ISRAEL JOURNAL OF MATHEMATICS **157** (2007), 193–207
DOI: 10.1007/s11856-006-0007-1

SMALL BALL PROBABILITY AND DVORETZKY'S THEOREM

BY

B. Klartag*

School of Mathematics, Institute for Advanced Study
Princeton, NJ 08540, USA
e-mail: klartag@ias.edu

AND

R. Vershynin†

Department of Mathematics, University of California
One Shields Ave, Davis, CA 95616, USA
e-mail: vershynin@math.ucdavis.edu

ABSTRACT

Large deviation estimates are by now a standard tool in Asymptotic Convex Geometry, contrary to small deviation results. In this note we present a novel application of a small deviations inequality to a problem that is related to the diameters of random sections of high dimensional convex bodies. Our results imply an unexpected distinction between the lower and upper inclusions in Dvoretzky's Theorem.

1. Introduction

In probability theory, the large deviation theory (or the tail probabilities) and the small deviation theory (or the small ball probabilities) are in a sense two complementary directions. The large deviation theory, which is a more classical direction, seeks to control the probability of deviation of a random variable X from its mean M, i.e. one looks for upper bounds on $\text{Prob}(|X - M| > t)$. The small deviation theory seeks to control the probability of X being very small,

* Supported by NSF grant DMS-0111298 and the Bell Companies Fellowship.
† Supported by NSF grant 0401032 and the Sloan Research Fellowship.
Received September 30, 2004

i.e. it looks for upper bounds on $\mathrm{Prob}(|X| < t)$. There are a number of excellent texts on large deviations, see e.g. the recent books [DZ] and [dH]. A recent exposition of the state of the art in small deviation theory can be found in [LS].

A modern powerful approach to large deviations is via the celebrated concentration of measure phenomenon. One of the early manifestations of this idea was V. Milman's proof of Dvoretzky's Theorem in the 1970s. Recall that Dvoretzky's Theorem entails that any n-dimensional convex body has a section of dimension $c \log n$ which is approximately a Euclidean ball. Since Milman's proof, the concentration of measure philosophy has played a major role in geometric functional analysis and in many other areas. A recent book by M. Ledoux [L] gives an account of many ramifications of this method. A standard instance of the concentration of measure phenomenon is the case of a Lipschitz function on the unit Euclidean sphere S^{n-1}. In view of the geometric applications, we shall state it for a norm $\|\cdot\|$ on $\mathbb{R}^n$, or equivalently for its unit ball, which is a centrally-symmetric convex body $K \subset \mathbb{R}^n$. We equip the sphere S^{n-1} with the unique rotation invariant probability measure σ. Two parameters are responsible for many geometric properties of the convex body K, the maximal and the average values of the norm on the sphere S^{n-1}:

$$b = b(K) = \sup_{x \in S^{n-1}} \|x\|, \quad M = M(K) = \int_{S^{n-1}} \|x\| d\sigma(x). \tag{1}$$

The concentration of measure inequality, which appears e.g. in the first pages of [MS], states that the norm is close to its mean M on most of the sphere. For any $t > 1$,

$$\sigma\{x \in S^{n-1} : |\|x\| - M| > tM\} < \exp(-ct^2 k) \tag{2}$$

where

$$k = k(K) = n\Big(\frac{M(K)}{b(K)}\Big)^2.$$

Here and thereafter the letters $c, C, c', \tilde{c}, c_1, c_2$ etc. denote some positive universal constants, whose values may be different in various appearances. The symbol $\asymp$ denotes equivalence of two quantities up to an absolute constant factor, i.e. $a \asymp b$ if $ca \le b \le Ca$ for some absolute constants $c, C > 0$.

The concentration of measure inequality can of course be interpreted as a large deviation inequality for the random variable $\|x\|$, and the connection to probability theory becomes even more sound when one recalls an analogous inequality for Gaussian measures; see [L]. The quantity $k(K)$ plays a crucial role in high dimensional convex geometry, as it is the critical dimension in Dvoretzky's Theorem. We will call this dimension $k(K)$ the **Dvoretzky dimension**. Milman's proof of Dvoretzky's Theorem [M1] (see also [MS1, 5.8]) provides accurate information regarding the dimension of the almost spherical sections of K. Milman's argument shows that if $l < ck(K)$, then with probability larger than $1 - e^{-c'l}$, a random l-dimensional subspace $E \in G_{n,l}$ satisfies

$$\frac{c}{M}(D^n \cap E) \subset K \cap E \subset \frac{C}{M}(D^n \cap E), \tag{3}$$

where $M = M(K)$, D^n denotes the unit Euclidean ball in $\mathbb{R}^n$, and the randomness is induced by the unique rotation invariant probability measure on the grassmanian $G_{n,l}$ of l-dimensional subspaces in $\mathbb{R}^n$.

