Random version of Dvoretzky's theorem in ℓ_p^n

Grigoris Paouris[*], Petros Valettas and Joel Zinn[†]

October 27, 2015

Abstract

We study the dependence on ε in the critical dimension $k(n, p, \varepsilon)$ that one can find random sections of the ℓ_p^n-ball which are $(1+\varepsilon)$-spherical. For any fixed n we give lower estimates for $k(n, p, \varepsilon)$ for all eligible values p and ε, which agree with the sharp estimates for the extreme values $p = 1$ and $p = \infty$. In order to do so, we provide bounds for the gaussian concentration of the ℓ_p-norm.

1 Introduction

The fundamental theorem of A. Dvoretzky from [5] in geometric language states that every centrally symmetric convex body on $\mathbb{R}^n$ has a central section of large dimension which is almost spherical. The optimal form of the theorem, which was proved by V. Milman in [17], reads as follows. For any $\varepsilon \in (0, 1)$ there exists a function $\eta(\varepsilon) > 0$ with the following property: for every n-dimensional symmetric convex body T there exists $k \geq \eta(\varepsilon) \log n$ and k-dimensional subspace F such that

$$\frac{1-\varepsilon}{M} B_F \subseteq T \cap F \subseteq \frac{1+\varepsilon}{M} B_F, \tag{1.1}$$

where B_F denotes the Euclidean ball in F and $M = M(T) = \int_{S^{n-1}} \|\theta\|_T \, d\sigma(\theta)$ and σ stands for the uniform probability measure on the n-dimensional sphere S^{n-1}. The example of the cube $T = B_\infty^n$ shows that this result is best possible with respect to n (see [1] or [25] for the details). The approach of [17] is probabilistic in nature and shows that the vast majority of k-dimensional sections are $(1 + \varepsilon)$-spherical. Here the vast majority means in terms of the Haar probability measure $\nu_{n,k}$ on the Grassmannian manifold $G_{n,k}$. Furthermore, provides an asymptotic formula, in terms of the global parameters $M(T)$ and $b(T) = \max_{\theta \in S^{n-1}} \|\theta\|_T$ of the body T, for which the random k-dimensional section is $(1+\varepsilon)$-spherical as long as $k \leq c(\varepsilon)k(T)$ (here $c(\varepsilon)$ stands for the function of ε in the probabilistic formulation). Then, we find a good

[*]Supported by the A. Sloan foundation, BSF grant 2010288 and the NSF CAREER-1151711 grant;

[†]Partially supported by NSF grants DMS-1208962 and DMS-1205781.

linear image of the body T for which the $k(T)$ is large enough with respect to n (see [18] for further details). It has been proved in [19] that this formulation is optimal with respect to the dimension $k(T)$ in the sense that the k-dimensional sections which are 4-Euclidean with probability greater than $\frac{n}{n+k}$ cannot exceed $Ck(T)$ for some absolute constant $C > 0$. (Here and everywhere else C and c stand for absolute constants).

The probabilistic approach of [17] was providing $c(\varepsilon) \geq c\varepsilon^2/\log\frac{1}{\varepsilon}$ and this was improved to $c(\varepsilon) \geq c\varepsilon^2$ by Gordon in [7] and alternatively by Schechtman in [23]. This dependence is known to be optimal. The recent works of Schectman in [24] and Tikhomirov in [29] established that the dependence on ε in the randomized Dvoretzky for B_∞^n is of the exact order $\varepsilon/\log\frac{1}{\varepsilon}$.

As far as the dependence on ε in the existential version of Dvoretzky's theorem is concerned, Schechtman proved in [24] that one can always $(1+\varepsilon)$-embed ℓ_2^k in any n-dimensional normed space X with $k \geq c\varepsilon \log n/(\log\frac{1}{\varepsilon})^2$. Tikhomirov in [30] proved that for 1-symmetric space X we may have $k \geq c\log n/\log\frac{1}{\varepsilon}$ and this was subsequently extended by Fresen in [6] for permutation invariant spaces. In this note we will not deal with the existential Dvoretzky theorem. Related results for ℓ_p spaces are presented in [12]. For more detailed information on the subject, explicit statements and historical remarks the reader is consulted in the recent monograph [1].

In this note we study the random version for the spaces ℓ_p^n and we give bounds on the function $k(n, p, \varepsilon)$. These bounds are continuous with respect to p and coincide with the known bounds in the extreme cases $p = 1$ and $p = \infty$. In order to do that first we study the concentration phenomenon for the ℓ_p norms and we prove the following result:

Theorem 1.1. *Let $n \geq 2$. Then, for any $1 \leq p \leq \infty$ one has:*

$$P\left(\left|\|X\|_p - \mathbb{E}\|X\|_p\right| > \varepsilon\mathbb{E}\|X\|_p\right) \leq C_1\exp(-c_1\beta(n, p, \varepsilon)), \tag{1.2}$$

for $0 < \varepsilon < 1$ and $C_1, c_1 > 0$ absolute constants where $\beta(n, p, \varepsilon)$ is defined as follows:

i. If $1 \leq p \leq 2$, then

$$\beta(n, p, \varepsilon) \simeq \varepsilon^2 n, \tag{1.3}$$

ii. for $2 < p \leq \varepsilon_0 \log n$,

$$\beta(n, p, \varepsilon) \simeq \begin{cases} C^{-2p}\varepsilon^2 n, & 0 < \varepsilon < C^p n^{-\frac{p-2}{2(p-1)}} \\ \varepsilon^{2/p} n^{2/p}, & C^p n^{-\frac{p-2}{2(p-1)}} \leq \varepsilon < 1/p \\ \varepsilon k_p, & 1/p \leq \varepsilon < 1 \end{cases}, \tag{1.4}$$

where $C > 0$ is suitable absolute constant. In addition we have:

$$\beta(n, p, \varepsilon) \simeq \log(\varepsilon n^c), \tag{1.5}$$

for some absolute constant $c > 0$.

iii. For $p > \varepsilon_0 \log n$ we have:

$$\beta(n,p,\varepsilon) \simeq \varepsilon k_p, \tag{1.6}$$

where $\varepsilon_0 > 0$ is suitable absolute constant.

The bound we retrieve in the case of fixed p is not new. It has been appeared before in the literature by Naor [20] in an even more general probabilistic context. Also, for $p = \infty$ we recover the same bounds proved by Schechtman in [24]. Therefore, the above concentration result interpolates between the sharp concentration estimates for fixed $1 \leq p < \infty$ and $p = \infty$ and is derived in a unified way. However, our methods are different from the techniques used in [20] and [24] and utilize Gaussian functional inequalities. Actually, following the same ideas as in [23] we will prove a distributional inequality for Gaussian random matrices similar to the concentration inequality described above. Using this inequality and a chaining argument we prove the second main result of this note: The dependence on ε in randomized Dvoretzky for B_p^n balls.

Theorem 1.2. *Let $1 \leq p \leq \infty$. Then, for each n and for any $0 < \varepsilon < 1$ the random k-dimensional section of B_p^n with dimension $k \leq k(n,p,\varepsilon)$ is $(1+\varepsilon)$-Euclidean with probability greater than $1 - C\exp(-ck(n,p,\varepsilon))$, where $k(n,p,\cdot)$ is defined as:*

i. If $1 \leq p < 2$, then

$$k(n,p,\varepsilon) \simeq \varepsilon^2 n, \ 0 < \varepsilon < 1. \tag{1.7}$$

ii. If $2 < p < \varepsilon_0 \log n$, then

$$k(n,p,\varepsilon) \simeq \begin{cases} (Cp)^{-p}\varepsilon^2 n, & 0 < \varepsilon \leq (Cp)^{p/2} n^{-\frac{p-2}{2(p-1)}} \\ p^{-1}\varepsilon^{2/p} n^{2/p}, & (Cp)^{p/2} n^{-\frac{p-2}{2(p-1)}} < \varepsilon \leq 1/p \\ \varepsilon p n^{2/p}/\log\frac{1}{\varepsilon}, & \frac{1}{p} < \varepsilon < 1 \end{cases}. \tag{1.8}$$

In fact for $p < \varepsilon_0 \log n$ and $p \simeq \log n$ we have:

$$k(n,p,\varepsilon) \simeq \log n / \log\frac{1}{\varepsilon}. \tag{1.9}$$

iii. If $p \geq \varepsilon_0 \log n$, then

$$k(n,p,\varepsilon) \simeq \varepsilon \log n/\log\frac{1}{\varepsilon}, \ 0 < \varepsilon < 1. \tag{1.10}$$

where $C, c, \varepsilon_0 > 0$ are absolute constants.

As one observes the dependence on ε in $1 \leq p \leq 2$ is ε^2 as predicted by V. Milman's proof (and its improvement by [8] and [23]). However, for $p > 2$ the dependence on ε is much better than ε^2 for all values of p. This permits us to find sections of B_p^n of polynomial dimension which are closer to the Euclidean ball than previously obtained. Observe that Theorem 1.2 retrieves the right dependence on $c(\varepsilon)$ when $p = 1$ (actually when p is fixed) and at $p = \infty$.

The rest of the paper is organized as follows: In Section 2 we fix the notation, we give the required background material and we include some basic probabilistic inequalities. Gaussian functional inequalities as logarithmic Sobolev inequality, Talagrand's $L_1 - L_2$ inequality and Pisier's gaussian inequality are also included.

Instead of proceeding and proving Theorem 1.1 we prefer to deal with an easier problem first; the problem of determining the right order of the gaussian variance of the ℓ_p norm. We study this problem in Section 3. This is a warm-up for the concentration result we will investigate in Section 4. The main techniques that we will use, as well as the main problems we have to resolve, will be apparent already in this Section 3. This estimate will be used to obtain the dependence $\log n/\log\frac{1}{\varepsilon}$ for $p \simeq \log n$ in Theorem 1.2.

In Section 4 we provide a proof of Theorem 1.1. Moreover, efforts has been made to provide lower estimates in the probability described in Theorem 1.1.

In Section 5 we prove Theorem 1.2 and we show that in several cases the result is best possible up to constants.

We conclude in Section 6 with further remarks and open questions.

2 Notation and background material

We work in $\mathbb{R}^n$ equipped with the standard inner product $\langle x, y\rangle = \sum_{i=1}^n x_i y_i$ for $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$ in $\mathbb{R}^n$. The p-norm in $\mathbb{R}^n$ $(1 \leq p < \infty)$ is defined as:

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}, \quad x = (x_1, \dots, x_n) \tag{2.1}$$

and for $p = \infty$ as:

$$\|x\|_\infty = \max_{1\leq i\leq n} |x_i|, \; x = (x_1, \dots, x_n). \tag{2.2}$$

The Euclidean sphere is defined as: $S^{n-1} = \{x \in \mathbb{R}^n : \|x\|_2 = 1\}$. The normed space $(\mathbb{R}^n, \|\cdot\|_p)$ is denoted by ℓ_p^n, for $1 \leq p \leq \infty$. The unit ball of ℓ_p^n is denoted by B_p^n, i.e. $B_p^n = \{x \in \mathbb{R}^n : \|x\|_p \leq 1\}$. For $1 \leq p < q \leq \infty$ we have:

$$\|x\|_q \leq \|x\|_p \leq n^{1/p-1/q}\|x\|_q, \tag{2.3}$$

for all $x \in \mathbb{R}^n$. We write $\|\cdot\|$ for an arbitrary norm on $\mathbb{R}^n$ and $\|\cdot\|_K$ if the norm induced by the centrally symmetric convex body K on $\mathbb{R}^n$. For any subspace F of $\mathbb{R}^n$ we write: $S_F := S^{n-1} \cap F$ and $B_F := B_2^n \cap F$. For any linear operator $T : X \to Y$ between normed space we write $\|T\|_{X\to Y}$ for the operator norm. If $X = \ell_p$ we simply write $\|T\|_{p\to Y}$.

The random variables or vectors in some probability space (Ω, P) are denoted by ξ, η or $X = (X_1, \dots, X_n)$ or Y, W, Z. For the expectation we write $\mathbb{E}$ and for the variance Var. We shall make frequent use of Paley-Zygmund inequality:

Lemma 2.1. *Let X be non-negative random variable in some probability space (Ω, P) with $X \in L_2$. Then,*

$$P\left(X \geq t\mathbb{E}X\right) \geq (1-t)^2 \frac{(\mathbb{E}X)^2}{\mathbb{E}X^2}, \tag{2.4}$$

for all $0 < t < 1$.

For a proof see [2].

Also the multivariate version of Chebyshev's association inequality due to Harris will be useful:

Proposition 2.2 (Harris). *Let $\zeta_1, \ldots, \zeta_k$ be i.i.d. random variables taking values almost surely in $\Omega \subseteq \mathbb{R}$. If $F, G : \Omega^k \subseteq \mathbb{R}^k \to \mathbb{R}$ are coordinatewise non-decreasing[1] functions, then we have:*

$$\mathbb{E}F(Z)G(Z) \geq \mathbb{E}F(Z)\mathbb{E}G(Z), \tag{2.5}$$

where $Z = (\zeta_1, \ldots, \zeta_k)$.

Harris' inequality can be derived from consecutive applications of Chebyshev's association inequality and conditioning. For the detailed proof the reader is consulted in [2]. For some measure space $(\Omega, \mathcal{E}, \mu)$ we write

$$\|f\|_{L_p(\mu)} := \left(\int_\Omega |f|^p \, d\mu\right)^{1/p}, \ 1 \leq p < \infty. \tag{2.6}$$

for any measurable function $f : \Omega \to \mathbb{R}$. If μ is Borel probability measure on $\mathbb{R}^n$ and K centrally symmetric convex body on $\mathbb{R}^n$ we also use the notation

$$I_r(\mu, K) := \left(\int_{\mathbb{R}^n} \|x\|_K^r \, d\mu(x)\right)^{1/r}, \ -n < r \neq 0, \tag{2.7}$$

while

$$I_0(\mu, K) = \exp\left(\int_{\mathbb{R}^n} \log \|x\|_K \, d\mu(x)\right). \tag{2.8}$$

If σ is the uniform probability measure on S^{n-1} which is invariant under orthogonal transformations and T is centrally symmetric convex body on $\mathbb{R}^n$ then we write:

$$M_q(T) := \left(\int_{S^{n-1}} \|\theta\|_T^q \, d\sigma(\theta)\right)^{1/q}, \quad q \neq 0. \tag{2.9}$$

For the random version of Dvoretzky's theorem recall V. Milman's formulation from [17] (see also [18] or [1]) and see [8] and [23] for the dependence on ε:

---

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

Theorem 2.3. *Let T be centrally symmetric convex body on $\mathbb{R}^n$. Define the critical dimension $k(T)$ of T as follows:*

$$k(T) = \frac{\mathbb{E}\|g\|_T^2}{b^2(T)} \simeq n\left(\frac{M(T)}{b(T)}\right)^2, \tag{2.10}$$

where $b(T)$ is the Lipschitz constant of the map $x \mapsto \|x\|_T$, i.e. $b = \max_{\theta\in S^{n-1}}\|\theta\|_T$. Then, the random k-dimensional subbace F of $X_T := (\mathbb{R}^n, \|\cdot\|_T)$ satisfies:

$$\frac{1}{(1+\varepsilon)M}B_F \subseteq T\cap F \subseteq \frac{1}{(1-\varepsilon)M}B_F \tag{2.11}$$

with probability greater than $1-e^{-ck}$ provided that $k \leq k(n,\varepsilon)$, where $k(n,\varepsilon)\simeq \varepsilon^2 k(T)$.

Here the probability is considered with respect to the Haar probability measure $\nu_{n,k}$ on the Grassmann manifold $G_{n,k}$ the set of all k-dimensional subspaces of $\mathbb{R}^n$, which is invariant under the orthogonal group action.

With some abuse of terminology for a subspace F of a normed space $X = (\mathbb{R}^n, \|\cdot\|)$ (or equivalently for a section $T\cap F$ of a centrally symmetric convex body T on $\mathbb{R}^n$) we say that is $(1+\varepsilon)$-*spherical* (or *euclidean*) if:

$$\max_{\theta\in S_F}\|\theta\| / \min_{\theta\in S_F}\|\theta\| < 1+\varepsilon \quad \text{or} \quad \max_{z\in S_F}\|z\|_T / \min_{z\in S_F}\|z\|_T < 1+\varepsilon. \tag{2.12}$$

Thus, previous theorem states that the random k-dimensional subspace of X_T is $\frac{1+\varepsilon}{1-\varepsilon}$-spherical with probability greater than $1-e^{-ck}$ as long as $k \leq \varepsilon^2 k(X_T)$. For the ℓ_p^n spaces we abbreviate by k_p the critical dimension $k(\ell_p^n)$. Next Section provides asymptotic estimates for k_p in terms of n and p.

§1. **Gaussian averages of ℓ_p norms.** If g_1 is standard gaussian random variable we set $\sigma_p^p := \mathbb{E}|g_1|^p$ for every $p > 0$. The next asymptotic estimate follows easily by Stirling's formula:

$$\sigma_p^p = \mathbb{E}|g_1|^p = \frac{2^{p/2}}{\sqrt{\pi}}\Gamma\left(\frac{p+1}{2}\right) \sim \sqrt{2/e}\left(\frac{p+1}{e}\right)^{p/2}, \quad p\to\infty. \tag{2.13}$$

The n-dimensional standard Gaussian measure with density $(2\pi)^{-n/2}e^{-\|x\|_2^2/2}$ is denoted by γ_n. Next Proposition is a special case of a more general result from [26].

Proposition 2.4. *Let $1\leq p\leq\infty$. Then, we have:*

$$\mathbb{E}_{\gamma_n}\|g\|_p = \int_{\mathbb{R}^n}\|x\|_p\, d\gamma_n(x) \simeq \begin{cases} n^{1/p}\sqrt{p}, & p < \log n \\ \sqrt{\log n}, & p \geq \log n \end{cases}. \tag{2.14}$$

We shall need Gordon's lemma for the Mill's ratio from [7]:

Lemma 2.5 (Gordon, 1941). *For any $x > 0$ we have:*

$$\frac{x}{1+x^2} \leq e^{x^2/2}\int_x^\infty e^{-t^2/2}\,dt \leq \frac{1}{x}. \tag{2.15}$$

Equivalently, we have:

$$1 \le \frac{\phi(x)}{x\Phi(-x)} \le 1 + \frac{1}{x^2}, \tag{2.16}$$

for $x > 0$, where $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2}\, dt$ and $\phi = \Phi'$.

