arXiv:1611.01723v1 [math.PR] 6 Nov 2016

A small deviation inequality for convex functions

Grigoris Paouris* and Petros Valettas†

November 8, 2016

Abstract

Let Z be an n-dimensional Gaussian vector and let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function. We prove that:

$$\mathbb{P}\left(f(Z) \leq \mathbb{E}f(Z) - t\sqrt{\mathrm{Var} f(Z)}\right) \leq \exp(-ct^2),$$

for all $t > 1$ where $c > 0$ is an absolute constant. As an application we derive refinements of the Johnson-Lindenstrauss flattening lemma and the random version of Dvoretzky's theorem.

1 Introduction

The purpose of this note is to establish a sharp distributional inequality for convex functions on Gauss' space $(\mathbb{R}^n, \gamma_n)$. Our goal and motivation stems from the attempt to strengthen the classical Gaussian concentration for special cases that are of interest in high-dimensional geometry. The Gaussian concentration phenomenon (see [1] and [14]) states that for any L-Lipschitz map $f : \mathbb{R}^n \to \mathbb{R}$ one has

$$\mathbb{P}\left(\left|f(Z) - M\right| > t\right) \leq \exp\left(-\tfrac{1}{2}t^2/L^2\right), \tag{1.1}$$

for all $t > 0$, where Z is n-dimensional standard Gaussian random vector and M is a median for $f(Z)$. In turn this implies bounds on the variance $\mathrm{Var} f(Z)$ for any Lipschitz map f in terms of the Lipschitz constant L:

$$\mathrm{Var}[f(Z)] \leq 2L^2. \tag{1.2}$$

*Supported by the NSF CAREER-1151711 grant;

†Supported in part by the NSF grant DMS-1612936.

2010 Mathematics Subject Classification. Primary: 60D05, Secondary: 52A21, 52A23

Keywords and phrases. Ehrhard's inequality, Concentration, Small ball probability, Johnson-Lindenstrauss lemma, Dvoretzky's theorem.

One may check that the above inequalities are sharp (up to constants) for linear functionals. However, one can easily construct examples of convex functions (see Section 2) for which the above estimates are far from being optimal. On the other hand it is known (see [18] and [21, Proposition 2.9]) that for positively homogeneous, convex functions the estimate (1.1) is sharp (up to absolute constants) in the large deviation regime $t > M$. Therefore, in this paper we focus on the one-sided small deviation inequality:

$$\mathbb{P}\left(f(Z) - M < -t\right) \leq \frac{1}{2} \exp\left(-\tfrac{1}{2}t^2/L^2\right),$$

which holds for all $t > 0$ and for any L-Lipschitz map f. Our main result reads as follows: For any convex map $f \in L_2(\gamma_n)$ one has

$$\mathbb{P}\left(f(Z) - M < -t\right) \leq \frac{1}{2} \exp\left(-\tfrac{\pi}{256}t^2/\mathrm{Var}[f(Z)]\right), \tag{1.3}$$

for all $t > 0$. In view of (1.2) this obviously improves the one-sided concentration inequality in the small deviation regime. We want to emphasize the fact that the inequality we propose does not require the Lipschitz condition; instead it is valid for any convex function $f \in L_2(\gamma_n)$ (in fact we may even consider $f \in L_{1,\infty}(\gamma_n)$).

As a consequence of this distributional inequality one can get refined versions of classical results such as: the Johnson-Lindenstrauss dimension reduction lemma from [7] and the random version of Dvoretzky's theorem due to V. Milman [16] (see also [17]).

The rest of the paper is organized as follows: In Section 2 we present a proof of the main result. The key ingredient in our argument is Ehrhard's inequality [3], inspired by the approach of Kwapien in [13]. The applications of our main result in asymptotic geometric analysis are presented in Section 3. We conclude in Section 4 with further comments.

2 Proof of the main result

Let Φ be the cumulative distribution function of a standard Gaussian random variable, i.e.

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-z^2/2}\, dz, \quad x \in \mathbb{R}.$$

Our goal is to prove the next:

Theorem 2.1. *Let Z be an n-dimensional standard Gaussian vector. Let f be a convex function on $\mathbb{R}^n$ with $f \in L_1(\gamma_n)$ and let M be a median for $f(X)$. Then, we have:*

$$\mathbb{P}\left(f(Z) - M < -t\mathbb{E}\left|f(Z) - M\right|\right) \leq \Phi\left(-\frac{\sqrt{2\pi}}{16}t\right),$$

for all $t > 0$.

Proof. We may assume that $\mathbb{E}|f(Z) - M| > 0$ otherwise we have nothing to prove. Note that since f is convex, the level sets $\{f < t\}$, $t \in \mathbb{R}$ are convex sets and the function $F(t) := P(f(Z) < t)$ is log-concave. The latter follows by the following inclusion:

$$(1-\lambda)\{f < t\} + \lambda\{f < s\} \subseteq \{f < (1-\lambda)t + \lambda s\},$$

for $t, s \in \mathbb{R}$ and $0 \leq \lambda \leq 1$ and the fact that γ_n is log-concave measure (see [1, Section 1.8] for the related definition). Now, we may use Ehrhard's inequality from [3] (see also [1, Theorem 4.2.1.]) to get that the map $s \mapsto \Phi^{-1} \circ F(s)$, $s \in \mathbb{R}$ is concave (for a proof see [1, Theorem 4.4.1.]). Therefore, we obtain:

$$\begin{aligned} (\Phi^{-1} \circ F)(s+M) &= (\Phi^{-1} \circ F)(s+M) - (\Phi^{-1} \circ F)(M) \\ &\leq s(\Phi^{-1} \circ F)'(M) = s\sqrt{2\pi}F'(M), \quad s \in \mathbb{R}. \end{aligned} \tag{2.1}$$

Now we give a lower bound for $F'(M)$ in terms of the standard deviation of $f(X)$.

Claim. We have the following:

$$F'(M) \geq \frac{1}{16\mathbb{E}|f(X) - M|}.$$

Proof of Claim. Fix $\delta > 0$ (that will be chosen appropriately later). Using the log-concavity of F we may write:

$$\begin{aligned} \delta\frac{F'(M)}{F(M)} \geq \log F(M+\delta) - \log F(M) &= \log\left(1 + 2\mathbb{P}(M \leq f(Z) < M+\delta)\right) \\ &\geq \mathbb{P}(M \leq f(Z) < M+\delta) \\ &= \frac{1}{2} - \mathbb{P}(f(Z) \geq M+\delta), \end{aligned}$$

where we have used the elementary inequality $\log(1+u) \geq u/2$ for all $0 < u \leq 1$. Now we apply Markov's inequality to get:

$$\mathbb{P}(f(Z) \geq M+\delta) \leq \frac{\mathbb{E}|f(Z) - M|}{\delta}.$$

Combing the above we conclude that:

$$F'(M) \geq \frac{1}{\delta}\left(\frac{1}{2} - \frac{\mathbb{E}|f(Z) - M|}{\delta}\right).$$

The choice $\delta = 4\mathbb{E}|f(Z) - M|$ yields the assertion of the Claim.