The Dvoretzky dimension $k(K)$ was proved in [MS2] to be the exact critical dimension for a random section to satisfy (3), in the following strong sense. If a random l-dimensional subspace $E \in G_{n,l}$ satisfies (3) with probability larger than, say, $1 - 1/n$, then necessarily $l < Ck(K)$. Thus a random section of dimension $l < ck(K)$ is close to Euclidean with high probability, and a random section of dimension $l > Ck(K)$ is typically far from Euclidean. These arguments resolve the question of the dimensions in which random sections of a given convex body are close to Euclidean. Once $b(K)$ and $M(K)$ are calculated, the behavior of a random section is known. For instance, the Dvoretzky dimension of the cube is $\asymp \log n$, while the cross polytope $K = \{x \in \mathbb{R}^n : \sum |x_i| \le 1\}$ has Dvoretzky dimension as large as $k(K) \asymp n$.

In this note we investigate Dvoretzky's Theorem from a different direction, which does not involve the standard large deviations inequality (2). The second-named author conjectured that a phenomenon similar to the concentration of measure should also occur for the small ball probability, and he proved a weaker statement. The conjecture has been recently proved by R. Latała and K. Oleszkiewitz [LO], using the solution to the B-conjecture by Cordero, Fradelizi and Maurey [CFM]:

THEOREM 1.1 (Small ball probability): *For every* $0 < \varepsilon < \frac{1}{2}$,

$$\sigma\{x \in S^{n-1} : \|x\| < \varepsilon M\} < \varepsilon^{ck(K)}$$

where $c > 0$ *is a universal constant.*

This theorem is related to the small ball probability (as a direction of the probability theory) in exactly the same way as the concentration of measure is related to large deviations. Here we apply Theorem 1.1 to study questions arising from Dvoretzky's Theorem. We show that for some purposes, it is possible to relax the Dvoretzky dimension $k(K)$, replacing it by a quantity *independent of the Lipschitzness* of the norm (which is quantified by the Lipschitz constant $b(K)$). Precisely, we wish to replace $k(K)$ by

$$d(K) = \min\Big\{ -\log \sigma\Big\{x \in S^{n-1} : \|x\| \le \frac{1}{2}M\Big\}, n\Big\},$$

where log stands for the natural logarithm. Selecting $t = \frac{1}{2}$ in the concentration of measure inequality (2), we conclude that $d(K)$ must be at least of the same order of magnitude as Dvoretzky's dimension $k(K)$:

$$d(K) \ge Ck(K).$$

The small ball Theorem 1.1 indeed holds with $d(K)$ (this is a part of the argument of Latała and Oleszkiewicz, reproduced below). The resulting inequality can be viewed as a Kahane–Khinchine type inequality for negative exponents:

Proposition 1.2 (Negative moments of a norm): *Assume that* $0 < \ell < cd(K)$. *Then*

$$cM < \Big(\int_{S^{n-1}} \|x\|^{-\ell} d\sigma(x) \Big)^{-\frac{1}{\ell}} < CM$$

where $c, C > 0$ *are universal constants.*

For positive exponents, this inequality was proved in [LMS]: for $0 < k < ck(K)$,

$$cM < \Big(\int_{S^{n-1}} \|x\|^{k} d\sigma(x) \Big)^{1/k} < CM. \tag{4}$$

For negative exponents $-1 < \ell < 0$, inequality (4) follows from results of Guédon [G] that generalize the Lovász–Simonovits inequality [LoSi]. Proposition 1.2 extends (4) to the range $[-cd(K), ck(K)]$ (which of course includes the range $[-ck(K), ck(K)]$).

In Proposition 1.2, $\|x\|^{-1}$ can be regarded as the radius of the one-dimensional section of the body K. Combining this with the recent inequality for diameters of sections due to the first-named author [K], we are able to lift the dimension of the section and thus compute the *average diameter of l-dimensional sections* of any centrally-symmetric convex body K.

Theorem 1.3 (Diameters of random sections): *Assume that* $0 < \ell < cd(K)$. *Select a random* ℓ*-dimensional subspace* $E \in G_{n,\ell}$. *Then with probability larger than* $1 - e^{-c'\ell}$,

$$K \cap E \subset \frac{C}{M}(D^n \cap E). \tag{5}$$

Furthermore,

$$\frac{\bar{c}}{M} < \left(\int_{G_{n,\ell}} \operatorname{diam}(K \cap E)^{\ell} d\mu(E) \right)^{1/\ell} < \frac{\bar{C}}{M} \tag{6}$$

where μ is the unique rotation invariant probability measure on $G_{n,\ell}$, and $c, c', \bar{c}, C, \bar{C} > 0$ are universal constants.