§2. **Functional inequalities on Gauss' space.** First we refer to the logarithmic Sobolev inequality. In general, if μ is Borel measure on $\mathbb{R}^n$ it is said that μ satisfies log-Sobolev inequality with constant ρ if for any smooth function f we have:

$$\mathrm{Ent}_\mu(f^2) := \mathbb{E}_\mu(f^2 \log f^2) - \mathbb{E}_\mu f^2 \log(\mathbb{E}_\mu f^2) \le \frac{2}{\rho} \int \|\nabla f\|_2^2 \, d\mu. \tag{2.17}$$

It is well known (see [14]) that the standard n-dimensional Gaussian measure γ_n satisfies log-Sobolev inequality with $\rho = 1$. Next lemma, based on classical Herbst's argument, is a useful estimate which holds for any measure satisfying log-Sobolev inequality:

Lemma 2.6. *Let μ be a measure satisfying log-Sobolev inequality with constant $\rho > 0$. Then, for any Lipschitz map f and for $2 \le p < q$ we have:*

$$\|f\|_{L_q(\mu)}^2 - \|f\|_{L_p(\mu)}^2 \le \frac{\|f\|_{\mathrm{Lip}}^2}{\rho}(q-p). \tag{2.18}$$

In particular, we have:

$$\frac{\|f\|_{L_q(\mu)}}{\|f\|_{L_2(\mu)}} \le \sqrt{1 + \frac{q-2}{\rho k(f)}}, \tag{2.19}$$

for $q \ge 2$ where $k(f) := \|f\|_{L_2(\mu)}^2 / \|f\|_{\mathrm{Lip}}^2$. Furthermore,

$$\frac{\|f\|_{L_2(\mu)}}{\|f\|_{L_p(\mu)}} \le \exp\left(\frac{1/p - 1/2}{\rho k(f)}\right), \tag{2.20}$$

for $0 < p \le 2$.

Proof. The proof of the first estimate is essentially contained in [27]. The second one is direct application of the first for $p = 2$. For the last assertion, note that by Lyapunov's convexity theorem (see [9]) the map $p \overset{\phi}{\mapsto} \log \|f\|_p^p$ is convex. Moreover, we have:

$$p\phi'(p) - \phi(p) = \frac{\mathrm{Ent}_\mu(|f|^p)}{\int |f|^p \, d\mu}. \tag{2.21}$$

For any $0 < p < 2$, the convexity of ϕ and log-Sobolev inequality yield:

$$2\frac{\phi(2) - \phi(p)}{2-p} \le 2\phi'(2) = \frac{\mathrm{Ent}_\mu(f^2)}{\|f\|_2^2} + \phi(2) \le \frac{2}{2\rho k} + \phi(2),$$

which gives that:

$$\log\left(\frac{\|f\|_2}{\|f\|_p}\right) \leq \frac{2-p}{2pk\rho}, \tag{2.22}$$

where $k \equiv k(f)$. □

Note. The above two estimates imply that:

$$\frac{\|f\|_{L_q(\gamma_n)}}{\|f\|_{L_1(\gamma_n)}} \leq \sqrt{1 + c_1\frac{q-1}{k(f)}}, \quad q \geq 1, \tag{2.23}$$

as long as $k(f) \gg 1$. Furthermore, integration in polar coordinates yields:

$$I_r(\gamma_n, C) = c_{n,r} M_r(C), \tag{2.24}$$

for any centrally symmetric convex body C on $\mathbb{R}^n$, where $c_{n,r} := \sqrt{2}[\Gamma(\frac{n+r}{2})/\Gamma(\frac{n}{2})]^{1/r}$ and $M_r^r(C) := \int_{S^{n-1}} \|\theta\|_C^r \, d\sigma(\theta)$. Applying this for $C = B_2^n$ we readily see that $c_{n,r} = I_r(\gamma_n, B_2^n)$. Therefore, for $-n < s < r$ we obtain:

$$\max\left\{\frac{M_r(C)}{M_s(C)}, \frac{I_r(\gamma_n, B_2^n)}{I_s(\gamma_n, B_2^n)}\right\} \leq \frac{M_r(C) I_r(\gamma_n, B_2^n)}{M_s(C) I_s(\gamma_n, B_2^n)} = \frac{I_r(\gamma_n, C)}{I_s(\gamma_n, C)}. \tag{2.25}$$

It follows that:

$$M_q(C)/M_1(C) \leq \sqrt{1 + c_1\frac{q-1}{k(C)}}, \quad q \geq 1. \tag{2.26}$$

This estimate improves considerably upon the estimate presented in [15, Statement 3.1] or [14, Proposition 1.10, (1.19)] in the range $1 \leq q \leq k(C)$. For a purely probabilistic approach of this fact the reader is consulted in [21]. It is immediate that $I_r \lesssim \begin{cases} I_1, & 1 \leq r \leq k \\ \sqrt{r/k} I_1, & r \geq k \end{cases}$ for any Lipschitz function. In [15] it was proved that for norms this estimate can be reversed:

Lemma 2.7. *Let $\|\cdot\|$ be a norm on $\mathbb{R}^n$. Then, we have:*

$$I_r \simeq \begin{cases} I_1, & r \leq k \\ \sqrt{r/k} I_1, & r \geq k \end{cases}. \tag{2.27}$$

This result implies the next well known fact:

Proposition 2.8. *Let $\|\cdot\|$ be a norm on $\mathbb{R}^n$. Then, we have:*

$$c\exp(-Ct^2k) \leq P(\|X\| > (1+t)I_1) \leq C\exp(-ct^2k), \tag{2.28}$$

for $t \geq 1$. Moreover, one has:

$$\left(\mathbb{E}\big|\|X\| - \mathbb{E}\|X\|\big|^r\right)^{1/r} \simeq \sqrt{\frac{r}{k}}\mathbb{E}\|X\|, \tag{2.29}$$

for all $r \geq k$.

Proof. (Sketch). There exists $c_1 \in (0,1)$ such that $I_s \geq c_1 \sqrt{s/k} I_1$ for all $s > k$. Thus, for $t \geq 1$ if we choose $r > k$ by $c_1 \sqrt{r/k} = 4t$ we may write:

$$P\left(\|X\| > \frac{1}{2} I_r\right) \leq P\left(\|X\| > \frac{c_1}{2} \sqrt{r/k} I_1\right) \leq P(\|X\| \geq (1+t) I_1).$$

On the other hand Paley-Zygmund inequality (Lemma 2.1) yields:

$$P\left(\|X\| > \frac{1}{2} I_r\right) \geq (1 - 2^{-r})^2 (I_r / I_{2r})^{2r} \geq c_2 e^{-C_2 r} \geq c_2 \exp(-C_2' t^2 k),$$

where we have used also the fact that $I_r \simeq I_{2r}$ which follows by Lemma 2.7. For the second assertion we apply integration by parts and we use the first estimate. □

The above estimate shows that the large deviation estimates for norms with respect to γ_n are completely settled. Therefore for the concentration inequalities we are interested in, we may restrict ourselves from now on in the range $0 < \varepsilon < 1$.

Other important functional inequalities related with the concentration of measure phenomenon are Poincaré inequalities. Using a standard variational argument (see [14]) we can show that any measure satisfying log-Sobolev inequality with constant ρ also satisfies Poincaré inequality with constant ρ:

$$\rho \mathrm{Var}_\mu(f) \leq \int_{\mathbb{R}^n} \|\nabla f\|_2^2 \, d\mu, \tag{2.30}$$

for any smooth function f.

A refinement of Poincaré inequality was proved by Talagrand in [28] for the discrete cube $\{-1,1\}^n$ (see also [2] for a recent exposition) and its continuous version, in the Gaussian context, was presented in [4] (see also [3]):

Theorem 2.9 (Talagrand's $L_1 - L_2$ bound). *Let $f : \mathbb{R}^n \to \mathbb{R}$ be smooth function. If $A_i := \|\partial_i f\|_{L_2(\gamma_n)}$ and $a_i := \|\partial_i f\|_{L_1(\gamma_n)}$, then one has:*

$$\mathrm{Var}_{\gamma_n}(f) \leq C \sum_{i=1}^n \frac{A_i^2}{1 + \log(A_i/a_i)}. \tag{2.31}$$

This inequality will be used in order to prove concentration for the ℓ_p norm when p is sufficiently large.

Pisier discovered (see [22]) another Gaussian inequality which contains (r,r)-Poincaré inequalities and the gaussian concentration inequality. Since, the proof is very elegant and useful for our approach we reproduce it below:

Theorem 2.10. *Let $\phi : \mathbb{R} \to \mathbb{R}$ be convex function and let $f : \mathbb{R}^n \to \mathbb{R}$ be C^1-smooth. Then, if X, Y are independent copies of a Gaussian random vector, then we have:*

$$\mathbb{E}\phi\left(f(X) - f(Y)\right) \leq \mathbb{E}\phi\left(\frac{\pi}{2}\langle \nabla f(X), Y\rangle\right). \tag{2.32}$$

Proof. (B. Maurey). Set $X_\theta := \cos\theta \cdot X + \sin\theta \cdot Y$ and $Y_\theta = \frac{d}{d\theta} X_\theta$. Note that (X_θ, Y_θ) has the same distribution as (X, Y) since $(X_\theta, Y_\theta) = \tilde{R}_\theta(X, Y)$ and $\tilde{R}_\theta$ is the matrix obtained by tensorization of $R_\theta := \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$ with I_n, i.e. $\tilde{R}_\theta = R_\theta \otimes I_n \in \mathbb{R}^{2n\times 2n}$ and is orthogonal. Furthermore, we may write:

$$\begin{aligned}\mathbb{E}\phi(f(Y) - f(X)) &= \mathbb{E}\phi\left(\int_0^{\pi/2} \frac{d}{d\theta} f(X_\theta)\, d\theta\right) = \mathbb{E}\phi\left(\int_0^{\pi/2} \langle \nabla f(X_\theta), Y_\theta\rangle\, d\theta\right) \\ &\leq \frac{2}{\pi}\int_0^{\pi/2} \mathbb{E}\phi\left(\frac{\pi}{2}\langle \nabla f(X_\theta), Y_\theta\rangle\right) = \mathbb{E}\phi\left(\frac{\pi}{2}\langle \nabla f(X), Y\rangle\right),\end{aligned}$$

where we have used Jensen's inequality, Fubini's theorem and previous remark. □

Remark 2.11. 1. *Poincaré inequality.* When $\phi(t) = t^2$ we have:

$$\mathrm{Var}(f(X)) \leq \frac{\pi^2}{8}\mathbb{E}|\langle \nabla f(X), Y\rangle|^2 = \frac{\pi^2}{8}\mathbb{E}\|\nabla f(X)\|_2^2, \tag{2.33}$$

which is Poincaré's inequality for Gaussian measure with non-optimal constant.
2. (r, r)-*Poincaré inequalities.* Even more generally, for $\phi(t) = |t|^r$, $r \geq 1$ we get:

$$\|f - \mathbb{E}f\|_{L_r(\gamma_n)} \simeq (\mathbb{E}|f(X) - f(Y)|^r)^{1/r} \leq \frac{\pi}{2}\sigma_r \left(\mathbb{E}\|\nabla f(X)\|_2^r\right)^{1/r}. \tag{2.34}$$

3. *Gaussian concentration.* For $\phi_\lambda(t) = \exp(\lambda t)$, $\lambda > 0$ we obtain:

$$\mathbb{E}\exp\left(\lambda(f(X) - f(Y))\right) \leq \mathbb{E}\exp\left(\lambda\frac{\pi}{2}\langle \nabla f(X), Y\rangle\right). \tag{2.35}$$

Note that if f is C^1-smooth and Lipschitz then, one has:

$$|\langle \nabla f(x), u\rangle| = |D_u f(x)| = \lim_{t\to 0}\left|\frac{f(x + tu) - f(x)}{t}\right| \leq \|f\|_{\mathrm{Lip}},$$

for all $u \in S^{n-1}$ and $x \in \mathbb{R}^n$. It follows that $\|\nabla f(x)\|_2 \leq \|f\|_{\mathrm{Lip}}$ for all $x \in \mathbb{R}^n$. Therefore, we have:

$$\mathbb{E}\exp\left(\lambda\frac{\pi}{2}\langle \nabla f(X), Y\rangle\right) = \mathbb{E}\exp\left(\frac{\lambda^2\pi^2}{2}\|\nabla f(X)\|_2^2\right) \leq \exp\left(\frac{\pi^2}{2}\lambda^2\|f\|_{\mathrm{Lip}}^2\right).$$

Let $t > 0$. It follows that:

$$\begin{aligned}P(f(X) - \mathbb{E}f > t) \leq e^{-\lambda t}\mathbb{E}\exp(\lambda(f(X) - \mathbb{E}f)) &\leq e^{-\lambda t}\mathbb{E}\exp(\lambda(f(X) - f(Y))) \\ &\leq e^{-\lambda t + \frac{\pi^2}{2}\lambda^2\|f\|_{\mathrm{Lip}}^2}.\end{aligned}$$

Optimizing over $\lambda > 0$ by choosing $\lambda = t/(\pi^2\|f\|_{\mathrm{Lip}}^2)$ we obtain:

$$P(f(X) - \mathbb{E}f > t) \leq \exp\left(-\frac{1}{2\pi^2}\frac{t^2}{\|f\|_{\mathrm{Lip}}^2}\right).$$

The same reasoning applied to $-f$ yields the concentration inequality:

$$P(|f(X) - \mathbb{E}f| > t) \leq 2\exp(-t^2/(2\pi^2\|f\|_{\mathrm{Lip}}^2)), \tag{2.36}$$

for all $t > 0$. Alternatively, we may conclude similar estimate by equations (2.34) and Markov's inequality.

§ 3. Negative moments of norms. Next Theorem is due Klartag and Vershynin from [11]:

Proposition 2.12. *Let T be a centrally symmetric convex body on $\mathbb{R}^n$. We define:*

$$d(T) := \min\left\{n, -\log\gamma_n\left(\frac{m}{2}T\right)\right\}, \tag{2.37}$$

where m is the median of $x \mapsto \|x\|_T$ with respect to γ_n. Then, one has:

$$\gamma_n\left(\{x : \|x\|_T \leq c\varepsilon\mathbb{E}\|X\|_T\}\right) \leq (C\varepsilon)^{cd(T)}, \tag{2.38}$$

for all $0 < \varepsilon < \varepsilon_0$ where $\varepsilon_0 > 0$ is an absolute constant.

Then, for all $0 < k < d(T)$ we have: $I_{-k}(\gamma_n, T) \geq cI_1(\gamma_n, T)$. Note that $d(T) > c_1k(T)$.

Note that this result inform us that the negative moments exhibit a stable behavior up to the point $d(T)$. Nevertheless one can show that up to the critical dimension the moments of any norm with respect to the Gaussian (or the uniform on the sphere) measure are almost constant, complementing the result of §2. We will need the next consequence of Proposition 2.12.

Lemma 2.13. *Let T be centrally symmetric convex body on $\mathbb{R}^n$ which satisfies the small ball probability estimate:*

$$\gamma_n(\varepsilon I_1 T) < (A\varepsilon)^{\alpha d}, \tag{2.39}$$

for all $0 < \varepsilon < \varepsilon_0$ ($A, \alpha > 0$). Then, for all $r, s > 0$ with $r + s < \alpha d/3$ we have:

$$I_{-r-s}^{-r-s}(\gamma_n, T) \leq \left(\frac{CA}{I_1}\right)^s I_{-r}^{-r}(\gamma_n, T), \tag{2.40}$$

where $C > 0$ is an absolute constant.

Proof. For any $0 < \varepsilon < \varepsilon_0$ we may write:

$$\begin{aligned} I_{-r-s}^{-r-s} = \int \frac{1}{\|x\|^{r+s}}\, d\gamma_n(x) &\leq \frac{1}{(\varepsilon I_1)^s}\int \frac{1}{\|x\|^r}\, d\gamma_n(x) + \int_{\varepsilon I_1 T} \frac{1}{\|x\|^{r+s}}\, d\gamma_n(x) \\ &\leq \frac{1}{(\varepsilon I_1)^s} I_{-r}^{-r} + (A\varepsilon)^{\alpha d/2} I_{-2(r+s)}^{-r-s}, \end{aligned}$$

by Cauchy-Schwarz inequality. Note that the small ball probability assumption implies that: $I_{-s} \geq c\varepsilon_0 I_1$ for all $0 < s < 2\alpha d/3$. Thus, if $r + s < \alpha d/3$ we get $I_{-2(r+s)} > c_1 I_{-(r+s)}$ and previous estimate yields:

$$I_{-r-s}^{-r-s} < \frac{1}{(\varepsilon I_1)^s} I_{-r}^{-r} + (A\varepsilon)^{\alpha d/2} c_1^{-r-s} I_{-r-s}^{-r-s}.$$

Choosing ε small enough so that $(A\varepsilon)^{\alpha d/2} < c_1^{r+s}/2$, say $0 < \varepsilon \leq c_1/(2A)$, then we conclude the result. □

Theorem 2.14. *Let T be a centrally symmetric convex body on $\mathbb{R}^n$ and let $k \equiv k(T)$ be the critical dimension of $X_T = (\mathbb{R}^n, \|\cdot\|_T)$. Then, one has:*

$$\frac{I_r(\gamma_n, T)}{I_{-r}(\gamma_n, T)} \leq 1 + \frac{Cr}{k}, \tag{2.41}$$

for all $0 < r < ck$, where $C, c > 0$ are absolute constants.

Proof. We present the argument in three steps:

Step 1. (positive moments). We use the log-Sobolev inequality, to estimate the growth of the moments. The basic observation is that:

$$\frac{d}{dr}\left(\log \|f\|_{L_r(\mu)}\right) = \frac{\mathrm{Ent}_\mu(|f|^r)}{r^2 \|f\|_{L_r(\mu)}^r}.$$

Applying this for the function $\|\cdot\|_T : (\mathbb{R}^n, \gamma_n) \to \mathbb{R}$ we find:

$$(\log I_r(T))' \leq \frac{1}{2I_r^r} \mathbb{E}\|X\|_T^{r-2} \| \nabla \|X\|_T \|_2 \leq \frac{b^2}{2I_r^r} I_{r-2}^{r-2}, \tag{2.42}$$

for all $r > 0$. It is easy to see that $(\log I_r)' \leq \frac{1}{2k(T)}$ for $r \geq 2$, while for $0 < r < 2$ we may write:

$$(\log I_r)' \leq \frac{b^2}{2I_{-(2-r)}^2} \leq \frac{C_1 b^2}{I_1^2} = \frac{C_1}{k(T)}, \tag{2.43}$$

where we have used Proposition 2.12.

Step 2. (negative moments). As before, using log-Sobolev inequality, for all $0 < r < c_1 d(T)$ we may write:

$$(\log I_{-r})' \geq -\frac{b^2}{2I_{-r}^{-r}} I_{-r-2}^{-r-2} \geq -\frac{C_2 b^2}{I_1^2} = -\frac{C_2}{k(T)}, \tag{2.44}$$

where we have used Lemma 2.13.

Step 3. Using (2.43) we may write:

$$\log(I_r/I_0) = \int_0^r (\log I_t)' \, dt \leq \int_0^r \frac{C_1}{k} \, dt = \frac{C_1 r}{k}, \tag{2.45}$$

for all $r > 0$. The same reasoning applied to (2.44) shows that:

$$\log(I_{-r}/I_0) \geq -\frac{C_2 r}{k}, \tag{2.46}$$

for all $0 < r < c_1 d(T)$. Adding these two estimates up and restricting $0 < r < c_2 k(T)$ we conclude that:

$$I_r/I_{-r} \leq \exp\left(\frac{C_3 r}{k}\right) \leq 1 + \frac{C_4 r}{k}, \tag{2.47}$$

as claimed. □

3 The variance of ℓ_p norm

A standard way that provides upper estimates for the variance is concentration inequality (2.36), e.g. see [15] or [14, Proposition 1.9]. An integration by parts argument implies that if $f : \mathbb{R}^n \to \mathbb{R}$ is L-Lipschitz function, then:

$$\mathrm{Var}(f) \leq CL^2, \tag{3.1}$$

for some absolute constant $C > 0$. In particular, if $f(x) = \|x\|_p$ this estimate yields:

$$\mathrm{Var}\|X\|_p \leq Cb^2(B_p^n) \simeq \max\{n^{2/p-1}, 1\}. \tag{3.2}$$

For $1 \leq p \leq 2$ this estimate turns out to be the correct one. But, for $2 < p \leq \infty$ this method gives bounds which are far away from the actual ones. The purpose of this paragraph is to compute the correct order of the variance of the ℓ_p norm with respect to the Gaussian measure.

