# Dichotomies, structure, and concentration in normed spaces

Grigoris Paouris[*] and Petros Valettas[†]

May 21, 2018

### Abstract

We use probabilistic, topological and combinatorial methods to establish the following deviation inequality: For any normed space $X = (\mathbb{R}^n, \|\cdot\|)$ there exists an invertible linear map $T : \mathbb{R}^n \to \mathbb{R}^n$ with

$$\mathbb{P}\left(\left|\|TG\| - \mathbb{E}\|TG\|\right| > \varepsilon \mathbb{E}\|TG\|\right) \leq C \exp\left(-c \max\{\varepsilon^2, \varepsilon\} \log n\right), \quad \varepsilon > 0,$$

where $G$ is the standard $n$-dimensional Gaussian vector and $C, c > 0$ are universal constants. It follows that for every $\varepsilon \in (0, 1)$ and for every normed space $X = (\mathbb{R}^n, \|\cdot\|)$ there exists a $k$-dimensional subspace of $X$ which is $(1 + \varepsilon)$-Euclidean and $k \geq c\varepsilon \log n / \log \frac{1}{\varepsilon}$. This improves by a logarithmic on $\varepsilon$ term the best previously known result due to G. Schechtman.

## 1 Introduction

The concentration inequality in Gauss' space states that for any Lipschitz map $f : \mathbb{R}^n \to \mathbb{R}$ with $|f(x) - f(y)| \leq L\|x - y\|_2$ for all $x, y \in \mathbb{R}^n$ one has

(1.1) $$\mathbb{P}\left(|f(G) - \mathbb{E}f(G)| > t\right) \leq 2 \exp(-\tfrac{1}{2}t^2/L^2), \quad t > 0,$$

where $G$ is the standard $n$-dimensional Gaussian vector (for a proof the reader is referred to [Pis86]; see [Mau91] for the precise constants). This inequality is the prototype of what is called nowadays the concentration of measure phenomenon, one of the most important ideas in modern probability theory. This fundamental tool

---

was put forward in the local theory of normed spaces in early 70's by V. Milman. Applying (1.1) for a norm $\| \cdot \|$ on $\mathbb{R}^n$ we get

$$(1.2) \qquad \mathbb{P}\left( \left| \|G\| - \mathbb{E}\|G\| \right| > t\mathbb{E}\|G\| \right) \leq 2\exp(-\tfrac{1}{2}t^2 k), \quad t > 0,$$

where $k = k(X) = k(B_X) := (\mathbb{E}\|G\|/b)^2$ is referred to as the *critical dimension* (or *Dvoretzky number*) of the normed space $X = (\mathbb{R}^n, \| \cdot \|)$ and $b = b(X) = b(B_X)$ is the Lipschitz constant of the norm $\| \cdot \|$, i.e. $b = \max\{\|\theta\| : \|\theta\|_2 = 1\}$. It is well known that the above estimate is sharp in the large deviation regime, namely

$$(1.3) \qquad \mathbb{P}(\|G\| \geq (1 + t)\mathbb{E}\|G\|) \geq c\exp(-Ct^2 k), \quad t \geq 1,$$

where $c, C > 0$ are universal constants[1] (see e.g. [LT91, Corollary 3.2], [LMS98, Statement 3.1] and [PVZ17, Proposition 2.10]). In the small deviation regime $0 < t < 1$ there exist many important examples which show that the obtained bounds are suboptimal; see [PVZ17] and [Val17] for a detailed discussion. Ideally one would like to know what properties of the underlying function improve the concentration. An example of such a result was recently obtained by the authors in [PV18] where they proved that a one-sided, variance-sensitive Gaussian small deviation inequality is valid for all convex functions.

This work is also concerned with optimal forms of the Gaussian concentration but the main focus will be on norms. Before stating the main problem of study, let us try to motivate the question which describes it. It is known (see e.g. [Mil71], [Pis86]) that for any norm $\| \cdot \|$ on $\mathbb{R}^n$, there exists a $T \in GL(n)$ such that

$$(1.4) \qquad \mathbb{P}\left( \left| \|TG\| - \mathbb{E}\|TG\| \right| > t\mathbb{E}\|TG\| \right) \leq C\exp(-ct^2 \log n), \quad t > 0,$$

where $G \sim N(\mathbf{0}, I_n)$. This follows from the fact that there exists a *position* (i.e. an invertible linear image) $T^{-1}(B_X)$ of the unit ball $B_X = \{x \in \mathbb{R}^n : \|x\| \leq 1\}$ for which the critical dimension $k(T^{-1}B_X)$ is at least as large as $\log n$ (note that by definition one has $c \leq k(X) \leq n$, hence a linear transformation is required to avoid degeneracy) and in turn this is combined with the general estimate (1.2). Traditionally, this is achieved for John's position, that is the maximal volume ellipsoid inscribed in $B_X$ is the standard Euclidean ball $B_2^n = \{x \in \mathbb{R}^n : \|x\|_2 \leq 1\}$. Let us note that the $\log n$ factor is optimal since for the $\ell_\infty$ norm we have that the cube $B_\infty^n = \{x \in \mathbb{R}^n : \|x\|_\infty \leq 1\}$ is in John's position and $k(\ell_\infty^n) \simeq \log n$. On the other hand, the $\ell_\infty$ norm exhibits exponential concentration (see [Tal91] and [Sch07]),

$$(1.5) \qquad ce^{-C\varepsilon \log n} \leq \mathbb{P}\left( \left| \|G\|_\infty - \mathbb{E}\|G\|_\infty \right| > \varepsilon\mathbb{E}\|G\|_\infty \right) \leq Ce^{-c\varepsilon \log n}, \quad 0 < \varepsilon < 1.$$

In view of the above remarks the following question arises naturally:

---

[1]Here and everywhere else $C, c, C_1, c_1, \ldots$ stand for positive universal constants whose values may change from line to line. For any two quantities $A, B$ depending on the dimension, on the parameters of the problem, etc. we write $A \simeq B$ if there exists a universal constant $C > 0$ -independent of everything- such that $A \leq CB$ and $B \leq CA$.

