VARIANCE ESTIMATES AND ALMOST EUCLIDEAN STRUCTURE

GRIGORIS PAOURIS AND PETROS VALETTAS

Abstract. We introduce and initiate the study of new parameters associated with any norm and any log-concave measure on $\mathbb{R}^n$, which provide sharp distributional inequalities. In the Gaussian context this investigation sheds light to the importance of the statistical measures of dispersion of the norm in connection with the local structure of the ambient space. As a byproduct of our study, we provide a short proof of Dvoretzky's theorem which not only supports the aforementioned significance but also complements the classical probabilistic formulation.

1. Introduction

The main focus of this note is to establish new distributional inequalities for convex functions with respect to log-concave measures. The new ingredient in these inequalities is that the controlling parameter is the variance rather than the Lipschitz constant or some moment of the length of the gradient of the function. The study of these inequalities has been motivated by the need to quantify efficiently the almost Euclidean subspaces of high-dimensional normed spaces. In particular, we would like to understand the dependence on ε in the almost isometric version of Dvoretzky's theorem [Dvo61]. In fact in this note, we are interested in the randomized Dvoretzky's theorem (as has been established in the seminal work of Milman [Mil71]) which states that for any $\varepsilon \in (0,1)$ there exists a constant $c(\varepsilon) > 0$ with the following property: for any normed space $X = (\mathbb{R}^n, \|\cdot\|)$ the random (with respect to the Haar measure on the Grassmannian) k-dimensional subspace F satisfies

$$(1-\varepsilon)M\|z\|_2 \leqslant \|z\| \leqslant (1+\varepsilon)M\|z\|_2 \tag{1.1}$$

for all $z \in F$ as long as $k \leqslant c(\varepsilon)k(X)$, where M is the average of the norm on the unit Euclidean sphere with respect to the uniform probability measure (see §2 for the related definitions). The parameter $k(X)$ is referred to as the *critical dimension* (or the *Dvoretzky number*) of X given by $k(X) = n(M(X)/b(X))^2$, where $b(X) = \max\{\|\theta\| : \|\theta\|_2 = 1\}$. We will write $k(X,\varepsilon)$ for the maximal dimension for which (1.1) holds with probability at least 1/2. Thus, Milman's formulation yields $k(X,\varepsilon) \geqslant c(\varepsilon)k(X)$. (For historical remarks, required background material and further extensions the reader may consult [MS86], [Pis89], [TJ89] and [AAGM15]). Concerning the dependence on ε, let us mention that Milman's proof yields $c(\varepsilon) \simeq \varepsilon^2/\log\frac{1}{\varepsilon}$. This was subsequently improved by Gordon in [Gor85] showing that one can always have $c(\varepsilon) \simeq \varepsilon^2$ and an alternative approach was presented by Schechtman in [Sch89]. Hence, one has $k(X,\varepsilon) \geqslant c\varepsilon^2 k(X)$. The latter asymptotic formula is optimal up to universal constants as the example of the ℓ_1^n shows. However, there are (several) examples of spaces which show that the function $c(\varepsilon)$ can be significantly improved.

The investigation of the randomized Dvoretzky's theorem on the almost isometric level for special cases of normed spaces was initiated by the works of Schechtman [Sch07] and Tikhomirov [Tik14] who determined asymptotically the dimension $k(\ell_\infty^n,\varepsilon)$. Their estimate $k(\ell_\infty^n,\varepsilon) \simeq \varepsilon|\log\varepsilon|^{-1}\log n$ is better than $\varepsilon^2\log n$ as was predicted by the optimal form of Milman's formula due Gordon's [Gor85]. This result has been extended by Zinn and the authors in [PVZ15] for all ℓ_p^n spaces. Once again the dimension $k(\ell_p^n,\varepsilon)$ is much larger than the estimate derived by the classical proof when $2 < p < \infty$. The proof is based on the fact that the ℓ_p norm is much more concentrated on Gauss' space than the usual "concentration of measure" suggests. In that case we will say that the norm is "overconcentrated".

Date: October 2016; last changes: March 2017. Preliminary version.

2010 *Mathematics Subject Classification.* Primary 46B07, 46B09; Secondary 52A21, 52A23.

Key words and phrases. Almost Euclidean sections, Grassmannian, Concentration of norms, Small ball estimates, log-concave measures.

The first author was supported by the NSF CAREER-1151711 grant.

The second author was supported by the NSF grant DMS-1612936.

This phenomenon is not only observed in the ℓ_p^n spaces, but is apparent in many other cases. In [PV15] we have shown that for all $2 < p < \infty$ and for any n-dimensional subspace X of L_p, in Lewis' position [Lew78], either X almost spherical sections of proportional dimension (that is $k(X) \simeq n$) or X is overconcentrated. Furthermore, among all n-dimensional subspaces X of L_p, $2 < p < \infty$ the worst $k(X, \varepsilon)$ occurs for the ℓ_p^n's. Recently Tikhomirov [Tik17] proved that for every 1-unconditional n-dimensional normed space X, in ℓ-position (see e.g. [TJ89] for the definition of the ℓ-position), either $k(X)$ is polynomial with respect to n or X is overconcentrated. His approach shows that the worst $k(X, \varepsilon)$ in this class of spaces is attained for the ℓ_∞^n.

In some of the above cases, the new observed phenomenon is due to superconcentration (as defined by Chatterjee in [Cha14]) and can be quantified by employing Talagrand's $L_1 - L_2$ bound for gaussian measure (see e.g. [Tal94] and [CEL12]). That was the crucial tool in the investigation in [PVZ15] as well as in [Tik17]. However, the superconcentration is not the only reason that causes the norm of a space to be more concentrated than expected. All the aforementioned cases share the common feature that $\mathrm{Var}\|Z\| \ll \mathrm{Lip}(\|\cdot\|)^2$ where Z is an n-dimensional standard gaussian vector (recall that the classical gaussian concentration yields the bound $\mathrm{Var}\|Z\| \leqslant \mathrm{Lip}(\|\cdot\|)^2$).

Our aim in this note is to put on display the importance of the (normalized) variance in the study of the almost Euclidean structure in high-dimensional normed spaces and to initiate a systematic investigation of the concentration properties of convex functions with respect to this measure of dispersion. This investigation has two main directions. Firstly, we show that this new parameter can be used to provide a short proof of Dvoretzky's theorem which can be viewed as the probabilistic and quantitative version of the topological proof due to Figiel [Fig76] and Szankowski's analytic proof from [Sza74]. Further study of this parameter is also considered and is compared with the classical Dvoretzky number.

Secondly, we work on distributional inequalities in the context of log-concave measures. We show that (one-sided) deviation estimates, where the variance is involved, can be proved for any pair of a convex function and a log-concave probability measure on Euclidean space, by extending a machinery developed in [PV16].

The rest of the paper is organized as follows: In Section 2 we fix the notation and we recall standard background material. In Section 3 we present a simple proof that every infinite dimensional Banach space contains ℓ_2^n's uniformly. The local version of our approach uses the parameter of the normalized variance and yields optimal dependence on the size of the distortion. In Section 4 we study further the aforementioned parameter and we compare it with the classical Dvoretzky number. Our approach uses concentration estimates for the mean width of random projections of any convex body in $\mathbb{R}^n$. We postpone the somewhat more systematic study of this topic to §6. In Section 5 we study, in the general context of log-concave measures small deviation and small ball estimates for norms whose tightness is quantified in terms of the parameter of the normalized variance. Finally, in Section 6 we study independently distributional inequalities for the mean width of random projections of any convex body in $\mathbb{R}^n$ in terms of the Haar measure on the Grasmannian.

2. Notation and background material

We work in $\mathbb{R}^n$ equipped with a Euclidean structure $\langle \cdot, \cdot \rangle$. The Euclidean norm is given by $\|x\|_2 = \sqrt{\langle x, x \rangle}$, $x \in \mathbb{R}^n$. For the $(n-1)$-dimensional Euclidean sphere we write $S^{n-1} = \{x \in \mathbb{R}^n : \langle x, x \rangle = 1\}$. Let also σ be the uniform probability measure on S^{n-1}.

The orthogonal group on $\mathbb{R}^n$ is denoted by $O(n)$ and consists of all matrices which preserve the angles, i.e. $\langle Ux, Uy \rangle = \langle x, y \rangle$ for all $x, y \in \mathbb{R}^n$. Thus,

$$O(n) = \{U \in \mathbb{R}^{n \times n} : UU^* = I\}.$$

The action of $O(n)$ on itself generates the Haar probability measure which we will denote by μ_n.

Let $SO^{\pm}(n)$ be the collection of all matrices which are volume preserving,

$$SO^{\pm}(n) = \{U \in O(n) : \det U = \pm 1\}.$$

Similarly, we may define the Haar probability measure on the special orthogonal group $SO(n)$. This is nothing more than the restriction of μ_n in $SO(n)$ (see e.g. [Mec14] for background details), thus we will still denote it by μ_n. Note that although $SO^-(n)$ is not a group itself we will refer to the restricted measure, with some abuse of terminology, as the Haar measure on $SO^-(n)$ since it is still invariant (within $SO^-(n)$) under the action of $SO^+(n)$. One may check that if U is Haar distributed in $SO^+(n)$ and M is a fixed matrix in $SO^-(n)$, then UM is Haar distributed in $SO^-(n)$. The Hilbert-Schmidt norm of a matrix T is denoted by $\|T\|_{\rm HS}$ and the operator norm of T with $\|T\|_{\rm op}$.

The Grassmann space $G_{n,k}$ is the set of all k-dimensional subspaces of $\mathbb{R}^n$. We consider the Haar measure $\nu_{n,k}$ on $G_{n,k}$ which is invariant under the action of the orthogonal group $O(n)$. For an arbitrary norm on $\mathbb{R}^n$ we write $\|\cdot\|$ and for the normed spaces $(\mathbb{R}^n, \|\cdot\|)$ we will use the letters X, Y.

The random variables will be denoted by $\zeta, \eta, \xi, \ldots$. For random vectors on $\mathbb{R}^n$, distributed according to some law μ, we write Z, W or $W = (w_1, \ldots, w_n)$. For any random variable ξ on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with $\xi \in L_2(\Omega)$ and $\mathbb{E}\xi \neq 0$ we define the parameter $\beta(\xi)$ as follows:

$$\beta(\xi) := \frac{\mathrm{Var}(\xi)}{(\mathbb{E}\xi)^2}, \tag{2.1}$$

where $\mathbb{E}$ stands for the expectation and Var for the variance. For any normed space $X = (\mathbb{R}^n, \|\cdot\|)$ and for any Borel probability measure μ on $\mathbb{R}^n$ we define the parameter:

$$\beta_\mu(X) := \beta(\|Z\|), \tag{2.2}$$

where Z is a random vector distributed according to μ.

Recall the well known fact that if ξ is a random variable in some probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with $\mathbb{E}\xi^2 < \infty$ and $m = \mathrm{med}(\xi)$ is a median of ξ, then

$$|\mathbb{E}\xi - m| \leqslant \mathbb{E}|\xi - m| \leqslant \sqrt{\mathrm{Var}(\xi)}.$$

If $\mathbb{E}\xi \neq 0$ then we readily get:

$$\left| \frac{m}{\mathbb{E}\xi} - 1 \right| \leqslant \sqrt{\beta(\xi)}.$$

This says that the parameter β quantifies how close are the measures of central tendency, median and expectation.

Let us note that for probability space $(\mathbb{R}^n, \mu)$, where μ is a log-concave probability measure and the variable ξ is $\|Z\|$, where $\|\cdot\|$ is a norm on $\mathbb{R}^n$ and Z is distributed according to μ, Borell's lemma [Bor75] (see also §5) implies $m \simeq \mathbb{E}\|Z\|$. Thus, if we are only interested in the order of magnitude of the parameter β, the expectation can be sufficiently replaced by the median. For general facts concerned with gaussian measures we refer the reader to [Bog98].

A convex body K in $\mathbb{R}^n$ is a compact, convex set with non-empty interior. The convex bodies will be denoted with $A, K, L, \ldots$. A convex body K is said to be symmetric when $x \in K$ if and only if $-x \in K$. For any symmetric convex body K we write $\|\cdot\|_K$ for the norm (gauge function) induced by K. Therefore, very frequently we identify notationally the normed space, whose norm is generated by a convex body, with the convex body itself. The volume (i.e. the Lebesgue measure) of a convex body K in $\mathbb{R}^n$ is denoted by $|K|$. We define the circumradius of K, as the number $R(K) := \max_{x \in B_2^n} \|x\|_K$. That is the smallest $R > 0$ for which $K \subseteq RB_2^n$. For any symmetric convex body K in $\mathbb{R}^n$ we define the following averages:

$$J_q(K) := \left(\frac{1}{|K|} \int_K \|x\|_2^q \, dx \right)^{1/q}, \quad -n < q \neq 0.$$

The next lemma is essentially from [Kla04]. The idea of the proof goes back to Rudelson (see [Kla07]):

Lemma 2.1. *Let K be a symmetric convex body in $\mathbb{R}^n$. Then, for all $q > 0$ we have:*

$$J_q(K) \geqslant a_{n,q}^{-1} R(K), \quad a_{n,q}^{-q} := \frac{q}{2} B(q, n+1).$$

Remarks 2.2. (i) Let us note that by using standard asymptotic estimates for the Beta function we have:

$$a_{n,q} \leqslant \exp\left(\frac{cn}{q} \log\left(\frac{eq}{n} \right) \right) \tag{2.3}$$

for all $q \geqslant n$.

(ii) The symmetry is not essential for the proof.

Next, for any symmetric convex body K in $\mathbb{R}^n$ we define:

$$\mathrm{vrad}(K) := \left(\frac{|K|}{|B_2^n|}\right)^{1/n}, \quad M_q^q(K) := \int_{S^{n-1}} \|\theta\|_K^q \, d\sigma(\theta), \quad q \neq 0.$$

The proof of the next lemma can be found in [Kla04].