§ 1. The variance of ℓ_p norm for $1 \leq p < \infty$. Our first approach lies in determining the limit distribution of the sequence of variables $(\|g\|_{\ell_p^n})_{n=1}^\infty$ we use the next Proposition known in Statistics as "Delta Method" (for a proof see [10, Chapter 5]):

Proposition 3.1. *Let (Y_n) be a sequence of random variables that satisfies $n^{1/2}(Y_n - \theta) \xrightarrow{d} N(0, \sigma^2)$ in distribution. For the differentiable function h assume that $h'(\theta) \neq 0$. Then,*

$$n^{1/2}(h(Y_n) - h(\theta)) \xrightarrow{d} N(0, \sigma^2(h'(\theta))^2) \tag{3.3}$$

in distribution.

Now we may prove the next asymptotic estimate:

Theorem 3.2. *Let $1 \leq p < \infty$. Let $(\xi_j)_{j=1}^\infty$ be sequence of i.i.d random variables with $m_{3p}^{3p} := \mathbb{E}|\xi_1|^{3p} < \infty$. Then, there exist positive constants c_p, C_p depending only on p and the distribution of (ξ_j) such that:*

$$c_p n^{\frac{2}{p}-1} \leq \mathrm{Var}\|\xi\|_{\ell_p^n} \leq C_p n^{\frac{2}{p}-1}, \tag{3.4}$$

for all n.

Proof. Let $Y_n := \frac{1}{n}\sum_{j=1}^n |\xi_j|^p$. Then by the Central Limit Theorem we know that:

$$\sqrt{n}(Y_n - m_p^p) \xrightarrow{d} N(0, v_p^2) \tag{3.5}$$

in distribution, where $v_p^2 := \mathrm{Var}|\xi_1|^p$. Consider the function $h(t) = t^{1/p}$, $t > 0$ and apply Proposition 3.1 to get:

$$\sqrt{n}(n^{-1/p}\|\xi\|_p - m_p) \xrightarrow{d} N\left(0, \frac{v_p^2}{p^2} m_p^{2(1-p)}\right), \tag{3.6}$$

in distribution. Thus, the sequence of any moment of ξ_i's converges to the corresponding moment of the limit distribution, in particular we conclude:

$$n^{1-\frac{2}{p}}\mathrm{Var}(\|\xi\|_p) = \mathrm{Var}\left(n^{\frac{1}{2}-\frac{1}{p}}\|\xi\|_p\right) = \mathrm{Var}\left[\sqrt{n}(n^{-1/p}\|\xi\|_p - m_p)\right] \to \frac{v_p^2}{p^2} m_p^{2(1-p)},$$

as $n \to \infty$ and the result follows. □

Remark. The reader should notice that for fixed $p \geq 1$ the dependence on the dimension is independent of the randomness we choose for the underlying variables (ξ_i) and the argument is crucially based on the stochastic independence. Moreover, in the case that (ξ_i) are normally distributed with mean zero and variance one, the above limit value is estimated as:

$$\frac{v_p^2}{p^2} m_p^{2(1-p)} \sim \frac{1}{e\sqrt{2}} \frac{2^p}{p}, \quad p \to \infty.$$

This suggests that the constants c_p, C_p should depend exponentially on p.

§2. The variance of ℓ_∞ norm. Of course the variance in that case can be computed by employing the tail estimates proved in [24] for the ℓ_∞-norm with respect to the Gaussian measure. We prefer to give a proof here of more "probabilistic flavor". Actually, the argument we present below works for any i.i.d. random variables with exponential tails, but we shall focus on Gaussians. Let $(g_i)_{i=1}^\infty$ be independent, standard Gaussian random variables and let $Y_n := \max_{i\leq n}|g_i|$, $n \geq 2$. We set $a_n := -\Phi^{-1}(\frac{1}{2n}) > 0$ and let $b_n := 1/a_n$. Note that $a_n \to \infty$ and Gordon's inequality (2.16) shows that $a_n \sim \sqrt{2\log n}$ as $n \to \infty$. We define $W_n := \frac{Y_n - a_n}{b_n}$ and we have the next well known fact whose proof is included for the sake of completeness:

Proposition 3.3. *Let η be Gumbel random variable, that is the cumulative distribution function of η is given as:*

$$F_\eta(t) := \exp(-e^{-t}), \ t \in \mathbb{R}. \tag{3.7}$$

If (W_n) is sequence defined as above then, for every $t \in \mathbb{R}$ we have:

$$\mathbb{P}(Z_n \leq t) \to \exp(-e^{-t}), \tag{3.8}$$

that is W_n converges to the Gumbel variable in distribution, i.e. $W_n \xrightarrow{d} \eta$.

Proof. For fixed t we may write:

$$\mathbb{P}(W_n \le t) = \mathbb{P}(Y_n \le tb_n + a_n) = [\mathbb{P}(|g_1| \le tb_n + a_n)]^n = [1 - 2\Phi(-tb_n - a_n)]^n .$$

Taylor's theorem yields:

$$\log \Phi(-tb_n - a_n) = \log \Phi(-a_n) - tb_n \frac{\phi(-a_n)}{\Phi(-a_n)} + O(t^2 b_n^2), \quad n \to \infty$$

or equivalently,

$$\log 2\Phi(-tb_n - a_n) = \log \frac{1}{n} - t\frac{\phi(a_n)}{a_n\Phi(-a_n)} + O\left(\frac{t^2}{\log n}\right).$$

Finally, we get:

$$\mathbb{P}(W_n \le t) = \left[1 - \frac{1}{n}\exp\left(-t\frac{\phi(a_n)}{a_n\Phi(-a_n)} + O\left(\frac{t^2}{\log n}\right)\right)\right]^n \to \exp(-e^{-t}), \tag{3.9}$$

as $n \to \infty$ where we have used the fact $\frac{\phi(a_n)}{a_n\Phi(-a_n)} \to 1$, which follows by Gordon's inequality (Lemma 2.5). □

It is known for the random variable η one has $\mathbb{E}\eta = \gamma$ (the Euler-Mascheroni constant) and $\mathrm{Var}\eta = \pi^2/6$. Therefore, we obtain:

$$a_n^2 \mathrm{Var}(Y_n) = \mathrm{Var}(W_n) \to \mathrm{Var}\eta, \tag{3.10}$$

as $n \to \infty$. This proves the following:

Theorem 3.4. *If g is Gaussian n-dimensional random vector, then we have:*

$$\mathrm{Var}\|g\|_\infty = \mathrm{Var}_{\gamma_n}\|x\|_\infty \simeq \frac{1}{\log n}. \tag{3.11}$$

The reader will observe that the dependence on dimension we get for fixed $1 \le p < \infty$ is polynomial in n while for $p = \infty$ is logarithmic in n. Moreover, the variance as a function of p decreases exponentially fast and at the very end becomes logarithmically small (on n). As we have already explained this "skew" behavior lies on the fact that as p grows the constants in the equivalence should be expected to be exponential in p. In the rest of the paragraph we try to quantify this phenomenon and to give as sharp bounds as possible describing the behavior of p along n, too.

§3. Tightening the bounds. The purpose of this subsection is to provide continuous bounds in terms of p for the variance of the ℓ_p norm when dimension n is fixed and p varies from 1 to ∞. One can easily see that:

$$c_1 p \le n^{1-2/p}\mathrm{Var}\|X\|_p \le c_2 p \mathrm{Var}|g_1|^p \simeq p(2p/e)^p, \tag{3.12}$$

by comparing with the variance of the ℓ_2 norm and the p-th power of the ℓ_p norm. Below, we show that one can always have better estimates. In order to prove these estimates we will use the following fact:

Fact. Let $4 \leq p \leq \infty$. Then, for all n one has:

$$I_r(\gamma_n, B_p^n)/I_{-r}(\gamma_n, B_p^n) \leq \exp\left(\frac{C_1 r}{k_p \log n}\right), \quad 0 < r < c_1\sqrt{k_p \log n}. \tag{3.13}$$

This will be proved in Section 4 (Theorem 4.10).

1. Upper bound: An approach through Talagrand's inequality. For $p > 1$ we have: $\partial_i \|x\|_p = \frac{|x_i|^{p-1}}{\|x\|_p^{p-1}} \operatorname{sgn}(x_i)$ a.s. Thus, one has:

(3.14)
$$A^2 := \left\|\partial_i \|\cdot\|_p\right\|_{L_2}^2 \leq \sigma_{2p-2}^{2p-2} I_{-2(p-1)}^{-2(p-1)}(\gamma_{n-1}, B_p^{n-1}), \ a := \left\|\partial_i \|\cdot\|_p\right\|_{L_1} \leq \sigma_{p-1}^{p-1} I_{-(p-1)}^{-(p-1)}(\gamma_{n-1}, B_p^{n-1}).$$

Set $I_s(\gamma_{n-1}, B_p^{n-1}) \equiv I_s$. Then, direct application of Theorem 2.9 yields:

$$\operatorname{Var}(\|X\|_p) \leq Cn \; \frac{\sigma_{2p-2}^{2p-2} I_{-2(p-1)}^{-2(p-1)}}{1 + \log\left(\frac{\sigma_{2p-2}^{p-1}}{\sigma_{p-1}^{p-1}} \frac{I_{-(p-1)}^{p-1}}{I_{-2(p-1)}^{p-1}}\right)} \leq C_1 n \frac{\sigma_{2p-2}^{2p-2}/I_{-2(p-1)}^{2(p-1)}}{p}, \tag{3.15}$$

where we have used the fact that: $\left(\sigma_{2p-2}/\sigma_{p-1}\right)^{p-1} \simeq 2^p$, follows by (2.13). Now as long as $2p < c_1\sqrt{k_p \log n}$, which is satisfied when $p \leq c_0 \log n$ for some sufficiently small absolute constant $c_0 > 0$ by Proposition 2.4, we may apply the Fact to get:

$$I_{-2(p-1)}^{2(p-1)} \geq e^{-\frac{c' p^2}{k_p \log n}} I_p^{2(p-1)} \geq c_1' \sigma_p^{2(p-1)} (n-1)^{2-2/p}. \tag{3.16}$$

Plug this estimate in (3.15) we derive the upper bound:

$$\operatorname{Var}\|X\|_p \leq C_2 \frac{\sigma_{2p-2}^{2p-2}}{\sigma_p^{2p-2} p} n^{\frac{2}{p}-1} \simeq \frac{2^p}{p} n^{2/p-1}. \tag{3.17}$$

Note that this is exactly of the same order as the one we obtained at the limit value using the delta method.

2. The lower bound. Here we will use the next numerical result:

Lemma 3.5. *Let $a, b > 0$ and $0 < \theta \leq 1$. Then, we have:*

$$\theta|a-b|\left(\frac{2}{a+b}\right)^{1-\theta} \leq |a^\theta - b^\theta| \leq \theta|a-b|\frac{a^{\theta-1}+b^{\theta-1}}{2}. \tag{3.18}$$

Proof. We may assume without loss of generality that $0 < a < b$ and $0 < \theta < 1$. If we set $f(t) = t^{\theta-1}$, $t > 0$, then note that f is convex in $[a, b]$ and:

$$\frac{a^\theta - b^\theta}{a-b} = \frac{\theta}{b-a}\int_a^b f(t)\, dt.$$

Therefore, Hermite-Hadamard inequality (see [9]):

$$f\left(\frac{a+b}{2}\right) \le \frac{1}{b-a}\int_a^b f \le \frac{f(a)+f(b)}{2}, \tag{3.19}$$

yields the assertion. □

Applying the lower bound of Lemma 3.5 for $a = \|X\|_p^p$, $b = \|Y\|_p^p$ and $\theta = 1/p$ we obtain:

(3.20)

$$2\mathrm{Var}\|X\|_p = \mathbb{E}(\|X\|_p - \|Y\|_p)^2 \ge \frac{2^{2/q}}{p^2}\mathbb{E}\frac{(\|X\|_p^p - \|Y\|_p^p)^2}{(\|X\|_p^p + \|Y\|_p^p)^{2/q}} \ge \frac{1}{p^2}\mathbb{E}\left|\sum_{i=1}^n \frac{|X_i|^p - |Y_i|^p}{S^{1/q}}\right|^2,$$

where q is the conjugate exponent of p, i.e. $1/p + 1/q = 1$ and

$$S := \|X\|_p^p + \|Y\|_p^p = \|Z\|_p^p, \quad Z = (Z_1, \dots, Z_{2n}) \sim N(\mathbf{0}, I_{2n}). \tag{3.21}$$

Now we observe that the variables $\eta_j := \frac{|X_j|^p - |Y_j|^p}{S^{1/q}}$ have the same distribution and satisfy $\mathbb{E}(\eta_i\eta_j) = 0$ for $i \neq j$. Therefore, we have:

$$\mathbb{E}\left|\sum_{i=1}^n \frac{|X_i|^p - |Y_i|^p}{S^{1/q}}\right|^2 = \mathbb{E}\left|\sum_{i=1}^n \eta_i\right|^2 = \sum_{i=1}^n \mathbb{E}\eta_i^2 = n\mathbb{E}\eta_1^2 \tag{3.22}$$

Hence, estimate (3.20) becomes:

$$\mathrm{Var}\|X\|_p \ge \frac{n}{2p^2}\mathbb{E}\frac{(|X_1|^p - |Y_1|^p)^2}{S^{2/q}} = \frac{n}{p^2}\left(\mathbb{E}\frac{|X_1|^{2p}}{S^{2/q}} - \mathbb{E}\frac{|X_1|^p|Y_1|^p}{S^{2/q}}\right). \tag{3.23}$$

Let $T := \sum_{i>1}|X_i|^p + \sum_{i>1}|Y_i|^p$ and note that $T \le S$, so we obtain:

$$\mathrm{Var}\|X\|_p \ge \frac{n}{p^2}\left[\mathbb{E}\frac{|Z_1|^{2p}}{S^{2/q}} - \sigma_p^{2p}\mathbb{E}(T^{-2/q})\right]. \tag{3.24}$$

Now we apply Hölder's inequality to get:

$$\left(\mathbb{E}\frac{|Z_1|^{2p}}{S^{2/q}}\right)^{3/4} = \left(\mathbb{E}\left(\frac{|Z_1|^{3p/2}}{\|Z\|_p^{3(p-1)/2}}\right)^{4/3}\right)^{3/4} \ge \frac{\sigma_{3p/2}^{3p/2}}{(\mathbb{E}\|Z\|_p^{6(p-1)})^{1/4}}. \tag{3.25}$$

An application of the Fact again yields the further bound:

$$\left(\mathbb{E}\frac{|Z_1|^{2p}}{S^{2/q}}\right)^{3/4} \gtrsim \frac{\sigma_{3p/2}^{3p/2}}{(\sigma_p n^{1/p})^{3(p-1)/2}} \Longrightarrow \mathbb{E}\frac{|Z_1|^{2p}}{S^{2/q}} \gtrsim \frac{\sigma_{3p/2}^{2p}}{\sigma_p^{2p-2}n^{2-2/p}}. \tag{3.26}$$

Again the Fact yields $\mathbb{E}(T^{-2/q}) \lesssim \frac{1}{\sigma_p^{2p-2}(n-1)^{2-2/p}}$ as long as $p \le c_0 \log n$. Putting them all together we obtain:

$$\mathrm{Var}\|X\|_p \gtrsim \frac{1}{p^2}\frac{\sigma_p^{2p}}{\sigma_p^{2p-2}n^{1-2/p}}\left(\frac{\sigma_{3p/2}^{2p}}{\sigma_p^{2p}} - C\right) \simeq \frac{(3/2)^p}{p}n^{2/p-1}, \tag{3.27}$$

for p larger than some sufficiently large absolute constant.

Finally, for larger values of p, namely for $p \geq c_0 \log n$, we employ Theorem 2.9 again. Recall that for $f(x) = \|x\|_p$ we have:

$$\partial_i f(x) = \frac{|x_i|^{p-1} \operatorname{sgn}(x_i)}{\|x\|_p^{p-1}},$$

almost everywhere for $i = 1, 2, \dots, n$. Therefore, if $a_i := \|\partial_i f\|_{L_1(\gamma_n)}$ we may write:

$$a_i = \int_{\mathbb{R}^n} \frac{|x_i|^{p-1}}{\|x\|_p^{p-1}} \, d\gamma_n(x) = \frac{1}{n} \int_{\mathbb{R}^n} \left(\frac{\|x\|_{p-1}}{\|x\|_p} \right)^{p-1} d\gamma_n(x) \leq \frac{n^{1/p}}{n} = n^{-1/q}, \tag{3.28}$$

where in the last step we have used estimate (2.3) and q is the conjugate exponent of p, i.e. $1/p + 1/q = 1$. Moreover, we have:

$$A_i^2 := \|\partial_i f\|_{L_2(\gamma_n)}^2 = \int_{\mathbb{R}^n} \frac{|x_i|^{2p-2}}{\|x\|_p^{2p-2}} = \frac{1}{n} \int_{\mathbb{R}^n} \left(\frac{\|x\|_{2p-2}}{\|x\|_p} \right)^{2p-2} d\gamma_n(x),$$

thus by estimates (2.3) again, it follows that:

$$n^{-1/q} \leq A_i \leq n^{-1/2}. \tag{3.29}$$

Plug estimate (3.28) into the inequality of Theorem 2.9 we derive:

$$\operatorname{Var}(\|g\|_p) \leq C \sum_{i=1}^n \frac{A_i^2}{1 + \frac{1}{q} \log n + \log A_i}.$$

Set $B_n := 1 + \frac{1}{q} \log n$ and consider the function $F_n : (e^{-(B_n - 1/2)}, \infty) \to \mathbb{R}$ defined as: $F_n(t) = \frac{t^2}{B_n + \log t}$. Differentiating we find: $F_n'(t) = \frac{2t}{(B_n + \log t)^2}(B_n - 1/2 + \log t) > 0$. Note that since $p > 2$ we have $e^{-(B_n - 1/2)} < n^{-1/q}$ hence, estimates (3.29), imply that $F_n(A_i) \leq F_n(n^{-1/2})$. Therefore we obtain:

$$\operatorname{Var}(\|g\|_p) \leq C n F_n(n^{-1/2}) = C \frac{1}{1 + (\frac{1}{2} - \frac{1}{p}) \log n} \simeq \frac{1}{\log n},$$

since $p \geq c_0 \log n$.