The relation between Theorem 1.3 and Dvoretzky's Theorem is clear. We show that for dimensions which may be much larger than $k(K)$, the upper inclusion in Dvoretzky's Theorem (3) holds with high probability. This reveals an intriguing point in Dvoretzky's Theorem. Milman's proof of Dvoretzky's Theorem focuses on the left-most inclusion in (3). Once it is proved that the left-most inclusion in (3) holds with high probability, the right-most inclusion follows almost automatically in his proof.

Furthermore, Milman–Schechtman's argument [MS2] implies in fact that the left-most inclusion does not hold (with large probability) for dimensions larger than the Dvoretzky dimension. The reason that a random l-dimensional section is far from Euclidean when $l > ck(K)$ is that a typical section does not contain a sufficiently large Euclidean ball. In comparison, we observe that the upper inclusion in (3) holds for a much wider range of dimensions.

There are cases, such as the case of the cube, where the Dvoretzky dimension satisfies $k(K) \asymp \log n$, while $d(K)$ is a polynomial in n. Hence, while sections of the cube of dimension n^c are already contained in the appropriate Euclidean ball (for any fixed $c < 1$, independent of n), only when the dimension is $\asymp \log n$, the sections start to "fill from inside", and an isomorphic Euclidean ball is observed. The specific case of the cube is contained, using different terminology, in [LMT]. The fact that $d(K)$ is typically larger than $k(K)$ is a little unexpected. It implies that the correct upper bound for random sections of a convex body appears sometimes in much larger dimensions than those for which we have the lower bound.

In the past decade, diameters of random lower-dimensional sections of convex bodies attracted a considerable amount of attention; see in particular [GM1, GM2, GMT]. Theorem 1.3 is a significant addition to this line of results. It implies that diameters of random sections are equivalent for a wide range of dimensions — starting from dimension one, when the random diameter simply equals $1/M(K)$, and up to the critical dimension $d(K)$. See also Remark 1 right after the proof of Theorem 1.3.

Remark: The proof shows that $d(K)$ can be further relaxed in all of our results. For any fixed $u > 1$, the parameter $d(K)$ can be replaced by

$$d_u(K) = \min\Big\{ -\log\sigma\Big\{x \in S^{n-1} : \|x\| \le \frac{1}{u}M\Big\}, n\Big\}.$$

The rest of the paper is organized as follows. In Section 2 we discuss the negative moments of the norm, proving Proposition 1.2 and Theorem 1.1 by the Latała–Oleszkiewicz argument. In Section 3 we perform the "dimension lift" and compute the average diameters of random sections, proving Theorem 1.3.

2. Concentration of measure and the small ball probability

We begin by proving Proposition 1.2. This proposition is a reformulation of the "small ball probability conjecture" due to the second-named author. It was recently deduced by R. Latała and K. Oleszkiewicz [LO] from the B-conjecture proved by Cordero, Fradelizi and Maurey [CFM]. We will reproduce the Latała–Oleszkiewicz argument here. We start with a standard and well-known lemma, on the close relation between the uniform measure σ on the sphere S^{n-1} and the standard gaussian measure γ on $\mathbb{R}^n$. For the reader's convenience, we include its proof.

LEMMA 2.1: *For every centrally-symmetric convex body $K \subset \mathbb{R}^n$,*

$$\frac{1}{2}\sigma(S^{n-1} \cap \frac{1}{2}K) \le \gamma(\sqrt{n}K) \le \sigma(S^{n-1} \cap 2K) + e^{-cn}$$

where $c > 0$ is a universal constant.

Proof: We will use the following two estimates on the Gaussian measure of the Euclidean ball,

$$\gamma(2\sqrt{n}D^n) > \frac{1}{2}, \quad \gamma(\frac{1}{2}\sqrt{n}D^n) < e^{-cn}.$$

The first estimate is simply Chebychev's inequality, and the second follows from standard large deviation inequalities, e.g. Cramer's Theorem [V]. Since K is star-shaped,

$$\begin{aligned}\gamma(\sqrt{n}K) &\ge \gamma(2\sqrt{n}D^n \cap \sqrt{n}K)\\ &\ge \gamma(2\sqrt{n}D^n)\sigma_1(2\sqrt{n}S^{n-1} \cap \sqrt{n}K)\end{aligned}$$

where σ_1 denotes the probability rotation invariant measure on the sphere $2\sqrt{n}S^{n-1}$,

$$\ge \frac{1}{2}\sigma(S^{n-1} \cap \frac{1}{2}K).$$

This proves the lower estimate in the lemma.