Going back to (2.1) we readily see that (for $s = -t\mathbb{E}|f(Z) - M|$):

$$\Phi^{-1}\left[\mathbb{P}\Big(f(Z) - M < -t\mathbb{E}|f(Z) - M|\Big)\right] \leq -t\frac{\sqrt{2\pi}}{16},$$

as required. □

Remarks 2.2. 1. The advantage of this one-sided concentration inequality is that it can be applied for the wide class of convex functions which are not necessarily (globally) Lipschitz or which are not even in $L_2(\gamma_n)$; e.g. the function $f(t) = \exp(-t + t^2/2)$ is (logarithmically) convex, belongs to $L_1(\gamma_1)$ but $f \notin L_2(\gamma_1)$. Moreover, a careful inspection of the argument shows that it is enough to have $f \in L_{1,\infty}(\gamma_n)$ (see e.g. [5]) and the conclusion still holds:

$$\mathbb{P}\left(f(Z) < M - t\|(f-M)_+\|_{1,\infty}\right) \leq \Phi(-ct), \quad t > 0, \tag{2.2}$$

where $c > 0$ is an absolute constant.[1]

2. Note that (2.1) implies that the variable $f(Z)$ is stochastically dominated by the normal random variable $\zeta := M + a \cdot g$, where g is a standard normal variable and $1/a := (2\pi)^{1/2} F'(M) > 0$, i.e.

$$\mathbb{P}(f(Z) \leq s) \leq \mathbb{P}(\zeta \leq s),$$

for all $s \in \mathbb{R}$. Hence one gets: $\mathbb{E} f(Z) \geq \mathbb{E}\zeta = M$. This result is due to Kwapien [13]. In fact our proof steps on the same starting line as in [13].

3. Taking into account the fact that $\mathbb{E}|f(Z) - M| \leq \sqrt{\operatorname{Var} f(Z)}$ and that

$$1 - \Phi(u) = \Phi(-u) \leq \frac{1}{2} e^{-u^2/2} \tag{2.3}$$

for all $u > 0$ (for a proof see [15, Lemma 1]) we immediately get:

$$\mathbb{P}\left(f(Z) - M < -t\sqrt{\operatorname{Var} f(Z)}\right) \leq \Phi\left(-t\frac{\sqrt{2\pi}}{16}\right) \leq \frac{1}{2}\exp\left(-\frac{\pi}{256}t^2\right),$$

for all $t > 0$, which is the announced estimate (1.3) provided that $f \in L_2$.

Furthermore, using the fact $M \geq \mathbb{E} f(Z) - \sqrt{\operatorname{Var} f(Z)}$ once more, we may conclude the following "Central Limit type" normalization in Theorem 2.1: For any convex function f on $\mathbb{R}^n$ with $f \in L_2(\gamma_n)$ one has the following distributional inequality:

$$\mathbb{P}\left(f(Z) - \mathbb{E} f(Z) < -t\sqrt{\operatorname{Var} f(Z)}\right) \leq \frac{1}{2}\exp\left(-\frac{\pi}{256}(t-1)^2\right) < e^{-t^2/1000}, \tag{2.4}$$

for all $t > 1$.

4. It turns out that one can prove a similar inequality for the n-dimensional exponential measure but for 1-unconditional functions f, i.e. functions which satisfy $f(x_1, \ldots, x_n) = f(|x_1|, \ldots, |x_n|)$ for all $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$.

We fix W for an n-dimensional exponential random vector, i.e. $W = (\xi_1, \ldots, \xi_n)$, where $(\xi_i)_{i=1}^n$ are independent identically distributed according to the measure ν_1 with density function $\frac{d\nu_1(x)}{dx} = \frac{1}{2}e^{-|x|}$. Note that if g_1, g_2 are i.i.d. standard normals and ξ is independent exponential random variable then $|\xi|$ and $\frac{g_1^2+g_2^2}{2}$ have the same distribution (follows easily by examining the moment generating functions). Based on this remark we have the following consequence of Theorem 2.1:

[1] Here and everywhere else C and c stand for absolute constants whose values may change from line to line.

Theorem 2.3. *Let f be 1-unconditional and convex function on $\mathbb{R}^n$. If W is an exponential random vector on $\mathbb{R}^n$, then one has:*

$$\mathbb{P}\left(f(W) - M \leq -t\mathbb{E}|f(W) - M|\right) \leq 1 - \Phi(ct) \leq \exp(-ct^2),$$

for all $t > 0$.

Proof. Consider the function $F : \mathbb{R}^{2n} \to \mathbb{R}$ defined as:

$$F(x_1, \ldots, x_n, y_1, \ldots, y_n) := f\left(\frac{x_1^2 + y_1^2}{2}, \ldots, \frac{x_n^2 + y_n^2}{2}\right).$$

Since f is convex and 1-unconditional it follows that f is convex and coordinatewise non-decreasing[2] in the octant $\mathbb{R}^n_+ = \{z = (z_1, \ldots, z_n) : z_i \geq 0\}$. Hence F is convex on $\mathbb{R}^{2n}$. Therefore a direct application of Theorem 2.1 yields:

$$\mathbb{P}\left(f(\tilde{W}) - M \leq -t\mathbb{E}|f(\tilde{W}) - M|\right) \leq \exp(-ct^2),$$

for all $t > 1$, where $\tilde{W} = (|\xi_1|, \ldots, |\xi_n|)$ and ξ_i are i.i.d. exponential random variables. The fact that $f(x_1, \ldots, x_n) = f(|x_1|, \ldots, |x_n|)$ completes the proof. □

5. In the above argument it is clear that we may also consider longer sums of the form $g_1^2 + \ldots + g_k^2$. That is, if $f : \mathbb{R}^n_+ \to \mathbb{R}$ is a coordinatewise non-decreasing and convex function, then

$$\mathbb{P}\left(f(\chi) < M - t\mathbb{E}|f(\chi) - M|\right) \leq \Phi(-t/20),$$

for all $t > 0$, where $\chi \sim \chi^2(k)$ is a chi squared random variable with k degrees of freedom.

6. Finally, let us note that in all statements above, one can derive the reverse distributional inequality for concave functions, i.e. if f is a *concave* function on $\mathbb{R}^n$ then,

$$\mathbb{P}\left(f(Z) - M > t\mathbb{E}|f(Z) - M|\right) \leq \Phi(-ct),$$

for all $t > 0$, where M is a median for $f(Z)$.