Finally, let us note that the variance of the ℓ_p norm stabilizes for $p > (\log n)^2$ in the following sense:

Proposition 3.6. *Let $p > (\log n)^2$. Then, we have:*

$$\operatorname{Var}\|X\|_p \simeq \frac{1}{\log n} \tag{3.30}$$

One way to verify that is to show that the concentration for the ℓ_p norm with $p > (\log n)^2$ is the same with the one of the ℓ_∞ norm (see [24] for the concentration of the ℓ_∞ norm):

Lemma 3.7. *Let $p > (\log n)^2$. Then, we have:*

$$ce^{-C\varepsilon \log n} \leq P\left(\left|\|X\|_p - \mathbb{E}\|X\|_p\right| > \varepsilon \mathbb{E}\|X\|_p\right) \leq Ce^{-c\varepsilon \log n}, \tag{3.31}$$

for all $0 < \varepsilon < 1$, where $C, c > 0$ are absolute constant.

Proof. Consider $\frac{2}{\log n} < \varepsilon < 1$ and write:

$$\begin{aligned} P\left(\left|\|X\|_p - \mathbb{E}\|X\|_p\right| > \varepsilon \mathbb{E}\|X\|_p\right) &\geq P(\|X\|_p > (1+\varepsilon)\mathbb{E}\|X\|_p) \\ &\geq P(\|X\|_\infty > (1+\varepsilon)n^{1/p}\mathbb{E}\|X\|_\infty) \\ &\geq P(\|X\|_\infty > (1+2\varepsilon)\mathbb{E}\|X\|_\infty) \\ &> ce^{-C\varepsilon \log n}, \end{aligned}$$

where we have used (2.3) and at the last step the concentration from [24]. It follows that:

$$P\left(\left|\|X\|_p - \mathbb{E}\|X\|_p\right| > \varepsilon \mathbb{E}\|X\|_p\right) \geq c' e^{-C\varepsilon \log n}, \tag{3.32}$$

for all $0 < \varepsilon < 1$. □

Proof of Proposition 3.6. We may write:

$$\begin{aligned} \mathrm{Var}(\|X\|_p) &= 2(\mathbb{E}\|X\|_p)^2 \int_0^\infty tP\left(\left|\|X\|_p - \mathbb{E}\|X\|_p\right| > t\mathbb{E}\|X\|_p\right) dt \\ &\geq 2c'(\mathbb{E}\|X\|_p)^2 \int_0^1 te^{-Ct\log n}\, dt \gtrsim \frac{(\mathbb{E}\|X\|_p)^2}{(\log n)^2}. \end{aligned}$$

The result follows by Proposition 2.4. □

We close this section with some discussion on the methods used for bounding the variance. If we are interested in giving sufficient upper estimates, we may use Poincaré inequality which bounds the variance by the L_2 average of the euclidean norm of the gradient of f, which in principle is smaller than the Lipschitz constant: $\left\|\|\nabla f\|_2\right\|_{L_2(\gamma_n)} \leq \left\|\|\nabla f\|_2\right\|_{L_\infty(\gamma_n)} = L$. The reader may check that for $2 < p < \infty$ we have:

$$\int_{\mathbb{R}^n} \left\|\nabla f(x)\right\|_2^2 d\gamma_n(x) = \int_{\mathbb{R}^n} \frac{\|x\|_{2p-2}^{2p-2}}{\|x\|_p^{2p-2}}\, d\gamma_n(x) \simeq_p \frac{1}{n^{1-\frac{2}{p}}} \ll 1 = b(B_p^n) \equiv \mathrm{Lip}(f).$$

In case $p = \infty$ we have that $\left\|\nabla\|x\|_\infty\right\|_2 \equiv 1$ almost everywhere, hence:

$$\int_{\mathbb{R}^n} \left\|\nabla\|x\|_\infty\right\|_2^2 d\gamma_n(x) = 1 = b(B_\infty^n).$$

Thus, Poincaré inequality also fails to give the sharp upper bound for the variance in this case. The recovery of the correct upper bound is promised by the different order of magnitude for the $L_1 - L_2$ norms of the partial derivatives of $x \mapsto \|x\|_\infty$ and

Talagrand's inequality: for $f(x) = \|x\|_\infty$ we have $\partial_i f(x) = \mathrm{sgn}(x_i)\mathbf{1}_{\{x\,:\,|x_j|\le|x_i|\,\forall j\}}(x)$ almost everywhere. Therefore, we get:

$$\|\partial_i f\|_{L_2(\gamma_n)}^2 = \|\partial_i f\|_{L_1(\gamma_n)} = 1/n. \tag{3.33}$$

Plug these estimates in Theorem 2.9 we obtain the same upper bound as in Theorem 3.4.

The results of this paragraph can be summarized in the next:

Theorem 3.8. *There exist absolute constants $c_0, c_1, C_1 > 0$ with the following property: For all n and for any $1 \le p \le c_0 \log n$ we have:*

$$c_1 \frac{(3/2)^p}{p} \le n^{1-\frac{2}{p}} \mathrm{Var}\|X\|_{\ell_p^n} \le C_1 \frac{2^p}{p}. \tag{3.34}$$

If $p > c_0 \log n$ then we have:

$$\mathrm{Var}\|X\|_{\ell_p^n} \le \frac{C_1}{\log n}, \tag{3.35}$$

whereas for $p \ge (\log n)^2$ we also have:

$$\mathrm{Var}\|X\|_{\ell_p^n} \ge \frac{c_1}{\log n}. \tag{3.36}$$

Note. The choice of the exponent 3/4 and 1/4 in the applications of Hölder's inequality for the lower bound of the variance $(3/2)^p/p$ was done for simplicity. Different choices will still give exponential lower bounds with base arbitrary close to 2 at the cost of the range of p.

4 Gaussian concentration for the ℓ_p norm

In this paragraph we study the Gaussian concentration for the ℓ_p-norms for $1 \le p \le \infty$. First we show how we may employ log-Sobolev inequality in order to get concentration results.

§ 1. An argument via log-Sobolev inequality. Note that for the ℓ_p norm with $1 \le p \le 2$ the estimate (2.23) implies:

$$\frac{I_r(\gamma_n, B_p^n)}{I_1(\gamma_n, B_p^n)} \le \sqrt{1 + \frac{C_1 r}{k_p}} \le \exp\left(\frac{C_2 r}{n}\right), \tag{4.1}$$

for all $r \ge 1$. Therefore, for any $0 < \varepsilon < 1$ we apply Markov's inequality to get:

$$P(\|X\|_p > (1+\varepsilon)I_1) \le P(\|X\|_p > e^{\varepsilon/2} I_1) \le e^{-\varepsilon r/2}(I_r/I_1)^r \le \exp(-\varepsilon r/2 + C_2 r^2/n).$$

Choosing $r = \varepsilon n/(4C_2)$ (as long as $\varepsilon > 4C_2/n$) we obtain:

$$P(\|X\|_p > (1+\varepsilon)I_1) \le \exp\left(-\frac{1}{16C_2}\varepsilon^2 n\right).$$

Taking into account Theorem 2.14 and arguing similarly we find:

$$P(\|X\|_p < (1-\varepsilon)I_1) \leq \exp(-c_2\varepsilon^2 n).$$

Combining those two estimates we arrive at the next concentration result:

$$P\left(\left|\|X\|_p - I_1\right| > \varepsilon I_1\right) \leq C_3 \exp(-c_3\varepsilon^2 n), \tag{4.2}$$

for all $0 < \varepsilon < 1$. This estimate is sharp, as we will show later, but as we have already explained the same method fails for the ℓ_p norm, when $2 < p \leq \infty$, to give the correct concentration estimate. By carefully inspecting the proof of the estimates we used before we see that we have bound the L_2 norm of the gradient by the L_∞ norm, i.e. the Lipschitz constant. A first attempt to improve the estimates, it would be to improve the bound on that quantity. To this end, we restrict ourselves in the range $2 < p < \log n$ and we use log-Sobolev inequality. We have the following:

Proposition 4.1. *Let $2 < p < c\log n$. Then, for every $r > 0$ we have:*

$$\frac{d}{dr}(\log I_r) \leq \frac{C^p}{n}\left(1 + \frac{r}{k_{2p-2}}\right)^{p-1} \leq \begin{cases} C_1^p/n, & 0 < r \leq k_{2p-2} \\ \frac{1}{r}(\frac{C_1 r}{k_p})^p, & k_{2p-2} \leq r < k_p/C_1 \end{cases}, \tag{4.3}$$

while for $0 < r < cd_p$ we have:

$$-\frac{d}{dr}(\log I_{-r}) \leq \frac{C^p}{n}, \tag{4.4}$$

where $c, C, C_1 > 0$ are absolute constants and $I_s \equiv I_s(\gamma_n, B_p^n)$.

Proof. First we prove the growth condition on the positive moments. Our starting point is the next estimate:

$$\frac{d}{dr}(\log I_r) = \frac{1}{r^2 I_r^r}\mathrm{Ent}_{\gamma_n}(\|x\|_p^r) \leq \frac{2}{r^2 I_r^r}\mathbb{E}\left\|\nabla(\|X\|_p^{r/2})\right\|_2^2 = \frac{1}{2I_r^r}\mathbb{E}\|X\|_{2p-2}^{2p-2}\|X\|_p^{r-2p}, \tag{4.5}$$

where we have used log-Sobolev inequality. We distinguish two cases:

Case 1: $0 < r \leq 2p$. We may write:

$$\frac{d}{dr}(\log I_r) \leq \frac{n}{\mathbb{E}\|X\|_{\ell_p^n}^r}\mathbb{E}\frac{|X_1|^{2p-2}}{\|X\|_p^{2p-r}} \leq \frac{n\sigma_{2p-2}^{2p-2}}{\mathbb{E}\|X\|_p^r}\mathbb{E}\frac{1}{\|X\|_p^{2p-r}} \leq \frac{n(cp)^{p-1}}{I_r^r(B_p^n)I_{-(2p-r)}^{2p-r}(B_p^n)} \leq \frac{n(cp)^p}{I_{-2p}^{2p}(B_p^n)},$$

by Proposition 2.2 and Hölder's inequality. By Proposition 2.12 for $0 < s < c_1k_p$ we have: $I_{-s} \geq c_2 I_1$. Since, $p < c_1 k_p$ for $p \lesssim \log n$ we get: $(\log I_r)' \leq C_2^p/n$.

Case 2: $r > 2p$. We may write:

$$\frac{d}{dr}(\log I_r) \leq \frac{1}{2I_r^r}\mathbb{E}\|X\|_{2p-2}^{2p-2}\|X\|_p^{r-2p} \leq \frac{I_r^{2p-2}(\gamma_n, B_{2p-2}^n)}{2I_r^{2p}},$$

by Hölder's inequality. By Lemma 2.6 we get:

$$\frac{d}{dr}(\log I_r) \leq \frac{I_{2p-2}^{2p-2}(\gamma_n, B_{2p-2}^n)}{2I_p^{2p}}\left(1+\frac{r}{k_{2p-2}}\right)^{p-1} = \frac{\sigma_{2p-2}^{2p-2}/\sigma_p^{2p}}{2n}\left(1+\frac{r}{k_{2p-2}}\right)^{p-1}$$
$$\leq \frac{C_3^p}{n}\left(1+\frac{r}{k_{2p-2}}\right)^{p-1},$$

for some absolute constant $C_3 > 0$.

Now we turn in providing bounds for the negative moments. Here the argument is simpler. Using log-Sobolev inequality again and Proposition 2.2 we have:

$$\frac{d}{dr}(\log I_{-r}) \geq -\frac{1}{2I_{-r}^{-r}}\mathbb{E}\|X\|_{2p-2}^{2p-2}\|X\|_p^{-r-2p} \geq -\frac{1}{2I_{-r}^{-r}}\mathbb{E}\|X\|_{2p-2}^{2p-2}\mathbb{E}\|X\|_p^{-r-2p}$$
$$= -\frac{1}{2I_{-r}^{-r}}\mathbb{E}\|X\|_{2p-2}^{2p-2}I_{-r-2p}^{-r-2p} \geq -C_2^p\frac{\sigma_{2p-2}^{2p-2}n}{I_1^{2p}} \geq -C_3^p/n,$$

for $r \leq c_4 d_p$, where in the last step we have used Lemma 2.13. The result easily follows. □

Now we are ready to prove the next concentration inequality. Note that the dependence we get on ε is better than the one we get if we employ (2.36).

Proposition 4.2. *Let $4 \leq p < \varepsilon_0 \log n$. Then, one has:*

$$P\left(\left|\|X\|_p - \mathbb{E}\|X\|_p\right| > \varepsilon\mathbb{E}\|X\|_p\right) \leq C_1\exp\left(-c_1\varepsilon^{1+\frac{1}{p}}k_p\right), \tag{4.6}$$

for all $0 < \varepsilon < 1$. Moreover, we have:

$$P\left(\|X\|_p \leq (1-\varepsilon)\mathbb{E}\|X\|_p\right) \leq C_2\exp(-c_2\varepsilon k_p), \tag{4.7}$$

for $0 < \varepsilon < 1$.

Proof. Let $4 \leq p \leq c\log n$, where $c > 0$ is the constant from Proposition 4.1. Then, for each $0 < \varepsilon < 1$ using Markov's inequality we may write:

$$P(\|X\|_p > (1+\varepsilon)I_0) \leq e^{-\varepsilon r/2}\exp(r\log(I_r/I_0)) = \exp\left[-r\left(\frac{\varepsilon}{2} - \log(I_r/I_0)\right)\right], \tag{4.8}$$

for all $r > 0$. Using Proposition 4.1 we obtain:

$$\log(I_r/I_0) \leq \frac{C^p}{n}\int_0^r\left(1+\frac{s}{k_{2p-2}}\right)^{p-1}ds < \frac{C^p k_{2p-2}}{pn}\left(1+\frac{r}{k_{2p-2}}\right)^p < \frac{(2C)^p k_{2p-2}}{pn}\left(\frac{r}{k_{2p-2}}\right)^p,$$

for $r > k_{2p-2}$. Therefore, (4.8) becomes:

$$P(\|X\|_p > (1+\varepsilon)I_0) \leq \exp\left(-\frac{\varepsilon r}{2} + \frac{(2C)^p}{pnk_{2p-2}^{p-1}}r^{p+1}\right), \tag{4.9}$$

for $r > k_{2p-2}$. Minimizing the right-hand side with respect to r we find that $r_{\min} = r_0$ satisfies:

$$\frac{(2C)^p}{pnk_{2p-2}^{p-1}}(p+1)r_0^p - \frac{\varepsilon}{2} = 0 \Longrightarrow r_0 := (2C)^{-1}\left(\frac{\varepsilon}{2}\right)^{1/p}\left(\frac{p}{p+1}\right)^{1/p} n^{1/p} k_{2p-2}^{\frac{p-1}{p}} \simeq \varepsilon^{1/p} k_p, \tag{4.10}$$

and in order this value to be admissible we ought to have $r_0 > k_{2p-2}$. Hence, the value r_0 is admissible if ε satisfies:

$$r_0 > k_{2p-2} \Longleftrightarrow (2C)^{-p}\frac{\varepsilon n}{2}\frac{p}{p+1} > k_{2p-2} \Longleftrightarrow \varepsilon > (2C)^p \frac{2(p+1)}{pn} k_{2p-2}.$$

Note that Proposition 2.4 implies that:

$$k_q \leq c_2 q n^{2/q}, \quad \forall\, 2 \leq q \leq \log n. \tag{4.11}$$

Since $p \geq 4$ it suffices to have $\varepsilon > (2C)^p 8 c_2 p n^{-\frac{p-2}{p-1}}$ or equivalently to have $\varepsilon > (16ec_2C)^p n^{-\frac{p-2}{p-1}}$.

First consider the case that $k_p^{-\frac{p}{p+1}} < \varepsilon < 1$. In this case the above restriction is satisfied as long as $p \leq c_3 \log n$ for some sufficiently small absolute constant $c_3 > 0$. Indeed; one needs to check that:

$$k_p^{-\frac{p}{p+1}} > (16ec_2C)^p n^{-\frac{p-2}{p-1}},$$

and by taking into account (4.11) again it suffices to have $\frac{n^{\frac{p-2}{p-1}}}{(c_2 p n^{2/p})^{\frac{p}{p+1}}} > (16ec_2C)^p$ or it's enough $n^{\frac{p^2-3p}{p^2-1}} > (16e^2c_2^2C)^p$. Thus, if $\varepsilon_0 := \min\{c_3, c\} > 0$ and $4 \leq p \leq \varepsilon_0 \log n$ we have all requirements and we conclude that:

$$\begin{aligned} P(\|X\|_p > (1+\varepsilon)I_0) \leq \exp\left(-\frac{\varepsilon r_0}{2} + \frac{(2C)^p}{pnk_{2p-2}^{p-1}} r_0^{p+1}\right) &\overset{(4.10)}{=} \exp\left(-\frac{\varepsilon}{2} r_0 + \frac{\varepsilon r_0}{2(p+1)}\right) \\ &= \exp\left(-\frac{p}{2(p+1)}\varepsilon r_0\right) \\ &\leq \exp\left(-c_0 \varepsilon^{1+\frac{1}{p}} k_p\right), \end{aligned}$$

for $4 \leq p \leq \varepsilon_0 \log n$ and for all $k_p^{-\frac{p}{p+1}} < \varepsilon < 1$. By adjusting the constants we conclude that:

$$P(\|X\|_p > (1+\varepsilon)I_0) \leq C_0' \exp\left(-c_0 \varepsilon^{1+\frac{1}{p}} k_p\right), \tag{4.12}$$

for the whole range $0 < \varepsilon < 1$ and for $4 \leq p \leq \varepsilon_0 \log n$.

Now we turn in bounding the probability $P(\|X\| \leq (1-\varepsilon)I_0)$. Proposition 4.1 shows that $(\log I_{-r})' \geq -C^p/n$ for $0 < r \leq c_1 k_p$. Hence, we get:

$$P(\|X\|_p \leq (1-\varepsilon)I_0) \leq P(\|X\|_p \leq e^{-\varepsilon} I_0) \leq e^{-\varepsilon r}\left(\frac{I_0}{I_{-r}}\right)^r \leq \exp(-\varepsilon r + r^2 C^p/n),$$

for all $0 < r < c_1 k_p$, where we have used the bound:

$$\log(I_0/I_{-r}) = -\int_0^r (\log I_{-s})'\,ds \leq \frac{C^p}{n} r,$$

for $0 < r < c_1 k_p$. Finally, choosing $r \simeq k_p$ we see that $C^p k_p^2/n < (2eC)^p n^{4/p-1} \leq C'$ as long as $4 \leq p \leq c_1' \log n$, hence we conclude:

$$P(\|X\|_p \leq (1-\varepsilon)I_0) \leq C' \exp(-c'\varepsilon k_p), \tag{4.13}$$

for $0 < \varepsilon < 1$. □

Although this concentration result improves upon the one we get by just using (2.36), it is still suboptimal. It turns out that although the L_2 average of the euclidean norm of the gradient is the proper quantity to be estimated for the concentration result still it shouldn't be used in order to bound the growth of the high moments of the norm, i.e. with the log-Sobolev inequality.