Lemma 2.3. *Let K be a symmetric convex body in $\mathbb{R}^n$. Then, for any $p > -n$ we have:*

$$\mathrm{vrad}(K)^n J_p^p(K) = \frac{n}{n+p} M_{-(n+p)}^{-(n+p)}(K).$$

We denote by $b(K) = \max_{\theta \in S^{n-1}} \|\theta\|_K$. Note that $r(K) = 1/b(K)$ is the maximal radius of the centered inscribed ball in K. The polar body K° of a convex body K is defined as $K^\circ = \{y : \langle x, y\rangle \leqslant 1 \ \forall x \in K\}$. For any symmetric convex body K the support function of K at $\theta \in S^{n-1}$ is the half width of the body in direction θ, i.e. $h_K(\theta) = \max\{\langle x, \theta\rangle : x \in K\}$. Note that h_K is the dual norm of $\|\cdot\|_K$. One may check that $b(K) = R(K^\circ)$ and $R(K)r(K^\circ) = 1$. The Banach-Mazur distance between two isomorphic normed spaces X, Y is denoted by $d(X, Y)$ with

$$d(X, Y) := \inf\{\|T\| \cdot \|T^{-1}\| : T : X \to Y, \text{ linear isomorphism}\}.$$

The reader may consult the monographs [MS86, Pis89, TJ89, AAGM15] for a detailed background information on the local theory of normed spaces.

Throughout the text we make use of $C, c, C_1, c_1, \ldots$ for positive absolute constants whose values may change from line to line. We also introduce the notation $Q_1 \lesssim Q_2$ for any two quantities Q_1 and Q_2 (which may depend on dimension or some geometric parameter of the normed space or the convex body) if there exists absolute constant $C > 0$ such that $Q_1 \leqslant CQ_2$. We write $Q_1 \simeq Q_2$ is $Q_1 \lesssim Q_2$ and $Q_2 \lesssim Q_1$. We write the above signs with a subscript $\lesssim_p, \simeq_p$, if the involved constant depends on the parameter p.

3. A probabilisitic proof of Dvoretzky's theorem without concentration: The variance method

In this section we provide a probabilistic proof that ℓ_2^n's embed uniformly into any infinite dimensional Banach space X without utilizing the concentration of measure. What is crucial in our approach is the use of the new critical parameter $\beta_{\gamma_n}(X)$ defined in (2.2). We investigate this parameter further in §3.

In order to prove the result we will use some standard lemmas such as the Dvoretzky-Rogers lemma and a net argument on the sphere which are rather standard in all probabilistic proofs. However, even if our approach is more elementary leads to better dependence on ε in several interesting cases (see §3.3). Our proposed local form of Dvoretzky's theorem reads as follows.

Theorem 3.1. *For every $\varepsilon \in (0, 1)$ and for any finite-dimensional normed space X, there exists k-dimensional subspace F of X with $k \geqslant c \log \frac{1}{\beta} / \log \frac{1}{\varepsilon}$, such that $d(F, \ell_2^k) < 1 + \varepsilon$, where $\beta = \beta(X)$.*

Later this formulation will be used to compare the new parameter $\beta(X)$ with the Dvoretzky number $k(X)$ of the normed space $X = (\mathbb{R}^n, \|\cdot\|)$. For the proof we will need the following standard lemma (whose proof is included for the sake of completeness).

Lemma 3.2. *Let X be a k-dimensional normed space.*

(i) *For any $\delta \in (0, 1)$ there exists a δ-net $\mathcal{N}$ on S_X with* $\mathrm{card}(\mathcal{N}) \leqslant \left(1 + \frac{2}{\delta}\right)^k$.

(ii) *Let Y be a normed space and let $T : X \to Y$ be a linear mapping with the property:*

$$1 - \varepsilon \leqslant \|Tz\| \leqslant 1 + \varepsilon$$

for all z in a δ-net $\mathcal{N}$ of S_X where $0 < \delta, \varepsilon < 1$. Then, we have:

$$\frac{1 - \varepsilon - 2\delta}{1 - \delta} \leqslant \|T\theta\| \leqslant \frac{1 + \varepsilon}{1 - \delta},$$

for all $\theta \in S_X$.

Proof. The proof of the first assertion can be found in [MS86]. Let us proceed with (ii). Fix $\theta \in S_X$. Then, there exists $z \in \mathcal{N}$ with $\|z - \theta\| < \delta$. Thus, we may write:

$$\|T\theta\| \leqslant \delta\|T\| + \|Tz\| \leqslant \delta\|T\| + (1+\varepsilon). \tag{3.1}$$

Since θ was arbitrary, it follows that $\|T\| \leqslant \delta\|T\| + (1+\varepsilon)$, or equivalently

$$\|T\| \leqslant (1-\delta)^{-1}(1+\varepsilon). \tag{3.2}$$

Plug this back in (3.1) we obtain the right-hand side estimate. For the left hand side we argue as follows:

$$\|T\theta\| \geqslant \|Tz\| - \|T(\theta - z)\| \geqslant (1-\varepsilon) - \delta\|T\| \geqslant (1-\varepsilon) - \frac{\delta(1+\varepsilon)}{1-\delta},$$

by the estimate (3.2). □

Now we are ready to prove the aforementioned form of Dvoretzky's theorem:

Proof of Theorem 3.1. Let $\{g_{ij}(\omega)\}_{i,j=1}^{n,k}$ be i.i.d. standard Gaussian random variables in some probability space $(\Omega, \mathcal{A}, \mathbb{P})$. We consider the random Gaussian operator: $G : \ell_2^k \to X$ with

$$G_\omega z = \sum_{i=1}^{n} \sum_{j=1}^{k} g_{ij}(\omega) z_j e_i, \quad z = (z_1, \dots, z_k) \in \ell_2^k, \ \omega \in \Omega.$$

Note that for fixed $\theta \in S^{k-1}$ if we apply Chebyshev's inequality we obtain:

$$\mathbb{P}\left(\big|\|G\theta\| - \mathbb{E}\|Z\|\big| > \varepsilon \mathbb{E}\|Z\|\right) = \mathbb{P}\left(\big|\|Z\| - \mathbb{E}\|Z\|\big| > \varepsilon \mathbb{E}\|Z\|\right) \leqslant \beta(X)/\varepsilon^2,$$

for $\varepsilon > 0$ and Z standard Gaussian random vector. Now fix $\varepsilon \in (0, 1/3)$ and employ Lemma 3.2 with $\delta = \varepsilon/2$, to get:

$$\mathbb{P}\left(\exists z \in \mathcal{N} : \big|\|Gz\| - \mathbb{E}\|Z\|\big| > \varepsilon \mathbb{E}\|Z\|\right) \leqslant \left(\frac{6}{\varepsilon}\right)^k \frac{\beta}{\varepsilon^2} \leqslant (6/\varepsilon)^{3k}\beta.$$

Therefore, as long as $(6/\varepsilon)^{3k}\beta < 1/2$ or $k \leqslant c \log \frac{1}{\beta} / \log \frac{1}{\varepsilon}$, there exists $\omega \in \Omega$ with the property:

$$\big|\|G_\omega z\| - \mathbb{E}\|Z\|\big| \leqslant \varepsilon \mathbb{E}\|Z\|,$$

for all $z \in \mathcal{N}$. Fix the operator $G = G_\omega$ and let $T : \ell_2^k \to X$ with $T := (\mathbb{E}\|Z\|)^{-1} G$. Then, by Lemma 3.2 and the choice of δ we conclude that:

$$1 - 2\varepsilon < \frac{1 - \varepsilon - 2\delta}{1-\delta} \leqslant \|T\theta\| \leqslant \frac{1+\varepsilon}{1-\delta} < 1 + 2\varepsilon,$$

for all $\theta \in S^{n-1}$. The result follows; for $F := T(\ell_2^k)$ we get $d(F, \ell_2^k) < 1 + 11\varepsilon$. □

3.1. **Quantitative form of Theorem 3.1.** Using John's position [Joh48] and the classical Dvoretzky-Rogers Lemma from [DR50] we can show the following estimate:

Proposition 3.3. *Let $X = (\mathbb{R}^n, \|\cdot\|)$ be a finite-dimensional normed space and assume that the Euclidean ball B_2^n is the ellipsoid of maximal volume inside B_X. Then, we have:*

$$\beta(X) \leqslant \frac{C}{\log n}. \tag{3.3}$$

Therefore, taking into account Proposition 3.3 we readily get the finite representability of ℓ_2 in any infinite dimensional Banach space X:

Corollary 3.4. *For every $\varepsilon \in (0,1)$ and for any n-dimensional normed space X there exists $k \geqslant c \log\log n / \log \frac{1}{\varepsilon}$ and k-dimensional subspace F of X with $d(F, \ell_2^k) < 1 + \varepsilon$.*

Proof of Proposition 3.3. (Sketch). In order to prove the above estimate we need the next lemma from [DR50]:

Lemma 3.5. *Let $X = (\mathbb{R}^n, \|\cdot\|)$ be n-dimensional normed space and let B_2^n be the ellipsoid of maximal volume inside B_X. Then, there exist orthonormal vectors $v_1, \dots, v_m$ with $m \simeq n$ such that $\|v_j\| \geqslant 1/2$ for $j = 1, 2, \dots, m$.*

A proof of this result can be found in [MS86]. Using Lemma 3.5 and an averaging argument over signs (see [MS86] for details) we arrive at the following estimate:

Claim 1. If B_2^n is the ellipsoid of maximal volume inside B_X we have:

$$\mathbb{E}\|Z\| \geqslant c\sqrt{\log n}, \quad Z \sim N(\mathbf{0}, I_n).$$

One more ingredient is needed:

Claim 2. If $\|\cdot\|$ is a norm on $\mathbb{R}^n$ and $b = \max_{\theta \in S^{n-1}} \|\theta\|$, then

$$\mathrm{Var}\|Z\| \leqslant b^2.$$

For the proof of Claim 2 we may employ the (Gaussian) Poincaré inequality:

$$\mathrm{Var}[f(Z)] \leqslant \mathbb{E}\big\|\nabla f(Z)\big\|_2^2. \tag{3.4}$$

The fact that $\langle \nabla\|x\|, v\rangle \leqslant b$ for all $x \in \mathbb{R}^n$ and $v \in S^{n-1}$ proves the claim.

The assertion follows if we combine the previous two claims. □

3.2. **Probabilistic estimate.** The argument in Theorem 3.1 in fact implies that for $k \leqslant c\log(1/\beta)/\log(1/\varepsilon)$ the *random* k-dimensional subspace F of X is $(1+\varepsilon)$-Euclidean with probability $> 1 - e^{-3k}$. Let us first recall the definition:

Let $X = (\mathbb{R}^n, \|\cdot\|)$ be a normed space. A subspace F of X is said to be $(1+\varepsilon)$-*spherical* if

$$\max_{z \in S_F} \|z\| / \min_{z \in S_F} \|z\| < 1 + \varepsilon.$$

In order to verify the aforementioned probabilistic statement, note that the proof of Theorem 3.1 yields the following estimate:

$$P\left(\big|\|G\theta\| - \mathbb{E}\|Z\|\big| \leqslant \varepsilon\mathbb{E}\|Z\|, \ \forall \theta \in S^{k-1}\right) \geqslant 1 - e^{-3k},$$

provided that $k \leqslant \frac{1}{3}\log(1/\beta)/\log(6e/\varepsilon)$. Note that:

$$\left\{\big|\|G\theta\| - \mathbb{E}\|Z\|\big| \leqslant \varepsilon\mathbb{E}\|Z\|, \ \forall \theta \in S^{k-1}\right\} \subseteq \left\{\max_\theta \|G\theta\| / \min_\theta \|G\theta\| < \frac{1+\varepsilon}{1-\varepsilon}\right\} \equiv \mathcal{A}_1$$

and similarly

$$\left\{\forall \theta \in S^{k-1}\,;\, \big|\|G\theta\|_2 - \mathbb{E}\|Z\|_2\big| \leqslant \varepsilon\mathbb{E}\|Z\|_2\right\} \subseteq \left\{\max_\theta \|G\theta\|_2 / \min_\theta \|G\theta\|_2 < \frac{1+\varepsilon}{1-\varepsilon}\right\} \equiv \mathcal{A}_2.$$

Furthermore, the event $\mathcal{A}_1 \cap \mathcal{A}_2$ satisfies:

$$\mathcal{A}_1 \cap \mathcal{A}_2 \subseteq \left\{G(\mathbb{R}^k) \text{ is } \left(\frac{1+\varepsilon}{1-\varepsilon}\right)^2 - \text{spherical}\right\}.$$

Finally, recall that $P(\omega : G_\omega(\mathbb{R}^k) \in \mathcal{B}) = \nu_{n,k}(F \in G_{n,k} : F \in \mathcal{B})$ for any Borel set $\mathcal{B}$ in $G_{n,k}$ and that

$$P\left(\left\{\big|\|G\theta\|_2 - \mathbb{E}\|Z\|_2\big| \leqslant \varepsilon\mathbb{E}\|Z\|_2, \ \forall \theta \in S^{k-1}\right\}\right) > 1 - Ce^{-c\varepsilon^2 n}.$$

Thus, we get:

$$\begin{aligned}\nu_{n,k}\left(\left\{F \text{ is } \left(\frac{1+\varepsilon}{1-\varepsilon}\right)^2 - \text{spherical}\right\}\right) &\geqslant P(\mathcal{A}_1 \cap \mathcal{A}_2) \geqslant 1 - e^{-3k} - Ce^{-c\varepsilon^2 n} \\ &\geqslant 1 - C'e^{-c'k},\end{aligned}$$

since $k \lesssim \varepsilon^2 n$ as long as $\varepsilon \gg n^{-1/2}$ (see §4).

3.3. **Concentration vs Chebyshev's inequality.** The function $c(\varepsilon) \simeq |\log \varepsilon|$ that appears in this simple argument of Theorem 3.1 is a nice feature. In particular, if it is combined with the fact that there exist n-dimensional normed spaces X with critical dimension $k(X) \simeq \log n$ whereas $\beta(X) \lesssim n^{-\alpha}$ for some absolute constant $\alpha > 0$ (see Proposition 4.2), then Theorem 3.1 yields the existence of almost Euclidean subspaces of the same dimension as Milman's formulation [Mil71] provides, but with better dependence on ε. Moreover, for those spaces we conclude that the k-dimensional *random* subspace is $(1+\varepsilon)$-Euclidean with probability $> 1 - e^{-ck}$ as long as $k \leqslant c \log n / \log \frac{1}{\varepsilon}$.

In the light of the above comments, the random version of Dvoretzky's theorem can be complemented in the following way:

Corollary 3.6. *Let X be an n-dimensional normed space. Then, for any $\varepsilon \in (0,1)$ there exists $k \geqslant c \max\left\{\varepsilon^2 k(X), (\log \frac{1}{\varepsilon})^{-1} \log \frac{1}{\beta(X)}\right\}$ such that the random k-dimensional subspace F satisfies:*

$$\frac{1-\varepsilon}{M} B_F \subseteq B_X \cap F \subseteq \frac{1+\varepsilon}{M} B_F,$$

with probability greater than $1 - e^{-ck}$.