**Question 1.1.** Is it true that for every norm $\|\cdot\|$ on $\mathbb{R}^n$ there exists a $T \in GL(n)$ with

$$\mathbb{P}\left(\Big|\|TG\| - \mathbb{E}\|TG\|\Big| > t\mathbb{E}\|TG\|\right) \le C \exp(-c \max\{t^2, t\} \log n), \quad t > 0?$$

One of the main difficulties to establish good small deviation estimates, is that the problem is "isometric" in nature rather than "isomorphic".

It is easy to see that two equivalent norms have Dvoretzky's number of the same order and therefore, by (1.3), they exhibit the same large deviation estimate. However, one may find two norms on $\mathbb{R}^n$ which are 2-equivalent and the variance of one is polynomially small while the variance of the other is only logarithmically small (with respect to the dimension), see for example [PVZ17], [PV17] and [LT17]. We should mention that when the norm under consideration is close to the $\ell_2$-norm then it automatically exhibits the optimal concentration in terms of $\varepsilon$ and $n$ (Section 5, §1). On the other hand, for norms close to the $\ell_\infty$ norm this is no longer true: there exists a norm on $\mathbb{R}^n$ which is 2-equivalent to $\|\cdot\|_\infty$ and the optimal concentration is as in (1.2) or (1.4), see e.g. [PV17]. In fact for any norm, one can construct a 5-equivalent norm for which (1.2) is sharp, see [Val17, Section 3]. After these observations it seems quite discouraging to tackle Question 1.1 and leaves only the hope that the appropriate selection of the linear transformation will fix the problem.

The above discussion leads naturally to the question of how we successfully select the position to establish improved concentration estimates. It turns out that John's position, which was widely used to attack problems lying in concentration estimates in the local theory of normed spaces, is not the proper one. It is possible to construct $n$-dimensional normed spaces whose unit ball is in John's position, the critical dimension can be of any order in the admissible range for this position and yet the optimal concentration is as in (1.2) (see Section 5 for the exact formulation). In the closely related setting of almost Euclidean subspaces, it was emphasized by the authors in [PV15] that the choice of the position is crucial for improving the estimates and subsequently, it was established by K. Tikhomirov in [Tik18] that John's position may give suboptimal bounds.

To the positive direction, there are recent results which indicate that better concentration may be obtained for special classes of normed spaces by selecting a different position than John's; see [PV15] and [Tik18]. In particular, in [PV15] for subspaces of $L_p$, $1 \le p < \infty$ the authors showed that in Lewis' position (see [TJ89] for the related definition) the concentration is at least as good as for the $\ell_p$ norms and in view of [PVZ17] this is best possible. Thus, it follows that for every $n$, for every $1 \le p \le c \log n$ and for every $n$-dimensional subspace $X$ of $L_p$ the answer in Question 1.1 is affirmative. In [Tik18] K. Tikhomirov proved that for 1-unconditional norms in $\ell$-position (see [TJ89] for the definition) the concentration is at least as good as for the $\ell_\infty$ norm, thus establishing Question 1.1 in affirmative for those spaces. Let us mention that in all the above approaches, Gaussian functional inequalities are used rather than the classical Gaussian concentration in terms of the Lipschitz constant.

The first main step to tackle Question 1.1, is to show that the latter has an affirmative answer when the normed space $X = (\mathbb{R}^n, \|\cdot\|)$ is not *extremal* with respect to the unconditional structure. In order to formulate the result, we have to recall some standard terminology. Following the notation used in Figiel and Johnson

[FJ80], for any basis $\{b_i\}_{i=1}^n$ in a normed space $X = (\mathbb{R}^n, \|\cdot\|)$ the *unconditional basis constant* is given by

$$\mathrm{unc}\{b_i\}_{i=1}^n = \sup\left\{\left\|\sum_{i=1}^n \varepsilon_i t_i b_i\right\| : \left\|\sum_{i=1}^n t_i b_i\right\| = 1\right\},$$

where the supremum is over all choices of signs $\varepsilon_i = \pm 1$ and all scalars $(t_i)$. Next, we define

$$\mathrm{unc}X = \inf\left\{\mathrm{unc}\{b_i\}_{i=1}^n : \{b_i\}_{i=1}^n \text{ is a basis for } X\right\}.$$

For any normed space $X = (\mathbb{R}^n, \|\cdot\|)$ one has $\mathrm{unc}X \leq \sqrt{n}$. Indeed; note that $\mathrm{unc}X$ is nothing more than the Banach-Mazur distance of $X$ to the class of all $n$-dimensional 1-unconditional normed spaces and hence, $\mathrm{unc}X \leq d(X, \ell_2^n) \leq \sqrt{n}$, by John's theorem [Joh48]. At this point we should mention that this bound is optimal (up to universal constants), since there exists $n$-dimensional normed space $E$ (in fact any "typical" subspace of $\ell_\infty^n$ of proportional dimension) which satisfies $\mathrm{unc}E \geq c\sqrt{n}$ (see e.g. [FKP77] and [FJ80]). With this notation we have the following:

**Theorem 1.2.** *Let $X = (\mathbb{R}^n, \|\cdot\|)$ be a normed space. Then, there exists a $T \in GL(n)$ such that*

$$\mathbb{P}\left(\left|\|TG\| - \mathbb{E}\|TG\|\right| > t\mathbb{E}\|TG\|\right) \leq C\exp\left(-c\max\{t^2, t\}\log\left(\frac{en}{(\mathrm{unc}X)^2}\right)\right), \quad t > 0,$$

*where $G$ is the standard n-dimensional Gaussian vector.*

The proof of Theorem 1.2 uses probabilistic and topological tools. For the probabilistic part we use Talagrand's improvement upon the classical Gaussian Poincaré inequality.