§ 2. Estimating centered moments. We distinguish three cases:

§ 2.1. The case $1 \leq p \leq 2$. We have the next theorem:

Theorem 4.3. *Let* $1 \leq p \leq 2$. *Then, one has:*

$$c_1 \exp(-C_1\varepsilon^2 n) \leq \gamma_n\left(\left\{x : \left|\|x\|_p - \mathbb{E}\|x\|_p\right| > \varepsilon \mathbb{E}\|x\|_p\right\}\right) \leq C_2 \exp(-c_2\varepsilon^2 n), \tag{4.14}$$

for $0 < \varepsilon < 1$, *where* $C_1, c_1, C_2, c_2 > 0$ *are absolute constants.*

Proof. (sketch). The rightmost inequality also follows by the gaussian concentration inequality (2.36), Proposition 2.4 and the fact that $\mathrm{Lip}(\|\cdot\|_p) = b(B_p^n) = n^{1/p-1/2}$ for $1 \leq p \leq 2$. Now we focus on the left-hand side inequality. We have the next:

Theorem 4.4. *Let* $1 \leq p \leq 2$. *Then, we have:*

$$\left(\mathbb{E}\left|\|X\|_p - \mathbb{E}\|X\|_p\right|^r\right)^{1/r} \simeq \sqrt{\frac{r}{n}}\mathbb{E}\|X\|_p, \tag{4.15}$$

for all $r \geq 1$.

Proof. Indeed; the estimate

$$\left(\mathbb{E}\left|\|X\|_p - \mathbb{E}\|X\|_p\right|^r\right)^{1/r} \leq C_3 \sqrt{\frac{r}{n}}\mathbb{E}\|X\|_p, \quad r \geq 1 \tag{4.16}$$

is well known and follows by integration by parts combined with the right-hand side estimate we just mentioned (or follows immediately by the (r,r)-Poincaré inequalities (2.34) - this approach will be used for the case $2 < p < \infty$, too). For the estimate

$$\left(\mathbb{E}\left|\|X\|_p - \mathbb{E}\|X\|_p\right|^r\right)^{1/r} \geq c_3 \sqrt{\frac{r}{n}}\mathbb{E}\|X\|_p \tag{4.17}$$

we may apply the triangle inequality, Lemma 3.5 and finally Cauchy-Schwarz inequality to write:

$$2\left(\mathbb{E}\left|\|X\|_p-\mathbb{E}\|X\|_p\right|^r\right)^{1/r}\geq\left(\mathbb{E}\left|\|X\|_p-\|Y\|_p\right|^r\right)^{1/r}\geq\frac{1}{2p}\frac{\left(\mathbb{E}\left|\|X\|_p^p-\|Y\|_p^p\right|^{r/2}\right)^{2/r}}{\left(\mathbb{E}\|X\|_p^{r(p-1)}\right)^{1/r}}.$$

Note that (4.16) already implies $\left(\mathbb{E}\|X\|_p^s\right)^{1/s}\leq 2C_3\mathbb{E}\|X\|_p\simeq n^{1/p}$ for all $1\leq s\leq n$. Moreover, we have:

$$\left(\mathbb{E}\left|\|X\|_p^p-\|Y\|_p^p\right|^s\right)^{1/s}\geq\mathbb{E}\left||X_1|^p-|Y_1|^p\right|\cdot\left(\mathbb{E}_\varepsilon\left|\sum_{i=1}^n\varepsilon_i\right|^s\right)^{1/s}\simeq\sqrt{sn},$$

where we have used the fact that the joint distribution of $(\varepsilon_i||X_i|^p-|Y_i|^p|)_i$ is the same with $(|X_i|^p-|Y_i|^p)_i$, Jensen's inequality and at the last step, the well-known fact that $\left(\mathbb{E}_\varepsilon\left|\sum_{i=1}^n\varepsilon_i\right|^s\right)^{1/s}\simeq\sqrt{sn}$ for $1\leq s\leq n$ (see e.g. [16]). Putting them all together we see:

$$\left(\mathbb{E}\left|\|X\|_p-\mathbb{E}\|X\|_p\right|^r\right)^{1/r}\geq c_4\frac{\sqrt{rn}}{n^{1-1/p}}\simeq\sqrt{\frac{r}{n}}\mathbb{E}\|X\|_p, \tag{4.18}$$

which completes the proof. □

Now we turn in the lower bound of the probabilistic estimate (4.14): For every $n^{-1/2}<\varepsilon<2c_3$ consider $r\in[1,n]$ so that $\varepsilon=2c_3\sqrt{r/n}$ to write:

$$\begin{aligned}P\left(\left|\|X\|_p-\mathbb{E}\|X\|_p\right|>\varepsilon\mathbb{E}\|X\|_p\right)&\geq P\left(\left|\|X\|_p-\mathbb{E}\|X\|_p\right|\geq\frac{1}{2}\left(\mathbb{E}\left|\|X\|_p-\mathbb{E}\|X\|_p\right|^r\right)^{1/r}\right)\\&=P(\zeta\geq 2^{-r}\mathbb{E}\zeta)\geq(1-2^{-r})^2\frac{(\mathbb{E}\zeta)^2}{\mathbb{E}\zeta^2},\end{aligned}$$

by Lemma 2.1, where $\zeta:=\left|\|X\|_p-\mathbb{E}\|X\|_p\right|^r$. Employing the estimates (4.16) and (4.17) we conclude:

$$P\left(\left|\|X\|_p-\mathbb{E}\|X\|_p\right|>\varepsilon\mathbb{E}\|X\|_p\right)\geq c_5e^{-C_5r},$$

as required. □

§ 2.2. The case $2<p<\infty$. It is clear from the argument of the previous paragraph that in order to obtain sharp concentration inequalities it is enough to get sharp estimates for the centered moments: $\left(\mathbb{E}\left|\|X\|_p-\|Y\|_p\right|^r\right)^{1/r}$. We have the next theorem:

Theorem 4.5. *Let $2<p<\infty$. Then, for each n we have:*

$$\left(\mathbb{E}\left|\|X\|_p-\|Y\|_p\right|^r\right)^{1/r}\simeq_p\max\left\{\left(\frac{r}{n}\right)^{1/2},\ \frac{r^{p/2}}{n}\right\}\mathbb{E}\|X\|_p, \tag{4.19}$$

for $4\leq r\leq n^{2/p}$.

In view of Lemma 3.5 it is clear that estimates for the centered moments $\left(\mathbb{E}\left|\|X\|_p^p - \|Y\|_p^p\right|^r\right)^{1/r}$ will provide estimates for the moments $\left(\mathbb{E}\left|\|X\|_p - \|Y\|_p\right|^r\right)^{1/r}$. This approach was used (in a more general probabilistic context) in [20] in order to establish the rightmost inequality in Theorem 4.5. Here, we keep the same idea for proving the lower estimate in Theorem 4.5 but for the upper estimate we employ the (r,r)-Poincaré inequalities instead. The reason we argue in two different ways is because the main tool for the upper bound will be exploited further in the next Section. There, will be used in order to obtain the optimal dependence on ε in the random version of Dvoretzky's theorem. Therefore, we think it is more instructive to present it here.

Proof of Theorem 4.5 (upper bound). Direct application of the inequality (2.34) yields:

$$\left(\mathbb{E}\left|\|X\|_p - \|Y\|_p\right|^r\right)^{1/r} \leq \frac{\pi}{2}\sigma_r \left(\mathbb{E}\left\|\nabla\|X\|_p\right\|_2^r\right)^{1/r} = \frac{\pi}{2}\sigma_r \left(\mathbb{E}\frac{\|X\|_{2p-2}^{r(p-1)}}{\|X\|_p^{r(p-1)}}\right)^{1/r}.$$

Now we focus in bounding $\left(\mathbb{E}\frac{\|X\|_{2p-2}^{r(p-1)}}{\|X\|_p^{r(p-1)}}\right)^{1/r}$. Note that since $\|\cdot\|_{2p-2} \leq \|\cdot\|_p$ one has the trivial bound $\left(\mathbb{E}\frac{\|X\|_{2p-2}^{r(p-1)}}{\|X\|_p^{r(p-1)}}\right)^{1/r} \leq 1$. Moreover, Cauchy-Schwarz inequality yields:

$$\left(\mathbb{E}\left\|\nabla\|X\|_p\right\|_2^r\right)^{1/r} = \left(\mathbb{E}\frac{\|X\|_{2p-2}^{r(p-1)}}{\|X\|_p^{r(p-1)}}\right)^{1/r} \leq \frac{I_{2r(p-1)}^{p-1}(\gamma_n, B_{2p-2}^n)}{I_{-2r(p-1)}^{p-1}(\gamma_n, B_p^n)}.$$

A standard application of Lemma 2.6 (2.34)

$$\frac{I_{2r(p-1)}^{p-1}(\gamma_n, B_{2p-2}^n)}{I_{2p-2}^{p-1}(\gamma_n, B_{2p-2}^n)} \leq \left(1 + \frac{2p(r-1)}{\sigma_{2p-2}^2 n^{\frac{1}{p-1}}}\right)^{\frac{p-1}{2}} \leq \left(1 + \frac{C_1(r-1)}{n^{\frac{1}{p-1}}}\right)^{\frac{p-1}{2}}.$$

Moreover, from Proposition 2.12 we see that:

$$\frac{I_{-2r(p-1)}^{p-1}(\gamma_n, B_p^n)}{I_p^{p-1}(\gamma_n, B_p^n)} \geq C_2^{-p},$$

for $rp \leq c_1 k_p$. Plug these estimates we find:

$$\begin{aligned}
\left(\mathbb{E}\frac{\|X\|_{2p-2}^{r(p-1)}}{\|X\|_p^{r(p-1)}}\right)^{1/r} &\leq \left(1 + \frac{C_1 r}{n^{\frac{1}{p-1}}}\right)^{\frac{p-1}{2}} C_2^p \frac{I_{2p-2}^{p-1}(\gamma_n, B_{2p-2})}{I_p^{p-1}(\gamma_n, B_p^n)} \\
&= \left(1 + \frac{C_1 r}{n^{\frac{1}{p-1}}}\right)^{\frac{p-1}{2}} C_2^p \frac{\sigma_{2p-2}^{p-1} n^{1/2}}{\sigma_p^{p-1} n^{1-1/p}} \\
&\leq C_3^p \max\left\{\frac{1}{n^{1/2-1/p}}, \frac{r^{\frac{p-1}{2}}}{n^{1-1/p}}\right\},
\end{aligned}$$

for $1 \leq r \leq c_2 n^{2/p}$ (since $n^{\frac{1}{p-1}} < n^{2/p}$ for $p > 2$), where $C_3 > 0$ is large absolute constant. Thus, we get the bound:

$$\left(\mathbb{E}\left\|\nabla\|X\|_p\right\|_2^r\right)^{1/r} \leq \min\left\{1,\ C_3^p \max\left\{\frac{1}{n^{1/2-1/p}},\ \frac{r^{\frac{p-1}{2}}}{n^{1-1/p}}\right\}\right\},$$

for all $r \geq 1$. It follows that:

$$\left(\mathbb{E}\left|\|X\|_p - \|Y\|_p\right|^r\right)^{1/r} \leq C_5 I_1(\gamma_n, B_p^n) \min\left\{\left(\frac{r}{k_p}\right)^{1/2},\ C_4^p \max\left\{\left(\frac{r}{n}\right)^{1/2},\ \frac{r^{p/2}}{n}\right\}\right\},$$

for all $r \geq 1$ where $C_5, C_4 > 0$ are absolute constants. □

Now we turn in the preparation needed in order to obtain the lower bound. Next result is contained essentially in [20]:

Proposition 4.6. *Let $1 \leq p < \infty$. Then, we have:*

$$\left(\mathbb{E}\left|\|X\|_p^p - \|Y\|_p^p\right|^r\right)^{1/r} \simeq \left(\frac{p}{e}\right)^{p/2} \max\left\{2^{p/2}(rn)^{1/2},\ r^{p/2} n^{1/r}\right\}, \tag{4.20}$$

for all $r \geq 2$.

Naor in [20] uses the following result of Latala from [13] in order to prove the aforementioned estimate:

Theorem 4.7. *If $r \geq 2$ and $\xi, \xi_1, \ldots, \xi_n$ are i.i.d. symmetric random variables, then we have:*

$$\left(\mathbb{E}\left|\sum_{i=1}^n \xi_i\right|^r\right)^{1/r} \simeq \sup\left\{\frac{r}{s}\left(\frac{n}{r}\right)^{1/s} (\mathbb{E}|\xi|^s)^{1/s} : \max\left\{2, \frac{r}{n}\right\} \leq s \leq r\right\}. \tag{4.21}$$

Note that if $X = (X_1, \ldots, X_n)$ is a gaussian random vector and Y an independent copy of it, the variables $\xi_i := |X_i|^p - |Y_i|^p$ are i.i.d. and symmetric. For applying Theorem 4.7 we need asymptotic estimates for the moments $(\mathbb{E}|\xi_1|^s)^{1/s}$. One can check that for any $s \geq 1$ we have:

$$\mathbb{E}|\xi_1|^s = \mathbb{E}\left||X_1|^p - |Y_1|^p\right|^s = \sqrt{\frac{2}{\pi}} \sigma_{1+ps}^{1+ps} \int_0^{\pi/2} |\cos^p \theta - \sin^p \theta|^s \, d\theta. \tag{4.22}$$

Next lemma provides bounds on the trigonometric integral.

Lemma 4.8. *For $p \geq 1$ and $s \geq 1$ we have:*

$$\left(\int_0^{\pi/2} |\cos^p \theta - \sin^p \theta|^s \, d\theta\right)^{1/s} \simeq p^{-1/(2s)}. \tag{4.23}$$

Proof. We may write:

$$\frac{1}{2}\int_0^{\pi/2} |\cos^p\theta - \sin^p\theta|^s \, d\theta = \int_0^{\pi/4} (\cos^p\theta - \sin^p\theta)^s \, d\theta =: J(p,s).$$

First note that one has the next simple estimate:

Claim. For $\beta \in [\frac{\pi}{6}, \frac{\pi}{4}]$ and for every $\alpha > 1$ we have:

$$\int_0^{\beta} \cos^{\alpha}\theta \, d\theta \simeq \alpha^{-1/2}. \tag{4.24}$$

Indeed; we may write:

$$\int_0^{\beta} \cos^{\alpha}\theta \, d\theta \leq \int_0^{\pi/2} \cos^{\alpha}\theta \, d\theta = \frac{1}{2}B\left(\frac{\alpha+1}{2}, \frac{1}{2}\right) \sim \sqrt{\pi/2}(\alpha+1)^{-1/2},$$

from standard approximation for the Beta function. For the lower bound we argue as follows:

$$\int_0^{\beta} \cos^{\alpha}\theta \, d\theta = \frac{1}{2}B\left(\frac{\alpha+1}{2}, \frac{1}{2}\right) - \int_{\beta}^{\pi/2} \cos^{\alpha}\theta \, d\theta \geq \frac{1}{2}B\left(\frac{\alpha+1}{2}, \frac{1}{2}\right) - \frac{2\cos^{\alpha+1}\beta}{\alpha+1},$$

which proves the claim. In order to estimate $J(p,s)$ from below note that for $\theta \in [0, \pi/6]$ one has $\sin\theta \leq 3^{-1/2}\cos\theta$, hence we may write:

$$\left(1 - 3^{-p/2}\right)^s \int_0^{\pi/6} (\cos\theta)^{ps} \, d\theta \leq \int_0^{\pi/6} (\cos^p\theta - \sin^p\theta)^s \, d\theta \leq J(p,s).$$

On the other hand, for $\theta \in (0, \pi/2)$ we have $\sin\theta > 0$, thus:

$$J(p,s) \leq \int_0^{\pi/4} (\cos\theta)^{ps} \, d\theta.$$

Now the assertion follows from the Claim and the previous estimates for $J(p,s)$. □

Proof of Proposition 4.6. Note that (2.13) implies that:

$$\sigma_{1+ps}^{1+ps} \simeq \left(\frac{ps+2}{e}\right)^{\frac{ps+1}{2}} \simeq (ps)^{1/2}\left(\frac{ps}{e}\right)^{\frac{ps}{2}}. \tag{4.25}$$

Thus, by taking into account (4.22) and Lemma 4.8 we get:

$$(\mathbb{E}|\xi_1|^s)^{1/s} \simeq p^{\frac{1}{2s}}\left(\frac{ps}{e}\right)^{p/2} p^{-\frac{1}{2s}} \simeq \left(\frac{ps}{e}\right)^{p/2}. \tag{4.26}$$

In view of Theorem 4.7 we have:

$$\begin{aligned}\left(\mathbb{E}\Big|\|X\|_p^p - \|Y\|_p^p\Big|^r\right)^{1/r} &\simeq \sup\left\{\frac{r}{s}\left(\frac{n}{r}\right)^{1/s}\left(\frac{ps}{e}\right)^{p/2} \,:\, \max\{2, r/n\} \leq s \leq r\right\} \\ &\simeq r\left(\frac{p}{e}\right)^{p/2} \sup\left\{s^{p/2-1}\left(\frac{n}{r}\right)^{1/s} \,:\, \max\{2, r/n\} \leq s \leq r\right\}.\end{aligned}$$

Set $f(s) := s^{p/2-1}\left(\frac{n}{r}\right)^{1/s}$, $s \geq 2$ and note that f has a unique extremum at $s_0 = (\frac{p}{2}-1)^{-1}\log(\frac{n}{r})$ which is actually a minimum. For $r \geq 2n$ we readily see that f is increasing, hence:

$$\left(\mathbb{E}\left|\|X\|_p^p - \|Y\|_p^p\right|^r\right)^{1/r} \simeq r\left(\frac{p}{e}\right)^{p/2} \max\{f(2),\ f(r)\}$$
$$\simeq \left(\frac{p}{e}\right)^{p/2} \max\left\{2^{p/2-1}(rn)^{1/2},\ r^{p/2}\left(\frac{n}{r}\right)^{1/r}\right\}, \tag{4.27}$$

which proves the assertion. □

Proof of Theorem 4.5 (lower bound). Using Lemma 3.5 for $a = \|X\|_p^p$, $b = \|Y\|_p^p$ and $\theta = 1/p$ we may write:

$$\left(\mathbb{E}\left|\|X\|_p - \|Y\|_p\right|^{2r}\right)^{1/(2r)} \geq \frac{1}{p}\left(\mathbb{E}\frac{\left|\|X\|_p^p - \|Y\|_p^p\right|^{2r}}{(\|X\|_p^p + \|Y\|_p^p)^{2r(p-1)/p}}\right)^{1/(2r)} \geq \frac{1}{p}\frac{\left(\mathbb{E}\left|\|X\|_p^p - \|Y\|_p^p\right|^r\right)^{1/r}}{\left(\mathbb{E}(\|X\|_p^p + \|Y\|_p^p)^{2r(p-1)/p}\right)^{1/(2r)}},$$

where in the last step we have applied Cauchy-Schwarz inequality. Applying the numerical inequality $(a+b)^s < 2^s(a^s+b^s)$ for $a, b, s > 0$ we obtain:

$$\left(\mathbb{E}\left|\|X\|_p - \|Y\|_p\right|^{2r}\right)^{1/(2r)} \gtrsim \frac{1}{p}\frac{\left(\mathbb{E}\left|\|X\|_p^p - \|Y\|_p^p\right|^r\right)^{1/r}}{(\mathbb{E}\|X\|_p^{2r(p-1)})^{1/(2r)}}.$$

Thus, Proposition 4.6 and Lemma 2.6 yield:

$$\begin{aligned}
\left(\mathbb{E}\left|\|X\|_p - \|Y\|_p\right|^{2r}\right)^{1/(2r)} &\gtrsim \frac{1}{p(1+\frac{C_1 r}{n^{2/p}})^{\frac{p-1}{2}} I_p^{p-1}}\left(\frac{p}{e}\right)^{p/2}\max\left\{2^{p/2}(rn)^{1/2},\ r^{p/2}n^{1/r}\right\} \\
&\gtrsim \frac{I_1}{p(1+\frac{C_1 r}{n^{2/p}})^{\frac{p-1}{2}}\sigma_p^p n}\left(\frac{p}{e}\right)^{p/2}\max\left\{2^{p/2}(rn)^{1/2},\ r^{p/2}n^{1/r}\right\} \\
&\geq \frac{C_2^{-p} I_1}{\left(\max\{1, \frac{r}{n^{2/p}}\}\right)^{\frac{p-1}{2}}}\max\left\{\left(\frac{r}{n}\right)^{1/2}, \frac{r^{p/2}}{n}\right\}, \\
&= C_2^{-p} I_1 \min\left\{\left(\frac{r}{n^{2/p}}\right)^{1/2}, \max\left\{\left(\frac{r}{n}\right)^{1/2}, \frac{r^{p/2}}{n}\right\}\right\}
\end{aligned}$$

for $r \geq 2$ and for some large absolute constant $C_2 > 0$. □

Now we may prove the two sided concentration inequality for $x \mapsto \|x\|_p$, $p > 2$.