For the upper estimate, note that no points of $\sqrt{n}K$ can lie outside both the ball $\frac{1}{2}\sqrt{n}D^n$ and the positive cone generated by $\frac{1}{2}S^{n-1}\cap\sqrt{n}K$. Adding the two measures together, we obtain

$$\gamma(\sqrt{n}K) \le \gamma(\frac{1}{2}\sqrt{n}D^n) + \sigma_2(\frac{1}{2}\sqrt{n}S^{n-1} \cap \sqrt{n}K)$$

where σ_2 denotes the probability rotation invariant measure on the sphere $\frac{1}{2}\sqrt{n}S^{n-1}$,

$$\le e^{-cn} + \sigma(S^{n-1} \cap 2K).$$

This completes the proof. ∎

Proof of Proposition 1.2: As usual, K will denote the unit ball of the norm $\|\cdot\|$. The B-conjecture, proved in [CFM], asserts that the function $t \mapsto \gamma(e^t K)$ is log-concave. This means that for any $a, b > 0$ and $0 < \lambda < 1$,

$$\gamma(a^\lambda b^{1-\lambda} K) \ge \gamma(aK)^\lambda \gamma(bK)^{1-\lambda}. \tag{7}$$

Let $Med = Med(K)$ be the median of the norm $\|\cdot\|$ on the unit sphere S^{n-1}. By Chebychev's inequality, $Med \le 2M(K)$. Set $L = Med \cdot \sqrt{n}K$. According to Lemma 2.1,

$$\gamma(2L) \ge \frac{1}{2}\sigma(S^{n-1} \cap Med \cdot K) \ge \frac{1}{4} \tag{8}$$

by the definition of the median. On the other hand, again by Lemma 2.1,

$$\begin{aligned}\gamma(\frac{1}{8}L) &\le \sigma\Big(S^{n-1} \cap \frac{1}{4}Med \cdot K\Big) + e^{-cn} \\ &= \sigma\Big(x \in S^{n-1} : \|x\| \le \frac{1}{4}Med\Big) + e^{-cn} \\ &\le \sigma\Big(x \in S^{n-1} : \|x\| \le \frac{1}{2}M(K)\Big) + e^{-cn} \\ &\le e^{-d(K)} + e^{-c'n} < 2e^{-Cd(K)}\end{aligned} \tag{9}$$

because $d(K) \le n$. Let $0 < \varepsilon < e^{-3}$, and apply (7) for $a = \varepsilon$, $b = 2$, $\lambda = 3/\log(1/\varepsilon)$. This yields

$$\gamma(\varepsilon L)^{3/\log(1/\varepsilon)} \gamma(2L)^{1-3/\log(1/\varepsilon)} \le \gamma(\varepsilon^{3/\log(1/\varepsilon)} 2^{1-3/\log(1/\varepsilon)} L) \le \gamma(\frac{1}{8}L).$$

Combining this with (8) and (9), we obtain that

$$\gamma(\varepsilon L) \le 8e^{C' d(K) \log \varepsilon} \le 8\varepsilon^{cd(K)} < (c'\varepsilon)^{cd(K)}$$

and according to Lemma 2.1 we can transfer this to the spherical measure, obtaining

$$\sigma(x \in S^{n-1} : \|x\| < \varepsilon M) < (C\varepsilon)^{cd(K)}.$$

By integration by parts, this yields that for any $0 < \ell < cd(K)/10$,

$$\left(\int_{S^{n-1}} \|x/M\|^{-\ell} d\sigma(x) \right)^{1/\ell} \leq C,$$

which implies the left hand side of the inequality in Proposition 1.2. The right hand side follows easily by Hölder's inequality. ∎

By Chebychev's inequality, Proposition 1.2 yields the desired tail inequality for the small ball probability:

Corollary 2.2 (The small ball probability): *For every* $0 < \varepsilon < \frac{1}{2}$,

$$\sigma\{x \in S^{n-1} : \|x\| < \varepsilon M\} < \varepsilon^{cd(K)} < \varepsilon^{c'k(K)},$$

where $c, c' > 0$ *are universal constants.*

Theorem 1.1 is contained in Corollary 2.2. Let us offer some interpretation of the expression in Proposition 1.2. For a subspace $E \subset \mathbb{R}^n$, let $S(E) = S^{n-1} \cap E$ and let σ_E be the unique rotation invariant probability measure on the sphere $S(E)$. We will use the fact that $\mathrm{Vol}(K) = \mathrm{Vol}(D^n) \int_{S^{n-1}} \|x\|^{-n} d\sigma(x)$. The **volume radius** of a k-dimensional set T is defined as

$$\text{v.rad.}(T) = \left(\frac{\mathrm{Vol}(T)}{\mathrm{Vol}(D^k)} \right)^{1/k}.$$