**Theorem 1.3** (Talagrand's $L_1 - L_2$ bound). *For any absolutely continuous function $f : \mathbb{R}^n \to \mathbb{R}$, we have*

$$(1.6) \qquad \mathrm{Var}_{\gamma_n}(f) \leq C\sum_{i=1}^n \frac{\|\partial_i f\|_{L_2}^2}{1 + \log\left(\|\partial_i f\|_{L_2}/\|\partial_i f\|_{L_1}\right)},$$

*where $\partial_i f$ stands for the i-th partial derivative of $f$.*

Talagrand proved the above theorem for the Hamming cube [Tal94] and the Gaussian version of this inequality was presented in [CEL12]. It is known that this inequality is also related to the *super-concentration* phenomenon. Following Chatterjee, a function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be $\varepsilon_n$-super-concentrated if

$$\mathrm{Var}[f(G)] \leq \varepsilon_n \mathbb{E}\|\nabla f(G)\|_2^2.$$

The reader is referred to [Cha14] for a detailed exposition of this very interesting subject and further applications. The use of Talagrand's inequality in the study of

4

the asymptotic theory of finite-dimensional normed spaces was put forward by the authors and J. Zinn in [PVZ17]. The authors there, use the inequality to prove sharp concentration for the $\ell_p$ norms when $p$ grows along with $n$. Additionally, it is proved that the $\ell_p$ norms are super-concentrated for $p > c \log n$. This inequality also played a central role in the aforementioned work of K. Tikhomirov [Tik18] for the case of 1-unconditional norms.

As long as the choice of the position is concerned, we should mention that most canonical positions that are in use in the geometry of finite-dimensional normed spaces arise as solution of an extremum value problem, where a geometric functional is optimized subject to a constraint. For a detailed exposition of this positions the reader is referred to [GM00]. However, in our approach, we prove the existence of a position in Theorem 1.2 by employing a topological tool, namely the Borsuk-Ulam antipodal theorem [Bor33], [Mat03]. The Borsuk-Ulam theorem has already found many fruitful applications in geometric (linear and nonlinear) functional analysis, see e.g. [Lin64], [Ros83], [GM97], [KL08] just to name a few.

The unconditional constant in Theorem 1.2 is naturally involved, since inequality (1.6) is sensitive with respect to the coordinate structure. One can find good local unconditional structure based on a fundamental result of Alon and Milman [AM83]. The latter states that for every $\varepsilon \in (0,1)$ there exists a constant $C(\varepsilon) > 0$ such that every normed space $X = (\mathbb{R}^n, \|\cdot\|)$ satisfies the following dichotomy:

- Either there exists a subspace $E$ with $\dim E = m \geq e^{\sqrt{\log n}}$ and $d(E, \ell_2^m) < 1 + \varepsilon$,

- Or there exists a subspace $F$ with $\dim F = k \geq e^{\sqrt{\log n} - e^{C(\varepsilon)}}$ and $d(F, \ell_\infty^k) < 1 + \varepsilon$.

The key tool, the authors prove in [AM83], for establishing the above dichotomy is a combinatorial result for locating $\ell_\infty$-structure (see Section 4 for the precise formulation). Combining this with Theorem 1.2 we prove the following probabilistic dichotomy for the Gaussian concentration.

**Theorem 1.4.** *Let $X = (\mathbb{R}^n, \|\cdot\|)$ be a normed space whose unit ball $B_X$ is in John's position and let $0 < \delta < 1/2$. Then, we have the following dichotomy:*

- *Either the random[2] subspace $E$ with $\dim E = k \geq n^{1/2-\delta}$ satisfies*

$$\mathbb{P}\left(\left|\|G\|_{E \cap B_X} - \mathbb{E}\|G\|_{E \cap B_X}\right| > t\mathbb{E}\|G\|_{E \cap B_X}\right) \leq Ce^{-ct^2k}, \quad t > 0,$$

- *Or there exists a subspace $F$ with $\dim F = m \geq cn^{1/2}$ and an invertible linear map $T : F \to F$ such that*

$$\mathbb{P}\left(\left|\|TG\|_{F \cap B_X} - \mathbb{E}\|TG\|_{F \cap B_X}\right| > t\mathbb{E}\|TG\|_{F \cap B_X}\right) \leq Ce^{-c\delta \max\{t^2, t\} \log m}, \quad t > 0,$$

*where $G$ is the standard Gaussian vector and $c, C > 0$ are universal constants.*

---

[2]Here the randomness is considered with respect to the unique probability measure on the Grassmannian $G_{n,k}$ which is invariant under the orthogonal group action.

The above result can be interpreted as the probabilistic aspect of Alon-Milman theorem for the Gaussian measure. In turn, this is sufficient to imply an affirmative answer to Question 1.1.

**Theorem 1.5.** *Let $X = (\mathbb{R}^n, \|\cdot\|)$ be a normed space. There exists a $T \in GL(n)$ such that for all $t > 0$,*

$$\mathbb{P}\left(\left|\|TG\| - \mathbb{E}\|TG\|\right| > t\mathbb{E}\|TG\|\right) \leq C \exp(-c \max\{t^2, t\} \log n),$$

*where $G$ is the standard Gaussian vector.*

We would like to emphasize the fact that the position involved in the above theorem is not one of the standard positions used in the local theory of Banach spaces. It is rather a position derived by the combination of all the aforementioned techniques. It would be interesting to know that Theorem 1.5 holds true for some classical position, such as the position of minimal $M$ or the $\ell$-position (see Section 2 for the related definitions; see also Section 5 for the fact that this cannot be achieved in John's position).