Theorem 4.9. *Let $2 < p < \infty$. Then, for all n we have:*

$$c\exp(-Cp\alpha_1(n,p,\varepsilon)) \leq P\left(\left|\|X\|_p - \mathbb{E}\|X\|_p\right| > \varepsilon\mathbb{E}\|X\|_p\right) \leq C\exp\left(-c\alpha_2(n,p,\varepsilon)\right), \tag{4.28}$$

for $0 < \varepsilon < 1$, where $\alpha_i(n,p,\cdot)$ are defined as:

$$\alpha_1(n,p,t) \simeq \min\left\{C^{2p}t^2 n, (tn)^{2/p}\right\} \quad , \quad \alpha_2(n,p,t) \simeq \min\left\{C^{-2p}t^2 n, (tn)^{2/p}\right\} \tag{4.29}$$

and $C, c > 0$ are absolute constants.

Proof. Let $\psi(n,p,r) := \max\{(r/n)^{1/2}, r^{p/2}/n\},\ r>0$. Then, Theorem 4.5 yields:

$$\left(\mathbb{E}\Big|\|X\|_p - \mathbb{E}\|X\|_p\Big|^r\right)^{1/r} \le C_1^p \psi(n,p,r)\mathbb{E}\|X\|_p, \tag{4.30}$$

for all $1 \le r \le n^{2/p}$ and some absolute constant $C_1 > 0$. Note that for $0<\varepsilon<1$ and for any $1 \le r \le n^{2/p}$ Markov's inequality implies:

$$P\left(\Big|\|X\|_p - \mathbb{E}\|X\|_p\Big| > \varepsilon \mathbb{E}\|X\|_p\right) \le \left(\frac{C_1^p \psi(n,p,r)}{\varepsilon}\right)^r$$

Assume that $e^{-1}C_1^{-p}\varepsilon \ge n^{-1/2}$. Then, we may choose $1 \le r \le n^{2/p}$ such that $\psi(n,p,r) = e^{-1}C_1^{-p}\varepsilon$, i.e. $r = \psi^{-1}(n,p,e^{-1}C_1^{-p}\varepsilon)$ and taking into account the fact that $\psi^{-1}(n,p,s) = \min\{s^2 n, (sn)^{2/p}\},\ s>0$, we obtain:

$$P\left(\Big|\|X\|_p - \mathbb{E}\|X\|_p\Big| > \varepsilon \mathbb{E}\|X\|_p\right) \le \exp(-\psi^{-1}(p,n,e^{-1}C_1^{-p}\varepsilon)),$$

while for $e^{-1}C_1^{-p}\varepsilon < n^{-1/2}$ we have $\psi^{-1}(p,n,e^{-1}C_1^{-p}\varepsilon) < 1$, therefore we get:

$$\begin{aligned} P\left(\Big|\|X\|_p - \mathbb{E}\|X\|_p\Big| > \varepsilon \mathbb{E}\|X\|_p\right) &\le 3\exp(-\psi^{-1}(n,p,e^{-1}C_1^{-p}\varepsilon)) \\ &\le 3\exp\left(-\min\left\{e^{-2}C_1^{-2p}\varepsilon^2 n,\ e^{-2/p}C_1^{-2}(\varepsilon n)^{2/p}\right\}\right) \\ &\le 3\exp\left(-e^{-2}C_1^{-2}\min\left\{C_1^{-2p}\varepsilon^2 n,\ (\varepsilon n)^{2/p}\right\}\right), \end{aligned}$$

for every $0<\varepsilon<1$. For the lower estimate we argue as in Theorem 4.3: From Theorem 4.5 (lower bound) we know that there exists an absolute constant $C_2 > 0$ such that:

$$\left(\mathbb{E}\Big|\|X\|_p - \mathbb{E}\|X\|_p\Big|^r\right)^{1/r} \ge C_2^{-p} \psi(n,p,r)\mathbb{E}\|X\|_p \tag{4.31}$$

for $4 \le r \le n^{2/p}$, hence for $C_2^{-p}n^{-1/2} < \varepsilon < 1/(2C_2^p)$ choose $s \in [4, n^{2/p}]$ such that $\psi(n,p,s) = 2\varepsilon C_2^p$ to write:

$$\begin{aligned} P\left(\Big|\|X\|_p - \mathbb{E}\|X\|_p\Big| > \varepsilon \mathbb{E}\|X\|_p\right) &\ge P\left(\Big|\|X\|_p - \mathbb{E}\|X\|_p\Big| > \frac{1}{2}\left(\mathbb{E}\Big|\|X\|_p - \mathbb{E}\|X\|_p\Big|^s\right)^{1/s}\right) \\ &\ge (1-2^{-s})^2 \frac{\left(\mathbb{E}\Big|\|X\|_p - \mathbb{E}\|X\|_p\Big|^s\right)^2}{\mathbb{E}\Big|\|X\|_p - \mathbb{E}\|X\|_p\Big|^{2s}} \\ &\ge \frac{1}{4}(2^{-p/2}C_1^{-p}C_2^{-p})^{2s} \\ &\ge c_3 e^{-C_3 ps} = c_3\exp\left(-C_3 p \psi^{-1}(n,p,2C_2^p\varepsilon)\right), \end{aligned}$$

by the estimates (4.30) and (4.31). The result follows. □

§ 2.3. The case $\log n < p \le \infty$.

Theorem 4.10. *Let $4 < p \leq \infty$. Then, for any $0 < r < s \leq c_1 \sqrt{k_p \log n}$ we have:*

$$\frac{I_s(\gamma_n, B_p^n)}{I_r(\gamma_n, B_p^n)} \leq \exp\left(\frac{c_2(2s-r)}{k_p \log n}\right), \tag{4.32}$$

and

$$\frac{I_{-s}(\gamma_n, B_p^n)}{I_{-r}(\gamma_n, B_p^n)} \geq \exp\left(-\frac{c_2(2s-r)}{k_p \log n}\right) \tag{4.33}$$

where $c_1, c_2 > 0$ are absolute constants.

Proof. Apply Theorem 2.9 for $f(x) := \|x\|_p^r$, $r \neq 0$. Then,

$$\partial_i f(x) = r\|x\|_p^{r-p}|x_i|^{p-1}\operatorname{sgn}(x_i), \;\; x \neq 0.$$

Set $I_s \equiv I_s(\gamma_n, B_p^n)$ and for $a = a_i := \|\partial_i f\|_{L_1(\gamma_n)}$ we get:

$$a_i = \frac{|r|}{n}\int_{\mathbb{R}^n} \|x\|_{p-1}^{p-1}\|x\|_p^{r-p}\, d\gamma_n(x) \leq \frac{|r|}{n^{1/q}} I_{r-1}^{r-1}, \tag{4.34}$$

where we have used (2.3). In similar fashion, for $A = A_i := \|\partial_i f\|_{L_2(\gamma_n)}$ we have that:

$$\frac{|r|}{n^{1/q}} I_{2r-2}^{r-1} \leq A = \frac{|r|}{n^{1/2}}\left(\int_{\mathbb{R}^n} \|x\|_p^{2r-2p}\|x\|_{2p-2}^{2p-2}\, d\gamma_n(x)\right)^{1/2} \leq \frac{|r|}{n^{1/2}} I_{2r-2}^{r-1}, \tag{4.35}$$

by (2.3) again. Therefore we obtain:

$$\operatorname{Var}_{\gamma_n}(f) \leq C_1 n \frac{A^2}{1+\log(A/a)}. \tag{4.36}$$

The function $F(t) = \frac{t^2}{1+\log(t/a)}$, $t > a$ is increasing and since (4.34) and (4.35) imply that $a \leq |r| n^{-1/q} I_{r-1}^{r-1} \leq A \leq |r| n^{-1/2} I_{2r-2}^{r-1}$ for all $r \neq 0$ we obtain:

$$I_{2r}^{2r} - I_r^{2r} = \operatorname{Var}_{\gamma_n}(f) \leq C_1 r^2 \frac{I_{2r-2}^{2r-2}}{1+\log\left(n^{1/q-1/2}\left(\frac{I_{2r-2}}{I_{r-1}}\right)^{r-1}\right)} \leq C_2 r^2 \frac{I_{2r-2}^{2r-2}}{\log n}, \tag{4.37}$$

since $1 \leq q < 4/3$ and $\log(I_{2r-2}/I_{r-1}) \geq 0$.

Claim. For $r > -k_p$, $r \neq 0$ we have:

$$I_{2r-2}^{2r-2} \leq C_3 I_{2r}^{2r}/k_p.$$

We distinguish three cases:

- For $0 < r < 1$ we have: $I_{2r-2}^{2r-2} = \frac{I_{2r-2}^{2r}}{I_{2r-2}^2} \leq c_2' I_1^{-2} I_{2r-2}^{2r} \leq \frac{c_2'}{k_p} I_{2r}^{2r}$.
- For $r \geq 1$ we may write: $I_{2r-2}^{2r-2} = \frac{I_{2r-2}^{2r}}{I_{2r-2}^2} \leq c_3 \frac{I_{2r}^{2r}}{I_1^2} = \frac{c_3}{k_p} I_{2r}^{2r}$, since $I_1 \simeq I_0$.

- Finally, for $-k_p < r < 0$ we have: $I_{2r-2}^{2r-2} \le \frac{c_4}{I_1^2} I_{2r}^{2r} = \frac{c_4}{k_p} I_{2r}^{2r}$, by Lemma 2.13.

So, (4.37) becomes:

$$I_{2r}^{2r} - I_r^{2r} \le Cr^2 \frac{I_{2r}^{2r}}{k_p \log n}. \tag{4.38}$$

for $r > -k_p,\ r \neq 0$. First we prove the stability for the positive means:

(a) Consider the case $r > 0$. As long as $0 < r < \sqrt{k_p \log n / C}$ we may write:

$$I_{2r}^{2r} \le \left(1 + \frac{Cr^2}{k_p \log n}\right) I_r^{2r}. \tag{4.39}$$

Iterating the last one we find:

$$\frac{I_{2^m r}}{I_r} \le \exp\left(C \sum_{j=0}^{m-1} \frac{2^j r}{k_p \log n}\right) \le \exp\left(\frac{Cr(2^m - 1)}{k_p \log n}\right), \tag{4.40}$$

for $m = 1, 2, \ldots$ as long as $2^m r \le \sqrt{k_p \log n / C}$. It follows that for $0 < r_1 < r_2 < \sqrt{k_p \log n / C}$ we get:

$$\frac{I_{r_2}}{I_{r_1}} \le \exp\left(\frac{C(2r_2 - r_1)}{k_p \log n}\right). \tag{4.41}$$

(b) The case $r < 0$ is treated similarly. Set $r = -s,\ s > 0$. Then, inequality (4.38) is written as:

$$I_{-2s}^{-2s} - I_{-s}^{-s} \le Cs^2 \frac{I_{-2s}^{-2s}}{k_p \log n},$$

for $s < k_p$. In particular, for $s < \sqrt{k_p \log n / C} < k_p$ we get:

$$I_{-2s} \ge \exp\left(-\frac{Cs}{k_p \log n}\right) I_{-s}.$$

Arguing as before we conclude that for $0 < s_1 < s_2 < \sqrt{k_p \log n / C}$ we have:

$$I_{-s_2} \ge \exp\left(-\frac{C(2s_2 - s_1)}{k_p \log n}\right) I_{-s_1}. \tag{4.42}$$

The proof of the Theorem is complete. □

Remark 4.11. The above argument also shows that for $0 < r, s < c_1 \sqrt{K},\ K := k_p \log n$ we get:

$$I_r \le I_0 \exp(c_2 r / K), \quad I_{-s} \ge I_0 \exp(-c_2 s / K), \tag{4.43}$$

where $I_0 := \exp(\int \log \|x\|_p \, d\gamma_n(x))$. In particular,

$$\frac{I_r}{I_{-s}} \leq \exp\left(\frac{c_2(s+r)}{K}\right). \tag{4.44}$$

Moreover, this allows us to conclude the next concentration inequality:

Corollary 4.12. *Let $c \log n < p \leq \infty$. Then, one has:*

$$P\left(\left|\|X\|_p - \mathbb{E}\|X\|_p\right| > \varepsilon \mathbb{E}\|X\|_p\right) \leq C \exp\left(-c \max\{\varepsilon, \varepsilon^2\} \log n\right), \tag{4.45}$$

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

Proof. Recall that by Proposition 2.8 we only have to consider the range $0 < \varepsilon < 1$. Then, using Markov's inequality and Remark 4.11 we may write:

$$P\left(\|X\|_p \geq (1+\varepsilon)I_1\right) \leq P\left(\|X\|_p \geq e^{\varepsilon/2} I_1\right) \leq e^{-\varepsilon r/2} \left(\frac{I_r}{I_1}\right)^r \leq \exp(-\varepsilon r/2 + c_2 r^2/K),$$

for all $0 < r < c_1\sqrt{K}$. The choice $r \simeq \sqrt{K}$ yields the one-sided estimate:

$$P\left(\|X\|_p > (1+\varepsilon)I_1\right) \leq C_1 \exp\left(-c_1' \varepsilon \sqrt{K}\right).$$

Working similarly for the probability $P(\|X\|_p < (1-\varepsilon)I_1)$ we conclude the concentration inequality:

$$P\left(\left|\|X\|_p - I_1\right| > \varepsilon I_1\right) < C \exp\left(-c\varepsilon \sqrt{K}\right),$$

for all $0 < \varepsilon < 1$. The fact that $k_p \simeq \log n$ for $p \gtrsim \log n$ and the standard Gaussian concentration complete the proof. □

Taking into account the results of this paragraph we may have a concentration result which interpolates between the concentration estimates for fixed $p \geq 1$ and $p = \infty$:

Theorem 4.13. *Let $n \geq 2$. Then, for any $1 \leq p \leq \infty$ one has:*

$$P\left(\left|\|X\|_p - \mathbb{E}\|X\|_p\right| > \varepsilon \mathbb{E}\|X\|_p\right) \leq C_1 \exp(-c_1 \beta(n, p, \varepsilon)), \tag{4.46}$$

for $0 < \varepsilon < 1$ and $C_1, c_1 > 0$ absolute constants where $\beta(n, p, \varepsilon)$ is defined as follows:

i. If $1 \leq p \leq 2$, then

$$\beta(n, p, \varepsilon) \simeq \varepsilon^2 n, \tag{4.47}$$

ii. for $2 < p \leq \varepsilon_0 \log n$,

$$\beta(n, p, \varepsilon) \simeq \begin{cases} C^{-2p}\varepsilon^2 n, & 0 < \varepsilon < C^p n^{-\frac{p-2}{2(p-1)}} \\ \varepsilon^{2/p} n^{2/p}, & C^p n^{-\frac{p-2}{2(p-1)}} \leq \varepsilon < 1/p \\ \varepsilon k_p, & 1/p \leq \varepsilon < 1 \end{cases}, \tag{4.48}$$

where $C > 0$ is suitable absolute constant, and

iii. for $p > \varepsilon_0 \log n$ we have:

$$\beta(n, p, \varepsilon) \simeq \varepsilon k_p, \tag{4.49}$$

where $\varepsilon_0 > 0$ is suitable absolute constant.

5 The dependence on ε in Dvoretzky's theorem for ℓ_p^n

In this paragraph we study the dependence on ε in random version of Dvoretzky's theorem for the ℓ_p^n spaces. Our argument is inspired by Schecthman's approach in [23]. The key point in his proof is a distributional inequality for rectangular matrices with independent standard gaussian entries. In particular, he proves that if $G = (g_{ij})_{i,j=1}^{n,k}$ is such a Gaussian matrix then the process $(Gx)_{x \in S^{k-1}}$ indexed by the k-dimensional sphere is sub-gaussian:

Theorem 5.1 (Schechtman's distributional inequality). *Let $f : \mathbb{R}^n \to \mathbb{R}$ be L-Lipschitz map (with respect to the Euclidean metric) and let $a, b \in S^{k-1}$. Then, for any random matrix $G = (g_{ij})_{i,j=1}^{n,k}$ (where g_{ij} are iid standard Gaussian rv's) we have:*

$$\text{Prob}\left(\left| f(G(a)) - f(G(b)) \right| > t \right) \leq 2 \exp\left(-\frac{2}{\pi^2} \frac{t^2}{L^2 \|a-b\|_2^2} \right), \tag{5.1}$$

for all $t > 0$.

Having proved this inequality, then a standard chaining argument gives the main result in [23]. The proof of Theorem 5.1 is based on an orthogonal splitting, combined with a conditioning argument and inequality (2.36). Here we use these ideas to prove a functional inequality which generalizes (5.1). Once again, the advantage of this new inequality is that involves $\|\nabla f\|_2$ instead of the Lipschitz constant of f. Our result reads as follows:

Theorem 5.2. *Let $a, b \in S^{k-1}$ and $G = (g_{ij})_{i,j=1}^{n,k}$ be random matrix with standard i.i.d. gaussians entries. If $f : \mathbb{R}^n \to \mathbb{R}$ is C^1-smooth, then we have:*

$$\left(\mathbb{E} \left| f(Ga) - f(Gb) \right|^r \right)^{1/r} \leq C \sqrt{r} \|a-b\|_2 \left(\mathbb{E} \|\nabla f(W)\|_2^r \right)^{1/r}, \tag{5.2}$$

for all $r \geq 1$, where $W \sim N(\mathbf{0}, I_n)$ and $C > 0$ is an absolute constant.