7. Probabilistic inequalities similar to (1.3), in the context of log-concave measures, will be presented elsewhere [19].

3 Applications in asymptotic convex geometry

The concentration of measure phenomenon has two famous applications in asymptotic geometric analysis: the Johnson-Lindenstrauss flattening lemma [7] and the

[2] A real valued function H defined on $U \subseteq \mathbb{R}^k$ is said to be *coordinatewise non-decreasing* if it is non-decreasing in each variable while keeping all the other variables fixed at any value.

random version of Dvoretzky's theorem due to V. Milman [16]. Being the concentration inequality a two sided estimate, provides sufficient information for low distortion embeddings of any point-set on the Euclidean space to an arbitrary normed space.

On the other hand, if one only focus on the one-sided inequality on the isomorphic version of the randomized Dvoretzky then the dimension dependence can be considerably improved. Klartag and Vershynin in [11], building on ideas of Latala and Oleszkiewicz from [15], introduced a new parameter $d(A)$, associated with any convex body A on $\mathbb{R}^n$ to study this phenomenon. They showed that this parameter replaces the critical dimension $k(A)$ and one may still has the one-sided inclusion (see below for the related definitions). They also proved that this parameter is at least as big as the critical dimension. However, there was no connection with other global, computable parameter associated with the body (or the norm); neither an almost isometric version of their result was available. The motivation to prove 2.1 is to provide answers to the above questions.

In the rest of the Section we start with proving a small ball probability estimate for norms. Then, we shall apply this result to get refined one-sided versions of the almost isometric version of Dvoretzky's theorem and the J-L lemma.

3.1 Small ball probabilities

Let A be a centrally symmetric convex body on $\mathbb{R}^n$ and let M be the median for the map $x \mapsto \|x\|_A$ with respect to the Gaussian measure. Recall the Klartag-Vershynin parameter from [11]:

$$d(A) := \min\left\{n, -\log \gamma_n\left(\frac{M}{2}A\right)\right\}.$$

Using the main argument of Latala and Oleszkiewicz from [15], Klartag and Vershynin prove in [11] that this parameter is responsible for small ball probability estimates. Namely, they show that $P(\|Z\|_A \leq c\varepsilon\mathbb{E}\|Z\|_A) \leq (C\varepsilon)^{cd(A)}$ for all $\varepsilon \in (0, \varepsilon_0)$, where $c, C > 0$ and $0 < \varepsilon_0 < 1$ are absolute constants and Z is an n-dimensional standard Gaussian vector. Furthermore, they show that $d(A) \geq ck(A)$ where $k(A)$ is the critical dimension of the body A defined as $k(A) := \mathbb{E}\|Z\|_A^2/b(A)^2$, where $b(A) = \max_{\theta\in S^{n-1}} \|\theta\|_A$. Thus, they recover the main result of [15]. Also, they comment that there exist convex bodies A for which $d(A) \gg k(A)$ (e.g. A being the n-dimensional cube B_∞^n; see [11]).

Here we suggest another global parameter $\beta(A)$, associated with any convex body A, which is defined as follows:

$$\beta(A) := \frac{\mathrm{Var}\|Z\|_A}{M^2}, \tag{3.1}$$

where Z is n-dimensional standard Gaussian vector and M stands for the median of $\|Z\|_A$. In view of (1.2) and by taking into account the fact that $M \simeq \mathbb{E}\|Z\|_A$ we may easily deduce the bound $\beta(A) \leq c/k(A)$. We should mention that there exist many classical examples (e.g. the ℓ_p balls B_p^n for $2 < p \leq \infty$) for which β is dramatically smaller compared to $1/k$ (see [21, Section 3]; see also [20] for an extension of this result for any subspace of L_p, $2 < p < \infty$). Below we prove that $1/\beta(A)$ serves as a general lower bound for $d(A)$. Namely, we have the following:

Theorem 3.1. *Let A be a centrally symmetric convex body on $\mathbb{R}^n$. Then, one has the one-sided concentration estimate:*

$$\mathbb{P}\left(\|Z\|_A \leq (1-\varepsilon)M\right) \leq \frac{1}{2}\exp\left(-c\varepsilon^2/\beta(A)\right), \tag{3.2}$$

for all $0 < \varepsilon < 1$ where Z is n-dimensional standard Gaussian random vector. In particular, one has:

$$d(A) \geq c_1/\beta(A) \geq c_2 k(A). \tag{3.3}$$

Moreover, we have the following small ball probability estimate:

$$\mathbb{P}\left(\|Z\|_A \leq \varepsilon\mathbb{E}\|Z\|_A\right) \leq \frac{1}{2}\varepsilon^{c/\beta(A)}, \tag{3.4}$$

for all $\varepsilon \in (0, 1/2)$.

Proof. The bound $d(A) \geq c/\beta(A)$ will follow by the definition of d and once we will have proved (3.2) by plugging $\varepsilon = 1/2$. The estimate (3.2) follows from the next general:

Claim. Let f be a convex map with $f \in L_1(\gamma_n)$ and let $M = \mathrm{med} f(Z) \neq 0$. Then, we have:

$$\mathbb{P}\left(f(Z) \leq (1-\varepsilon)M\right) \leq \frac{1}{2}\exp\left(-c\varepsilon^2/\beta(f)\right),$$

for all $\varepsilon \in (0,1)$, where $\beta(f) = \mathrm{Var}[f(Z)]/M^2$.

Proof of Claim. Fix $\varepsilon \in (0,1)$. Apply Theorem 2.1 for $t_\varepsilon := \varepsilon/\sqrt{\beta}$ to get:

$$\mathbb{P}\left(f(Z) \leq (1-\varepsilon)M\right) = \mathbb{P}\left(f(Z) - M \leq -t_\varepsilon\sqrt{\mathrm{Var}[f(Z)]}\right) \leq \Phi\left(-t_\varepsilon\frac{\sqrt{2\pi}}{16}\right) \leq \frac{1}{2}e^{-c\varepsilon^2/\beta},$$

where in the last step we have used (2.3) again. This proves the claim.