The Gaussian concentration for norms is closely related to the local almost Euclidean structure. In his seminal work [Mil71], V. Milman establishes a random version of the celebrated result of Dvoretzky [Dvo61] on the almost spherical sections of convex bodies; see also [FLM77], [MS86]. V. Milman uses (1.2) to prove that for any $\varepsilon \in (0,1)$ the random $m$-dimensional subspace $E$ of $X$ (with respect to the Haar measure on the Grassmannian) is $(1 + \varepsilon)$-spherical, i.e.

$$(1 - \varepsilon)aB_E \subset B_X \cap E \subset (1 + \varepsilon)aB_E$$

for some appropriate constant $a > 0$ depending only on $X$, with probability greater than $1 - e^{-c\varepsilon^2 k(X)}$, as long as $m \leq c\varepsilon^2 k(X)$ (see [Gor85] and [Sch89] for the dependence $\varepsilon^2$). Thus, if we define $k_r(X, \varepsilon)$ to be the maximal $k$ for which the *random $k$-dimensional subspace* of $X$ is $(1 + \varepsilon)$-spherical with probability at least 2/3 say, Milman's argument shows that $k_r(X, \varepsilon) \geq c\varepsilon^2 k(X)$. Theorem 1.5 then, combined with a standard net argument, implies the following.

**Corollary 1.6.** *For every normed space $X = (\mathbb{R}^n, \|\cdot\|)$ there exists a position $B$ of $B_X$ such that for every $\varepsilon \in (0,1)$ one has $k_r(B, \varepsilon) \geq c\varepsilon \log n / \log(1/\varepsilon)$. That is, the random $k$-dimensional section of $B$ with $k \leq c\varepsilon \log n / \log(1/\varepsilon)$ is $(1+\varepsilon)$-spherical with probability greater than $1 - n^{-c\varepsilon}$.*

It is a question of Grothendieck [Gro53, §7] to determine the largest possible $k = k(n, \varepsilon)$ for which *every* $n$-dimensional normed space $X$ admits a $k$-dimensional subspace which is $(1 + \varepsilon)$-Euclidean. More precisely, for given $\varepsilon \in (0,1)$ and $X = (\mathbb{R}^n, \|\cdot\|)$ we denote by $k(X, \varepsilon)$ the largest $k$ so that *there exists* a $k$-dimensional subspace of $X$ which is $(1 + \varepsilon)$-Euclidean. Then, for $0 < \varepsilon < 1$ we set

$$k(n, \varepsilon) = \inf\{k(X, \varepsilon) : \dim X = n\}.$$

6

It is easy to show (see e.g. [Sch13]) that

$$k(n, \varepsilon) \leq k(\ell_\infty^n, \varepsilon) \leq C \log n / \log(1/\varepsilon),$$

while the fundamental fact that the function $k(n, \varepsilon) \to \infty$, for $\varepsilon = \varepsilon_n \to 0$ as $n \to \infty$ has first been established by Dvoretzky in [Dvo61, Theorem 1], who showed the quantitative estimate $k(n, \varepsilon) \geq c\varepsilon \sqrt{\log n} / \log \log n$. The aforementioned randomized version of Dvoretzky's theorem by V. Milman [Mil71] improved the bound to $k(n, \varepsilon) \geq c\varepsilon^2 \log n / \log \frac{1}{\varepsilon}$ and then Gordon in [Gor85] showed that $k(n, \varepsilon) \geq c\varepsilon^2 \log n$ (see also [Sch89] for an alternative proof of this estimate). Schechtman proved in [Sch06] that one can always have $k(n, \varepsilon) \geq c\varepsilon \log n / (\log \frac{1}{\varepsilon})^2$. Corollary 1.6 also gives the best known estimate for this question up-to-date, but the best possible lower estimate for the function $k(n, \varepsilon)$ remains a fundamental open problem.

Our approach shares common points with Schechtman's argument but is essentially different. In both cases the Alon-Milman theorem is a crucial tool. Schechtman invokes an iteration scheme based on James' distortion lemma [Jam64] to find further a subspace which is sufficiently close to $\ell_\infty$. This strategy is followed because $\ell_\infty$ admits finer dependence on $\varepsilon$ for the existential Dvoretzky; see [Sch13] and [Sch06]. This procedure yields a redundant logarithmic term of $\varepsilon$ compared to Corollary 1.6. To the contrary, we use Alon-Milman theorem to determine the linear map $T$ for which the norm exhibits at least as good concentration as in the $\ell_\infty$ case. Thus, we obtain the dependence on $\varepsilon$ that holds true for the random version of Dvoretzky's theorem in the case of $\ell_\infty$, see e.g. [Sch07] and [Tik14].

Concluding, we would like to point out that V. Milman in [Mil88] had observed the connection of topological tools with the problem of the dependence on $\varepsilon$ in Dvoretzky's theorem, yet it hadn't been exploited until now.

The rest of the paper is organized as follows: In Section 2 we use Talagrand's $L_1 - L_2$ bound to establish a two-level Gaussian deviation inequality for Lipschitz functions, where the Lipschitz condition is considered in both $\ell_2$ and $\ell_\infty$ sense. In Section 3 we present the proof of Theorem 1.2. In Section 4 we employ Theorem 1.2 and the Alon-Milman theorem to obtain Theorem 1.5. Finally, in Section 5 we conclude with remarks and questions that arise from our work. For background material on the geometry of Banach spaces the reader may consult the monographs [MS86, TJ89, Pis89, AAGM15].