3.4. **Two examples on the dependence on $k(X,\varepsilon)$.** The classical results of Milman [Mil71], Gordon [Gor85] and Schechtman [Sch89] predict that the dependence on ε should be ε^2. Here we show that this is always the case after a small perturbation.

Let $f : \mathbb{R}^{n+1} \to \mathbb{R}$ be the mapping:

$$f(t,x) = |t| + \|x\|_\infty, \quad (t,x) \in \mathbb{R} \times \mathbb{R}^n. \tag{3.5}$$

Then, we have the following properties:

i. f is 2-equivalent norm to $\|\cdot\|_{\ell_\infty^{n+1}}$.
ii. $\mathrm{Var}[f(Z)] \simeq 1$.
iii. Moreover, we have: $k(f) \simeq k(\ell_\infty^n) \simeq \log n$.
iv. $\mathbb{E}[f(Z)] \simeq \sqrt{\log n}$.

We also have the following:

Proposition 3.7. *For any $r \geqslant 1$ we have:*

$$\left(\mathbb{E}\,|f(g_1,Z_1) - f(g_2,Z_2)|^r\right)^{1/r} \simeq \sqrt{r},$$

where Z_1, Z_2 are independent, standard Gaussian random vectors on $\mathbb{R}^n$.

Proof. Using the triangle inequality we may write:

$$\begin{aligned}\left(\mathbb{E}\,|f(g_1,Z_1) - f(g_2,Z_2)|^r\right)^{1/r} &\geqslant \left(\mathbb{E}\,||g_1| - |g_2||^r\right)^{1/r} - \left(\mathbb{E}\,|\|Z_1\|_\infty - \|Z_2\|_\infty|^r\right)^{1/r} \\ &\geqslant c_2\sqrt{r} - C_2 \frac{r}{\sqrt{\log n}} \geqslant c_2'\sqrt{r}\end{aligned}$$

for all $1 \leqslant r \leqslant c_3 \log n$. The assertion follows. □

Corollary 3.8. *Let f be as above. Then, we have:*

$$\mathbb{P}\left(\left|f(Z) - \mathbb{E}f(Z)\right| > \varepsilon \mathbb{E}f(Z)\right) \geqslant c_1 e^{-C_1 \varepsilon^2 \log n},$$

for all $\varepsilon > 0$, where $Z \sim N(\mathbf{0}, I_{n+1})$.

Proof. It follows from Proposition 3.7 and the Paley-Zygmund inequality. □

Arguing as in [PVZ15, Section 5] and using Corollary 3.8 we may conclude the following:

Theorem 3.9. *For every n there exists an n-dimensional, 1-unconditional normed space X which is 2-isomorphic to ℓ_∞^n and has the following property: If $k(X,\varepsilon)$ is the maximal dimension k for which the k-dimensional random subspace of X is $(1+\varepsilon)$-Euclidean with probability greater than $1 - e^{-k}$, then $k(X,\varepsilon) \simeq \varepsilon^2 \log n$.*

At this point let us note that recently Tikhomirov [Tik17] showed that there exists 1-unconditional normed space X whose ball is in John's position and the critical dimension $k(X,\varepsilon)$ in the randomized Dvoretzky is of the order $\varepsilon^2 \log n$.

In the classical paper of Figiel, Lindenstrauss and Milman [FLM77], it is proven that for spaces with cotype $2 \leqslant q < \infty$ (see e.g. [MS86] for the related definition) the critical dimension in the randomized Dvoretzky, in John's position, is at least of the order $\simeq_q \varepsilon^2 n^{2/q}$ (after employing Gordon's result from [Gor85]). The next example is concerned with the dependence on ε for spaces having the cotype property. It shows that there exists an 1-unconditional n-dimensional normed space X with cotype q, $2 \leqslant q < \infty$ which has Dvoretzky number $\simeq_q n^{2/q}$, though the critical dimension $k(X,\varepsilon)$ in the randomized Dvoretzky is of the exact order $\simeq_q \varepsilon^2 n^{2/q}$.

Let $2 < q < \infty$. As before, we consider the map $f : \mathbb{R}^{n+1} \to \mathbb{R}$ with

$$f(t,x) = |t| + \|x\|_q, \quad (t,x) \in \mathbb{R} \times \mathbb{R}^n. \tag{3.6}$$

We have the following properties:

i. $\|(t,z)\|_q \leqslant f(t,z) \leqslant 2\|(t,z)\|_q$ for all $t \in \mathbb{R}$ and $z \in \mathbb{R}^n$. Therefore, the q-cotype constant $C_q(X)$ of X satisfies:

$$C_q(X) \simeq C_q(\ell_q^{n+1}) \simeq \sqrt{q},$$

e.g. see [AK16].

ii. If $X = (\mathbb{R}^{n+1}, f(\cdot))$, then we have: $k(X) \simeq q n^{2/q}$.

iii. If Z is a standard Gaussian vector on $\mathbb{R}^{n+1}$, we have: $\mathbb{E}[f(Z)] \simeq \mathbb{E}\|Z\|_q \simeq \sqrt{q} n^{1/q}$.

Recall the following fact proved in [PVZ15]:

Lemma 3.10. *Let $2 < q < \infty$. Then, for all large enough n we have:*

$$\left(\mathbb{E}\big|\|Z_1\|_q - \|Z_2\|_q\big|^r\right)^{1/r} \leqslant c_1(q)\mathbb{E}\|Z_1\|_q \max\left\{\sqrt{\frac{r}{n}}, \frac{r^{q/2}}{n}\right\},$$

for all $r \geqslant 1$, where $c_1(q) > 0$ is a constant depending only on q and Z_1, Z_2 are independent standard Gaussian random vectors on $\mathbb{R}^n$.

Using that we obtain the following:

Proposition 3.11. *Let f be as above. Then, we have:*

$$\left(\mathbb{E}\big|f(g_1,Z_1) - f(g_2,Z_2)\big|^r\right)^{1/r} \geqslant c_2(q)\sqrt{\frac{r}{k(X)}}\mathbb{E}[f(g_1,Z_1)],$$

for all $r \geqslant 1$, where g_1, g_2 are i.i.d. standard normals and Z_1, Z_2 are independent standard Gaussian vectors on $\mathbb{R}^n$.

The proof is similar to that of Proposition 3.7, thus it is omitted. Finally, we get:

Corollary 3.12. *Let f as above. Then, we have:*

$$\mathbb{P}\left(|f(Z) - \mathbb{E}f(Z)| > \varepsilon \mathbb{E}f(Z)\right) \geqslant c e^{-c_3(q)\varepsilon^2 n^{2/q}},$$

for all $\varepsilon > 0$, where $Z \sim N(\mathbf{0}, I_{n+1})$.

The next Theorem is the analogue of Theorem 3.9 in the case of spaces with cotype:

Theorem 3.13. *Let $2 < q < \infty$ there exists a constant $C(q) \gg 1$ with the following property: For any $n \geqslant C(q)$ there exists an n-dimensional, 1-unconditional normed space X which satisfies:*

i. *X has cotype q with $C_q(X) \simeq \sqrt{q}$ and,*

ii. *If for any $\varepsilon \in (0,1)$ we denote by $k(X,\varepsilon)$ the largest dimension k for which the random subspace of X is $(1+\varepsilon)$-Euclidean with probabability greater than $1 - e^{-ck}$, then $k(X,\varepsilon) \simeq_q \varepsilon^2 n^{2/q}$.*

A more detailed study of the instability of the concentration will appear in [PV17].

4. On the parameter β

In this Section we study the parameter β and we show connections with the one-sided inclusion in the randomized Dvoretzky theorem. Let us recall the definition of β for a normed space $X = (\mathbb{R}^n, \|\cdot\|)$ equipped with some log-concave probability measure μ:

$$\beta_\mu(X) = \frac{\mathrm{Var}_\mu \|Z\|}{(\mathbb{E}_\mu \|Z\|)^2}.$$

If the prescribed measure is the Gaussian we will often omit the subscript. The first result shows that the extremal space for the parameter $\beta_{\gamma_n}(X)$ is the Euclidean:

Lemma 4.1. *Let X be n-dimensional normed space. Then, one has:*

$$\beta_{\gamma_n}(X) \geqslant \beta_{\gamma_n}(\ell_2^n) \simeq 1/n.$$

Proof. Using polar coordinates we see that:

$$\mathbb{E}\|Z\|^p = \mathbb{E}\|Z\|_2^p \int_{S^{n-1}} \|\theta\|^p \, d\sigma(\theta),$$

for all $p > 0$. Thus, we may write:

$$\begin{aligned} 1 + \beta(X) = \frac{\mathbb{E}\|Z\|^2}{(\mathbb{E}\|Z\|)^2} &= \frac{\mathbb{E}\|Z\|_2^2}{(\mathbb{E}\|Z\|_2)^2} \frac{\int \|\theta\|^2 \, d\sigma(\theta)}{(\int \|\theta\| \, d\sigma(\theta))^2} \\ &= (1 + \beta(\ell_2^n))(1 + \beta_\sigma(X)). \end{aligned}$$

Since $\beta_\sigma \geqslant 0$, the assertion readily follows. □

The following estimates for the parameter β of the classical spaces can be found in [PVZ15]:

Proposition 4.2. *There exist absolute constants $0 < c < 1 < C$ with the following property: For all $n \geqslant 1$ sufficiently large, one has*

$$\beta_{\gamma_n}(\ell_p^n) \simeq \begin{cases} \frac{2^p}{p^2 n}, & 1 \leqslant p \leqslant c \log n \\ \frac{1}{(\log n)^2}, & C \log n \leqslant p \leqslant \infty \end{cases}.$$

Another class of spaces with this property consists of those which "sit" inside L_p and p might be moderately growing along with the dimension of the underlying space. Namely, if we take into account the estimate for the variance for the subspaces of L_p, proved in [PV15], it follows that for all $1 \leqslant p < \infty$ and for all subspaces X of L_p with $\dim X = n$, there exists a linear image $\tilde{X}$ of X such that:

$$\beta(\tilde{X}) \leqslant \frac{C^p}{n} \quad \text{or} \min_{T \in GL(n)} \beta(TX) \leqslant \frac{C^p}{n}. \tag{4.1}$$

General upper and lower bounds for the global parameter β are given in the next:

Proposition 4.3. *Let X be finite-dimensional normed space. Then, we have:*

$$\log \frac{1}{\beta(X)} \lesssim k(X) \lesssim \frac{1}{\beta(X)}.$$

Note that the upper bound follows from Claim 2 in §3.1. and the definition of $k(X)$. For the lower bound we give a proof which lies on results of independent interest.

For any norm $\|\cdot\|$ (or any symmetric convex body K) on $\mathbb{R}^n$ we define the following collections of subspaces: for $1 \leqslant k \leqslant n-1$ and $0 < \varepsilon, \delta < 1$ let

$$\mathcal{S}_{k,\varepsilon} := \{F \in G_{n,k} \, : \, F \text{ is } (1+\varepsilon)-\text{spherical subpace of } X = (\mathbb{R}^n, \|\cdot\|)\} \tag{4.2}$$

and

$$\mathcal{F}_{k,\delta} := \{E \in G_{n,k} : (1-\delta)M \leqslant \|\phi\| \leqslant (1+\delta)M, \ \forall \phi \in S_E\}. \tag{4.3}$$

Next Proposition shows that these two descriptions are essentially equivalent:

Proposition 4.4. *Let $X = (\mathbb{R}^n, \|\cdot\|)$, let $0 < \varepsilon < 1$ and let $1 \leqslant m \leqslant n-1$. Then, we have:*

i. $\mathcal{F}_{m,\varepsilon/2} \subset \mathcal{S}_{m,2\varepsilon}$.

ii. $\nu_{n,m}(\mathcal{S}_{m,\varepsilon/3}) \leqslant C\exp(-c\varepsilon^2 mk(X)) + \nu_{n,m}(\mathcal{F}_{m,\varepsilon})$.

For the proof we shall use a concentration result for the map $F \mapsto M_F$ which is defined as follows: Fix A a symmetric convex body on $\mathbb{R}^n$. Then $F \mapsto M_F(A) = M(A \cap F)$ for any $F \in G_{n,k}$, $1 \leqslant k \leqslant n-1$. Let us mention that a large deviation estimate for this mapping was proved by Klartag and Vershynin in [KV07, Lemma 3.2]. Their result reads as follows:

Lemma 4.5. *Let A be a symmetric convex body on $\mathbb{R}^n$ and let $1 \leqslant k \leqslant n-1$. Then, we have:*

$$\nu_{n,k}\left(\{F \in G_{n,k} : M_F(A) > ctM(A)\}\right) \leqslant Ce^{-ct^2 k},$$

for all $t > 1$.

Note that this estimate, when $k = 1$, is reduced to the large deviation estimate of the norm $\theta \mapsto \|\theta\|_A$. But in that case the above estimate loses by a term $k(A)$ on the exponent. In fact, one can recover this missing factor and moreover, prove a concentration inequality for this mapping:

Theorem 4.6. *Let A be a symmetric convex body in $\mathbb{R}^n$ and let $1 \leqslant s \leqslant n-1$. Then, we have:*

$$\nu_{n,s}\left(\{F \in G_{n,s} : |M_F(A) - M(A)| > \varepsilon M(A)\}\right) \leqslant C\exp(-cs\varepsilon^2 k(X)),$$

for all $\varepsilon > 0$ where $C, c > 0$ are absolute constants.

We believe that distributional inequalities for functionals on the Grassmannian are interesting on their own right, thus we discuss this topic separately. In fact, we provide two proofs of the latter probabilistic estimate in §6. Taking this for granted we may proceed with the proof of Proposition 4.4.