Thus

$$\text{v.rad.}(K) = \left(\int_{S^{n-1}} \|x\|^{-n} d\sigma(x) \right)^{1/n},$$

for an n-dimensional body K. By the rotation invariance of all the measures (as in [K]), we conclude that

$$\begin{aligned} \int_{S^{n-1}} \|x\|^{-k} d\sigma(x) &= \int_{G_{n,k}} \int_{S(E)} \|x\|^{-k} d\sigma_E(x) d\mu(E) \\ &= \int_{G_{n,k}} \text{v.rad.}(K \cap E)^k d\mu(E), \end{aligned} \tag{10}$$

where, as before, μ is the unique rotation invariant probability measure on $G_{n,k}$. Thus Proposition 1.2 asymptotically computes the average volume radius of random sections. This perfectly fits the estimates for diameters of sections in [K], to be applied next.

3. Diameters of random sections

In this section we prove the main result of the paper, Theorem 1.3. We regard $\|x\|^{-1}$ as the radius of the one-dimensional section spanned by x; thus Proposition 1.2 is an asymptotically sharp bound on the diameters of random one-dimensional sections. Theorem 1.3 extends this bound to k-dimensional sections, for all k up to the critical dimension $d(K)$. We start with a "dimension lift", which is based on the "low M estimate", Proposition 3.9 in [K] (the case $\lambda = \frac{1}{2}$ there). Here we estimate the L_k norm, rather than only the tail probability as in Proposition 3.9 in [K].

Proposition 3.1 (Dimension lift for diameters of sections): *Let $1 \le k_0 < n$. Then for any integer $k < k_0/4$,*

$$\left(\int_{G_{n,k}} \operatorname{diam}(K \cap E)^k d\mu(E) \right)^{1/k} \le CM(K) \left(\int_{S^{n-1}} \|x\|^{-k_0} d\sigma(x) \right)^{2/k_0}.$$

Remark: Note the special normalization in Proposition 3.1 (compare with Proposition 3.9 in [K]). In fact, as follows from the proof, for any $\lambda > 0$, the right hand side of Proposition 3.1 may be replaced by

$$C(\lambda) M(K)^{\lambda} \left(\int_{S^{n-1}} \|x\|^{-k_0} d\sigma(x) \right)^{(1+\lambda)/k_0}$$

at the cost of replacing the requirement $k < k_0/4$ by $k < c(\lambda)k_0$ (simply replace the Cauchy–Schwartz inequality with the appropriate Hölder inequality in the proof). However, we cannot have $\lambda = 0$, as is demonstrated by the example where $K = \mathbb{R}^{n-1}$. The average diameter of one-dimensional sections of K is zero, while the diameter of any section of dimension larger than one is infinity. Thus there cannot be any formal dimension lift unless one has an extra factor which must be infinity for "flat" bodies. Such a factor is $M(K) = \int_{S^{n-1}} \|x\| d\sigma(x)$.

In order to prove Theorem 1.2 we need a standard lemma on the stability of the average norm M. We are unaware of a reference for the exact statement we need (a similar result appears e.g. in Lemma 6.6 of [M2]), so a proof is provided. The average norm on a subspace $E \in G_{n,k}$ is denoted by $M_E = \int_{S(E)} \|x\| d\sigma_E(x)$.

Lemma 3.2: *For every norm $\|\cdot\|$ on $\mathbb{R}^n$ and every integer $0 < k < n$,*

$$cM < \left(\int_{G_{n,k}} (M_E)^{2k} d\mu(E) \right)^{1/2k} < CM \tag{11}$$

where $c, C > 0$ are universal constants.

Proof: The left hand size inequality in (11) follows easily from Hölder's inequality. In the proof of the right hand side inequality, we will use a variant of Raz's argument (see [R, MW]). We normalize so that $M = 1$. Let $X_1, \dots, X_k$ be k independent random vectors, distributed uniformly on S^{n-1}. It is well-known that a norm of a random vector on the sphere has a subgaussian tail (e.g. [LMS, 3.1]. It actually follows from (2) above):

$$\mathbb{E}\exp(s\|X_i\|) < \exp(cs^2) \quad \text{for all } i \text{ and all } s > 1$$

which by independence implies

$$\mathbb{E}\exp\Big(s \cdot \frac{1}{k}\sum_{i=1}^{k} \|X_i\|\Big) < \exp\Big(\frac{Cs^2}{k}\Big) \quad \text{for } s > 1.$$