In order to prove the probabilistic estimate (3.4) we may apply directly the result from [11]. However, we prefer to present a proof which provides explicit constants and range for ε. To this end we may argue as in [15] and apply a theorem of Cordero-Erausquin, Fradelizi and Maurey from [2] known as B-Theorem. The latter states that for any centrally symmetric convex body K on $\mathbb{R}^n$ one has:

$$\gamma_n(a^{1-\lambda}b^\lambda K) \geq [\gamma_n(aK)]^{1-\lambda}[\gamma_n(bK)]^\lambda,$$

for all $a, b > 0$ and $0 < \lambda < 1$. To the contrary of the approach in [15] instead of using the Gaussian isoperimetry, we use Ehrhard's inequality which is formally stronger. We roughly give the details: Let $\varepsilon \in (0, 1/2)$. We choose $\lambda \in (0,1)$ such that $1/2 = \varepsilon^{1-\lambda}$, i.e. $1 - \lambda = \frac{\log 2}{\log(1/\varepsilon)}$. We employ B-Theorem to obtain:

$$\gamma_n\left(\frac{M}{2}A\right) \geq [\gamma_n(\varepsilon MA)]^{1-\lambda}[\gamma_n(MA)]^\lambda,$$

or equivalently

$$[2\gamma_n(\varepsilon MA)]^{1-\lambda} \leq 2\gamma_n\left(\frac{M}{2}A\right). \tag{3.5}$$

Now we estimate $\gamma_n(\frac{M}{2}A)$ from above by using Ehrhard's inequality. Set for simplicity $F_A(t) = \gamma_n(\{x : \|x\|_A \leq t\}) = \gamma_n(tA),\ t > 0$; we may write:

$$\Phi^{-1} \circ F_A(M/2) \leq -\frac{M}{2}\sqrt{2\pi}F'(M) \leq -\frac{\sqrt{2\pi}}{32}\frac{M}{\sqrt{\mathrm{Var}\|Z\|_A}},$$

where we have used the same reasoning as in Theorem 2.1. It follows that:

$$\gamma_n\left(\frac{M}{2}A\right) = F_A(M/2) \leq \Phi\left(-\frac{\sqrt{2\pi}}{32}\beta^{-1/2}\right) \leq \frac{1}{2}\exp\left(-\frac{\pi}{10000}\beta^{-1}\right). \tag{3.6}$$

Plug (3.6) into (3.5) and taking into account the value of λ we obtain:

$$2\gamma_n(\varepsilon MA) \leq \exp\left(-\frac{\pi\beta^{-1}}{10000(1-\lambda)}\right) = \varepsilon^{c/\beta},$$

as required. □

Remark. In [2] it is also proved that any 1-unconditional log-concave measure μ and 1-unconditional convex body K on $\mathbb{R}^n$ has the B-property, that is $t \mapsto \mu(e^t K)$ is concave. Therefore, in view of Theorem 2.3 we readily get the following:

Proposition 3.2. *If K is 1-unconditional convex body on $\mathbb{R}^n$, then one has:*

$$\nu_1^n(\varepsilon mK) \leq \frac{1}{2}\varepsilon^{c/\beta}, \quad \varepsilon \in (0, 1/2),$$

where m is a median for $x \mapsto \|x\|_K$ with respect to ν_1^n and $\beta = \sqrt{\mathrm{Var}\|W\|_K}/m$.

We also have the following reverse Hölder inequality for negative moments due to the small deviation (3.2) and the small ball probability (3.4):

Corollary 3.3. *Let K be a centrally symmetric convex body on $\mathbb{R}^n$. Then, one has:*

$$\mathbb{E}\|Z\|_K \left(\mathbb{E}\|Z\|_K^{-q}\right)^{1/q} \leq \exp\left(C\sqrt{\beta} + Cq\beta\right),$$

for all $0 < q < c/\beta(K)$ where $C, c > 0$ are absolute constants and Z is n-dimensional standard Gaussian vector.

Proof. We know that:

$$\mathbb{P}(\|Z\|_K \leq \varepsilon M) \leq \frac{1}{2}\varepsilon^{c_1/\beta}, \quad \mathbb{P}(\|Z\|_K \leq (1-\varepsilon)M) \leq \frac{1}{2}e^{-c_2\varepsilon^2/\beta},$$

for all $\varepsilon \in (0, 1/2)$, where M is the median for $\|Z\|_K$ and $Z \sim N(\mathbf{0}, I_n)$. Therefore, we may write:

$$\begin{aligned}
\mathbb{E}\|Z\|_K^{-q} &= M^{-q} \int_0^\infty \mathbb{P}(\|Z\|_K \le tM) \frac{q}{t^{q+1}}\, dt \\
&\le M^{-q} \left(\frac{q}{2} \int_0^{1/2} \varepsilon^{\frac{c_1}{\beta} - q - 1}\, d\varepsilon + \int_{1/2}^1 \frac{q}{t^{q+1}} P(\|Z\|_K \le tM)\, dt + 1 \right) \\
&\le M^{-q} \left(\left(\frac{1}{2}\right)^{\frac{c_1}{\beta} - q} \frac{q\beta}{c_1 - q\beta} + q \int_0^{1/2} \frac{1}{(1-\varepsilon)^{q+1}} e^{-c_2 \varepsilon^2/\beta}\, d\varepsilon + 1 \right) \\
&\le M^{-q} \left(1 + c_3 q\beta + q \int_0^{1/2} \exp(2(q+1)\varepsilon - c_2\varepsilon^2/\beta)\, d\varepsilon \right),
\end{aligned}$$

for all $0 < q < c_4/\beta$, where we have also used the the elementary inequality $1 - v \ge e^{-2v}$ for $0 \le v \le 1/2$. It is easy to check that the last integral can be bounded as:

$$\int_0^{1/2} \exp(2(q+1)\varepsilon - c_2\varepsilon^2/\beta)\, d\varepsilon \lesssim \sqrt{\beta} \exp(c_5 q^2 \beta),$$

for[3] all $0 < q \lesssim 1/\beta$. The result follows. □

3.2 The Johnson-Lindenstrauss flattening lemma

The J-L lemma from [7] (see also [8]) asserts that: if $\varepsilon \in (0,1)$ and $x_1, \ldots, x_N \in \ell_2$ then there exists a linear mapping (which can be chosen to be an orthogonal projection) $P : \ell_2 \to F$, where F is a subspace of ℓ_2 with $\dim F \le c\varepsilon^{-2} \log N$ such that

$$(1-\varepsilon)\|x_i - x_j\|_2 \le \|Px_i - Px_j\|_2 \le (1+\varepsilon)\|x_i - x_j\|_2,$$

for all $i, j = 1, \ldots, N$.

This dimension reduction principle has found many applications in mathematics and computer science, in addition to the original application in [7] for the Lipschitz extension problem. We refer the interested reader to [6, 12, 25] and the references therein for a partial list of its many applications.

Below we suggest a form of the J-L Lemma for an arbitrary target space, as was formulated in [24], which exhibits an improved one-sided scaling.

Proposition 3.4 (J-L lemma). *Let $E = (\mathbb{R}^n, \|\cdot\|)$ be a normed space and let T be a finite subset of the metric space ℓ_2^k. Let $\varepsilon \in (0,1)$ and assume that $\log|T| \lesssim \varepsilon^2 k(E)$. Then, there exists a map $P : T \to E$ with the following property:*

$$(1 - s\varepsilon) \cdot \|u - v\|_2 \le \|Pu - Pv\| \le (1+\varepsilon) \cdot \|u - v\|_2,$$

for all $u, v \in T$, where $s \equiv s(E) := \sqrt{\mathrm{Var}\|Z\|}/b(E) < 1$ and $b(E) = \max_{\theta \in S^{n-1}} \|\theta\|$.