Proof of Proposition 4.4. We write $k = k(X)$ for short. The first inclusion is trivial. For the second part, note the following inclusion:

$$\mathcal{S}_{m,\varepsilon} \subset \mathcal{A}_{m,\varepsilon} := \{E \in G_{n,m} : (1+\varepsilon)^{-1} M_E \leqslant \|z\| \leqslant (1+\varepsilon)M_E,\ \forall z \in S_E\}.$$

If we define

$$\mathcal{C}_{m,\delta} := \{E \in G_{n,m} : |M_E - M| > \delta M\},\ 0 < \delta < 1,$$

then we have:

$$\begin{aligned}\nu_{n,m}(\mathcal{S}_{m,\varepsilon/3}) \leqslant \nu_{n,m}(\mathcal{A}_{m,\varepsilon/3}) &\leqslant \nu_{n,m}(\mathcal{C}_{m,\varepsilon/3}) + \nu_{n,m}(\mathcal{A}_{m,\varepsilon/3} \setminus \mathcal{C}_{m,\varepsilon/3}) \\ &\leqslant Ce^{-c\varepsilon^2 mk} + \nu_{n,m}(\mathcal{F}_{m,\varepsilon}),\end{aligned}$$

where we have used Theorem 4.6 and the fact that $\mathcal{A}_{m,\varepsilon/3} \setminus \mathcal{C}_{m,\varepsilon/3} \subset \mathcal{F}_{m,\varepsilon}$. □

Proof of Proposition 4.3. (Lower bound). We set $\beta = \beta(X)$. Note that Theorem 3.1 implies that the set

$$\mathcal{S}_{s,1/6} = \{F \in G_{n,s} : F \text{ is } 7/6-\text{spherical subpace of } X\},$$

has $\nu_{n,s}(\mathcal{S}_{s,1/6}) \geqslant 1 - e^{-s}$ for $s \simeq |\log \beta|$. Furthermore, Proposition 4.4 implies that:

$$1 - e^{-c_1 s} \leqslant \nu_{n,s}\left(\left\{E \in G_{n,s} : \frac{M}{2} \leqslant \|z\| \leqslant 2M,\ \forall z \in S_E\right\}\right).$$

Now we may employ a result of Milman and Schechtman from [MS97] (see also [HW16] for a recent development) to conclude. □

Note. Let us mention that the bounds we derive in Proposition 4.3 are sharp (up to constants): For $X = \ell_1^n$ note that we have $k(\ell_1^n) \simeq 1/\beta(\ell_1^n)$. The example of ℓ_p^n with $p = c_0 \log n$ for some sufficiently small absolute constant $c_0 > 0$ (see Proposition 4.2) guarantees the existence of n-dimensional normed spaces X with critical dimension $k(X) \simeq \log n$ and $\beta(X) \lesssim n^{-\alpha}$.

4.1. **Dvoretzky's theorem revisited.** In [PV16] we prove refined random versions of the classical dimension reduction lemma of Johnson and Lindenstrauss [JL84] and of Dvoretzky's theorem due to V. Milman [Mil71], in terms of the parameter β and the critical dimension k. More precisely, given a finite dimensional normed space $X = (\mathbb{R}^n, \|\cdot\|)$ we may define the parameter $\mathbf{ov}(X)$ associated with X as follows:

$$\mathbf{ov}(X)^2 := \beta(X)k(X) = \frac{\mathrm{Var}\|Z\|}{b(X)^2},$$

where $Z \sim N(\mathbf{0}, I_n)$. The applications in [PV16] exhibit improved one-sided behavior which takes into account the parameter $\mathbf{ov}(X)$. It turns out that the parameter $\mathbf{ov}(X)$ measures the sharpness of the concentration of the norm $\|\cdot\|_X$ in terms of the Lipschitz constant; in particular, it measures how "over-concentrated" is the norm (see [PV17]).

Below we give a more detailed formulation of the random version of Dvoretzky's theorem in terms of this parameter. We also provide one-sided, almost isometric inclusion in large dimensions proportional to $1/\beta$. The latter can be viewed as the almost isometric version of the result of Klartag and Vershynin in [KV07].

Theorem 4.7 (Randomized Dvoretzky)**.** *For any symmetric convex body A we set $k_* := k_*(A)$, $\beta_* := \beta_*(A)$ and $[\mathbf{ov}_*(A)]^2 = \beta_*(A)k_*(A)$. Then, for every $\varepsilon \in (0,1)$ we have:*

i. *There exists a set $\mathcal{A}$ in $G_{n,k}$ with $k \simeq \varepsilon^2 k_*$ and $\nu_{n,k}(\mathcal{A}) \geqslant 1 - e^{-c\varepsilon^2 k_*}$ such that*

$$P_E(A) \subseteq (1+\varepsilon)w(A)B_E,$$

for all $E \in \mathcal{A}$.

ii. *There exists a set $\mathcal{B}$ in $G_{n,\ell}$ with $\ell \simeq \frac{\varepsilon^2}{\beta_* \log(1/\varepsilon)}$ and $\nu_{n,\ell}(\mathcal{B}) \geqslant 1 - e^{-c\varepsilon^2/\beta_*}$ such that*

$$P_F(A) \supseteq (1-\varepsilon)w(A)B_F,$$

for all $F \in \mathcal{B}$.

iii. *There exists a set $\mathcal{C}$ in $G_{n,m}$ with $m \simeq \frac{\varepsilon^2 k_*}{\log(1/\varepsilon)}$ and $\nu_{n,m}(\mathcal{C}) \geqslant 1 - e^{-c\varepsilon^2 k_*}$ such that*

$$w(A)\left(1 - \frac{\varepsilon \mathbf{ov}_*(A)}{\log(\frac{1}{\varepsilon})}\log\left(\frac{e}{\varepsilon \mathbf{ov}_*(A)}\right)\right)B_G \subseteq P_G(A) \subseteq \left(1 + \frac{\varepsilon}{\sqrt{\log(1/\varepsilon)}}\right)w(A)B_G,$$

for all $G \in \mathcal{C}$,

where $c > 0$ is an absolute constant.

We shall need some auxiliary results. The first one follows from Klartag-Vershynin estimate (Lemma 4.5) in the dual setting.

Lemma 4.8. *Let A be a symmetric convex body on $\mathbb{R}^n$ and let $1 \leqslant k \leqslant n-1$. Then, we have:*

$$\left(\int_{G_{n,k}} (w(P_F A))^k \, d\nu_{n,k}(F)\right)^{1/k} \leqslant Cw(A),$$

where $C > 0$ is an absolute constant.

The next Lemma is essentially from [KV07]. However, the next formulation, takes into account the magnitude of the constants involved.

Lemma 4.9 (dimension lift)**.** *Let A be a symmetric convex body on $\mathbb{R}^n$ and let $1 \leqslant s \leqslant n-1$. Then, for any $p \geqslant s$ we have:*

$$\left(\int_{G_{n,s}} [r(P_E A)]^{-p/2} \, d\nu_{n,s}(E)\right)^{2/p} \leqslant \frac{e^{cs/p} a_{s,p}}{w(A)} \left(\frac{w(A)}{w_{-2p}(A)}\right)^2,$$

where $c > 0$ is an absolute constant and $a_{s,p}$ is defined in Lemma 2.1.

Proof. Recall that Lemma 2.1 yields:

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

Also Lemma 2.3 implies:

$$\mathrm{vrad}(K)^n J_q^q(K) = M_{-(n+q)}^{-(n+q)}(K).$$

Recall that $1/r(K) = R(K^\circ)$, hence if we apply the above for $K = (P_E A)^\circ$ we get:

$$\begin{aligned}
\int_{G_{n,s}} \frac{1}{r(P_E A)^{p/2}}\, d\nu_{n,s}(E) &\leqslant a_{s,p}^{p/2} \int_{G_{n,s}} J_p^{p/2}((P_E A)^\circ)\, d\nu_{n,s}(E) \\
&= \frac{p a_{s,p}^{p/2}}{s+p} \int_{G_{n,s}} w_{-(s+p)}^{-(s+p)/2}(P_E A) \cdot [\mathrm{vrad}((P_E A)^\circ)]^{-s/2}\ d\nu_{n,s}(E) \\
&\leqslant a_{s,p}^{p/2} \left(\int_{G_{n,s}} w_{-(s+p)}^{-(s+p)}(P_E A)\, d\nu_{n,s}(E) \right)^{1/2} \cdot \left(\int_{G_{n,s}} [\mathrm{vrad}((P_E A)^\circ)]^{-s}\, d\nu_{n,s}(E) \right)^{1/2} \\
&\leqslant a_{s,p}^{p/2} w_{-(s+p)}^{-(s+p)/2}(A) \left(\int_{G_{n,s}} [w(P_E A)]^s\, d\nu_{n,s}(E) \right)^{1/2} \\
&\leqslant a_{s,p}^{p/2} w_{-(s+p)}^{-(s+p)/2}(A) w(A)^{s/2} e^{cs},
\end{aligned}$$

where we have also used inequality $w(K)\mathrm{vrad}(K^\circ) \geqslant 1$ (following by Hölder's inequality), and Lemma 4.8. The result follows. □

Proposition 4.10. *Let A be a symmetric convex body on $\mathbb{R}^n$. Then, we have:*

$$\frac{w(A)}{w_{-q}(A)} \leqslant \frac{I_1(\gamma_n, A^\circ)}{I_{-q}(\gamma_n, A^\circ)} \leqslant \exp\left(C \max\{\sqrt{\beta_*}, q\beta_*\} \right),$$

for all $0 < q < c/\beta_*$, where $\beta_* = \beta_*(A) := \beta(A^\circ)$.

For a proof of the estimate on the Gaussian averages the reader may consult [PV16]. In order to pass to the sphere, recall the well known formula (which follows by integration in polar coordinates):

$$I_r(\gamma_n, A^\circ) = I_r(\gamma_n, B_2^n) w_r(A),$$

for all $-n < r \neq 0$. Thus, Hölder's inequality yields the asserted bound.

Now we turn to proving Theorem 4.7.

Proof of Theorem 4.7. Fix $1 \leqslant s \leqslant n-1$. Using Lemma 4.9 and restricting ourselves to the range $1 \leqslant s \leqslant p \leqslant c_1/\beta_*$ we obtain:

$$\left(\int_{G_{n,s}} [r(P_E A)]^{-p}\, d\nu_{n,s}(E) \right)^{1/p} \leqslant \frac{1}{w(A)} \exp\left(\frac{cs}{p} \log\left(\frac{ep}{s}\right) + c\sqrt{\beta_*} + cp\beta_* \right),$$

where we have also used Proposition 4.10.

(ii) Given $\varepsilon \in (0,1)$ with $\varepsilon \gtrsim \sqrt{\beta_*}$ (otherwise there is nothing to prove) we choose $p \simeq \frac{\varepsilon}{\beta_*}$ and $s \simeq \varepsilon p / \log(1/\varepsilon)$. Then, we obtain:

$$\left(\int_{G_{n,\ell}} [r(P_E A)]^{-p}\, d\nu_{n,\ell}(E) \right)^{1/p} \leqslant \frac{1}{w(A)} \exp\left(C\varepsilon \right),$$

with $\ell = s \simeq \varepsilon^2/(\beta_* \log\frac{1}{\varepsilon})$. The asserted set $\mathcal{B}$ is defined as:

$$\mathcal{B} := \left\{ F \in G_{n,\ell} : r(P_F A) \geqslant w(A) e^{-2C\varepsilon} \right\}.$$

Note that by Markov's inequality we have: $\nu_{n,\ell}(\mathcal{B}^c) \leqslant e^{-c_1 \varepsilon p} \leqslant e^{-c_2 \varepsilon^2/\beta_*}$

(iii) Set $\tau_* = [\mathbf{ov}_*(A)]^2$. Given $\varepsilon \in (0,1)$ (with $\varepsilon \gtrsim 1/\sqrt{k_*}$) we choose $p \simeq \varepsilon\sqrt{\tau_*}/\beta_*$ and $s \simeq p\varepsilon\sqrt{\tau_*}/\log\frac{1}{\varepsilon}$ and we argue as before to get:

$$\mathcal{C} := \left\{ G \in G_{n,m} : r(P_G(A)) \geqslant w(A) e^{-2C\varepsilon\sqrt{\tau_*}\log(\frac{e}{\varepsilon\tau_*})/\log\frac{1}{\varepsilon}} \right\},$$

where $m = s \simeq \varepsilon^2 k_*/\log\frac{1}{\varepsilon}$ and $\nu_{n,m}(\mathcal{C}^c) \leqslant e^{-c_1 p \frac{\varepsilon\sqrt{\tau_*}}{\log(1/\varepsilon)} \log(\frac{e}{\varepsilon\sqrt{\tau_*}})} \leqslant e^{-c_2\varepsilon^2 k_*}$. The proof is complete. □

5. Small deviation estimates and the parameter ϑ

The aim of this paragraph is to prove tight small deviation and small ball probability estimates and reveal the critical parameters that govern such estimates. We will not restrict in Gaussian measure rather we will extend our work in the more general setting of log-concave probability measures. Our starting point is the following one-sided small deviation inequality proved in [PV16]:

Theorem 5.1. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex map with $f \in L_1(\gamma_n)$. Then,*

$$\mathbb{P}\left(f(Z) < m - t\right) \leqslant \frac{1}{2}\exp(-ct^2/\|f-m\|_{L_1}^2), \tag{5.1}$$

for all $t > 0$, where m is a median of $f(Z)$ and $Z \sim N(\mathbf{0}, I_n)$.

The latter can be viewed as a strengthening of the one-sided classical Gaussian concentration, inside the class of convex functions, since one can replace the Lipschitz constant by the a priori smaller quantity of the standard deviation. The main ingredients of the proof can be summarized in the following two observations:

- When f is a convex function, the map $t \mapsto \Phi^{-1} \circ \gamma_n(f \leqslant t)$ is concave, (this follows by Ehrhard's inequality [Ehr83]) and,
- the derivative of the above function at the median m of f with respect to γ_n is relatively large in terms of the standard deviation of f.

Here we extend the above approach in the context of log-concave probability measures. Let μ be a log-concave, Borel probability measure on $\mathbb{R}^n$. Let Z be a random vector on $\mathbb{R}^n$ distributed according to μ, i.e. $\mathbb{P}(Z \in B) = \mu(B)$ for any Borel set $B \subseteq \mathbb{R}^n$.

For any convex function $\psi : \mathbb{R}^n \to \mathbb{R}$ we define the random variable $\xi := \psi(Z)$. We write $F_\xi(t) := \mu(z \in \mathbb{R}^n : \psi(z) \leqslant t) = \mathbb{P}(\psi(Z) \leqslant t)$ for the cumulative distribution function of ξ. The first main step is to replace function Φ by a suitably chosen function that yields the concavity. The next standard Lemma (for a proof e.g. see [Bob99]) shows that in the context of log-concave measures this function can be at least the exponential.