## 2 A two-level Gaussian deviation inequality

It is well known that Poincaré inequalities imply exponential concentration for Lipschitz maps (see [Led01] and [BLM13]). Since Talagrand's $L_1 - L_2$ inequality is an improved version of the classical Poincaré inequality one gets straightforward improvements on the corresponding exponential concentration. In order to illustrate that, let us examine what is the corresponding deviation estimate we obtain by Talagrand's inequality, if we employ the standard method of bounding the variance of the moment generating function of a Lipschitz map $f$. To this end, we introduce the

following notation: For any Lipschitz map $f : \mathbb{R}^n \to \mathbb{R}$ let

$$b = b(f) := \inf\{t > 0 : |f(x) - f(y)| \leq t\|x - y\|_2, \ \forall \, x, y \in \mathbb{R}^n\}.$$

and similarly

$$a = a(f) := \inf\{t > 0 : |f(x) - f(y)| \leq t\|x - y\|_\infty, \ \forall \, x, y \in \mathbb{R}^n\}.$$

Note that in the light of $\|\cdot\|_\infty \leq \|\cdot\|_2 \leq \sqrt{n}\|\cdot\|_\infty$ one has

(2.1) $$b \leq a \leq b\sqrt{n}.$$

If $\|\cdot\|$ is an arbitrary norm on $\mathbb{R}^n$ and $|f(x) - f(y)| \leq L\|x - y\|$ for all $x, y \in \mathbb{R}^n$, then

(2.2) $$\|\nabla f(x)\|_* \leq L,$$

where $\|\cdot\|_*$ is the dual norm of $\|\cdot\|$, i.e.

$$\|y\|_* = \sup\{\langle x, y\rangle : \|x\| \leq 1\}, \quad y \in \mathbb{R}^n,$$

and the gradient of $f$ is defined almost everywhere by Rademacher's theorem, see e.g. [EG92].

Now we come to the aforementioned improvement of the exponential concentration via the $L_1 - L_2$ bound. In order to simplify considerably the computations let us assume that $f$ has some symmetries, i.e. $f$ is permutation invariant[3]. In that case we have that $h = e^{\lambda f}, \lambda > 0$ is also permutation invariant and

$$\partial_i h = \partial_i (h \circ P_\pi) = \langle P_\pi^* \circ \nabla h \circ P_\pi, e_i\rangle = (\partial_{\pi(i)} h) \circ P_\pi, \quad i = 1, 2, \ldots, n,$$

for any permutation $\pi$. It follows that

$$\lambda^p \mathbb{E} e^{p\lambda f} |\partial_i f|^p = \|\partial_i h\|_{L_p}^p = \|\partial_{\pi(i)} h\|_{L_p}^p,$$

for all $i \leq n$, for any permutation $\pi$ and for $p > 0$, since $P_\pi$ is orthogonal. In particular, the $L_2$-norm of all partial derivatives of $h$ are equal, thus

$$\|\partial_i h\|_{L_2}^2 = \frac{1}{n}\sum_{i=1}^n \|\partial_i h\|_{L_2}^2 = \frac{\lambda^2}{n}\mathbb{E} e^{2\lambda f}\|\nabla f\|_2^2 \leq \frac{\lambda^2 b(f)^2}{n}\mathbb{E} e^{2\lambda f}, \quad i = 1, \ldots, n.$$

Arguing similarly, we get

$$\|\partial_i h\|_{L_1} \leq \frac{\lambda a(f)}{n}\mathbb{E} e^{\lambda f}, \quad i = 1, \ldots, n.$$

Applying Theorem 1.3 for $h$ and taking into account the previous estimates we obtain

$$\mathbb{E} e^{2\lambda f} - (\mathbb{E} e^{\lambda f})^2 = \mathrm{Var}(e^{\lambda f}) \leq \frac{C\lambda^2 b^2}{1 + \log(nb^2/a^2)}\mathbb{E} e^{2\lambda f}, \quad \lambda > 0,$$

---

[3] For any permutation $\pi : [n] \to [n]$ we define the permutation matrix $P_\pi$ associated with $\pi$ as follows: $P_\pi(e_i) = e_{\pi(i)}$. Note that $P_\pi \circ P_\sigma = P_{\pi\sigma}$ and $P_\pi^{-1} = P_{\pi^{-1}} = P_\pi^*$ for all permutations $\pi, \sigma$. A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be permutation invariant if $f \circ P_\pi = f$ for any permutation $\pi$.

where $a = a(f)$ and $b = b(f)$. Next, we argue as in [BLM13, p.70]. Set $\rho^2 = \frac{Cb^2}{1+\log(nb^2/a^2)}$ and apply the previous estimate for $F = f - \mathbb{E}f$ and $\lambda = s/\rho$. Then, for $\psi(s) = \mathbb{E}e^{2sF/\rho}$, we obtain the following conditions:

$$\lim_{s \to 0} \frac{\psi(s) - 1}{s} = 0, \quad (1 - s^2)\psi(s) \leq (\psi(s/2))^2, \quad 0 < s < 1.$$

It is a calculus exercise to show that such a function satisfies $\psi(s) \leq (1 - s^2)^{-2}$ for all $0 < s < 1$. In particular, $\psi(1/\sqrt{2}) \leq 4$. Using Markov's inequality and combining with the classical Gaussian concentration (1.1) we conclude the following.