Proof. Fix $a, b \in S^{k-1}$ and assume without loss of generality that $a \neq \pm b$. Define $p := \frac{a+b}{2}$ and note that since $\|a\|_2 = \|b\|_2$ the vector $u := a - p$ is perpendicular to p. Set $X := G(u)$ and $Z := G(p)$ and note that X, Z are independent gaussian random vectors in $\mathbb{R}^n$ with $X \sim N(\mathbf{0}, \|u\|_2^2 I_n)$, $Z \sim N(\mathbf{0}, \|p\|_2^2 I_n)$ and $G(a) = Z + X$ while $G(b) = Z - X$. Thus, we may write:

$$\mathbb{E} \left| f(Ga) - f(Gb) \right|^r = \mathbb{E}_Z \mathbb{E}_X \left| f(Z+X) - f(Z-X) \right|^r.$$

For $x, z \in \mathbb{R}^n$ we define $F(x,z) := f(z+x) - f(z-x)$. Note that for fixed z we have $\mathbb{E}_X F(X,z) = 0$ since, X is symmetric random vector. Applying Theorem 2.10 for $\phi(t) = |t|^r$, $r \geq 1$ and $x \mapsto F(x,z)$ instead of f we derive:

$$\begin{aligned} \mathbb{E}|F(X,z)| = \mathbb{E}_X \left| f(z+X) - f(z-X) \right|^r &\leq \left(\frac{\pi}{2} \right)^r \mathbb{E}_{X,Y} \left| \langle \nabla f(z+X), Y \rangle + \langle \nabla f(z-X), Y \rangle \right|^r \\ &\leq \pi^r \mathbb{E}_{X,Y} \left| \langle \nabla f(z+X), Y \rangle \right|^r \\ &= \pi^r \|a-b\|_2^r \sigma_r^r \, \mathbb{E}_X \|\nabla f(z+X)\|_2^r. \end{aligned}$$

Moreover, note that $W := X + Z \sim N(\mathbf{0}, I_n)$, thus we get:

$$\begin{aligned}(5.3)\qquad \mathbb{E}\,|f(G(a)) - f(G(b))|^r = \mathbb{E}|F(X,Z)| &\le \pi^r \|a-b\|_2^r \sigma_r^r\, \mathbb{E}_{X,Z}\, \|\nabla f(Z+X)\|_2^r \\ &= \pi^r \|a-b\|_2^r\, \sigma_r^r\, \mathbb{E}\, \|\nabla f(W)\|_2^r\,,\end{aligned}$$

where W is standard Gaussian random vector in $\mathbb{R}^n$. □

Remark 5.3. 1. Assuming further that f is L-Lipschitz and applying Markov's inequality we conclude Schechtman's distributional inequality (5.1).

2. The same proof can provide the following variant of Theorem 2.10 which we state for future reference:

Theorem 5.4. *Let $\phi : \mathbb{R} \to \mathbb{R}$ be convex function and let $f : \mathbb{R}^n \to \mathbb{R}$ be C^1-smooth. If $G = (g_{ij})_{i,j=1}^{n,k}$ is Gaussian matrix and $a, b \in S^{k-1}$, then we have:*

$$(5.4)\qquad \mathbb{E}\phi\left(f(Ga) - f(Gb)\right) \le \mathbb{E}\phi\left(\frac{\pi}{2}\|a-b\|_2 \langle \nabla f(X), Y\rangle\right),$$

where X, Y are independent copies of a Gaussian n-dimensional random vector.

The proof is left as an exercise to the interested reader.

3. For $a, b \in S^{k-1}$ with $\langle a, b\rangle = 0$ the above statements are reduced to the inequalities we discussed in Section 2.

Now this refined form of (2.34) and the same argument as in Theorem 4.5 (upper bound) yields the following:

Corollary 5.5. *Let $2 < p < \infty$. Let $a, b \in S^{k-1}$ and let $G = (g_{ij})_{i,j=1}^{n,k}$ typical Gaussian random variables. Then,*

$$(5.5)\qquad \left(\mathbb{E}\left|\|G(a)\|_p - \|G(b)\|_p\right|^r\right)^{1/r} \le C^p \|a-b\|_2 \psi(n,p,r) \mathbb{E}\|X\|_p,$$

for $r \ge 1$. Moreover, for any $t > 0$ one has:

$$(5.6)\qquad P\left(\left|\|G(a)\|_p - \|G(b)\|_p\right| > tI_1\right) \le C \exp\left(-c\tau\left(n, p, \frac{t}{\|a-b\|_2}\right)\right),$$

where

$$(5.7)\qquad \tau(n,p,t) = \begin{cases} C^{-2p} t^2 n, & 0 < t \le C^p n^{-\frac{p-2}{2(p-1)}} \\ C^{-2} t^{2/p} n^{2/p}, & C^p n^{-\frac{p-2}{2(p-1)}} < t \le 1 \\ C^{-2} t^2 n^{2/p}, & t > 1 \end{cases},$$

and $C, c > 0$ are absolute constants.

The chaining method: Dudley-Fernique decomposition. For each $j = 1, 2, \dots$ consider δ_j-nets $\mathcal{N}_j$ on S^{k-1} with cardinality $|\mathcal{N}_j| \leq (3/\delta_j)^k$ (see [18, Lemma 2.6]). Note that for any $\theta \in S^{k-1}$ and for all j there exist $u_j \in \mathcal{N}_j$ with $\|\theta - u_j\|_2 \leq \delta_j$ and by triangle inequality it follows that $\|u_j - u_{j-1}\|_2 \leq \delta_j + \delta_{j-1}$. Moreover, if we assume that $\delta_j \to 0$ as $j \to \infty$ and (t_j) is sequence of numbers with $t_j \geq 0$ and $\sum_j t_j \leq 1$ then, for any $\varepsilon > 0$ we have the next:

Claim. If we define the following sets:

$$A := \left\{\omega \mid \exists \theta \in S^{k-1} : \left| \|G_\omega(\theta)\| - E \right| > \varepsilon E \right\}, \tag{5.8}$$
$$A_1 := \left\{\omega \mid \exists u_1 \in \mathcal{N}_1 : \left| \|G_\omega(u_1)\| - E \right| > t_1 \varepsilon E \right\}$$

and for $j \geq 2$

$$A_j := \left\{\omega \mid \exists u_j \in \mathcal{N}_j, u_{j-1} \in \mathcal{N}_{j-1} : \left| \|G_\omega(u_j)\| - \|G_\omega(u_{j-1})\| \right| > t_j \varepsilon E \right\}, \tag{5.9}$$

then the following inclusion holds:

$$A \subseteq \bigcup_{j=1}^{\infty} A_j. \tag{5.10}$$

Indeed; if $\omega \notin \bigcup_{j=1}^{\infty} A_j$ then for any j and any $u_j \in \mathcal{N}_j$ we have:

$$\left| \|G_\omega(u_1)\| - E \right| \leq t_1 E \quad \text{and} \quad \left| \|G_\omega(u_j)\| - \|G_\omega(u_{j-1})\| \right| \leq t_j E, \quad j = 2, 3, \dots.$$

For any θ there exist $\theta_j \in \mathcal{N}_j$ such that $\|\theta - \theta_j\|_2 < \delta_j$ for $j = 1, 2, \dots$ and for any $N \geq 2$ we may write:

$$\begin{aligned} \left| \|G_\omega(\theta)\| - E \right| &\leq \left| E - \|G_\omega(\theta_1)\| \right| + \sum_{j=2}^{N} \left| \|G_\omega(\theta_{j-1})\| - \|G_\omega(\theta_j)\| \right| + \left| \|G_\omega(\theta_N)\| - \|G_\omega(\theta)\| \right| \\ &\leq \sum_{j=1}^{N} \varepsilon t_j E + \delta_N \cdot \|G_\omega\|_{2 \to X}. \end{aligned}$$

Since, $N \geq 2$ is arbitrary the claim is proved.

Now we apply the above chaining method for the ℓ_p norm with $p > 2$ and we employ the distributional inequality of Corollary 5.5 to prove our second main result:

Theorem 5.6 (Random Dvoretzky for ℓ_p^n). *Let $1 \leq p \leq \infty$. Then, for each n and for any $0 < \varepsilon < 1$ the random subspace of ℓ_p^n with dimension $k \leq k(n, p, \varepsilon)$ is $(1+\varepsilon)$-Euclidean with probability greater than $1 - C\exp(-ck(n, p, \varepsilon))$, where $k(n, p, \cdot)$ is defined as follows:*

(i) *For $1 \leq p < 2$ we have:*

$$k(n, p, \varepsilon) \simeq \varepsilon^2 n, \tag{5.11}$$

(ii) *For $2 < p < \varepsilon_0 \log n$ we have:*

$$k(n,p,\varepsilon) \simeq \begin{cases} (Cp)^{-p}\varepsilon^2 n, & 0 < \varepsilon \le (Cp)^{p/2} n^{-\frac{p-2}{2(p-1)}} \\ \frac{1}{p}(\varepsilon n)^{2/p}, & (Cp)^{p/2} n^{-\frac{p-2}{2(p-1)}} < \varepsilon \le 1/p \\ \varepsilon p n^{2/p}, & 1/p < \varepsilon < 1. \end{cases} \tag{5.12}$$

Moreover, for $p < \varepsilon_0 \log n$ but proportional to $\log n$ we have:

$$k(n,p,\varepsilon) \simeq \log n / \log \frac{1}{\varepsilon}, \tag{5.13}$$

(iii) *For $\varepsilon_0 \log n < p \le \infty$ we have:*

$$k(n,p,\varepsilon) \simeq \varepsilon \log n / \log \frac{1}{\varepsilon}, \tag{5.14}$$

where $C, c, \varepsilon_0 > 0$ are absolute constants.

Proof. For $1 \le p < 2$ follows from Theorem 2.3 and the fact that $k_p \simeq n$ for $1 \le p < 2$. Let $2 < p < \infty$ and onsider $0 < \varepsilon < 1$. Choose $\delta_j = e^{-j}$, $t_j = s_p^{-1} j^{p/2} e^{-j}$, with $s_p := \sum_{j=1}^\infty j^{p/2} e^{-j}$. Then, according to the previous chaining method we may write:

$$\begin{aligned} P(A) &\le C|\mathcal{N}_1| \exp(-c_1 \tau(n,p,\varepsilon t_1)) + C \sum_{j=2}^\infty |\mathcal{N}_{j-1}| \cdot |\mathcal{N}_j| \exp(-c_1 \tau(n,p,\varepsilon s_p^{-1} t_j e^j/4)) \\ &\le C \sum_{j=1}^\infty (3e^j)^{2k} \exp(-c_2 \tau(n,p,s_p^{-1}\varepsilon j^{p/2})). \end{aligned}$$

Note that

$$\tau(n,p,t) \simeq \min\left\{ \frac{t^2 n}{C_1^p}, (tn)^{2/p} \right\},\ t > 0, \tag{5.15}$$

thus we get:

$$\tau(n,p,s_p^{-1}\varepsilon j^{p/2}) \gtrsim j \min\left\{ \frac{\varepsilon^2 n}{(Cp)^p}, \frac{(\varepsilon n)^{2/p}}{p} \right\} = jk(n,p,\varepsilon), \tag{5.16}$$

where we have used the fact that $s_p \lesssim \sqrt{p}(\frac{p}{2e})^{p/2}$. Therefore, we have:

$$\begin{aligned} P(A) &\le C \sum_{j=1}^\infty \exp\left(c_3 jk - c_4 jk(n,p,\varepsilon)\right) \\ &\le \sum_{j=1}^\infty \exp\left(-\frac{c_4}{2} jk(n,p,\varepsilon)\right) \le C' \exp\left(-\frac{c_4}{2} k(n,p,\varepsilon)\right). \end{aligned}$$

as long as $k \leq \frac{c_4}{2c_2} k(n,p,\varepsilon)$. In the case that $p < c_0 \log n$ and $p \gg 1$ for the range $1/p < \varepsilon < 1$ we have for any $\theta \in S^{k-1}$ the concentration inequality:

$$P\left(\left|\|G\theta\|_p - \mathbb{E}\|X\|_p\right| > \varepsilon \mathbb{E}\|X\|_p\right) \leq C \exp(-c\varepsilon k_p), \tag{5.17}$$

by Proposition 4.2. Thus, the classical net argument yields the estimate: $k(n,p,\varepsilon) \simeq \varepsilon k_p / \log \frac{1}{\varepsilon}$. Indeed; if $\mathcal{N}$ is δ-net on S^{k-1}, then previous distributional inequality implies:

(5.18)
$$P\left(\exists z \in \mathcal{N} : \left|\|Gz\|_p - \mathbb{E}\|X\|_p\right| > \varepsilon \mathbb{E}\|X\|_p\right) \leq C|\mathcal{N}| \exp(-c\varepsilon k_p) \leq C(3/\delta)^k \exp(-c\varepsilon k_p).$$

So, for $k \lesssim \frac{\varepsilon}{\log \frac{1}{\delta}} k_p$ we get that with probability greater than $1 - C\exp(-\varepsilon k_p / \log \frac{1}{\delta})$ we have:

$$\left|\|Gz\|_p - \mathbb{E}\|X\|_p\right| \leq \varepsilon \mathbb{E}\|X\|_p \tag{5.19}$$

for all $z \in \mathcal{N}$. A standard approximation argument and a suitable choice $\delta \simeq \varepsilon$ shows that:

$$\left|\|G\theta\|_p - \mathbb{E}\|X\|_p\right| \leq 3\varepsilon \mathbb{E}\|X\|_p \tag{5.20}$$

for all $\theta \in S^{k-1}$.

Moreover, for $p < c_0' \log n$ but $p \simeq \log n$ the main result of Section 2 shows that $\mathrm{Var}\|X\|_p \lesssim n^{-c_1}$ for some absolute constant $c_1 > 0$. Therefore, Chebyshev's probabilistic inequality yields:

$$\forall \theta \in S^{k-1} \quad P\left(\left|\|G\theta\|_p - \mathbb{E}\|X\|_p\right| > \varepsilon \mathbb{E}\|X\|_p\right) \lesssim \frac{1}{\varepsilon^2 n^{c_1}}. \tag{5.21}$$

The net argument as before implies in that case $k(n,p,\varepsilon) \simeq \log n / \log \frac{1}{\varepsilon}$. Finally, for $p \gtrsim \log n$ we employ Corollary 4.12 with the net argument again to get $k(n,p,\varepsilon) \simeq \varepsilon \log n / \log \frac{1}{\varepsilon}$. The result follows once we observe that the complement of the event $\{\omega \mid \exists \theta : \left|\|G_\omega \theta\|_p - E\right| > \varepsilon E\}$ is contained in the event that $G(\mathbb{R}^k)$ is k-dimensional subspace which is $\frac{1+\varepsilon}{1-\varepsilon}$-spherical and by recalling the known fact that $P(G(\mathbb{R}^k) \in \mathcal{B}) = \nu_{n,k}(\mathcal{B})$ for any Borel set $\mathcal{B}$ in $G_{n,k}$. □

Note.

Below we show that the dependence on ε we get for the randomized Dvoretzky in ℓ_p^n, for fixed $2 < p < \infty$ is essentially optimal. We have the following:

Theorem 5.7 (Optimality in Random Dvoretzky for ℓ_p^n). *Let $2 < p < \infty$. Assuming that with probability larger than $1 - e^{-\beta k}$, a k-dimensional subspace satisfies that the ratio between the ℓ_p^n norm and a multiple of the ℓ_2^n norm are $(1+\varepsilon)$ equivalent for all vectors in the subspace, with $n^{-\frac{p-2}{2(p-1)}} \lesssim_p \varepsilon < 1$, then $k \lesssim \beta^{-1} p (\varepsilon n)^{2/p}$.*

For the proof we will need the next lemma from [24]:

Lemma 5.8. *Let $1 \leq k \leq n-1$ and let $\mathcal{A} \subset G_{n,k}$ be a $\nu_{n,k}$-measurable set. Then, for $U_{\mathcal{A}} := \bigcup\{F \mid F \in \mathcal{A}\}$ we have:*

$$\nu_{n,k}(\mathcal{A}) \leq [\gamma_n(U_{\mathcal{A}})]^k. \tag{5.22}$$

Proof of Theorem 5.7. Let $0 < \varepsilon < 1/3$ and define the collection of all k-dimensional subspaces of the space $X = (\mathbb{R}^n, \|\cdot\|_p)$ for which the restricted norm there has distortion (with respect to the euclidean norm) at most $1+\varepsilon$:

$$\mathcal{A}_\varepsilon := \{F \in G_{n,k} \mid \exists \lambda_F : \lambda_F \leq \|\theta\| \leq (1+\varepsilon)\lambda_F,\ \forall \theta \in S_F\}. \tag{5.23}$$

Note that for $F \in \mathcal{A}_\varepsilon$ we have: $(1+\varepsilon)^{-1}M_F \leq \lambda_F \leq M_F$, thus instead of working with λ_F we may define $\mathcal{A}_\varepsilon$ using $M_F := M(F \cap B_X)$ namely, if

$$\mathcal{F}_\varepsilon := \left\{F \in G_{n,k} \mid (1+\varepsilon)^{-1}M_F \leq \|\theta\| \leq (1+\varepsilon)M_F \ \ \forall \theta \in S_F\right\},$$

then we get $\mathcal{A}_\varepsilon \subset \mathcal{F}_\varepsilon$. Define further:

$$\mathcal{B}_\varepsilon := \left\{F \in \mathcal{F}_\varepsilon \mid (1-2\varepsilon)\frac{\mathbb{E}\|g\|}{\mathbb{E}\|g\|_2} \leq M_F \leq (1+2\varepsilon)\frac{\mathbb{E}\|g\|}{\mathbb{E}\|g\|_2}\right\} \tag{5.24}$$

and note that $\mathcal{F}_\varepsilon$, $\mathcal{B}_\varepsilon$ are measurable.[2] Hence, an application of Lemma 5.8 yields:

$$\begin{aligned}\nu_{n,k}(\mathcal{F}_\varepsilon) &= \nu_{n,k}(\mathcal{F}_\varepsilon \setminus \mathcal{B}_\varepsilon) + \nu_{n,k}(\mathcal{B}_\varepsilon)\\ &\leq \left[\gamma_n\left(\left\{x : \|x\| \geq \frac{1+2\varepsilon}{1+\varepsilon}\frac{\mathbb{E}\|g\|}{\mathbb{E}\|g\|_2}\|x\|_2 \textbf{ or } \|x\| \leq (1+\varepsilon)(1-2\varepsilon)\frac{\mathbb{E}\|g\|}{\mathbb{E}\|g\|_2}\|x\|_2\right\}\right)\right]^k +\\ &\quad \left[\gamma_n\left(\left\{x : \frac{1-2\varepsilon}{1+\varepsilon}\|x\|_2\frac{\mathbb{E}\|g\|}{\mathbb{E}\|g\|_2} \leq \|x\| \leq (1+\varepsilon)(1+2\varepsilon)\frac{\mathbb{E}\|g\|}{\mathbb{E}\|g\|_2}\|x\|_2\right\}\right)\right]^k.\end{aligned}$$

Apply this argument for the ℓ_p norm with $2 < p < \infty$ and consider the next:

Claim. For every $C^p n^{-\frac{p-2}{2(p-1)}} < t < 1$ we have:

$$ce^{-Cp(tn)^{2/p}} \leq P\left(\|g\|_p \leq \frac{(1-t)\mathbb{E}\|g\|_p}{\mathbb{E}\|g\|_2}\|g\|_2 \textbf{ or } \|g\|_p \geq \frac{(1+t)\mathbb{E}\|g\|_p}{\mathbb{E}\|g\|_2}\|g\|_2\right) \leq Ce^{-c(tn)^{2/p}} \tag{5.25}$$

Proof of the Claim. Follows by Theorem 4.3 and Theorem 4.9. □

Now assume that $C^p n^{-\frac{p-2}{2(p-1)}} < \varepsilon < 1/3$ and take into account the previous claim to get:

$$\begin{aligned}\nu_{n,k}(\mathcal{F}_\varepsilon) &\leq C^k e^{-ck(\varepsilon n)^{2/p}} + (1 - ce^{-Cp(\varepsilon n)^{2/p}})^k\\ &\leq e^{-c'k(\varepsilon n)^{2/p}} + 1 - ce^{-Cp(\varepsilon n)^{2/p}}.\end{aligned}$$

[2] The map $F \mapsto M_F$ is Lipschitz continuous with respect to the unitarily invariant metric d on $G_{n,k}$ defined as: $d(E,F) = \inf\{\|I - U\|_{\rm op} : U(E) = F,\ U \in O(n)\}$, $E, F \in G_{n,k}$.