Lemma 5.2. *Let μ be a log-concave probability measure on $\mathbb{R}^n$ and let ψ be a convex map on $\mathbb{R}^n$. If $F(t) = \mathbb{P}(\psi(Z) \leqslant t)$, then we have the following:*

a. *For $t, s \in \mathbb{R}$ and $0 < \lambda < 1$ we have:*

$$F\left((1-\lambda)t + \lambda s\right) \geqslant [F(t)]^{1-\lambda}\,[F(s)]^{\lambda},$$

that is F is log-concave.

b. *If ψ is a semi-norm[1], then we also have:*

$$1 - F\left((1-\lambda)t - \lambda s\right) \geqslant (1-F(t))^{1-\lambda} F(s)^{\lambda}$$

for all $t, s > 0$ and $0 < \lambda < \frac{t}{t+s}$.

Sketch of proof. For b. we may check (using the subadditivity of the seminorm) that

$$(1-\lambda)\{\psi > t\} + \lambda\{\psi \leqslant s\} \subset \{\psi > (1-\lambda)t - \lambda s\},$$

and we use the log-concavity of μ. □

Remark 5.3. Note that (b) easily implies Borell's lemma from [Bor75]: Fix $s > 0$ (say $s = 1$). For any $t > s$ we choose $\lambda \in (0, \frac{t}{t+s})$ such that $(1-\lambda)t - \lambda s = s$, i.e. $\lambda = \frac{t-s}{t+s}$ and hence,

$$1 - F(t) \leqslant F(s)\left(\frac{1-F(s)}{F(s)}\right)^{\frac{t+s}{2s}}, \quad 0 < s < t.$$

The next lemma shows that $f(m)$ can be estimated in terms of the standard deviation of the function, thus fulfills the second main observation in the general context of log-concave measures.

[1] A seminorm $\psi : V \to [0, \infty)$ on a vector space V is a function which is positively homogeneous, i.e. $\psi(\lambda v) = |\lambda|\psi(v)$ for all scalars λ and $v \in V$ and sub-additive, that is $\psi(u+v) \leqslant \psi(u) + \psi(v)$ for all $u, v \in V$.

Lemma 5.4. *Let μ be a log-concave probability measure on $\mathbb{R}^n$ and let $\psi : \mathbb{R}^n \to \mathbb{R}$ be a convex map with $\psi \in L_1(\mu)$. If $F(t) = \mathbb{P}(\psi(Z) \leqslant t)$, $t \in \mathbb{R}$ and $f = F'$, then*

$$f(m) \geqslant \frac{1}{32\|(\psi - m)_+\|_{L_1(\mu)}},$$

where m is a median of ψ with respect toμ.

Proof. Since F is log-concave (Lemma 5.2) we have:

$$\begin{aligned} 2f(m) = (\log F)'(m) &\geqslant \frac{\log F(m+\delta) - \log F(m)}{\delta} \\ &= \frac{1}{\delta}\log\left[1 + 2\mathbb{P}\left(m < \psi(Z) \leqslant m+\delta\right)\right] \\ &\geqslant \frac{1}{\delta}\mathbb{P}\left(m < \psi(Z) \leqslant m+\delta\right) \\ &= \frac{1}{\delta}\left[\frac{1}{2} - \mathbb{P}(\psi(Z) > m+\delta)\right]. \end{aligned}$$

The choice $\delta = 4\|(\psi - m)_+\|_{L_1}$ and Markov's inequality yield the result. □

Note 5.5. In the above argument we may replace the L_1 norm by the difference $M_q - m$ for any q-quantile of $\psi(Z)$ with $q > 1/2$. That is, if $\mathbb{P}(\psi(Z) \leqslant M_q) = q > 1/2$ then $f(m) \geqslant \frac{q-1/2}{M_q - m}$. Note that if $\psi \in L_{p,\infty}(\mu)$ then $M_q - m \leqslant (1-q)^{-1/p}\|(\psi - M)_+\|_{p,\infty}$.

Theorem 5.6. *Let μ be a log-concave probability measure on $\mathbb{R}^n$. Let ψ be a convex function on $\mathbb{R}^n$ and let m be a median for ψ with respect to μ. Then,*

$$\mu\left(\left\{x : \psi(x) < m - t\int |\psi - m|\, d\mu\right\}\right) \leqslant \frac{1}{2}\exp(-t/16),$$

for all $t > 0$.

Proof. Let $F(t) = P(\psi(Z) \leqslant t)$, $t \in \mathbb{R}$ and $f = F'$. Then, by Lemma 5.2 we have:

$$\log F(m-t) - \log F(m) \leqslant -t\frac{f(m)}{F(m)} = -2tf(m).$$

It follows that:

$$F(m-t) \leqslant F(m)\exp(-2tf(m)) = \frac{1}{2}\exp(-2tf(m)), \tag{5.2}$$

for all $t > 0$. Using Lemma 5.4 we get:

$$F(t-m) \leqslant \frac{1}{2}\exp\left(-t/(16\mathbb{E}_\mu|\psi(Z) - m|)\right),$$

as required. □

For any symmetric convex body K in $\mathbb{R}^n$ and for any log-concave probability measure μ on $\mathbb{R}^n$ we write F_K for the cumulative distribution function of $\|Z\|_K$, where Z is distributed according to μ, i.e. $F_K(t) = \mu(x : \|x\|_K \leqslant t)$, $t > 0$. If f_K is the density of this random variable, we define:

$$\vartheta(\mu, K) \equiv \vartheta_\mu(K) = mf_K(m), \tag{5.3}$$

where m is a median for $\|Z\|_K$. With this notation Lemma 5.4 shows that

$$\vartheta_\mu(K) \gtrsim \frac{1}{\sqrt{\beta_\mu(K)}}. \tag{5.4}$$

Below, we also show that ϑ raises naturally in small deviation and small ball estimates for log-concave probability measures on $\mathbb{R}^n$. For this end, let us recall the *B-property.* A pair (μ, K) of a log-concave probability measure on $\mathbb{R}^n$ and a symmetric convex body K in $\mathbb{R}^n$ it is said to have the B-property if the map $t \mapsto \mu(e^t K)$ is log-concave. In [CEFM04], Cordero-Erausquin, Fradelizi and Maurey proved that (γ_n, K)

has the B-property for any symmetric convex body K in $\mathbb{R}^n$. They also proved that any pair (μ, K) of an 1-unconditional log-concave measure and 1-unconditional convex body on $\mathbb{R}^n$ also has the B-property. In view of Theorem 5.6, we have the following analogue of [PV16, Theorem 3.1] for log-concave measures:

Theorem 5.7. *Let K be a symmetric convex body on $\mathbb{R}^n$ and let μ be a log-concave probability measure. If m is the median of $x \mapsto \|x\|_K$ with respect to μ, then we have:*

a. *For every $\varepsilon \in (0,1)$,*

$$\mu(\{x : \|x\|_K \leqslant (1-\varepsilon)m\}) \leqslant \frac{1}{2}\exp\left(-2\varepsilon\vartheta_\mu(K)\right) \leqslant \frac{1}{2}\exp\left(-c\varepsilon/\sqrt{\beta_\mu(K)}\right).$$

b. *Furthermore, if the pair (μ, K) possesses the B-property and $\mu(K) \leqslant 1/2$,*

$$\mu(\varepsilon K) \leqslant \varepsilon^{2f_K(m)}\mu(K) \leqslant \varepsilon^{\frac{1}{16\mathbb{E}|\|Z\|_K - m|}}\mu(K),$$

for all $\varepsilon \in (0,1)$. In particular,

$$\mu(\{x : \|x\|_K \leqslant \varepsilon m\}) \leqslant \frac{1}{2}\varepsilon^{2\vartheta_\mu(K)} \leqslant \frac{1}{2}\varepsilon^{c/\sqrt{\beta_\mu(K)}},$$

for every $\varepsilon \in (0,1)$.

Proof. The first assertion follows by (5.2) and (5.4). For the second estimate recall the fact that the B-property implies that for any symmetric convex body A the map $t \mapsto [\mu(tA)/\mu(A)]^{1/\log(1/t)}$, $t \in (0,1)$ is non-decreasing (e.g. see [LO05]). Therefore, if we define

$$d := \sup_{0<\varepsilon<1} \frac{\log[\mu(\varepsilon K)/\mu(K)]}{\log(1/\varepsilon)},$$

the monotonicity shows that:

$$d = \lim_{\varepsilon \to 1^-} \frac{\log[\mu(\varepsilon K)/\mu(K)]}{\log(1/\varepsilon)} = -\frac{d}{d\varepsilon}\Big|_{\varepsilon=1} \log F_K(\varepsilon) = -(\log F_K)'(1)$$

On the other hand, the map $t \mapsto (\log F_K)'(t)$ is non-increasing (Lemma 5.2) and since $\mu(K) \leqslant 1/2$ we get that $m \geqslant 1$, which implies

$$(\log F_K)'(1) \geqslant (\log F_K)'(m) = 2f_K(m)$$

Combining the above we infer that:

$$\frac{\log[\mu(\varepsilon K)/\mu(K)]}{\log(1/\varepsilon)} \leqslant -2f_K(m) \Longrightarrow \mu(\varepsilon K) \leqslant \varepsilon^{2f_K(m)}\mu(K)$$

for every $\varepsilon \in (0,1)$ as required. □

Remark 5.8. Recently, in [ENT16], Eskenazis, Nayar and Tkocz proved that the pair (ν_1^n, K) has the B-property for all symmetric convex bodies K on $\mathbb{R}^n$. In view of Theorem 5.7 and their result one has:

$$\nu_1^n(\varepsilon K) \leqslant \varepsilon^{c/\mathbb{E}|\|Z\|_K - m|}\nu_1^n(K), \quad 0 < \varepsilon < 1, \tag{5.5}$$

for all symmetric convex bodies K on $\mathbb{R}^n$ with $\nu_1^n(K) \leqslant 1/2$. A different estimate was proved in [ENT16, Corollary 13] in terms of the inradius $r(K)$ of the body K, which is in the spirit of [LO05]. Let us mention that for the class of 1-unconditional convex bodies a better estimate than (5.5) is known for the exponential distribution; see [PV16, Proposition 3.2]. Furthermore, for the Gaussian distribution γ_n and all symmetric convex bodies K in $\mathbb{R}^n$ with $\gamma_n(K) \leqslant 1/2$, one has:

$$\gamma_n(tK) \leqslant (2t)^{cm/(\mathbb{E}|\|Z\|_K - m|)^2}\gamma_n(K),$$

for every $t \in [0,1]$, where m is the median of $\|Z\|_K$ with $Z \sim N(\mathbf{0}, I_n)$ (see [PV16]). One should compare the latter with [LO05, Theorem 2].

In order to illustrate the difference, between estimate (5.5) and the one proved in [ENT16, Corollary 13], let us consider the following example. Fix $1 \leqslant p \leqslant \infty$ and let W be an n-dimensional exponential random vector, i.e. $W \sim \nu_1^n$. Then, we have

$$\mathrm{Var}\|W\|_p \simeq_p \begin{cases} n^{\frac{2}{p}-1}, & 1 \leqslant p < \infty \\ 1, & p = \infty \end{cases},$$

for all sufficiently large n (for a proof of this fact the reader may consult [PVZ15, Theorem 3.2 & §3.2]). Furthermore,

$$\mathbb{E}\|W\|_p \simeq \begin{cases} pn^{1/p}, & 1 \leqslant p \leqslant \log n \\ \log n, & p \geqslant \log n \end{cases}.$$

For a proof of the latter fact the reader is referred to [SZ90]. Thus, one has:

$$\beta_{\nu_1^n}(B_p^n) \simeq_p \begin{cases} n^{-1}, & 1 \leqslant p < \infty \\ (\log n)^{-2}, & p = \infty \end{cases},$$

for fixed p and for all sufficiently large n. It follows that if $m = m_{p,n}$ is the median of $\|W\|_p$ with $2 < p \leqslant \infty$, then the small ball estimate (5.5) (applied for $K = mB_p^n$) yields

$$\mathbb{P}(\|W\|_p \leqslant \varepsilon m) \leqslant \frac{1}{2}\varepsilon^{c/\sqrt{\beta_{\nu_1^n}(B_p^n)}}, \tag{5.6}$$

while the corresponding small ball in terms of the inradius $r(B_p^n)$ yields

$$\mathbb{P}\left(\|W\|_p \leqslant \varepsilon m\right) \leqslant \frac{1}{2}\varepsilon^{c\sqrt{k_{\nu_1^n}(B_p^n)}},$$

where

$$k_{\nu_1^n}(B_p^n) := \big(r(B_p^n)\cdot \mathbb{E}\|W\|_p\big)^2 \simeq \begin{cases} n, & 1 \leqslant p \leqslant 2 \\ p^2 n^{2/p}, & 2 \leqslant p \leqslant \log n \\ (\log n)^2, & p \geqslant \log n \end{cases}.$$

Actually in view of [PV16, Proposition 3.2] one can have an even better estimate than (5.6) since the ℓ_p norms are 1-unconditional.

For further comparison purposes, recall that if μ satisfies a Poincaré (or spectral gap) inequality with constant $\lambda_1 = \lambda_1(\mu) > 0$, i.e.

$$\lambda_1 \mathrm{Var}_\mu(f) \leqslant \int \|\nabla f\|_2^2 \, d\mu,$$

for all smooth f, then

$$\sqrt{\lambda_1}\|f - m\|_{L_1(\mu)} \leqslant \big(\mathbb{E}_\mu\|\nabla f\|_2^2\big)^{1/2} \leqslant \|f\|_{\mathrm{Lip}},$$

for every Lipschitz map f, where m is a median of f. In the case that f is the norm $\|\cdot\|_K$ induced by the symmetric convex body K, note that $\|f\|_{\mathrm{Lip}} = 1/r(K)$, which indicates that the L_1 deviation is in general smaller than the inradius of the body.

