3.1 and (a), for every $n/4 < m < n/2$

$$C_2 \leq \mathbb{E} \sum_{i>m} \frac{g_i^*}{i} \leq C_3.$$

(c) For every $m \leq n/4$ and, again, by Lemma 3.1,

$$\mathbb{E} \sum_{m<i\leq n/2} \frac{g_i^*}{i} \sim \sum_{m<i\leq n/2} \frac{1}{i} \sqrt{\ln \frac{2n}{i}} \sim \left(\ln \frac{2n}{m} \right)^{3/2}.$$

Combining these estimates with Lemma 4.2, it follows that

$$\mathbb{E} \sup_{y \in K} \left\langle \sum_{i=1}^n g_i e_i, \, y \right\rangle \sim \rho \, \mathbb{E} \, ||| \sum_{i=1}^n g_i e_i ||| + \mathbb{E} \sum_{i>m} \frac{g_i^*}{i} \sim \left(\ln \frac{2n}{m} \right)^{3/2},$$

which completes the proof. □

6 Gelfand widths

The $(k+1)$-th Gelfand width of a given symmetric convex body $T \in \mathbb{R}^n$, $c_{k+1}(T)$, is defined as the smallest possible diameter (in the Euclidean metric) of k-codimensional section of K. The literature over the decades about Gelfand numbers is enormous. For classic results related to our applications see e.g. Chapter 5 of [Pi]. If $T \cap E \subset aB_2^n$ for "most" (in the sense of normalized Haar measure on the Grassmanian) k-codimensional subspaces E then we say that it is true for a "random" subspace E. We prefer not to discuss measure estimates here, i.e. not to specify the word "most" (usually it means that the Haar measure of such subspaces is larger than $1 - e^{-ck}$, where c is an absolute positive constant). The smallest a satisfying $T \cap E \subset aB_2^n$ for a "random" k-codimensional subspace E is called random Gelfand width and is denoted by $cr_{k+1}(T)$. The connection between c_k and cr_k was first investigated in [LT], [MaT] and then in recent works [GiMT, V, LPT].

Recall our notation. Given a symmetric body $T \subset \mathbb{R}^n$, let $T_\rho = T \cap \rho B_2^n$ and

$$\ell_*(T) = \mathbb{E} \sup_{t \in T} \left\langle \sum_{i=1}^n g_i e_i, t \right\rangle.$$

Set $1 - 1/(4\sqrt{k}) < \omega_k := \sqrt{\frac{2}{k}}\Gamma\left(\frac{k+1}{2}\right)/\Gamma\left(\frac{k}{2}\right) < 1$. A variant of so-called "Low M^*-estimate", already mentioned in the introduction, can be formulated as follows.

Theorem 6.1 *Let $1 \le k \le n$ and let T be a symmetric convex body in $\mathbb{R}^n$. Assume that $\rho > 0$ satisfies*

$$k > \left(\frac{\ell_*(T_\rho)}{\omega_k \rho}\right)^2. \tag{3}$$

Then $cr_{k+1}(T) \le \rho$.

Combining Theorem 6.1 with Theorems 5.1 and 5.2 we obtain the following corollaries.

Corollary 6.2 *There exist an absolute positive constant C such that for every $k < n$ and every $1 \le p \le 2 \le q \le \infty$ satisfying $1/p + 1/q = 1$ one has*

(i) if $q \ge \ln(2n)$ (that is, when B_p^n is equivalent to B_1^n) then

$$cr_k\left(B_p^n\right) \le C\ \left(\frac{1}{k}\ \ln\left(\frac{2n}{k}\right)\right)^{1/2},$$

(ii) if $q < \ln(2n)$ then

$$cr_k\left(B_p^n\right) \le C\ k^{-1/2}\ \sqrt{q} n^{1/q}.$$

This corollary is well known ([K], [GaG], [Gl], see also a recent work [Go2]). We provide a proof for completeness. For other related results see for example [GoGMP].

Proof: To simplify notation we denote B_p^n by T. We apply Theorem 5.1 with $q_0 = 2$, $q_1 = q$. Then $1/r = 1/2 - 1/q$ and $L = B_2^n \cap tB_p^n = tT_\rho$, where $\rho = 1/t$. Formally, we should check that $n^{-1/r} \le \rho \le 1$, but this condition will follow automatically, since one trivially has $n^{-1/r} \le cr_k(T) \le 1$.