[3] For any two quantities A, B depending on n, p, etc. we write $A \lesssim B$ if there exists numerical constant $C > 0$ - independent of everything - such that $A \le CB$. We write $A \gtrsim B$ if $B \lesssim A$ and $A \simeq B$ if $A \lesssim B$ and $B \lesssim A$. Accordingly we write $A \simeq_p B$ if the constants involved are depending only on p.

Proof. The classical Gaussian concentration estimate asserts that for any Lipschitz map $f : \mathbb{R}^n \to \mathbb{R}$ we have:

$$\mathbb{P}\left(f(Z) > t + M\right) \leq \frac{1}{2}\exp\left(-\frac{t^2}{2L^2}\right), \quad t > 0,$$

where $L = \mathrm{Lip}(f)$. Consider $Z_1, \ldots, Z_k$ i.i.d. standard Gaussian vectors on $\mathbb{R}^n$ and define the random matrix $G = [Z_1, \ldots, Z_k]$. Then, for any fixed $\theta \in S^{k-1}$ one has:

$$\mathbb{P}(\|G\theta\| > (1+t)M) = \mathbb{P}(\|Z_1\| > tM + M) \leq \frac{1}{2}\exp(-t^2k/2), \ t > 0,$$

where $k = k(E)$. If $T = \{u_1, \ldots, u_N\}$, consider the points $\Theta := \left\{\frac{u_i - u_j}{\|u_i - u_j\|_2} \ : \ 1 \leq i < j \leq N\right\}$ on S^{k-1}. Then, by the union bound we get:

$$\mathbb{P}(\exists\, \theta \in \Theta \ : \ \|G\theta\| > (1+t)M) < \frac{N^2}{4}\exp(-t^2k/2).$$

Arguing similarly and applying Theorem 3.1 (note that $x \mapsto \|x\|$ is convex) we get:

$$\mathbb{P}(\exists\, \theta \in \Theta \ : \ \|G\theta\| < (1-ts)M) < \frac{N^2}{4}\exp\left(-\frac{t^2k}{800}\right),$$

for all $t \in (0,1)$. Thus, for any $\varepsilon \in (0,1)$ assuming that $\log N \leq \varepsilon^2 k/4$ we have with positive probability (greater than 1/2) that there exists a Gaussian matrix G such that:

$$(1 - s\varepsilon)M \leq \|G\theta\| \leq (1+\varepsilon)M,$$

for all $\theta \in \Theta$. The required map P is given by $P := \frac{1}{M}G$. $\square$

Following the same idea of proof, but using (3.4) instead, we derive:

Proposition 3.5 (One-sided J-L). *Let $X = (\mathbb{R}^n, \|\cdot\|)$ be normed space and let $T \subseteq \ell_2$ be a finite set with $T = \{u_1, \ldots, u_N\}$. Let $\varepsilon \in (0,1)$ and assume that $\log|T| \lesssim \log(1/\varepsilon)/\beta(X)$. Then, the random Gaussian matrix $G = (g_{ij})_{i,j=1}^{N,n}$ satisfies:*

$$\|Gu_i - Gu_j\| \gtrsim \varepsilon \cdot \mathbb{E}\|Z\|_X \cdot \|u_i - u_j\|_2,$$

for all $i, j, = 1, \ldots, N$, where $Z \sim N(\mathbf{0}, I_n)$, with probability greater than $1 - c\varepsilon^{c/\beta(X)}$.

3.3 The random version of Dvoretzky's theorem

The classical Dvoretzky theorem asserts that every centrally symmetric convex body on $\mathbb{R}^n$ has a linear image which in turn has a central section of relatively large dimension which is almost Euclidean. The optimal form of the theorem, in terms of the dimension, was proved by V. Milman in [16]. His pioneering work put forth the concentration of measure phenomenon and established it as the main tool for the study of high-dimensional structures. Milman's random formulation reads as follows: For any $\varepsilon \in (0,1)$ there exists a $c(\varepsilon) > 0$ with the following property: For any centrally symmetric convex body A on $\mathbb{R}^n$, there exists $k \geq c(\varepsilon)k(A)$ such that the random k-dimensional subspace F satisfies

$$\frac{1}{(1+\varepsilon)M_1} B_F \subseteq A \cap F \subseteq \frac{1}{(1-\varepsilon)M_1} B_F,$$

provided that the *critical dimension* $k(A) = (\mathbb{E}\|Z\|_A)^2/b(A)^2$ is large enough and $M_1 \equiv M_1(A) = \int_{S^{n-1}} \|\theta\|_A \, d\sigma(\theta)$. Here the randomness is considered with respect to the Haar probability measure on the Grassmann space $G_{n,k}$ of all k-dimensional subspaces. Milman's approach yields the dependence $c(\varepsilon) \simeq \varepsilon^2/\log(1/\varepsilon)$. This is improved to $c(\varepsilon) \geq c\varepsilon^2$ by Gordon in [4] and an alternative approach is given by Schechtman in [22]. This dependence is known to be optimal.

Dvoretzky's theorem can also be viewed as the "continuous version" of the J-L principle (see [24] for a unified approach). Roughly speaking, if we consider the set $T \subseteq S^{k-1}$ to be dense enough then we can embed the whole sphere S^{k-1}. Since the mapping is linear this entails in an almost isometric embedding of ℓ_2^k into X.

Next, we state our version of the randomized Dvoretzky theorem with a refined one-sided estimate.[4]

Theorem 3.6 (Randomized Dvoretzky). *For every $\varepsilon \in (0,1)$ there exists $\eta = \eta(\varepsilon)$ with the following property: For any centrally symmetric convex body A on $\mathbb{R}^n$ (with $k_*(A) \gg 1$) there exists a set $\mathcal{F}$ in $G_{n,k}$ with $k := \lfloor \eta k_*(A) \rfloor$, which satisfies $\nu_{n,k}(\mathcal{F}) \geq 1 - e^{-c\varepsilon^2 k_*}$ and for any $E \in \mathcal{F}$ we have*

$$w(A)\left(1 - \frac{\varepsilon\sqrt{\beta_* k_*}}{\log(1/\varepsilon)} \log\left(\frac{e}{\varepsilon\sqrt{k_*\beta_*}}\right)\right) B_E \subseteq P_E(A) \subseteq \left(1 + \frac{\varepsilon}{\sqrt{\log(1/\varepsilon)}}\right) w(A) B_E.$$

We may take $\eta(\varepsilon) \simeq \varepsilon^2/\log\frac{1}{\varepsilon}$.