Examples 5.9. 1. (The parameter $\vartheta_{\gamma_n}(B_\infty^n)$). Let m be the median for $x \mapsto \|x\|_\infty$ with respect to γ_n. Then, we have:

$$\vartheta(\gamma_n, B_\infty^n) \simeq \log n.$$

Indeed; note that for any $s > 0$ we have:

$$F_{B_\infty^n}(s) = \gamma_n(sB_\infty^n) = (\gamma_1([-s,s]))^n = (2\Phi(s) - 1)^n.$$

Hence,

$$f_{B_\infty^n}(s) = 2n\phi(s)(2\Phi(s) - 1)^{n-1}.$$

In particular, if m is the median of $x \mapsto \|x\|_\infty$ with respect to γ_n, then

$$f_{B_\infty^n}(m) = n\frac{\phi(m)}{2\Phi(m) - 1} = n2^{1/n}\phi(m).$$

On the other hand we have:

$$2^{-1/n} = \int_{-m}^{m} \phi(t)\,dt = 1 - 2\int_{m}^{\infty} \phi(t)\,dt \Longrightarrow 2\int_{m}^{\infty} \phi = 1 - 2^{-1/n} \simeq 1/n$$

and from standard estimates on the error function we have:

$$\int_{m}^{\infty} \phi(t)\,dt \simeq \frac{1}{m}\phi(m), \quad m \to \infty.$$

Hence, we obtain $f_{B_\infty^n}(m) \simeq m$ or $\vartheta(\gamma_n, B_\infty^n) \simeq m^2 \simeq \log n$.

2. (The parameter $\vartheta_{\gamma_n}(B_2^n)$). Let m be a median for $x \mapsto \|x\|_2$ with respect to Gaussian. Then, we have:

$$\vartheta = \vartheta(\gamma_n, B_2^n) \simeq \sqrt{n}.$$

Indeed; it is known that $\beta(B_2^n) \simeq 1/n$ which implies that $\vartheta \gtrsim \sqrt{n}$. For the upper estimate we argue as follows:

$$e^{-c_1\vartheta^2} \geqslant \gamma_n\left(\frac{m}{2}B_2^n\right) \geqslant (2\pi)^{-n/2}\left|\frac{m}{2}B_2^n\right| e^{-m^2/8} \geqslant e^{-c_2 n},$$

where we have used the fact that $m \simeq \sqrt{n}$.

In fact the upper estimate holds for any centrally symmetric convex body in $\mathbb{R}^n$:

Corollary 5.10. *Let K be a symmetric convex body on $\mathbb{R}^n$. Then, one has:*

$$\vartheta(\gamma_n, K) \lesssim \sqrt{n}.$$

Proof. We shall need [KV07, Lemma 2.1] which states:

$$\frac{1}{2}\sigma(S^{n-1} \cap \frac{1}{2}L) \leqslant \gamma_n(\sqrt{n}L),$$

for any centrally symmetric convex body L. Using the dual Sudakov inequality (see e.g. [LT11]) we get:

$$\frac{|\sqrt{n}B_2^n|}{|\sqrt{n}B_2^n \cap \frac{m}{4}K|} \leqslant N(\sqrt{n}B_2^n, \frac{m}{4}K) \leqslant \exp\left(c_1(\mathbb{E}\|Z\|_K)^2(4\sqrt{n}/m)^2\right) \leqslant e^{c_2 n},$$

where m is a median for $\|Z\|_K$ and $Z \sim N(\mathbf{0}, I_n)$. On the other hand we have:

$$\frac{|\sqrt{n}B_2^n \cap \frac{m}{4}K|}{|\sqrt{n}B_2^n|} = \sigma(S^{n-1} \cap \frac{m}{4\sqrt{n}}K) \leqslant 2\gamma_n(\frac{m}{2}K) \leqslant \exp(-c_3\vartheta^2).$$

Combining all the above we get the result. □

5.1. **A remark on isoperimetry.** Let ν_1 be the 1-dimensional exponential measure with density $d\nu_1(x) = \frac{1}{2}e^{-|x|}\,dx$. Let F_{ν_1} be its cumulative distribution function, i.e.

$$F_{\nu_1}(x) = \nu_1((-\infty, x]) = \frac{1}{2}\int_{-\infty}^{x} e^{-|t|}\,dt, \quad x \in \mathbb{R}.$$

Let $X = (\mathbb{R}^n, \|\cdot\|)$ be a normed space and let μ be a log-concave measure on $\mathbb{R}^n$. Following [Bob99] we shall denote the induced measure by μ_X. That is the push forward of the measure μ under the map $x \mapsto \|x\|$, i.e.

$$\mu_X(A) = \mu(x \in \mathbb{R}^n : \|x\| \in A), \quad A \subseteq [0, \infty), \text{ Borel}.$$

Let $T : (0, \infty) \to \mathbb{R}$ with $T = F_{\nu_1}^{-1} \circ F_K$ be the map which transports μ_X to ν_1, where $K = \{x : \|x\| \leqslant 1\}$. Bobkov in [Bob99] proved the following:

Proposition 5.11. *Let μ be a log-concave probability measure on $X = (\mathbb{R}^n, \|\cdot\|)$. Then, one has:*

$$\mathrm{Is}(\mu_X) = \inf_{s>0}(F_{\nu_1}^{-1} \circ F_K)'(s) \gtrsim \frac{1}{\mathbb{E}_\mu\|Z\|}.$$

It is known that $\mathrm{Is}(\mu_X) \lesssim 1/\sqrt{\mathrm{Var}_\mu \|Z\|}$ (see e.g. [Bob99] and the references therein). In [Bob99], Bobkov shows that for 1-dimensional log-concave measures the reverse estimate also holds [Bob99, §4]. Moreover, he asks if a reverse estimate should be expected to be true for the measure μ_X. Note that even though the measure μ_X is not necessarily log-concave, it enjoys many properties, see e.g. Proposition 5.11. Here we observe that Lemma 5.4 partially answers the aforementioned question:

Corollary 5.12. *Let μ be a log-concave probability measure on $\mathbb{R}^n$ and let K be a symmetric convex body on $\mathbb{R}^n$. Then, we have:*

$$\inf_{0<s\leqslant m} (F_{\nu_1}^{-1} \circ F_K)'(s) = (F_{\nu_1}^{-1} \circ F_K)'(m) \gtrsim \frac{1}{\mathbb{E}_\mu \big| \|Z\|_K - m \big|},$$

where m is the median of the function $x \mapsto \|x\|_K$ with respect to μ.

Moreover, a lower estimate for the full range of s shouldn't be expected as the next argument shows: If $\inf_{s>0} T'(s) := 1/L$, then $T : (0,\infty) \to \mathbb{R}$ is homeomorphism and $1/L$-expansive, i.e. $|Tu - Tv| \geqslant |u-v|/L$ for all $u, v > 0$. Thus, the map $T^{-1} : \mathbb{R} \to (0,\infty)$ is L-Lipschitz and transports the measure ν_1 to μ_X, i.e. for any Borel set $A \subseteq \mathbb{R}$ we have:

$$\nu_1(A) = \mu_X(T^{-1}(A)) = \mu\left(x : \|x\|_K \in T^{-1}(A)\right). \tag{5.7}$$

Claim. For any $t > 0$ one has[2]:

$$T^{-1}([-t,t]) \subseteq [m - tL, m + tL].$$

In particular, for $A = (-t, t)$, $t > 0$ in (5.7) we have:

$$\begin{aligned} 1 - \mathbb{P}(|\|Z\|_K - m| > tL) &= \mu(x : \|x\|_K \in [m - tL, m + tL]) \\ &\geqslant \mu\left(x : \|x\|_K \in T^{-1}([-t,t])\right) = \nu_1([-t,t]). \end{aligned}$$

It follows that $\mathbb{P}(|\|Z\|_K - m| > tL) \leqslant e^{-t}$ for all $t > 0$, or equivalently

$$\mathbb{P}\left(\big|\|Z\|_K - m\big| > \varepsilon m\right) \leqslant \exp\left(-\varepsilon m/L\right),$$

for all $\varepsilon > 0$. Hence, if $L \simeq \sqrt{\mathrm{Var}_\mu \|Z\|}$, we would obtain:

$$\mathbb{P}\left(\big|\|Z\|_K - m\big| > \varepsilon m\right) \leqslant \exp\left(-c\varepsilon/\sqrt{\beta_\mu(K)}\right),$$

for all $\varepsilon > 0$. Note that for any fixed $2 < p < \infty$ one has:

$$c\exp(-C_p \min\{\varepsilon^2 n, (\varepsilon n)^{2/p}\}) \leqslant \mathbb{P}\left(\big|\|Z\|_p - M_{p,n}\big| > \varepsilon M_{p,n}\right),$$

for all $0 < \varepsilon < 1$, where $Z \sim N(\mathbf{0}, I_n)$ and $M_{p,n}$ is a median for $\|Z\|_p$. Thus, for $p = 5$ say, we get a contradiction for all $\varepsilon \in (n^{-1/6}, 1)$.

5.2. **Another small deviation inequality.** In this subsection we prove a small deviation inequality similar to (5.1) with the variance replaced by the $\mathbb{E}\|\nabla f\|_2^2$. The inequality is known to specialists and the method of proof goes at least back to [Sam03] and [BG99]. The inequality is weaker than (5.1) but holds for a larger class of measures: for all measures which satisfy a quadratic cost inequality á la Talagrand [Tal96] (see also [AS17] for a recent development on the related subject). First we recall the necessary definitions. For any two Borel probability measures μ and ν on $\mathbb{R}^n$ the Wasserstein distance $W_2(\mu,\nu)$ is defined as

$$W_2^2(\mu,\nu) = \inf_\pi \iint \|x - y\|_2^2 \, d\pi(x,y),$$

where the infimum is taken over all couplings (or matchings) π of μ and ν, that is π has marginals μ and ν respectively. The Kullback-Leibler divergence (or relative entropy) of ν with respect to μ is defined by

$$D(\nu||\mu) = \mathrm{Ent}_\mu\left(\frac{d\nu}{d\mu}\right) = \int \log \frac{d\nu}{d\mu} \, d\nu,$$

if ν is absolutely continuous with respect to μ with Radon-Nikodym derivative $\frac{d\nu}{d\mu}$ and ∞ otherwise.

[2]Indeed; if $|z| \leqslant t$ then, since $T(m) = 0$, we have: $|T^{-1}(z) - m| = |T^{-1}(z) - T^{-1}(0)| \leqslant L|z| \leqslant tL$.

A Borel probability measure μ on $\mathbb{R}^n$ it is said to satisfy the *quadratic transportation cost inequality* with constant $A > 0$ if

$$W_2(\mu, \nu) \leqslant \sqrt{AD(\nu||\mu)}, \tag{5.8}$$

for any Borel probability measure ν with $\nu \ll \mu$. It is known that measures with this property can be characterized in terms of infimum convolution inequalities with cost function $w(z) = \|z\|_2^2/(2A),\ z \in \mathbb{R}^n$ (e.g. see [Led01, Corollary 6.4.]). Furthermore, Otto and Villani in [OV00] showed that measures which satisfy a log-Sobolev inequality then also satisfy a quadratic transportation cost inequality. The main inequality of this subsection, which is in the same spirit as Theorem 5.6, reads as follows.

Theorem 5.13. *Let μ be any Borel probability measure on $\mathbb{R}^n$, which satisfies a quadratic transportation cost inequality* (5.8)*. Then, for any smooth, convex map $f : \mathbb{R}^n \to \mathbb{R}$ we have:*

$$\log \mathbb{E}_\mu e^{-f} \leqslant -\mathbb{E}_\mu(f) + \frac{A}{4}\mathbb{E}_\mu \|\nabla f\|_2^2.$$

In particular,

$$\mu\left(\left\{x : f(x) - \mathbb{E}_\mu(f) \leqslant -t\left(\mathbb{E}_\mu\|\nabla f\|_2^2\right)^{1/2}\right\}\right) \leqslant e^{-t^2/A},$$

for all $t > 0$.

Proof. Since f is convex and smooth, for any $x, y \in \mathbb{R}^n$ we may write:

$$f(x) - f(y) \leqslant \langle \nabla f(x), x - y\rangle \leqslant \|\nabla f(x)\|_2 \|x - y\|_2.$$

Fix any probability measure ν with $\nu \ll \mu$ and let π be any coupling of μ and ν. Thus, integration with respect to π yields:

$$\mathbb{E}_\mu f - \mathbb{E}_\nu f \leqslant \int \|\nabla f(x)\|_2 \cdot \|x - y\|_2 \, d\pi(x, y) \leqslant \left(\mathbb{E}_\mu \|\nabla f\|_2^2\right)^{1/2} \left(\int \|x - y\|_2^2 \, d\pi(x, y)\right)^{1/2}.$$

Since the left-hand side is fixed for any coupling π of μ and ν we infer:

$$\mathbb{E}_\mu f - \mathbb{E}_\nu f \leqslant \sqrt{A\mathbb{E}_\mu \|\nabla f\|_2^2 D(\nu||\mu)}, \tag{5.9}$$

where we have used the assumption on μ. Now we employ Gibb's variational formula (for a proof see [BLM13, Corollary 4.14]): For any μ-measurable map f one has

$$\log \mathbb{E}_\mu e^f = \sup_{\nu \ll \mu} \left\{\mathbb{E}_\nu f - D(\nu||\mu)\right\}. \tag{5.10}$$

Applying the latter for $-f$ and taking into account (5.9) we obtain:

$$\log \mathbb{E}_\mu e^{-f} \leqslant -\mathbb{E}_\mu f + \sup_{\nu \ll \mu} \left\{\sqrt{A\mathbb{E}_\mu \|\nabla f\|_2^2 D(\nu||\mu|)} - D(\nu||\mu)\right\} \leqslant -\mathbb{E}_\mu f + \frac{A}{4}\mathbb{E}_\mu \|\nabla f\|_2^2.$$

The result follows. □

6. Probabilistic estimates on the Grassmannian

In this Section we prove Theorem 4.6. We present two approaches to derive this probabilistic estimate. The first one uses Gaussian tools and is based on the new small deviation inequality (5.1), therefore yields better tails estimate in the one-sided small deviation regime, but restricts the range of t. The second approach overcomes this obstacle by working directly on the Grassmann space, but the tail estimate we obtain relies on the Lipschitz constant since we employ the Gromov-Milman theorem for $SO(n)$. A small ball probability estimate for the mapping $F \mapsto w(P_F A)$ is also provided.

6.1. **From Gauss' space to Grassmannian.** First we provide a Gaussian proof of Theorem 4.6. Let us recall the following:

Lemma 6.1. *Let A be symmetric convex body on $\mathbb{R}^n$. For any matrix $T=(t_{ij})_{i,j=1}^{k,n}$, $1\leqslant k\leqslant n$ we define the map $T\mapsto w(TA)$. Then, we have the Lipschitz condition:*

$$|w(TA)-w(SA)|\leqslant \frac{R(A)}{\sqrt{k}}\|T-S\|_{\mathrm{HS}}, \tag{6.1}$$

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

$$\mathbb{P}\left(w(GA)\geqslant \mathbb{E}[w(GA)]+t\right)\leqslant \exp(-ct^2k/R(A)^2), \tag{6.2}$$

and

$$\mathbb{P}\left(w(GA)\leqslant \mathbb{E}[w(GA)]-t\right)\leqslant \exp\left(-ct^2\max\left\{\frac{1}{\mathrm{Var}[h_A(Z)]},\frac{k}{R(A)^2}\right\}\right), \tag{6.3}$$

for all $t>0$, where $G=(g_{ij})$ is $k\times n$ matrix with i.i.d. standard Gaussian entries and $Z\sim N(\mathbf{0},I_n)$.