By Theorem 5.1 for $q \ge \ln(2n)$ we have

$$\ell_*(T_\rho) = (1/t)\ \ell_*(L) \le C_1\sqrt{2 + \ln(2n/t^2)} \le C_2\sqrt{\ln(2n\rho^2)},$$

where C_1 and C_2 are positive absolute constants. Therefore there exists a positive absolute constant C_3 such that the choice

$$\rho = C_3 k^{-1/2}\sqrt{\ln(2n/k)}$$

satisfies inequality (3) which shows the first estimate.

For the second estimate it is enough to use Remark 3.5: there exists an absolute constant C_4 such that

$$\ell_*(T_\rho) \le \ell_*(B_p^n) \le C_4\sqrt{q} n^{1/q}.$$

Therefore the choice

$$\rho = 2C_4 k^{-1/2}\sqrt{q} n^{1/q}$$

satisfies inequality (3). □

Corollary 6.3 *There exists an absolute positive constant C such that for every $k < n$ one has*

$$cr_k\left(B^n_{1\infty}\right) \le C\ k^{-1/2}\left(\ln\left(\frac{2n}{k}\right)\right)^{3/2}.$$

Proof: Denoting $T = B^n_{1\infty}$, by Theorem 5.2 we have

$$\ell_*(T_\rho) \le C_1\left(\ln\left(2n\rho^2\right)\right)^{3/2},$$

where C_1 is a positive absolute constant. Therefore there exists a positive absolute constant C such that the choice

$$\rho = C\ k^{-1/2}\left(\ln\left(\frac{2n}{k}\right)\right)^{3/2}$$

satisfies inequality (3), from which the desired result follows. □

Corollary 6.4 *There exists an absolute positive constant C such that for every $0 < p < 1$ and every. $k < n$ one has*

$$cr_k\left(B^n_{p\infty}\right) \le \left(\frac{C\ln\left(\frac{2n}{k(1-p)^2}\right)}{k(1-p)^2}\right)^{\frac{1}{p}-\frac{1}{2}}.$$

Proof: Denoting $T = B^n_{p\infty}$, by Theorem 5.2 we have

$$\ell_*(T_\rho) \le \frac{C_1}{1-p}\ \rho^{2\frac{1-p}{2-p}}\ \sqrt{\ln\left(2n\rho^\gamma\right)}$$

where C_1 is a positive absolute constant and $\gamma = 1/(1/p-1/2)$. Note that $1-2\frac{1-p}{2-p} = \gamma/2$. Therefore to satisfy inequality (3) it is enough to choose ρ such that

$$\rho^\gamma \ge \frac{2C_1}{(1-p)^2}\frac{\ln\left(2n\rho^\gamma\right)}{k}$$

Clearly, there exists a positive absolute constant C such that the choice

$$\rho = \left(\frac{C\ln\left(\frac{2n}{k(1-p)^2}\right)}{k(1-p)^2}\right)^{\frac{1}{p}-\frac{1}{2}}$$

works. □

7 The Approximate reconstruction problem

Finally, let us present an example of how these bounds can be used in the approximate reconstruction problem for an arbitrary convex, symmetric set $T \subset \mathbb{R}^n$.

Consider the set $T - T = \{t - s \mid t, s \in T\}$. Since T is convex and symmetric, $T - T \subset 2T$. Note that if $\Gamma = k^{-1/2} \sum_{i=1}^{k} \langle X_i, \cdot \rangle e_i$ and $t, s \in T$ for which $\Gamma t = \Gamma s$ then $t - s \in (T - T) \cap \ker(\Gamma)$. In particular, if $t_0 \in T$ is the unknown vector we wish to reconstruct and $\hat{t} \in T$ satisfies that $\langle X_i, t \rangle = \langle X_i, \hat{t} \rangle$, then

$$\hat{t} - t_0 \in 2T \cap \ker(\Gamma).$$

Hence, to estimate $\|t - \hat{t}\|$ in our case, it suffices to prove the following: that if μ is an isotropic, L-subgaussian measure on $\mathbb{R}^n$ and if $X_1, ..., X_k$ are independent, distributed according to μ, then with high probability

$$\operatorname{diam}(2T \cap \ker(\Gamma)) \leq c_1 r_k^*(\theta, T),$$

for $\theta = c_2/L^2$. This fact was proved in [MePT].