Now assuming that $\nu_{n,k}(\mathcal{F}_\varepsilon) \geq 1 - e^{-\beta k}$ for some absolute constant $\beta > 0$ and note that $\beta \ll (\varepsilon n)^{2/p}$, for this range of ε we obtain:

$$1 - ce^{-Cp(\varepsilon n)^{2/p}} \geq 1 - e^{-\beta k} - e^{-c' k (\varepsilon n)^{2/p}} \geq 1 - 2e^{-c''\beta k},$$

which implies that:

$$k \leq \frac{C'}{\beta} p (\varepsilon n)^{2/p},$$

as required. □

6 Further remarks and questions

1. Instability of the variance. It is worth mentioning that the variance is not isomorphic invariant. Note that Theorem 3.8 implies the following:

There exists absolute constant $0 < c_0 < 1$ with the following property: for every $n \geq 2$ there exist two 1-symmetric convex bodies K and L on $\mathbb{R}^n$ such that:

$$\mathrm{Var}_{\gamma_n}\|x\|_K = O(n^{-\delta}(\log n)^{-1}),\ \mathrm{Var}_{\gamma_n}\|x\|_L \simeq (\log n)^{-1} \text{ and } e^{-1/c_0} L \subseteq K \subseteq L,$$

where $\delta = 1 - c_0$.

Indeed; for $p_0 := c_0 \log n$, where $0 < c_0 < 1$ as in Theorem 3.8, we consider $K := B_{p_0}^n$, hence we have:

$$\mathrm{Var}\|X\|_K \leq C_1 \frac{2^{p_0}}{p_0 n^{1-2/p_0}} = \frac{C_1 e^{2/c_0}}{c_0 \log n} n^{-(1 - c_0 \log 2)},$$

whereas for $L := B_\infty^n$ we have $\mathrm{Var}\|X\|_L \simeq \frac{1}{\log n}$ and $\|x\|_L \leq \|x\|_K \leq n^{1/p_0}\|x\|_L$ for any $x \in \mathbb{R}^n$.

2. Non-centered moments. We know that for any centrally symmetric convex body T on $\mathbb{R}^n$ one has:

$$\frac{c_1 r}{n} \leq \left(\frac{I_r(\gamma_n, T)}{I_1(\gamma_n, T)}\right)^2 - 1 \leq \frac{c_2 r}{k(T)}, \tag{6.1}$$

for all $r \geq 2$, where $c_1, c_2 > 0$ are absolute constants. This follows from the lower estimate in (2.25) and Lemma 2.6. In particular, for $2 \leq r \leq k(T)$ we obtain:

$$\frac{c_1' r}{n} \leq \frac{I_r(\gamma_n, T)}{I_1(\gamma_n, T)} - 1 \leq \frac{c_2' r}{k(T)} \tag{6.2}$$

and when $k(T) \simeq n$ we readily see that this estimate is sharp up to constants. Furthermore, one can show that this is the case even for the ℓ_p norms for $2 < p < \infty$

even though the critical dimension k_p is much smaller than n: For $2 < p < \infty$ we have:

$$\frac{I_r(\gamma_n, B_p^n)}{I_1(\gamma_n, B_p^n)} \leq 1 + \frac{C^p}{n} r, \tag{6.3}$$

for all $1 \leq r \leq k_p/C$. In fact for the negative moments this is already clear if we take into account Theorem 2.14, Proposition 4.1 and Theorem 4.10. More precisely we have: For $1 \leq p < c \log n$ and for any $1 \leq r \leq ck_p$ we get:

$$\max\left\{ \frac{I_1(\gamma_n, B_p^n)}{I_{-r}(\gamma_n, B_p^n)}, \frac{I_r(\gamma_n, B_p^n)}{I_1(\gamma_n, B_p^n)} \right\} \leq 1 + \frac{C^p}{n} r \tag{6.4}$$

and for $p \geq c \log n$ and $1 \leq r \leq ck_p$ we have:

$$\max\left\{ \frac{I_1(\gamma_n, B_p^n)}{I_{-r}(\gamma_n, B_p^n)}, \frac{I_r(\gamma_n, B_p^n)}{I_1(\gamma_n, B_p^n)} \right\} \leq 1 + \frac{C}{(\log n)^2} r. \tag{6.5}$$

We should note here the next threshold phenomenon when $2 < p \leq \infty$:

- $2 < p \leq c \log n$: It is $I_r/I_1 - 1 \lesssim_p r/n = O_p(n^{2/p-1})$ for $1 \leq r \leq ck_p$ while for $r \geq c' k_p$ we have $I_r/I_1 - 1 \simeq 1$.
- $p > c \log n$: It is $I_r/I_1 - 1 \lesssim r/(\log n)^2 = O((\log n)^{-1})$ for $r \leq ck_p \simeq \log n$, while for $r \geq c' k_p$ we have $I_r/I_1 - 1 \simeq 1$,

for some absolute constant $c' > c$. The detailed study of this phenomenon will be presented elsewhere. Let us also note that although the behavior of the quantities $\frac{I_r}{I_1} - 1$, $\frac{I_1}{I_{-r}} - 1$ is completely determined for the ℓ_p norms – it is of the order r/n for $1 \leq r \leq ck_p$ – combining this information with Markov's inequality we still do not derive the optimal concentration inequality in the range $2 < p < \infty$.

3. The dependence on ε in randomized Dvoretzky. The gaussian concentration for any norm $\|\cdot\|_T$ states:

$$\gamma_n\left(x : \left|\|x\|_T - I_1\right| \geq t I_1\right) \leq C e^{-ct^2 k(T)}, \quad t > 0. \tag{6.6}$$

In case that $k(T) \simeq n$ we may prove that this estimate is best possible in the following sense:

Lemma 6.1. *Let T be centrally symmetric convex body on $\mathbb{R}^n$ with $k = k(T) \geq \alpha n$ for some constant $\alpha \in (0,1)$ with $\alpha \gg (\frac{\log n}{n})^{1/2}$. Then,*

$$\gamma_n\left(x : \left|\|x\|_T - I_1\right| \geq \varepsilon I_1\right) \geq c e^{-C\varepsilon^2 n/\alpha^2}, \tag{6.7}$$

for all $\sqrt{\log n/n} < \varepsilon < 1$ and $C, c > 0$ absolute constants.

Proof. We know that there exists $c_1 \in (0,1)$ such that $I_r > (1+\frac{c_1 r}{n})I_1$ for all $2 \le r \le k(T)$. Let $\sqrt{\log n/n} < \varepsilon < c_1\alpha/2$. Set $r_0 := \frac{2n\varepsilon}{c_1}$ and note that $2 < r_0 < \alpha n \le k(T)$. Then, using (6.2) we may write:

$$\gamma_n(x : \|x\| > (1+\varepsilon)I_1) \ge \gamma_n\left(x : \|x\| > \frac{1+\varepsilon}{1+\frac{c_1 r_0}{n}} I_{r_0}\right) = \gamma_n(x : \|x\| > \delta I_{r_0}) \ge (1-\delta^{r_0})^2 \frac{I_{r_0}^{2r_0}}{I_{2r_0}^{2r_0}},$$

by the Paley-Zygmund inequality, where $\delta := \frac{1+\varepsilon}{1+\frac{c_1 r_0}{n}}$. Recall that there exists absolute constant $C_2 > 1$ such that $I_{2r}/I_r \le C_2$ for all $r \ge 1$. Thus, we get:

$$\gamma_n(x : \|x\| > (1+\varepsilon)I_1) \ge (1-\delta)^2 \exp(-2r_0 \log C_2) \ge (\varepsilon/2)^2 e^{-C_3\varepsilon^2 n} > c_4 e^{-C_4\varepsilon^2 n}.$$

For the range $c_1\alpha/2 \le \varepsilon < 1$ we argue as follows: From Lemma 2.7 we have $I_r \ge c_5\sqrt{r/k}I_1$ for all $r > k$. Set $t := (1+\frac{2}{c_1\alpha})\varepsilon > 1$ and note that for $s := 4kt^2/c_5^2 > k$ we have:

$$\gamma_n(x : \|x\| > (1+\varepsilon)I_1) \ge \gamma_n(x : \|x\| > tI_1) \ge \gamma_n\left(x : \|x\| > \frac{1}{2}I_s\right) \ge \frac{1}{4}e^{-2s\log C_2},$$

where in the last step we have used Paley-Zygmund inequality again. The estimate $s \lesssim k\varepsilon^2/\alpha^2$ completes the proof. □

Although, the Gaussian concentration for such spaces is sharp, the argument provided in Section 5 fails to give the optimal dependence on ε in randomized Dvoretzky. The reason for that, roughly speaking, is that in the Gaussian setting norms with concentration estimate less than $e^{-\varepsilon^2 n}$ cannot be distinguished from the Euclidean norm. To this end is more appropriate to work with the uniform probability measure on S^{n-1} when we study Dvoretzky's theorem in normed spaces. Assume that $X = (\mathbb{R}^n, \|\cdot\|)$ satisfies:

$$\sigma\left(\left\{\theta \in S^{n-1} : \left|\|\theta\| - M_1\right| > \varepsilon M_1\right\}\right) \ge c e^{-C\varepsilon^2 k(X)}, \quad 0 < \varepsilon < 1. \tag{6.8}$$

If a random k-dimensional subspace satisfies that the ratio between the given norm and a multiple of the ℓ_2^n norm are $1+\varepsilon$ equivalent for all vectors in the subspace with probability $> 1 - e^{-\beta k}$, then $k \le C\beta^{-1}\varepsilon^2 k(X)$. The proof follows the same lines as in Section 5, but we skip the dilation passing from the sphere to the whole space: Consider the set

$$\mathcal{F}_\varepsilon = \{F \in G_{n,k} : M_F \le \|\phi\| \le (1+\varepsilon)M_F, \ \forall \phi \in S_F\}$$

and the set:

$$\mathcal{G}_\varepsilon := \{F \in \mathcal{F}_\varepsilon : (1-2\varepsilon)M \le M_F \le (1+\varepsilon)M\}.$$

Then, we have:

$$\begin{aligned}
\nu_{n,k}(\mathcal{F}_\varepsilon) &= \nu_{n,k}(\mathcal{F}_\varepsilon \setminus \mathcal{G}_\varepsilon) + \nu_{n,k}(\mathcal{G}_\varepsilon) \\
&\le \left[\sigma\left(\theta \in S^{n-1} : \left|\|\theta\| - M\right| > \varepsilon M\right)\right]^k + \left[\sigma\left(\theta : (1-2\varepsilon)M \le \|\theta\| < (1+3\varepsilon)M\right)\right]^k \\
&\le \exp(-ck\varepsilon^2 k(X)) + 1 - c\exp(-C\varepsilon^2 k(X)),
\end{aligned}$$

by the assumption and we have also used the fact that if $z_1, \dots, z_k$ are independent random vectors uniformly distributed over S^{n-1}, then $F = \mathrm{span}\{z_1, \dots, z_k\}$ is uniformly distributed in $G_{n,k}$ almost surely. It follows that $1 - e^{-\beta k} - e^{-ck\varepsilon^2 k(X)} \leq 1 - ce^{-C\varepsilon^2 k(X)}$ and assuming that $\beta \lesssim \varepsilon^2 k(X)$ the assertion follows.

4. Optimal Gaussian concentration and "new dimensions". The reader should notice that the refined form of the Gaussian concentration for $2 < p < \infty$ (Theorem 4.9) and moreover Theorem 5.6 provide random, almost Euclidean subspaces of relatively large dimensions in which the norm has very small distortion. Previously, that phenomenon couldn't be observed if one was using the classical concentration inequality in terms of the Lipschitz constant. In order to illustrate this let us consider an example, say the ℓ_p norm with $p = 5$. The classical setting yields the existence of k-dimensional sections of B_5^n which are $(1+\varepsilon)$-isomorphic to a mupltiple of B_2^k as long as $k \lesssim \varepsilon^2 n^{2/5}$ and this makes sense for $\varepsilon \in (n^{-1/5}, 1)$. Now, we may consider distortions smaller than $n^{-1/5}$, in fact as small as $n^{-1/2}$, since $\tau(n, 5, \varepsilon) \simeq \min\{\varepsilon^2 n, (\varepsilon n)^{2/5}\}$ for $0 < \varepsilon < 1$. For instance, there exists (for $\varepsilon \simeq n^{-2/5}$) a k-dimensional section (in fact with probability $> 1 - e^{-cn^{1/5}}$) of B_5^n with $k \simeq n^{1/5}$, which is $(1 + n^{-2/5})$-isomorphic to a multiple of B_2^k.

5. The existence of $\log(1/\varepsilon)$ **as** $p \to \infty$. Note that Theorem 4.10 and furthermore Corollary 4.12 suggest that the concentration of the ℓ_p norm with $p \gtrsim \log n$ is similar with the one we get for the ℓ_∞ norm. This means that the classical net argument yields random subspaces which are $(1+\varepsilon)$-spherical as long as $k \lesssim \varepsilon \log n / \log \frac{1}{\varepsilon}$. We do not know if this logarithmic on ε term is needed, for this range of p. Nevertheless, it is easy to check that this is the case when $p > (\log n)^2$.

Proposition 6.2. *Let* $p > (\log n)^2$ *and* $\varepsilon \in (0, 1/3)$. *If the random k-dimensional subspace of* ℓ_p^n *is* $(1+\varepsilon)$*-spherical with probability greater* $3/4$, *then* $k \leq C\varepsilon \log n / \log \frac{1}{\varepsilon}$, *where* $C > 0$ *is an absolute constant.*

Proof. It follows by reducing it to the case of the ℓ_∞^n and by applying Tikhomirov's main result from [29]. □

References

[1] S. Artstein-Avidan, A. Giannopoulos and V. D. Milman, *Asymptotic Geometric Analysis, Part I*, AMS-Mathematical Surveys and Monographs **202** (2015).

[2] S. Boucheron, G. Lugosi and P. Massart, *Concentration Inequalities: A non-asymptotic theory of independence*, Oxford University Press (2013).

[3] S. Chatterjee, *Superconcentration and related topics*, Springer Monographs in Mathematics (2013).

[4] D. Cordero-Erausquin and M. Ledoux, *Hypercontractive Measures, Talagrand's inequality, and Influences*, Geometric Aspects of Functional Analysis, Lecture Notes in Mathematics **2050**, (2012), 169–189.

[5] A. Dvoretzky, *Some results on convex bodies and Banach spaces*, Proc. Int. Symp. on linear spaces, Jerusalem (1961), 123–160.

[6] D. Fresen, *Explicit Euclidean embeddings in permutation invariant normed spaces*, Adv. Math. **266**, 1–16 (2014).

[7] R. D. Gordon, *Values of Mills' ratio of area bounding ordinate and of the normal probability integral for large values of the argument*, Ann. Math. Statistics **12**, no. 3, (1941), 364–366.

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

[9] G. Hardy, J. E. Littlewood and G. Pólya, *Inequalities*, Cambridge University Press 1988.

[10] D. Hunter, `http://sites.stat.psu.edu/~dhunter/asymp/fall2006/lectures`.

[11] B. Klartag and R. Vershynin, *Small ball probability and Dvoretzky theorem*, Israel J. Math., Vol. **157**, no. 1 (2007), 193–207.

[12] H. König, *Isometric imbeddings of Euclidean spaces into finite-dimensional ℓ_p-spaces*, Panoramas of Mathematics (Colloquia 93-94), 79–87, Banach Center Publ. **34**, Polish Acad. Sci., Warszawa, 1995.

[13] R. Latala, *Estimation of moments of sums of independent real random variables*, Ann. Probab. **25** (1997), no. 3, 1502–1513.

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

[15] A. Litvak, V. D. Milman and G. Schechtman, *Averages of norms and quasi-norms*, Math. Ann. **312** (1998), 95–124.

[16] S. J. Montogomery-Smith, *The distribution of Rademacher sums*, Proc. Amer. Math. Soc. **109** (1990), 517–522.

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

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

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

[20] A. Naor, *The Surface Measure and Cone Measure on the Sphere of ℓ_p^n* , Transactions of the American Mathematical Society **359** (2007), 1045–1079.

[21] G. Paouris, P. Pivovarov and P. Valettas, *Concentration and asymmetry measures of random variables*, in preparation.

[22] G. Pisier, *Probabilistic Methods in the Geometry of Banach Spaces*, CIME, Varenna, 1985, Springer, Lecture Notes in Mathematics **1206** (1986), 167–241.

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

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

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

[26] G. Schechtman and J. Zinn, *On the volume of the intersection of two L_p^n balls*, Proc. Amer. Math. Soc. **110**, No.1, (1990), 217–224.

[27] P. Stavrakakis and P. Valettas, *On the geometry of log-concave probability measures with bounded log-Sobolev constant*, Proceedings of the Asymptotic Geometric Analysis Programme, Fields Institute Communications **68** (2013), 359–380.

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

[29] K. E. Tikhomirov, *The Randomized Dvoretzky's theorem in ℓ_∞^n and the χ-distribution*, Geometric Aspects of Functional Analysis, Israel Seminar (GAFA) 2011–2013 (eds. B. Klartag and E. Milman), Lecture Notes in Mathematics **2116**, (2013) Springer.

[30] K. E. Tikhomirov, *Almost Euclidean sections in symmetric spaces and concentration of order statistics*, J. Funct. Anal. **265** (9), 2074–2088, (2013).

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

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

Joel Zinn: `jzinn@math.tamu.edu`
Department of Mathematics, Mailstop 3368
Texas A&M University
College Station, TX 77843-3368