Using Chebychev's inequality and optimizing over s (e.g. [MS1, 7.4]), we obtain

$$\text{Prob}\left\{\frac{1}{k}\sum_{i=1}^{k} \|X_i\| > Ct\right\} < \exp(-t^2 k) \quad \text{for } t > 1. \tag{12}$$

Let E be the linear span of $X_1, \dots, X_k$. Then E is distributed uniformly in $G_{n,k}$ (up to an event of measure zero). Since for any two events one has $\text{Prob}(A) \le \text{Prob}(B)/\text{Prob}(B|A)$, we conclude that

$$\text{Prob}\{M_E > 2ct\} \le \frac{\text{Prob}\{\frac{1}{k}\sum_{i=1}^{k} \|X_i\| > ct\}}{\text{Prob}\{\frac{1}{k}\sum_{i=1}^{k} \|X_i\| > ct | M_E > 2ct\}}. \tag{13}$$

The numerator in (13) is bounded by (12). To bound the denominator from below, note that $\|X_i\| < C\sqrt{k} M_E$ pointwise for all i; this is a consequence of a simple comparison inequality for the Gaussian analogs of M and M_E (see e.g. [MS1, 5.9]). Let us fix a subspace $E \in G_{n,k}$. Note that, conditioning on $E = \text{span}\{X_1, \dots, X_k\}$, each of the vectors X_i is distributed uniformly in $S(E)$. Next, we estimate the probability

$$P_E = \text{Prob}\left\{\frac{1}{k}\sum_{i=1}^{k} \|X_i\| > \frac{M_E}{2} | \,\text{span}\{X_1, \dots, X_k\} = E\right\}$$

via Chebychev's inequality as

$$M_E = \mathbb{E}\Big(\frac{1}{k}\sum_{i=1}^{k} \|X_i\| \Big| \,\text{span}\{X_1, \dots, X_k\} = E\Big) \le C\sqrt{k} M_E P_E + \frac{M_E}{2}(1 - P_E).$$

Hence $P_E \geq \bar{c}/\sqrt{k}$ for every $E \in G_{n,k}$. Thus,

$$\begin{aligned}\text{denominator in (13)} &\geq \operatorname{Prob}\left\{\frac{1}{k}\sum_{i=1}^{k}\|X_i\| > \frac{M_E}{2}\,\Big|\, M_E > 2ct\right\}\\ &= \frac{1}{\operatorname{Prob}\{E \in G_{n,k};\, M_E > 2ct\}}\int_{E\in G_{n,k};M_E>2ct} P_E\,d\mu(E)\\ &\geq \min_{E\in G_{n,k}} P_E \geq \bar{c}/\sqrt{k}.\end{aligned}$$

Combining this with (12) and (13) we get

$$\operatorname{Prob}\{M_E > 2ct\} < c\sqrt{k}e^{-t^2k} < e^{-Ct^2k} \quad \text{for } t > 1.$$

By integration by parts we obtain the desired estimate. ∎

Proof of Proposition 3.1: By the Hölder inequality, the right hand side increases with k_0, hence we may assume that $k_0 = 4k$. We shall rely on the main result in [K], which claims that for any centrally-symmetric convex body $T \subset \mathbb{R}^n$, and for all $0 < k \leq l < n$,

$$\text{(14)} \qquad \text{v.rad.}(T) > C\left(\int_{G_{n,k}} \text{v.rad.}(T\cap E)^l \operatorname{diam}(T\cap E)^{n-l} d\mu(E)\right)^{1/n}.$$

We are going to apply (14) to $T = K \cap E$, for subspaces $E \in G_{n,k_0}$. Denote by $G_{E,k}$ the grassmanian of all k-dimensional subspaces of E, equipped with the unique rotational invariant probability measure μ_E. Then by (10), (14) and the rotational invariance of all measures,

$$\begin{aligned}&\int_{E\in G_{n,k}} \text{v.rad.}(K\cap E)^{2k} \operatorname{diam}(K\cap E)^{2k} d\mu(E)\\ &= \int_{E\in G_{n,4k}}\int_{F\in G_{E,k}} \text{v.rad.}(K\cap F)^{2k}\operatorname{diam}(K\cap F)^{2k} d\mu_E(F)d\mu(E)\\ &\leq C^{k_0}\int_{E\in G_{n,k_0}} \text{v.rad.}(K\cap E)^{k_0} d\mu(E) = C^{k_0}\int_{S^{n-1}} \|x\|^{-k_0} d\sigma(x).\end{aligned}$$