**Proposition 2.1.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a permutation invariant function. If $|f(x) - f(y)| \leq b\|x - y\|_2$ and $|f(x) - f(y)| \leq a\|x - y\|_\infty$ for all $x, y \in \mathbb{R}^n$, then*

$$(2.3) \qquad \mathbb{P}(|f(G) - \mathbb{E}[f(G)]| > t) \leq 4\exp\left(-c\max\left\{\frac{t^2}{b^2}, \frac{t}{b}\sqrt{\log(nb^2/a^2)}\right\}\right),$$

*for all $t > 0$, where $G$ is the standard n-dimensional Gaussian vector and $c, C > 0$ are universal constants.*

The purpose of this section is to show that a similar concentration inequality can be proved regardless the symmetries of $f$. To this end, we will need the following consequence of Talagrand's $L_1 - L_2$ inequality (for a proof see e.g. [Cha14, Theorem 5.4]):

**Lemma 2.2.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be an absolutely continuous function and let*

$$R(f) = \frac{\mathbb{E}\|\nabla f(G)\|_2^2}{\sum_{i=1}^n (\mathbb{E}|\partial_i f(G)|)^2}.$$

*Then, we have*

$$(2.4) \qquad \mathrm{Var}[f(G)] \leq C \frac{\mathbb{E}\|\nabla f(G)\|_2^2}{1 + \log R(f)},$$

*where $G$ is the standard Gaussian vector on $\mathbb{R}^n$.*

Now we are ready to prove the aforementioned two-level deviation inequality. This inequality is in the spirit of Talagrand's two-level deviation inequality for the exponential distribution [Tal91] (see also [BL97] for an alternative proof).

**Proposition 2.3.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a Lipschitz map with*

$$|f(x) - f(y)| \leq b\|x - y\|_2, \quad |f(x) - f(y)| \leq a\|x - y\|_\infty, \quad x, y \in \mathbb{R}^n$$

*and $\|\partial_i f\|_{L_1} \leq A$ for all $i \leq n$. Then, if we set $F = f - \mathbb{E}f$, for all $\lambda > 0$ we have*

$$(2.5) \qquad \mathrm{Var}(e^{\lambda F}) \leq \frac{C\lambda^2 b^2}{\log(e + \frac{b^2}{aA})}\mathbb{E}e^{2\lambda F}.$$

9

*Moreover, we obtain*

$$(2.6) \qquad \mathbb{P}\left(|f(G) - \mathbb{E}[f(G)]| > t\right) \le 4 \exp\left(-c \max\left\{\frac{t^2}{b^2}, \frac{t}{b}\sqrt{\log\left(e + \frac{b^2}{aA}\right)}\right\}\right), \quad t > 0,$$

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

*Proof.* We fix $\lambda > 0$ and we apply Lemma 2.2 for the function $e^{\lambda F}$. Then, we obtain

$$\mathrm{Var}(e^{\lambda F}) \le C\lambda^2 \frac{\mathbb{E}(\|\nabla f\|_2^2 e^{2\lambda F})}{1 + \log\left(\frac{\mathbb{E}\|\nabla f\|_2^2 e^{2\lambda F}}{w}\right)},$$

where $w = \sum_{i=1}^n \left(\mathbb{E}|\partial_i f|e^{\lambda F}\right)^2$. Note that $\mathbb{E}\|\nabla f\|_2^2 e^{2\lambda F} \le b^2 \mathbb{E}e^{2\lambda F}$ and the function $z \mapsto \frac{z}{1+\log(z/w)}$ is non-decreasing for $z \ge w$, hence we get

$$\mathrm{Var}(e^{\lambda F}) \le \frac{C\lambda^2 b^2}{1 + \log\left(w^{-1} b^2 \mathbb{E}e^{2\lambda F}\right)}\mathbb{E}e^{2\lambda F}.$$

Finally, note that

$$w = \sum_{i=1}^n \left(\mathbb{E}|\partial_i f|e^{\lambda F}\right)^2 \le \sum_{i=1}^n \mathbb{E}e^{\lambda F}\mathbb{E}|\partial_i f|^2 e^{\lambda F} \le A\mathbb{E}e^{\lambda F}\mathbb{E}\|\nabla f\|_1 e^{\lambda F} \le aA(\mathbb{E}e^{\lambda F})^2,$$

where we have used the Cauchy-Schwarz inequality, the bound on the $L_1$ norm of the partial derivatives, and the pointwise bound $\|\nabla f(x)\|_1 \le a$. This completes the proof of the first assertion. The concentration estimate (2.6) can be proved in a standard fashion as before, by using (2.5). See e.g. [BLM13, p.70] for the details. $\square$

**Remark 2.4.** The logarithm appearing on the estimate (2.6) is almost the same as in (2.3) without the parameter $A$. Note that in general the least possible $A$ satisfies $A \le a$. In the next section we show that after composing the function with a suitable diagonal matrix we may bound $A \le a/n$ and hence derive exactly the estimate (2.3).

Next, we present an application of the previous distributional inequality in the context of 1-unconditional norms. To this end let us recall the definition of the *position of minimal M*: A norm $\|\cdot\|$ on $\mathbb{R}^n$ is said to be in position of minimal $M$ if for every $T \in SL(n)$ we have

$$\mathbb{E}\|G\| \le \mathbb{E}\|TG\|, \quad G \sim N(\mathbf{0}, I_n).$$

In this case the norm satisfies the following *isotropic condition*:

$$(2.7) \qquad \int_{\mathbb{R}^n} \langle\nabla\|x\|, \theta\rangle\langle x, \theta\rangle \, d\gamma_n(x) = \frac{\mathbb{E}\|G\|}{n}, \quad \theta \in S^{n-1}.$$

For more properties of this position the reader is referred to [GM00].