Before we proceed with the proof of the theorem we would like to illustrate its concept in two classical cases (for the related estimates in the examples the reader is consulted in [21, Section 2 & 3]):

i. *The ℓ_1 ball (crosspolytope).* We have $k_*(\ell_1^n) = k(\ell_\infty^n) \simeq \log n$ and $\beta_*(\ell_1^n) = \beta(\ell_\infty^n) \simeq (\log n)^{-2}$. The above theorem yields that the random $\log n$-dimensional subspace E satisfies:

$$\left(1 - c\frac{\log\log n}{\sqrt{\log n}}\right) w(B_1^n) B_E \subseteq P_E(B_1^n) \subseteq 2w(B_1^n) B_E,$$

with probability greater than $1 - Ce^{-c\log n}$.

[4] The $*$ quantities are referred to the corresponding ones for the polar body, i.e. $\beta_*(A) \equiv \beta(A^\circ)$, $k_*(A) \equiv k(A^\circ)$ and $w(A) = M_1(A^\circ)$.

ii. *The ℓ_p ball; $p = 4/3$.* We have $k_*(B_p^n) = k(B_4^n) \simeq n^{1/2}$ and $\beta_*(B_p^n) = \beta(B_4^n) \simeq 1/n$. Then Theorem 3.6 shows that the random $n^{1/2}$-dimensional subspace F satisfies:

$$\left(1 - c\frac{\log n}{n^{1/4}}\right) w(B_p^n) B_F \subseteq P_F(B_p^n) \subseteq 2w(B_p^n) B_F,$$

with probability greater than $1 - Ce^{-c\sqrt{n}}$.

For the proof of Theorem 3.6 we focus on the lower inclusion (the upper inclusion follows by the work of Gordon [4] – see Schechtman [22] for an alternative proof). To this end we shall need several auxiliary results. The first lemma, which can be proved by using integration in polar coordinates, goes back to [9]:

Lemma 3.7. *Let A be symmetric convex body on $\mathbb{R}^m$. Then, for any $p > 0$ we have:*

$$[\mathrm{v.rad}(A)]^m \, J_p^p(A) = \frac{m}{m+p} M_{-(m+p)}^{-(m+p)}(A),$$

where $J_p(A) := \left(\frac{1}{|A|}\int_A \|x\|_2^p \, dx\right)^{1/p}$ and $M_{-q}^{-q}(A) := \int_{S^{m-1}} \|\theta\|_A^{-q} \, d\sigma(\theta)$.

Next lemma is essentially from [9]. The proof we present here is due to Rudelson (see [10, Lemma 2.4]):

Lemma 3.8. *Let K be symmetric convex body on $\mathbb{R}^n$. Then, for all $q > 0$ we have:*

$$J_q(K) := \left(\frac{1}{|K|}\int_K \|x\|_2^q \, dx\right)^{1/q} \geq a_{n,q}^{-1} R(K), \quad a_{n,q}^{-q} := \frac{q}{2} B(q, n+1),$$

where $R(K) := \max_{x \in K} \|x\|_2$.

Proof. Let $R = R(K)$. Then we have the following:

Claim. For any $0 < \varepsilon < 1$ we have:

$$|\{x \in K : \|x\|_2 \geq \varepsilon R\}| \geq (1-\varepsilon)^n |K|/2.$$

Proof of Claim. (M. Rudelson). Let $x_0 \in K$ such that $R = \|x_0\|_2$. Then, there exists $x_0^* \in S^{n-1}$ such that $\langle x_0, x_0^* \rangle = \|x_0\|_2 = R$. We easily check that:

$$\varepsilon x_0 + (1-\varepsilon)\{x \in K : \langle x, x_0^* \rangle \geq 0\} \subseteq \{x \in K : \|x\|_2 \geq \varepsilon R\}.$$

Taking volumes we obtain:

$$|\{x \in K : \|x\|_C \geq \varepsilon R\}| \geq (1-\varepsilon)^n |K^+| = (1-\varepsilon)^n |K|/2,$$

where $K^+ = \{x \in K : \langle x, x_0^* \rangle \geq 0\}$. This proves the claim.
Now we may write:

$$J_q^q(K) = \frac{qR^q}{|K|}\int_0^\infty t^{q-1} |\{x \in K : \|x\|_2 \geq tR\}| \, dt \geq \frac{qR^q}{2}\int_0^1 t^{q-1}(1-t)^n \, dt,$$

where we have used the Claim. The proof is complete. □

Remark. Let us note that by using standard asymptotic estimates for the Beta function we have:

$$a_{n,q} \leq \exp\left(\frac{cn}{q}\log\left(\frac{eq}{n}\right)\right) \tag{3.7}$$

for all $q \geq n$.

Proposition 3.9. *Let A be a centrally symmetric convex body on $\mathbb{R}^n$. For any matrix $T = (t_{ij})_{i,j=1}^{k,n} \in \mathbb{R}^{k\times n}$ we define the map $T \mapsto \psi(T) := w(TA)$. Then, ψ enjoys the following properties:*

i. *The map ψ is Lipschitz on $(\mathbb{R}^{k\times n}, \|\cdot\|_{\mathrm{HS}})$ with:*

$$|\psi(T) - \psi(S)| \leq c_1 \frac{R(A)}{\sqrt{k}} \|T - S\|_{\mathrm{HS}}, \tag{3.8}$$

for all $T = (t_{ij}), S = (s_{ij}) \in \mathbb{R}^{k\times n}$.

ii. *If $G = (g_{ij}) \in \mathbb{R}^{k\times n}$ has i.i.d. standard Gaussian entries, then*

$$\mathbb{E}[\psi(G)] = \mathbb{E}[h_A(Z)] = w(A)\mathbb{E}\|Z\|_2, \tag{3.9}$$

where Z is an n-dimensional standard Gaussian vector.

iii. *For $G = (g_{ij})$ as before we have:*

$$\frac{(\mathbb{E}[\psi(G)]^r)^{1/r}}{\mathbb{E}[\psi(G)]} \leq \sqrt{1 + \frac{c_2 r}{kk_*(A)}}, \tag{3.10}$$

for any $r \geq 1$.

Proof. (i) Note that for any matrix $T = (t_{ij})$ we may write:

$$\psi(T)\mathbb{E}\|Y\|_2 = \mathbb{E}[h_{TA}(Y)] = \mathbb{E}[h_A(T^*Y)],$$

where $Y \sim N(\mathbf{0}, I_k)$. Thus, we get:

$$\begin{aligned}\mathbb{E}\|Y\|_2|\psi(T) - \psi(S)| \leq \mathbb{E}|h_A(T^*Y) - h_A(S^*Y)| &\leq R(A)\mathbb{E}\|(T-S)^*Y\|_2 \\ &\leq R(A)\cdot\|(T-S)^*\|_{\mathrm{HS}}.\end{aligned}$$

The assertion follows once we recall that $\mathbb{E}\|Y\|_2 \simeq \sqrt{k}$.