Also, by the Cauchy–Schwartz inequality,

$$\begin{aligned}&\left(\int_{E\in G_{n,k}} \operatorname{diam}(K\cap E)^k d\mu(E)\right)^{\frac{1}{k}}\\ &\leq \left(\int_{E\in G_{n,k}} \text{v.rad.}(K\cap E)^{2k}\operatorname{diam}(K\cap E)^{2k} d\mu(E)\right)^{\frac{1}{2k}}\\ &\quad\times\left(\int_{E\in G_{n,k}} \frac{1}{\text{v.rad.}(K\cap E)^{2k}} d\mu(E)\right)^{\frac{1}{2k}}.\end{aligned}$$

We will use the standard inequality

$$\frac{1}{v.rad.(K \cap E)} \leq M_E,$$

which follows directly from the Hölder inequality. Then,

$$\begin{aligned}&\Big(\int_{E\in G_{n,k}} \operatorname{diam}(K\cap E)^k d\mu(E)\Big)^{\frac{1}{k}}\\ &\leq C\Big(\int_{S^{n-1}} \|x\|^{-k_0} d\sigma(x)\Big)^{\frac{2}{k_0}}\Big(\int_{E\in G_{n,k}} (M_E)^{2k} d\mu(E)\Big)^{\frac{1}{2k}}\end{aligned}$$

and the proposition follows by Lemma 3.2. ∎

Proof of Theorem 1.3: It is sufficient to prove (6), since (5) follows by Chebychev's inequality. The left hand side inequality is clear. According to Proposition 3.1 and Proposition 1.2, the right hand side inequality of (6) follows, as

$$\Big(\int_{G_{n,\ell}} \operatorname{diam}(K\cap E)^\ell d\mu(E)\Big)^{\frac{1}{\ell}} < CM(K)\Big(\frac{C}{M(K)}\Big)^2 \leq \frac{C'}{M(K)}. \qquad ∎$$

Remarks. 1. (Optimality)*:* The estimate $\ell < cd(K)$ in Theorem 1.3 is essentially optimal, for any centrally-symmetric convex $K \subset \mathbb{R}^n$. Indeed, suppose that $\ell > d_u(K)$ for some $u \gg 1$. Then,

$$\Big(\int_{S^{n-1}} \|x\|^{-\ell} d\sigma(x)\Big)^{\frac{1}{\ell}} > \frac{u}{M}e^{-1} \gg \frac{1}{M}. \tag{15}$$

Since we always have $\operatorname{diam}(K\cap E) \geq 2\,\text{v.rad.}(K\cap E)$, then by (15) and (10),

$$\Big(\int_{G_{n,\ell}} \operatorname{diam}(K\cap E)^\ell d\mu(E)\Big)^{\frac{1}{\ell}} \gg \frac{1}{M}.$$

Thus, (6) cannot hold for $\ell > d_u(K)$ when $u \gg 1$.

2. (Example of the cube)*:* Suppose $K = B_\infty^n$ is a cube, the unit ball of l_∞^n, and let us estimate $d_u(K)$. It is well-known that

$$c\sqrt{\frac{\log n}{n}} < M(B_\infty^n) < C\sqrt{\frac{\log n}{n}}.$$

According to Lemma 2.1, we may equivalently carry out our computations in the Gaussian setting. Now, for $t > 0$,

$$\gamma(t\sqrt{\log n}B_\infty^n) = \prod_{i=1}^{n} \text{Prob}\{|X| \le t\sqrt{\log n}\} = (1 - 2\Phi(t\sqrt{\log n}))^n$$

where $X \sim N(0,1)$ is a standard normal random variable, and

$$\Phi(s) = \int_s^\infty \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du.$$

For $s > 1$, a crude estimate gives $e^{-c_1 s^2} < \Phi(s) < e^{-c_2 s^2}$. Thus, when $t > \frac{10}{\sqrt{\log n}}$,

$$\exp(-c_1' n^{1-c_2' t^2}) < \gamma(t\sqrt{\log n}B_\infty^n) < \exp(-c_1 n^{1-c_2 t^2}).$$

Therefore, by Lemma 2.1,

$$c_1 n^{1-c_1/u^2} < d_u(B_\infty^n) < c_2 n^{1-c_2/u^2}$$

for any $u > 1$. Theorem 1.3 implies that random sections of this (polynomial!) dimension $d_u(B_\infty^n)$ have a diameter $\asymp 1/M(B_\infty^n)$.