**Proposition 2.5.** *Let $\| \cdot \|$ be a 1-unconditional norm on $\mathbb{R}^n$ which is in position of minimal $M$. Then, we have the following distributional inequalities:*

(i) *For all $t > 0$,*

$$(2.8) \qquad \mathbb{P}\left( \big| \|G\| - \mathbb{E}\|G\| \big| > t\mathbb{E}\|G\| \right) \leq C \exp\left( -c \max\left\{ t^2 k, t\sqrt{k \log(en/k)} \right\} \right),$$

*where $k = k(X)$ and $X = (\mathbb{R}^n, \| \cdot \|)$.*

(ii) *In particular $k \geq c \log n$, hence*

$$(2.9) \qquad \mathbb{P}\left( \big| \|G\| - \mathbb{E}\|G\| \big| > t\mathbb{E}\|G\| \right) \leq C \exp\left( -c \max\{t^2, t\} \log n \right), \quad t > 0,$$

*where $G$ is the standard $n$-dimensional Gaussian vector and $C, c > 0$ are universal constants.*

*Proof.* (i). Set $f(x) = \|x\|$. The unconditionality and the convexity of $f$ implies that $x_i \mapsto \partial_i f(x)$ is nondecreasing function of $|x_i|$, Hence, Chebyshev's association inequality (see e.g. [BLM13, Section 2.10]) yields

$$\mathbb{E}|\partial_i f(G)| \cdot \mathbb{E}|g_i| \leq \mathbb{E}|g_i \partial_i f(G)| = \frac{\mathbb{E}f(G)}{n} \implies \|\partial_i f\|_{L_1(\gamma_n)} \leq \frac{c_1 \mathbb{E}f(G)}{n}, \quad i = 1, \ldots, n.$$

Furthermore we have the following:

*Claim.* Note that $a(f) = \max\{\|x\| : \|x\|_\infty \leq 1\}$ and $\mathbb{E}\|G\| \geq ca(f)$.

*Proof of Claim.* Indeed; we may write

$$\mathbb{E}\left\| \sum_{i=1}^n g_i e_i \right\| = \mathbb{E}_\varepsilon \mathbb{E}\left\| \sum_{i=1}^n \varepsilon_i |g_i| e_i \right\| \geq \mathbb{E}_\varepsilon \left\| \sum_{i=1}^n \varepsilon_i \mathbb{E}|g_i| e_i \right\| = \sqrt{\frac{2}{\pi}} \left\| \sum_{i=1}^n e_i \right\|,$$

by Jensen's inequality and the unconditionality of the norm. On the other hand we have

$$\|i : \ell_\infty^m \to X\| = \max_{\|x\|_\infty \leq 1} \|x\| = \max_{\varepsilon_i = \pm 1} \left\| \sum_{i=1}^n \varepsilon_i e_i \right\| = \left\| \sum_{i=1}^n e_i \right\|,$$

which proves the assertion. $\qquad \square$

Thus, a straightforward application of Proposition 2.3 combined with the above estimates yields (2.8).

(ii). From (2.8) applied for $t \simeq 1$ and compared with (1.3) we get $k \geq c \log(en/k)$ which yields the desired estimate. $\qquad \square$

**Note 2.6.** In [Tik18] K. Tikhomirov proves (2.9) using the $\ell$-position. The latter is a variant of the position we use here and satisfies an analogous isotropic condition, namely

$$(2.10) \qquad \int_{\mathbb{R}^n} \langle \nabla \|x\|, \theta \rangle \langle x, \theta \rangle \|x\| \, d\gamma_n(x) = \frac{\mathbb{E}\|G\|^2}{n}, \quad \theta \in S^{n-1}.$$

11

Using (2.10) instead, we can again show that $n\|\partial_i f\|_{L_1} \leq c\mathbb{E}f(G)$ for all $i \leq n$, where $f(x) = \|x\|$ is 1-unconditional. Note that, according to the argument of Proposition 2.5.(i), it suffices to obtain an upper bound of the form $\mathbb{E}|g_i \partial_i f(G)| \leq C\mathbb{E}f(G)/n$ for each $i \leq n$. We will show that indeed this follows from (2.10) and the unconditionality of $f$. To this end recall the known fact (see e.g. [Tik18, Proposition 16] for a proof) that $k(X) \geq c_1 \log n$ for any normed space $X = (\mathbb{R}^n, \|\cdot\|)$ whose unit ball $B_X$ is in $\ell$-position. Using the small ball probability estimate for norms (see e.g. [LO05], [KV07] and [PV18] for a refinement) we have

$$A := \{\|G\| \leq c_0 \mathbb{E}\|G\|\}, \quad \mathbb{P}(A) \leq e^{-c_2 k(X)} \leq 1/n^2.$$

Thus, using the unconditionality, the fact that

$$|\partial_i f(x)| \leq b \leq (\mathbb{E}[f(G)]^2)^{1/2} \leq C\mathbb{E}f(G) \quad \text{a.e.},$$

and the Cauchy-Schwarz inequality we may write

$$
\begin{aligned}
\mathbb{E}|g_i \partial_i f(G)| &= \mathbb{E}|g_i \partial_i f(G)| \mathbf{1}_{A^c} + \mathbb{E}|g_i \partial_i f(G)| \mathbf{1}_A \\
&\leq \frac{1}{c_0 \mathbb{E}f(G)} \mathbb{E}[|g_i \partial_i f(G)| f(G)] + b\sqrt{\mathbb{P}(A)} \\
&\leq \frac{1}{c_0 \mathbb{E}f(G)} \frac{\mathbb{E}f^2(G)}{n} + \frac{b}{n} \\
&\leq \frac{C_0 \mathbb{E}f(G)}{n},
\end{aligned}
$$

as required.