(ii) Follows by the invariance of the Gaussian measure under orthogonal transformations and integration in polar coordinates.

(iii) Follows by [21, Lemma 2.6] and the assertion (i). □

Proof of Theorem 3.6. Let $p > 0$ and $1 \le k \le n-1$ with $k + p \le n-1$. Using Lemma 3.8 and Lemma 3.7 we may write:

$$\begin{aligned}
\mathbb{E}[r(GA)]^{-p/2} &\le a_{k,p}^{p/2}\mathbb{E}[J_p^{p/2}((GA)^\circ)] \\
&= \frac{pa_{k,p}^{p/2}}{k+p}\mathbb{E}[w_{-(k+p)}^{-(k+p)/2}(GA)] \cdot [\mathrm{vrad}((GA)^\circ)]^{-k/2} \\
&\le a_{k,p}^{p/2}\left(\mathbb{E}[w_{-(k+p)}^{-(k+p)}(GA)]\right)^{1/2} \cdot \left(\mathbb{E}[\mathrm{vrad}((GA)^\circ)]^{-k}\right)^{1/2} \\
&\le a_{k,p}^{p/2}\left(\mathbb{E}[h_A(Z)]^{-(k+p)}\right)^{1/2}\left(\mathbb{E}[w(GA)]^{k}\right)^{1/2} \\
&\le a_{k,p}^{p/2}\left(\mathbb{E}[h_A(Z)]^{-(k+p)}\right)^{1/2}(\mathbb{E}[h_A(Z)])^{k/2}\, e^{c/k_*},
\end{aligned}$$

where we have applied the inequality $w(K)\mathrm{vrad}(K^\circ) \ge 1$ (following by Hölder's inequality) and Proposition 3.9. Using Corollary 3.3 and restricting ourselves in the case $1 \le k \le p \le c_1/\beta_*$ we obtain:

$$\begin{aligned}
\left(\mathbb{E}[r(GA)]^{-p/2}\right)^{2/p} &\le \frac{a_{k,p}}{\mathbb{E}h_A(Z)}\exp\left(\frac{c}{pk_*} + c\sqrt{\beta_*} + cp\beta_*\right) \\
&\le \frac{1}{\mathbb{E}h_A(Z)}\exp\left(\frac{ck}{p}\log\left(\frac{ep}{k}\right) + \frac{c}{pk_*} + c\sqrt{\beta_*} + cp\beta_*\right),
\end{aligned}$$

where in the last step we have also used the estimate (3.7). Given $\varepsilon \in (0,1)$ we choose $p := c_1\varepsilon\sqrt{\beta_* k_*}/\beta_*$ and $k := p\varepsilon\sqrt{\beta_* k_*}/\log\frac{e}{\varepsilon}$ to get:

$$\left(\mathbb{E}[r(GA)]^{-p/2}\right)^{2/p} \le \frac{e^D}{\mathbb{E}h_A(Z)},$$

where $D := \frac{c_2\varepsilon\sqrt{\beta_* k_*}}{\log(e/\varepsilon)}\log(\frac{e}{\varepsilon\sqrt{\beta_* k_*}})$. Markov's inequality shows that the set:

$$\mathcal{F}_1 := \left\{G = (g_{ij})_{i,j=1}^{k,n} \,:\, r(GA) \ge e^{-2D}\mathbb{E}h_A(Z)\right\},$$

with $k \simeq \varepsilon^2 k_*/\log\frac{1}{\varepsilon}$ satisfies

$$\mathbb{P}(\mathcal{F}_1^c) \le \exp\left(-c_1' pD\right) \le \exp(-c_2\varepsilon^2 k_*).$$

In order to translate the above conclusion to the Grassmann space $G_{n,k}$ we need one more step: First recall the fact that the random Gaussian matrix $G = (g_{ij})_{i,j=1}^{k,n}$ satisfies for any $t \in (0,1)$, with probability greater than $1 - e^{-c_3 t^2 n}$, that

$$(1-t)\|x\|_2 \le (\mathbb{E}\|Z\|_2)^{-1}\|G^* x\|_2 \le (1+t)\|x\|_2,$$

for all $x \in \mathbb{R}^k$. Apply this for $t = D$ we find that the set

$$\mathcal{F}_2 := \left\{G = (g_{ij})_{i,j=1}^{k,n} \,:\, \|G^* x\|_2 \le (1+D)\cdot\mathbb{E}\|Z\|_2 \cdot \|x\|_2 \;\forall x \in \mathbb{R}^k\right\}$$

satisfies

$$\mathbb{P}(\mathcal{F}_2^c) \leq \exp(-c_3 D^2 n) \leq \exp(-c_4 \varepsilon^2 k_*).$$

Moreover, we can easily see that:

$$\mathcal{F}_1 \cap \mathcal{F}_2 \subseteq \left\{G = (g_{ij})_{i,j=1}^{k,n} : h_A(z) \geq e^{-3D} w(A) \|z\|_2 \ \forall z \in \mathrm{Im} G^*\right\}.$$

Finally, the fact that the $\mathrm{Im}G^* = G^*(\mathbb{R}^k) = F$ is distributed over the $G_{n,k}$ with respect to the Haar probability measure $\nu_{n,k}$ (generated by the action of the orthogonal group $O(n)$ onto $G_{n,k}$), completes the proof. □

4 Further remarks

We end this note with some concluding comments that arise from our work.

1. We should stress the fact that in the statement of Theorem 2.1 we refer to convex functions in $L_{1,\infty}$. Thus it is pointless to ask about a similar upper estimate other than the weak L_1 estimate. However for all applications presented in Section 3 the functions under consideration are norms or more generally Lipschitz functions which are known to belong in $L_{\psi_2}(\gamma_n)$. In fact $\|f - M\|_{\psi_2} \lesssim \mathrm{Lip}(f)$. Moreover we have mentioned that there are many examples of norms for which β is much smaller than k. Therefore, it is natural to ask if there is one-sided concentration estimate (in the large regime) which takes into account both the parameter β and the Lipschitz constant. A naive approach which puts these remarks together is to combine Chebyshev's inequality with the concentration estimate in terms of the Lipschitz constant:

$$\mathbb{P}(|f(Z) - M| > t) \leq \exp\left(-\tfrac{1}{2} \max\left\{\log\left(t / \sqrt{\mathrm{Var} f(Z)}\right), t^2/L^2\right\}\right).$$

Even in the case of a norm with $k \simeq \log(1/\beta)$ this bound depends continuously on $t > 0$ and seems to be the right one. Example of such a norm is the ℓ_p norm on $\mathbb{R}^n$ with $p = c_0 \log n$, for sufficiently small absolute constant $c_0 > 0$ (see [21, Section 3]).