Proof. For a proof of (6.1) the reader may consult with [PV16]. Then, estimate (6.2) immediately follows from the concentration on Gauss' space. For proving (6.3) we need the next fact:

Claim. If $G=(g_{ij})_{i,j=1}^{k,n}$ is a Gaussian matrix then,

$$\mathrm{Var}[w(GA)]\leqslant \min\left\{\frac{R(A)^2}{k},\mathrm{Var}[h_A(Z)]\right\}, \tag{6.4}$$

where $Z\sim N(\mathbf{0},I_n)$.

Proof of Claim. The bound in terms of the circumradius follows from the Lipschitz condition and the Poincaré inequality (3.4). For bounding in terms of the variance we use the Cauchy-Schwarz inequality.

Finally, the estimate (6.3) follows from the small deviation inequality (5.1) applied for the convex function $T\mapsto w(TA)$. □

One more ingredient is the polar decomposition of any matrix $T\in\mathbb{R}^{k\times n}$. If $T\in\mathbb{R}^{k\times n}$ we may write: $T=SQ$ where $S=(TT^*)^{1/2}$ and Q is the orthogonal projection onto $F=\mathrm{Im}T^*$.

The next Lemma follows if we take into account the above decomposition and the ideal property of the ℓ-norm (see e.g. [TJ89]).

Lemma 6.2. *If $T\in\mathbb{R}^{k\times n}$ and A is a symmetric convex body on $\mathbb{R}^n$, then we have:*

$$\lambda_k((TT^*)^{1/2})w(P_FA)\leqslant w(TA)\leqslant \lambda_1((TT^*)^{1/2})w(P_FA),$$

where $\lambda_j((TT^*)^{1/2})\equiv s_j(T)$ is the j-th eigenvalue of $(TT^*)^{1/2}$ (or the j-th singular value of T) and $F=\mathrm{Im}T^*$.

Proof. First note that if $S:\ell_2^k\to\ell_2^k$ is a linear map which satisfies $0<a\leqslant\|S\theta\|_2\leqslant b$ for all $\theta\in S^{k-1}$, then

$$a\mathbb{E}[h_A(Y)]\leqslant \mathbb{E}[h_A(SY)]\leqslant b\mathbb{E}[h_A(Y)] \tag{6.5}$$

where $Y\sim N(\mathbf{0},I_k)$. This follows by the ideal property of the ℓ-norm, i.e. for any operator $u:\ell_2^n\to\ell_2^n$ and $v:\ell_2^n\to X$ we have $\ell(vu)\leqslant\ell(v)\|u\|_{\mathrm{op}}$. Now in our setting, set $S=(TT^*)^{1/2}$. If $Y\sim N(\mathbf{0},I_k)$ and $c_k:=\mathbb{E}\|Y\|_2$, we may write:

$$w(TA)=c_k^{-1}\mathbb{E}[h_{TA}(Y)]=c_k^{-1}\mathbb{E}[h_{P_FA}(S^*Y)]\leqslant c_k^{-1}\|S\|_{2\to2}\mathbb{E}[h_{P_FA}(Y)],$$

where we have used the right-hand side of (6.5). We work similarly for the lower estimate. □

In the Gaussian random setting the variables $\lambda_j((GG^*)^{1/2})$ and $w(P_FA)$ with $F=\mathrm{Im}G^*$ are independent each other. Namely, we have the following:

Lemma 6.3. *Let A be a symmetric convex body on $\mathbb{R}^n$ and let $G=(g_{ij})_{i,j=1}^{k,n}$ where g_{ij} are i.i.d. standard normals. Then, $F=\mathrm{Im}G^*$ is uniformly distributed over $G_{n,k}$ and the random variables $\lambda_j((GG^*)^{1/2})$ and $w(P_FA)$ are independent.*

For a proof of this fact the reader is referred to [PPZ14, Proposition 4.1]. Now we are ready to prove the following:

Theorem 6.4. *Let A be a symmetric convex body on $\mathbb{R}^n$. Fix $1 \leqslant k \leqslant n-1$. Then, we have:*

$$\nu_{n,k}\left(F \in G_{n,k} : w(P_F A) > (1+t)w(A)\right) \leqslant C \exp(-ct^2 k k_*(A)),$$

for all $t > c_1\sqrt{k/n}$. Furthermore,

$$\nu_{n,k}\left(F \in G_{n,k} : w(P_F A) \leqslant (1-t)w(A)\right) \leqslant C \exp\left(-ct^2 \max\left\{kk_*(A), \frac{1}{\beta_*(A)}\right\}\right),$$

for all $c_2\sqrt{k/n} < t < 1$, provided that $k \lesssim n$.

Proof. Note that for any Gaussian matrix $G = (g_{ij})_{i,j=1}^{k,n}$ we have:

$$\mathbb{E}[w(GA)] = \mathbb{E}[h_A(Z)] = \mathbb{E}\|Z\|_2 \cdot w(A),$$

where $Z \sim N(\mathbf{0}, I_n)$. Fix $t > 0$. From Lemma 6.2 and Lemma 6.3 we may write:

$$\begin{aligned}\nu_{n,k}(w(P_F A) > (1+t)w(A))\mathbb{P}\left(\lambda_k((GG^*)^{1/2}) > (1-\delta)\mathbb{E}\|Z\|_2\right) \\ \leqslant \mathbb{P}(w(GA) > (1+t/2)\mathbb{E}w(GA)),\end{aligned}$$

for $0 < \delta < \frac{t}{2(1+t)}$. Recall the following well known:

Fact. Let $\delta \in (0,1)$. Then the random Gaussian matrix $G = (g_{i,j})_{i,j=1}^{k,n}$ with $k \leqslant c\delta^2 n$ satisfies:

$$(1-\delta)\mathbb{E}\|Z\|_2 < \lambda_k((GG^*)^{1/2}) \leqslant \lambda_1((GG^*)^{1/2}) < (1+\delta)\mathbb{E}\|Z\|_2$$

with probability greater than $1 - e^{-c\delta^2 n}$, where $Z \sim N(\mathbf{0}, I_n)$.

Now Lemma 6.1 yields:

$$(1 - e^{-c\delta^2 n})\nu_{n,k}\left(w(P_F A) > (1+t)w(A)\right) \leqslant C \exp\left(-ct^2 kk_*(A)\right).$$

The choice $\delta \simeq \sqrt{k/n}$ yields the upper estimate. We work similarly for the lower estimate. □

6.1.1. *A small ball estimate.* Next, we prove the following:

Theorem 6.5 (small ball for the mean width of projections)**.** *Let A be a symmetric convex body on $\mathbb{R}^n$ and let $1 \leqslant k \leqslant n-1$. Then, we have:*

$$\nu_{n,k}\left(\{F \in G_{n,k} : w(P_F A) \leqslant c\varepsilon w(A)\}\right) \leqslant (C\varepsilon)^{c\max\{kk_*(A), \frac{1}{\beta_*(A)}\}},$$

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

Proof. First note that the function $T \mapsto w(TA)$ is indeed a norm on $\mathbb{R}^{k\times n}$. If $\mathcal{C}_A = \left\{T \in \mathbb{R}^{k\times n} : w(TA) \leqslant 1\right\}$ is its unit ball, then estimate (6.4) shows that $\beta(\mathcal{C}_A) \leqslant \min\left\{\frac{1}{kk_*(A)}, \beta_*(A)\right\}$. It follows by [PV16, Theorem 3.1] that:

$$\mathbb{P}\left(w(GA) \leqslant c\varepsilon\mathbb{E}[w(GA)]\right) \leqslant \frac{1}{2}\varepsilon^{c/\beta(\mathcal{C}_A)},$$

for all $\varepsilon \in (0, 1/2)$. Now we use Lemma 6.2 and Lemma 6.3 to get:

$$\begin{aligned}c_1\nu_{n,k}\left(w(P_F A) \leqslant c'\varepsilon w(A)\right) \leqslant \nu_{n,k}\left(w(P_F A) \leqslant c'\varepsilon w(A)\right)\mathbb{P}\left(\lambda_1((GG^*)^{1/2}) \leqslant C_1\mathbb{E}\|Y\|_2\right) \\ \leqslant \mathbb{P}\left(w(GA) \leqslant c\varepsilon\mathbb{E}[w(GA)]\right).\end{aligned}$$

The result readily follows. □

6.2. **Breaking the barrier.** Note that the argument of the preceding proof does not allow to consider very small $t > 0$. In order to retrieve the behavior for small enough t we have to work beyond the Gaussian setting. First we state Lipschitz estimates on the Grasmmannian for the map $F \mapsto w(P_F A)$ with respect to the normalized metrics:

$$\sigma_\infty(E,F) = \|P_E - P_F\|_{\mathrm{op}}, \quad \sigma_2(E,F) = \frac{\|P_E - P_F\|_{\mathrm{HS}}}{\sqrt{k}}.$$

Lemma 6.6. *Let A be a (symmetric) convex body on $\mathbb{R}^n$ and let $1 \leqslant k \leqslant n-1$. Then, for any $E, F \in G_{n,k}$ we have:*

$$|w(P_E A) - w(P_F A)| \leqslant c\sqrt{n/k}\, w(A)\sigma_\infty(E,F). \tag{6.6}$$

Furthermore, we have:

$$|w(P_E A) - w(P_F A)| \leqslant c' R(A)\sigma_2(E,F), \tag{6.7}$$

where $c, c' > 0$ are absolute constants. Equivalently, in the dual setting, we may write:

$$|M(A\cap E) - M(A \cap F)| \lesssim \min\left\{\sqrt{n/k}M(A)\sigma_\infty(E,F), b(A)\sigma_2(E,F)\right\}.$$

Proof. Using the formula

$$c_k w(P_F A) = c_k \int_{S_F} h_A(\phi)\, d\sigma_F(\phi) = \int_F h_A(x)\, d\gamma_F(x) = \int_{\mathbb{R}^n} h_A(P_F x)\, d\gamma_n(x), \tag{6.8}$$

where $c_k = \mathbb{E}\|Y\|_2 = \sqrt{2}\Gamma(\frac{k+1}{2})/\Gamma(\frac{k}{2}) \simeq \sqrt{k}$ with $Y \sim N(\mathbf{0}, I_k)$, we may write:

$$\begin{aligned} c_k|w(P_E A) - w(P_F A)| &\leqslant \int_{\mathbb{R}^n} |h_A(P_E x) - h_A(P_F x)|\, d\gamma_n(x) \\ &\leqslant \int_{\mathbb{R}^n} h_A(P_E x - P_F x)\, d\gamma_n(x). \end{aligned}$$

Now we proceed as follows. In order to prove the first estimate recall the ideal property of the ℓ-norm (e.g. see [TJ89]). Applying this for $u = P_E - P_F$, $v = id$ and $X = (\mathbb{R}^n, \|\cdot\|_{A^\circ})$ we obtain:

$$\int_{\mathbb{R}^n} h_A((P_E - P_F)x)\, d\gamma_n(x) \leqslant \int_{\mathbb{R}^n} h_A(x)\, d\gamma_n(x)\cdot \|P_E - P_F\|_{\mathrm{op}}.$$

Therefore we get:

$$|w(P_E A) - w(P_F A)| \leqslant \frac{c_n}{c_k} w(A)\|P_E - P_F\|_{\mathrm{op}}.$$

For the second assertion we proceed as follows:

$$\begin{aligned} \int_{\mathbb{R}^n} h_A(P_E x - P_F x)\, d\gamma_n(x) &\leqslant R(A)\int_{\mathbb{R}^n} \|(P_E - P_F)x\|_2\, d\gamma_n(x) \\ &\leqslant R(A)\|P_E - P_F\|_{\mathrm{HS}}, \end{aligned}$$

by the Cauchy-Schwarz inequality and the isotropicity of γ_n. □

Let us recall the concentration on $SO(n)$. The next result is due to Gromov and V. Milman from [GM83].

Theorem 6.7. *Let $f : SO(n) \to \mathbb{R}$ be L-Lipschitz map with respect to the Hilbert-Schmidt norm, i.e. $|f(U) - f(V)| \leqslant L\|U - V\|_{\mathrm{HS}}$ for any $U, V \in SO(n)$. Then we have:*

$$\mu_n\left(\{U \in SO(n) : |f(U) - \mathbb{E}f| \geqslant t\}\right) \leqslant C_1 \exp(-c_1 n t^2/L^2),$$

for all $t > 0$.

First note that the distribution of the map $F \mapsto w(P_F A)$ can be carried out over the orthogonal group $O(n)$. This follows directly from the Haar-invariance of $\nu_{n,k}$ under $O(n)$. Indeed,

$$\nu_{n,k}(\{F \in G_{n,k} : w(P_F A) \in B\}) = \mu_n(\{U \in O(n) : w(P_{UE} A) \in B\})$$

for some (any) fixed $E \in G_{n,k}$. That is $F \mapsto w(P_F A)$ can be equivalently viewed as function of $U \in O(n)$. So, in order to apply a concentration result for $F \mapsto w(P_F A)$ on $G_{n,k}$ it suffices to suitably apply the

Gromov-Milman theorem for $U \mapsto w(P_{UE}A)$ for some fixed $E \in G_{n,k}$. Toward this end we shall need the Lipschitz constant of $U \mapsto \psi(U) = w(P_{UE}A)$ with respect to the Hilbert-Schmidt norm. Then Lemma 6.6 implies:

$$|\psi(U_1) - \psi(U_2)| \lesssim \frac{R(A)}{\sqrt{k}} \|P_{U_1E} - P_{U_2E}\|_{\mathrm{HS}}.$$

On the other hand we have[3] $\|P_{U_1E} - P_{U_2E}\|_{\mathrm{HS}} \leqslant 2\|U_1 - U_2\|_{\mathrm{HS}}$, which proves the Lipschitz condition for ψ.