3. (Example of the l_p ball): Consider $K = B_p^n$, the unit ball of l_p^n, for some fixed $1 \le p < \infty$. In this case, the conclusion of Theorem 1.3 is well-known. For $1 \le p \le 2$, we know that $k(B_p^n) > cn$ (e.g. [MS1, 5.4]), hence also $d(B_p^n) > cn$ and the conclusion of the theorem follows from the classical Dvoretzky Theorem. In the case where $2 < p < \infty$, we have

$$B_p^n \subset \frac{c_p}{M(B_p^n)} D^n,$$

and thus the conclusion of Theorem 1.3 is obvious in this case (with constants depending on p), and $d_{c_p}(B_p^n) > C_p n$, for some constants c_p, C_p depending only on p.

References

[CFM] D. Cordero-Erausquin, M. Fradelizi and B. Maurey, *The (B) conjecture for the Gaussian measure of dilates of symmetric convex sets and related problems,* Journal of Functional Analysis **214** (2004), 410–427.

[DZ] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications,* Second edition, Applications of Mathematics (New York), 38, Springer-Verlag, New York, 1998.

[GM1] A. Giannopoulos and V. Milman, *On the diameter of proportional sections of a symmetric convex body,* International Mathematics Research Notices **1** (1997), 5–19.

[GM2] A. Giannopoulos and V. Milman, *Mean width and diameter of proportional sections of a symmetric convex body,* Journal für die reine und angewandte Mathematik **497** (1998), 113–139.

[GMT] A. Giannopoulos, V. Milman and A. Tsolomitis, *Asymptotic formulas for the diameter of sections of symmetric convex bodies,* Journal of Functional Analysis **223** (2005), 86–108.

[G] O. Guédon, *Kahane–Khinchine type inequalities for negative exponent,* Mathematika **46** (1999), 165–173.

[dH] F. den Hollander, *Large Deviations,* Fields Institute Monographs, 14, American Mathematical Society, Providence, RI, 2000.

[K] B. Klartag, *A geometric inequality and a low M estimate,* Proceedings of the American Mathematical Society **132** (2004), 2619–2628.

[LO] R. Latała and K. Oleszkiewitz, *Small ball probability estimates in terms of width,* Studia Mathematica **169** (2005), 305–314..

[L] M. Ledoux, *The Concentration of Measure Phenomenon,* Mathematical Surveys and Monographs, 89, American Mathematical Society, Providence, RI, 2001.

[LS] W. V. Li and Q. -M. Shao, *Gaussian processes: inequalities, small ball probabilities and applications, Stochastic processes: theory and methods,* in *Handbook of Statistics*, 19, North-Holland, Amsterdam, 2001, pp. 533–597.

[LoSi] L. Lovász and M. Simonovits, *Random walks in a convex body and an improved volume algorithm,* Random Structures and Algorithms **4** (1993), 359–412.

[LMS] A. E. Litvak, V. D. Milman, V. D. and G. Schechtman, *Averages of norms and quasi-norms,* Mathematische Annalen **312** (1998), 95–124.

[LMT] A. E. Litvak, P. Mankiewicz and N. Tomczak-Jaegermann, *Randomized isomorphic Dvoretzky theorem,* Comptes Rendus de l'Académie des Sciences, Paris **335** (2002), 345–350.

[M1] V. D. Milman, *A new proof of the theorem of A. Dvoretzky on sections of convex bodies,* Functional Analysis and its Applications **5** (1971), 28–38 (translated from Russian).

[M2] V. D. Milman, *Isomorphic symmetrizations and geometric inequalities,* in *Geometric Aspects of Functional Analysis (1986/87)*, Israel seminar, Lecture Notes in Mathematics **1317**, Springer-Verlag, Berlin, 1988, pp. 107–131.

[MS1] V. D. Milman and G. Schechtman, *Asymptotic Theory of Finite-dimensional Normed Spaces,* Lecture Notes in Mathematics **1200**, Springer-Verlag, Berlin, 1986.

[MS2] V. D. Milman and G. Schechtman, *Global versus local asymptotic theories of finite-dimensional normed spaces,* Duke Mathematical Journal **90** (1997), 73–93.

[MW] V. D. Milman and R. Wagner, *Some remarks on a lemma of Ran Raz,* in *Geometric Aspects of Functional Analysis*, Proceedings of the Israel seminar (GAFA) 2001–2002, Lecture Notes in Mathematics **1807**, 2003, pp. 158–168.

[R] R. Raz, *Exponential separation of quantum and classical communication complexity,* Annual ACM Symposium on Theory of Computing, Atlanta, GA, 1999, pp. 358–367.

[V] S. R. S. Varadhan, *Large Deviations and Applications,* CBMS-NSF Regional Conference Series in Applied Mathematics, Vol. 40, 1993.