2. (Non-optimality in ℓ_∞^n). Note that Theorem 3.1 for the case of $A = B_\infty^n$ only yields:

$$\mathbb{P}(\|Z\|_\infty < (1-\varepsilon)M) \leq \frac{1}{2} e^{-c\varepsilon^2 \log^2 n},$$

for all $0 < \varepsilon < 1$. This estimate is far from being the sharp one: It is known (see [23, Claim 3]) that one has:

$$\exp(-Ce^{c'\varepsilon \log n}) \leq \mathbb{P}(\|Z\|_\infty < (1-\varepsilon)M) \leq C \exp(-ce^{c\varepsilon \log n}),$$

for all $0 < \varepsilon < 1/2$.

3. (Optimality in ℓ_p^n, $1 \leq p < \infty$). In [21] it is proved that for any $1 \leq p < \infty$ one has $\beta(\ell_p^n) \lesssim_p 1/n$ (see also [20] for an extension of this result to any finite dimensional subspace of L_p). On other hand, for any norm $\|\cdot\|$ on $\mathbb{R}^n$ we can deduce that:

$$\mathbb{P}\left(\|Z\| < (1-\varepsilon)\mathbb{E}\|Z\|\right) \geq c \exp(-C\varepsilon^2 n),$$

for all $0 < \varepsilon < 1$. Therefore, we obtain:

$$\mathbb{P}\left(\|Z\|_p < (1-\varepsilon)M_{p,n}\right) \geq c' \exp\left(-C_p \varepsilon^2 \beta(B_p^n)\right),$$

where $M_{p,n}$ is a median for $\|Z\|_p$.

Acknowledgments. We are grateful to Ramon van Handel for useful discussions.

References

[1] V. Bogachev, *Gaussian measures*, Mathematical Surveys and Monographs **62**, American Mathematical Society (1998).

[2] D. Cordero-Erausquin, M. Fradelizi and B. Maurey, *The (B) conjecture for the Gaussian measure of dilates of symmetric convex sets and related problems*, J. Funct. Anal. **214** (2004), no. 2, 410–427.

[3] A. Ehrhard, *Symmétrisation dans l' espace de Gauss*, Math. Scand. **53**, (1983), 281–381.

[4] Y. Gordon, *Some inequalities for Gaussian processes and applications*, Israel J. Math. **50** (1985), 265–289.

[5] L. Grafakos, *Classical and Modern Fourier Analysis*, Pearson Education, Inc., (2004).

[6] P. Indyk, *Algorithmic applications of low-distortion geometric embeddings*, In 42nd IEEE Symposium on Foundations of Computer Science (Las Vegas, NV, 2001), 10–33. IEEE Computer Soc., Los Alamitos, CA (2001).

[7] W. B. Johnson and J. Lindenstrauss, *Extensions of Lipschitz mappings into a Hilbert space*, in Conference in modern analysis and probability (New Haven, Conn., 1982). Providence, RI: Amer. Math. Soc., (1984), 189–206.

[8] W. B. Johnson and A. Naor, *The Johnson-Lindenstrauss lemma almost characterizes Hilbert space, but not quite*, Discrete and Computational Geometry **43**, issue 3 (2010), 542–553.

[9] B. Klartag, *A geometric inequality and a low M-estimate*, Proc. Amer. Math. Soc., Vol. **132**, No. 9, (2004), 2919–2628.

[10] B. Klartag, *Uniform almost sub-gaussian estimates for linear functionals on convex sets*, Algebra i Analiz (St. Petersburg Math. Journal), Vol. **19**, no. 1 (2007), 109–148.

[11] B. Klartag and R. Vershynin, *Small ball probability and Dvoretzky's theorem*, Israel J. Math. **157** (2007), 193–207.

[12] E. Kushilevitz, R. Ostrovsky and Y. Rabani, *Efficient search for approximate nearest neighbor in high dimensional spaces*, SIAM J. Comput., **30** (2): 457–474 (electronic), (2000).

[13] S. Kwapien, *A remark on the median and the expectation of convex functions of Gaussian vectors*, Probability in Banach spaces **9** (Sandjberg, 1993), 271–272, Progr. Probab. **35**, Birkhäuser Boston, MA, 1994.

[14] M. Ledoux, *The concentration of measure phenomenon*, Mathematical Surveys and Monographs **89**, American Mathematical Society, Providence, RI, (2001).

[15] R. Latala and K. Oleszkiewicz, *Small ball probability estimates in terms of widths*, Studia Math. **169** (2005), no. 3, 305–314.

[16] V. D. Milman, *New proof of the theorem of A. Dvoretzky on sections of convex bodies*, (Russian), Funkcional. Anal. i Prilozen. **5** (1971) 28–37.

[17] V. D. Milman and G. Schechtman, *Asymptotic theory of finite dimensional normed spaces*, Lecture Notes in Math. **1200** (1986), Springer, Berlin.

[18] V. D. Milman and G. Schechtman, *Global versus Local asymptotic theories of finite-dimensional normed spaces*, Duke Math. Journal **90** (1997), 73–93.

[19] G. Paouris and P. Valettas, *Variance estimates and almost Euclidean structure*, in preparation.

[20] G. Paouris and P. Valettas, *On Dvoretzky's theorem for subspaces of L_p*, preprint (2015), available at `http://arxiv.org/abs/1510.07289`

[21] G. Paouris, P. Valettas and J. Zinn, *Random version of Dvoretzky's theorem in ℓ_p^n*, preprint (2015), available at `http://arxiv.org/abs/1510.07284`

[22] G. Schechtman, *A remark concerning the dependence on ε in Dvoretzky's theorem*, Geometric Aspects of Functional Analysis (1987-88), 274-277, Lecture Notes in Mathematics **1376**, Springer, Berlin (1989).

[23] G. Schechtman, *The random version of Dvoretzky's theorem in ℓ_∞^n*, GAFA Seminar 2004-2005, 265–270, Lecture Notes in Math., **1910**, Springer-Verlag (2007).

[24] G. Schechtman, *Two observations regarding embedding subsets of Euclidean spaces in normed spaces*, Advances in Mathematics **200** (2006) 125–135.

[25] S. S. Vempala, *The random projection method*, DIMACS Series in Discrete Mathematics and Theoretical Computer Science **65**, American Mathematical Society, Providence, RI (2004). With a foreword by Christos H. Papadimitriou.

Grigoris Paouris: `grigoris@math.tamu.edu`
Department of Mathematics, Mailstop 3368
Texas A & M University
College Station, TX 77843-3368

Petros Valettas: `valettasp@missouri.edu`
Mathematics Department
University of Missouri
Columbia, MO 65211