Now we explain how we derive the concentration estimate for ψ on the full orthogonal group by following an argument from [Mec14]. Let U_1 be Haar distributed on $SO(n)$ and let M_π be the permutation matrix[4] which corresponds to the permutation π such that $\pi(j) = j$ for $j \leqslant n-2$, $\pi(n-1) = n$ and $\pi(n) = n-1$. Then, $U_2 := U_1 M_\pi$ is the same as U_1 (as matrices) but the last two rows which are switched. Therefore, U_2 is Haar distributed on $SO^-(n)$. Define the orthogonal map U being U_1 with probability $1/2$ and U_2 with probability $1/2$. It follows that U is Haar distributed over $O(n)$. Moreover, note that since E can be considered as $E = [e_i : 1 \leqslant i \leqslant k]$ we have $M_\sigma(E) = E$ as long as $k \leqslant n-2$. In that case:

$$\mathbb{E}_U \psi(U) = \mathbb{E}_{U_1} \psi(U_1) = \mathbb{E}_{U_2} \psi(U_2) = \int_{G_{n,k}} w(P_F A)\, d\nu_{n,k}(F) = w(A).$$

Conditioning on whether $U = U_1$ or $U = U_2$ and taking into account Theorem 6.7 we get:

$$\mu_n \left(\left\{ U \in O(n) : \left| \psi(U) - \int_{O(n)} \psi \, d\mu_n \right| > t \right\} \right) \leqslant c_1 \exp(-c_2 n t^2 k / R(A)^2),$$

for all $t > 0$. The preceding discussion leads us to the next:

Theorem 6.8 (concentration for the mean width of projections)**.** *Let A be a symmetric convex body on $\mathbb{R}^n$ and let $1 \leqslant k \leqslant n-1$. Then, one has the following concentration inequality:*

$$\nu_{n,k}\left(\{ F \in G_{n,k} : |w(P_F A) - w(A)| \geqslant t w(A) \} \right) \leqslant c_1 \exp\left(-c_2 t^2 k k_*(A) \right),$$

for all $t > 0$, where $c_1, c_2 > 0$ are absolute constants.

Proof. The case $1 \leqslant k \leqslant n-2$ follows from the previous argument. We treat the case $k = n-1$ separately. Note that any $F \in G_{n,n-1}$ can be identified with $e^\perp$ for $e \in S^{n-1}$. Thus, we may write:

$$\int_{G_{n,n-1}} w(P_F A)\, d\nu_{n,n-1}(F) = \int_{S^{n-1}} w(P_{e^\perp} A)\, d\sigma(e) = w(A),$$

and as in (6.8) we have:

$$w(P_{e^\perp} A) = \frac{c_n}{c_{n-1}} \int_{S^{n-1}} h_A(P_{e^\perp}\theta)\, d\sigma(\theta) = \frac{c_n}{c_{n-1}} \int_{S^{n-1}} h_A(\theta - \langle \theta, e\rangle e)\, d\sigma(\theta).$$

Since $c_n / c_{n-1} \simeq 1$, it suffices to work with the map $f : S^{n-1} \to \mathbb{R}$ defined by $f(e) := \int_{S^{n-1}} h_A(\theta - \langle \theta, e \rangle e)\, d\sigma(\theta)$. We have the following:

Claim. The mapping $f : S^{n-1} \to \mathbb{R}$ is Lipschitz with:

$$\mathrm{Lip}(f) \leqslant R(A) \sqrt{\frac{2}{n}}.$$

[3] Set $U = U_1^* U_2$. Using the (right) invariance under the orthogonal group and the contractive property of the Hilbert-Schmidt norm, we may write:

$$\begin{aligned} \|P_{U_1E} - P_{U_2E}\|_{\mathrm{HS}} = \|P_E - U P_E U^*\|_{\mathrm{HS}} &\leqslant \|P_E(I - U^*)\|_{\mathrm{HS}} + \|(I-U)P_E U^*\|_{\mathrm{HS}} \leqslant \|I - U^*\|_{\mathrm{HS}} + \|(I-U)P_E\|_{\mathrm{HS}} \\ &\leqslant 2\|I - U\|_{\mathrm{HS}} = 2\|U_1 - U_2\|_{\mathrm{HS}}. \end{aligned}$$

[4] For any permutation π of the set $\{1, 2, \ldots, n\}$ the permutation matrix M_π acts as: $M_\pi(e_j) = e_{\pi(j)}$ for $j = 1, \ldots, n$.

Proof of Claim. For any $u, v \in S^{n-1}$, we have:

$$|f(u) - f(v)| \leqslant \int_{S^{n-1}} h_A \left(\langle u, \theta\rangle u - \langle v, \theta\rangle v\right) \, d\sigma(\theta)$$
$$\leqslant R(A) \left(\int_{S^{n-1}} \|\langle u, \theta\rangle u - \langle v, \theta\rangle v\|_2^2 \, d\sigma(\theta) \right)^{1/2}.$$

Note that:

$$\int_{S^{n-1}} \|\langle u, \theta\rangle u - \langle v, \theta\rangle v\|_2^2 \, d\sigma(\theta) = 2 \int_{S^{n-1}} \theta_1^2 \, d\sigma(\theta) - 2\langle u, v\rangle \int_{S^{N-1}} \langle u, \theta\rangle\langle v, \theta\rangle \, d\sigma(\theta)$$
$$= \frac{2}{n}(1 - \langle u, v\rangle^2) \leqslant \frac{2}{n}\|u - v\|_2^2.$$

The result now follows from the concentration on the sphere S^{n-1}. □

Acknowledgements. The authors are grateful to Peter Pivovarov for useful remarks.

References

[AAGM15] S. Artstein-Avidan, A. Giannopoulos, and V. D. Milman. *Asymptotic geometric analysis. Part I*, volume 202 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2015.

[AK16] F. Albiac and N. Kalton. *Topics in Banach space theory*, volume 233 of *Graduate Texts in Mathematics*. Springer, [Cham], second edition, 2016. With a foreword by Gilles Godefory.

[AS17] R. Adamczak and C. Strzeleki. On the convex Poincar'e inequality and weak transportation transportation inequalities. `https://arxiv.org/abs/1703.01765`, (2017). preprint.

[BG99] S. G. Bobkov and F. Götze. Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. *J. Funct. Anal.*, 163(1):1–28, 1999.

[BLM13] S. Boucheron, G. Lugosi, and P. Massart. *Concentration inequalities*. Oxford University Press, Oxford, 2013. A nonasymptotic theory of independence, With a foreword by Michel Ledoux.

[Bob99] S. G. Bobkov. Isoperimetric and analytic inequalities for log-concave probability measures. *Ann. Probab.*, 27(4):1903–1921, 1999.

[Bog98] V. I. Bogachev. *Gaussian measures*, volume 62 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1998.

[Bor75] C. Borell. Convex set functions in d-space. *Period. Math. Hungar.*, 6(2):111–136, 1975.

[CEFM04] 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(2):410–427, 2004.

[CEL12] D. Cordero-Erausquin and M. Ledoux. Hypercontractive measures, Talagrand's inequality, and influences. In *Geometric aspects of functional analysis*, volume 2050 of *Lecture Notes in Math.*, pages 169–189. Springer, Heidelberg, 2012.

[Cha14] S. Chatterjee. *Superconcentration and related topics*. Springer Monographs in Mathematics. Springer, Cham, 2014.

[DR50] A. Dvoretzky and C. A. Rogers. Absolute and unconditional convergence in normed linear spaces. *Proc. Nat. Acad. Sci. U. S. A.*, 36:192–197, 1950.

[Dvo61] A. Dvoretzky. Some results on convex bodies and Banach spaces. In *Proc. Internat. Sympos. Linear Spaces (Jerusalem, 1960)*, pages 123–160. Jerusalem Academic Press, Jerusalem; Pergamon, Oxford, 1961.

[Ehr83] A. Ehrhard. Symétrisation dans l'espace de Gauss. *Math. Scand.*, 53(2):281–301, 1983.

[ENT16] A. Eskenazis, P. Nayar, and T. Tkocz. Gaussian mixtures: entropy and geometric inequalities. `https://arxiv.org/abs/1611.04921`, (2016). preprint.

[Fig76] T. Figiel. A short proof of Dvoretzky's theorem on almost spherical sections of convex bodies. *Compositio Math.*, 33(3):297–301, 1976.

[FLM77] T. Figiel, J. Lindenstrauss, and V. D. Milman. The dimension of almost spherical sections of convex bodies. *Acta Math.*, 139(1-2):53–94, 1977.

[GM83] M. Gromov and V. D. Milman. A topological application of the isoperimetric inequality. *Amer. J. Math.*, 105(4):843–854, 1983.

[Gor85] Yehoram Gordon. Some inequalities for Gaussian processes and applications. *Israel J. Math.*, 50(4):265–289, 1985.

[HW16] H. Huang and F. Wei. Upper bound for the Dvoretzky dimension in Milman–Schechtman theorem. `https://arxiv.org/abs/1612.03572`, (2016). preprint.

[JL84] 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)*, volume 26 of *Contemp. Math.*, pages 189–206. Amer. Math. Soc., Providence, RI, 1984.

[Joh48] F. John. Extremum problems with inequalities as subsidiary conditions. In *Studies and Essays Presented to R. Courant on his 60th Birthday, January 8, 1948*, pages 187–204. Interscience Publishers, Inc., New York, N. Y., 1948.

[Kla04] B. Klartag. A geometric inequality and a low M-estimate. *Proc. Amer. Math. Soc.*, 132(9):2619–2628 (electronic), 2004.

[Kla07] B. Klartag. Uniform almost sub-Gaussian estimates for linear functionals on convex sets. *Algebra i Analiz*, 19(1):109–148, 2007.

[KV07] B. Klartag and R. Vershynin. Small ball probability and Dvoretzky's theorem. *Israel J. Math.*, 157:193–207, 2007.

[Led01] M. Ledoux. *The concentration of measure phenomenon*, volume 89 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2001.

[Lew78] D. R. Lewis. Finite dimensional subspaces of L_p. *Studia Math.*, 63(2):207–212, 1978.

[LO05] R. Latała and K. Oleszkiewicz. Small ball probability estimates in terms of widths. *Studia Math.*, 169(3):305–314, 2005.

[LT11] M. Ledoux and M. Talagrand. *Probability in Banach spaces*. Classics in Mathematics. Springer-Verlag, Berlin, 2011. Isoperimetry and processes, Reprint of the 1991 edition.

[Mec14] E. Meckes. Concentration of measure on the compact classical matrix groups. `http://www.case.edu/artsci/math/esmeckes/Haar_notes.pdf`, (2014). Lecture Notes.

[Mil71] V. D. Milman. A new proof of A. Dvoretzky's theorem on cross-sections of convex bodies. *Funkcional. Anal. i Priložen.*, 5(4):28–37, 1971.

[MS86] V. D. Milman and G. Schechtman. *Asymptotic theory of finite-dimensional normed spaces*, volume 1200 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1986. With an appendix by M. Gromov.

[MS97] V. D. Milman and G. Schechtman. Global versus local asymptotic theories of finite-dimensional normed spaces. *Duke Math. J.*, 90(1):73–93, 1997.

[OV00] F. Otto and C. Villani. Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.*, 173(2):361–400, 2000.

[Pis89] G. Pisier. *The volume of convex bodies and Banach space geometry*, volume 94 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1989.

[PPZ14] G. Paouris, P. Pivovarov, and J. Zinn. A central limit theorem for projections of the cube. *Probab. Theory Related Fields*, 159(3-4):701–719, 2014.

[PV15] G. Paouris and P. Valettas. On Dvoretzky's theorem for subspaces of L_p. `http://arxiv.org/abs/1510.07289`, (2015). preprint.

[PV16] G. Paouris and P. Valettas. A small deviation inequality for convex functions. `https://arxiv.org/abs/1611.01723`, (2016). preprint.

[PV17] G. Paouris and P. Valettas. On the sharpness of gaussian concentration for norms, (2017). in preparation.

[PVZ15] G. Paouris, P. Valettas, and J. Zinn. Random version of Dvoretzky's theorem in ℓ_p^n. Stochastic Processes and their applications (to appear), (2015).

[Sam03] P. M. Samson. Concentration inequalities for convex functions on product spaces. In *Stochastic inequalities and applications*, volume 56 of *Progr. Probab.*, pages 33–52. Birkhäuser, Basel, 2003.

[Sch89] G. Schechtman. A remark concerning the dependence on ϵ in Dvoretzky's theorem. In *Geometric aspects of functional analysis (1987–88)*, volume 1376 of *Lecture Notes in Math.*, pages 274–277. Springer, Berlin, 1989.

[Sch07] G. Schechtman. The random version of Dvoretzky's theorem in ℓ_∞^n. In *Geometric aspects of functional analysis*, volume 1910 of *Lecture Notes in Math.*, pages 265–270. Springer, Berlin, 2007.

[SZ90] G. Schechtman and J. Zinn. On the volume of the intersection of two L_p^n balls. *Proc. Amer. Math. Soc.*, 110(1):217–224, 1990.

[Sza74] A. Szankowski. On Dvoretzky's theorem on almost spherical sections of convex bodies. *Israel J. Math.*, 17:325–338, 1974.

[Tal94] M. Talagrand. On Russo's approximate zero-one law. *Ann. Probab.*, 22(3):1576–1587, 1994.

[Tal96] M. Talagrand. Transportation cost for Gaussian and other product measures. *Geom. Funct. Anal.*, 6(3):587–600, 1996.

[Tik14] K. E. Tikhomirov. The randomized Dvoretzky's theorem in l_∞^n and the χ-distribution. In *Geometric aspects of functional analysis*, volume 2116 of *Lecture Notes in Math.*, pages 455–463. Springer, Cham, 2014.

[Tik17] K. E. Tikhomirov. Superconcentration, and randomized Dvoretzky's theorem for spaces with 1-unconditional bases. `https://arxiv.org/abs/1702.00859`, (2017). preprint.

[TJ89] N. Tomczak-Jaegermann. *Banach-Mazur distances and finite-dimensional operator ideals*, volume 38 of *Pitman Monographs and Surveys in Pure and Applied Mathematics*. Longman Scientific & Technical, Harlow; copublished in the United States with John Wiley & Sons, Inc., New York, 1989.

Department of Mathematics, Mailstop 3368, Texas A&M University, College Station, TX, 77843-3368.
E-mail address: `grigorios.paouris@gmail.com`

Mathematics Department, University of Missouri, Columbia, MO, 65211.
E-mail address: `valettasp@missouri.edu`