Let us mention that in the language of the previous section, the approximate reconstruction problem can be solved using an estimate on the random $(k+1)$-Gelfand number of T, but with a different source of randomness – instead of a random element in the Grassman manifold, a random k-codimensional subspace which is given by the kernel of the random matrix Γ.

The particular example we consider here is when $T = B_{1\infty}^n$, the unit ball in weak-ℓ_1^n.

Theorem 7.1 *Fix any $t_0 \in B_{1\infty}^n$ and let μ be an isotropic, L-subgaussian measure on $\mathbb{R}^n$. Set $X_1, ..., X_k$ to be independent vectors selected according to μ, and put $\hat{t} \in B_{1\infty}^n$ to be a point which satisfies $\langle X_i, \hat{t} \rangle = \langle X_i, t_0 \rangle$ for all $1 \leq i \leq k$. Then with probability at least $1 - 2\exp(-c_L k)$*

$$\|\hat{t} - t_0\|_2 \leq \frac{C_L}{\sqrt{k}} \ln^{3/2}\left(C_L \frac{n}{k}\right),$$

where c_L and C_L are positive constants depending on L only.

Remark. Note that such a point $\hat{t}$ always exists because $t_0 \in T$ satisfies these conditions.

Proof: Recall that

$$r_k^*(\theta, T) := \inf\left\{\rho > 0 : \rho \geq 2\ell_*(T_\rho)/\theta\sqrt{k}\right\},$$

where $\theta = c/L^2$ and that $T_\rho = T \cap \rho B_2^n$. By estimates from [MePT], with probability at least $1 - 2\exp(-c_L k)$,

$$\|\hat{t} - t_0\| \leq C_0 r_k^*(\theta, B_{1\infty}^n),$$

where $c_L > 0$ depends only on L and $C_0 > 0$ is an absolute constant. Applying Theorem 5.2(*ii*),

$$\ell_*((B_{1\infty}^n)_\rho) \leq \left(\ln\left(2n\rho^2\right)\right)^{3/2},$$

and solving for ρ shows that

$$r_k^*(\theta, B_{1\infty}) \leq \frac{C_1 L^2}{\sqrt{k}} \ln^{3/2}\left(C_1 L^4 \frac{n}{k}\right),$$

where $C_1 > 0$ is an absolute constant. Therefore, with probability at least $1-2\exp(-c_L k)$,

$$\|\hat{t} - t_0\|_2 \leq \frac{C_0 C_1 L^2}{\sqrt{k}} \ln^{3/2}\left(C_1 L^4 \frac{n}{k}\right).$$

It proves the result with $C_L = \max\{C_0 C_1 L^2, C_1 L^4\}$. □

References

[CDS] Chen S. S., Donoho D. L. & Saunders M. A. Atomic decomposition by basis pursuit. SIAM J. Scientific Computing 20 (1999), 33-61.

[CT1] Candes, E. & Tao, T. Near optimal recovery from random projections: universal encoding strategies. Preprint.

[CT2] Candes, E. & Tao, T. Decoding by Linear Programming. IEEE Trans. Inform. Theory 51 (2005), no. 12, 4203–4215.

[CT3] Candes, E. & Tao, T. Error Correction via Linear Programming. Preprint.

[D] Donoho D. L. For most large undetermined systems of linear equations the minimal ℓ_1-norm solution is the sparsest solution, Commun. Pure Appl. Math. 59(2006), No. 6, 797–829.

[DE] Donoho D. L. & Elad, M. Optimally sparse representation in general (non-orthogonal) dictionaries via ℓ_1 minimization. Proc. Natl. Acad. Sci. USA 100(2003) 2197-2202.

[DET] Donoho D. L., Elad, M. & Temlyakov, V. Stable recovery of sparse overcomplete representations in the presence of noise, IEEE Trans. Inform. Theory 52 (2006), no. 1, 6–18.

[GaG] Garnaev A. Yu. & Gluskin E. D. On widths of the Euclidean ball. Sov. Math., Dokl. 30, 200-204 (1984); translation from Dokl. Akad. Nauk SSSR 277, 1048-1052 (1984).

[Gl] Gluskin E. D. Norms of random matrices and widths of finite-dimensional sets. Math. USSR, Sb. 48 (1984), 173-182.

[GiMT] Giannopoulos A., Milman V. D., & Tsolomitis A., Asymptotic formulas for the diameter of sections of symmetric convex bodies, J. of Funct. Analysis, 223 (2005), 86–108.

[Go1] Gordon, Y. On Milman's inequality and random subspaces which escape through a mesh in $\mathbb{R}^n$. Geometric aspects of functional analysis (1986/87), Lecture Notes in Math., 1317 (1988), 84–106.

[Go2] Gordon, Y. A note on an observation of G. Schechtman. Geometric aspects of functional analysis (2004/2005), Lecture Notes in Math., 1910 (2007), to appear.

[GoGMP] Gordon Y., Guédon O., Meyer M., & Pajor A. Random Euclidean sections of some classical Banach spaces, Math. Scand., 91 (2002), 247-268.

[GoLSW1] Gordon Y., Litvak A. E., Schütt C., & Werner E. Orlicz Norms of Sequences of Random Variables, Ann. of Prob., 30 (2002), 1833–1853.

[GoLSW2] Gordon Y., Litvak A. E., Schütt C., & Werner E. On the minimum of several random variables, Proc. Amer. Math. Soc. 134 (2006), no. 12, 3665–3675.

[K] Kashin, B. S. Diameters of some finite-dimensional sets and classes of smooth functions, Math. USSR, Izv. 11 (1977), 317-333.

[H] Holmstedt T. Interpolation of quasi-normed spaces, Math. Scand. 26 (1970), 177-199.

[LPT] Litvak A. E., Pajor A., & Tomczak-Jaegermann N. Diameters of Sections and Coverings of Convex Bodies, J. of Funct. Anal., 231 (2006), 438–457.

[LT] Litvak A. E. & Tomczak-Jaegermann N. Random aspects of high-dimensional convex bodies, GAFA, Israel Seminar, Lecture Notes in Math., 1745, Springer-Verlag, 2000, 169-190.

[MaT] Mankiewicz P. & N. Tomczak-Jaegermann N. Volumetric invariants and operators on random families of Banach spaces, Studia Math., 159 (2003), 315–335.

[MePT] Mendelson, S., Pajor A., & Tomczak-Jaegermann N. Reconstruction and subgaussian operators, Geometric and Functional Analysis, to appear.

[Mi1] Milman, V. Random subspaces of proportional dimension of finite dimensional normed spaces: approach through the isoperimetric inequality. Lecture Notes in Math., 1166 (1985), 106–115.

[Mi2] Milman, V. Almost Euclidean quotient spaces of subspaces of a finite-dimensional normed space. Proc. Amer. Math. Soc. 94 (1985), 445-449.

[PT1] Pajor, A. & Tomczak-Jaegermann, N. Subspaces of small codimension of finite-dimensional Banach spaces, Proc. Amer. Math. Soc. 97(4), (1986), 637-642.

[PT2] Pajor, A. & Tomczak-Jaegermann, N. Gelfand numbers and Euclidean sections of large dimension, Probability in Banach Spaces 6, Proc. of the VI Int. Conf in Probab. in Banach Spaces, Sandjberg, Denmark, 1986, Birkhäuser, 252-264.

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

[RV] Rudelson, M. & Vershynin, R. Geometric approach to error-correcting codes and reconstruction of signals. Int. Math. Res. Not. 2005, no. 64, 4019–4041.

[V] Vershynin R. Isoperimetry of waists and local versus global asymptotic convex geometries, (with an appendix by M. Rudelson and R. Vershynin), Duke Math. J. 131(2006), 1-16.

Y. Gordon, Department of Mathematics, Technion, Haifa 32000, Israel.
e-mail: gordon@techunix.technion.ac.il

A.E. Litvak, Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, AB, Canada, T6G 2G1.
e-mail: alexandr@math.ualberta.ca

S. Mendelson, Centre for Mathematics and its Applications, The Australian National University, Canberra, ACT 0200, Australia, and Department of Mathematics, Technion, Haifa 32000, Israel.
e-mail: shahar.mendelson@anu.edu.au

A. Pajor, Equipe d'Analyse et Mathématiques Appliquées, Université de Marne-la-Vallée, 5, boulevard Descartes, Champs sur Marne, 77454 Marne-la-Vallée Cedex 2, France.
e-mail : Alain.Pajor@univ-mlv.fr