# 3 Concentration for norms with moderate unconditional structure

In this section we study Question 1.1 and we prove that it has an affirmative answer for normed spaces which do not have extremal unconditional basis constant, by establishing Theorem 1.2. In fact our argument takes into account a slightly weaker notion, that of the *left random unconditional constant* (see Theorem 3.4 for the details). The approach we present uses Proposition 2.3 and the Borsuk-Ulam theorem [Bor33] (see also [Mat03]). We start with the following:

**Lemma 3.1** (Balancing the partial derivatives). *Let* $f : \mathbb{R}^m \to \mathbb{R}$ *be a* $C^1$-*smooth function with bounded partial derivatives and* $q > 0$. *Then there exists a diagonal matrix* $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_m)$ *with*

(a) $\|\Lambda\|_{\mathrm{HS}} = 1$ *and*

(b) $\|\partial_i(f \circ \Lambda)\|_{L_q(\gamma_m)} = \|\partial_j(f \circ \Lambda)\|_{L_q(\gamma_m)}$ *for* $i, j = 1, \ldots, m$.

12

*Proof.* For each $1 \leq j < m$ consider the functions $h_j : S^{m-1} \to \mathbb{R}$ defined by

$$h_j(\lambda) := \lambda_j \|(\partial_j f) \circ \Lambda\|_{L_q} - \lambda_{j+1}\|(\partial_{j+1}f) \circ \Lambda\|_{L_q},$$

where $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_m)$. The dominated convergence theorem and the continuity of $\partial_j f$ imply the continuity of $h_j$, while the symmetry of $\gamma_m$ implies that $h_j$ is odd, that is $h_j(-\lambda) = -h_j(\lambda)$ for all $\lambda \in S^{m-1}$. Hence, if we consider the mapping $H : S^{m-1} \to \mathbb{R}^{m-1}$ defined by

$$H(\lambda_1, \ldots, \lambda_m) := (h_1(\lambda), \ldots, h_{m-1}(\lambda)),$$

we readily see that it is continuous and odd. Therefore, by the Borsuk-Ulam antipodal theorem [Mat03] we obtain $\lambda \in S^{m-1}$ such that $H(\lambda) = 0$, that is

(3.1) $$\lambda_i \|(\partial_i f) \circ \Lambda\|_{L_q} = \lambda_j \|(\partial_j f) \circ \Lambda\|_{L_q}, \quad i, j = 1, 2, \ldots, m.$$

In particular $\|\partial_i(f \circ \Lambda)\|_{L_q} = \|\partial_j(f \circ \Lambda)\|_{L_q}$ for all $i, j$ which proves the assertion. $\square$

**Remarks 3.2.** 1. Note that if $f$ is not constant in any proper subspace, then $\lambda_i > 0$ for all $i$. Indeed; note that the set $\sigma := \{i : \lambda_i \neq 0\}$ is non empty. Assuming that $\sigma^c \neq \emptyset$, by (3.1) we get $\|(\partial_i f) \circ \Lambda\|_{L_q} = 0$ for all $i \in \sigma$. Note that $(\partial_i f) \circ \Lambda \equiv 0$ for all $i \in \sigma$, by the continuity. It follows that $\partial_i f \equiv \mathbf{0}$ on $\Lambda(\mathbb{R}^m) = \mathbb{R}^\sigma \equiv [e_i : i \in \sigma]$ for all $i \in \sigma$, i.e. $f|_{\mathbb{R}^\sigma}$ is constant. Moreover, (3.1) implies that all $\lambda_j$ have the same sign. Since $H(\lambda) = H(-\lambda) = 0$ we may assume that $\lambda_j \geq 0$ for all $j$.

2. Note that the proof of Lemma 3.1 can also be applied on the boundary $\partial K$ of any symmetric convex body $K$ in $\mathbb{R}^m$, thus we may also have $\|\lambda\|_K = 1$ instead of $\|\lambda\|_2 = 1$. However, this is not crucial for our purposes.

**Lemma 3.3.** *Let $f : \mathbb{R}^m \to \mathbb{R}$ be a $C^1$-smooth Lipschitz map which is not a constant in any proper subspace. Then, there exist $\lambda_1, \ldots, \lambda_m > 0$ such that $\sum_{j=1}^m \lambda_j^2 = 1$ and*

$$\|\partial_j(f \circ \Lambda)\|_{L_1(\gamma_m)} \leq \frac{1}{m} a(f \circ \Lambda), \quad j = 1, 2, \ldots, m,$$

*where $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_m)$.*

*Proof.* Since $f$ is $C^1$-smooth we may consider the diagonal matrix $\Lambda$ from Lemma 3.1 and by taking into Remark 3.2.1 we also have $\lambda_j > 0$ for all $j$. Note that,

$$\|\partial_j(f \circ \Lambda)\|_{L_1(\gamma_m)} = \frac{1}{m} \sum_{j=1}^m \|\partial_j(f \circ \Lambda)\|_{L_1(\gamma_m)} = \frac{1}{m} \int_{\mathbb{R}^m} \|\nabla(f \circ \Lambda)\|_1 \, d\gamma_m \leq \frac{a(f \circ \Lambda)}{m},$$

for all $j \leq m$, as required. $\square$

Now we are ready to prove the main result of this section. To this end, let us recall a variant of a one-sided unconditional constant. The *random unconditional*

*divergence constant* of a normed space $X = (\mathbb{R}^m, \|\cdot\|)$, denoted by $\mathrm{rud}(X)$, is the least $L > 0$ for which there exists a basis $(x_i)_{i=1}^m$ in $X$ such that

$$(3.2) \qquad \left\| \sum_{i=1}^m \alpha_i x_i \right\| \leq L \mathbb{E}_\varepsilon \left\| \sum_{i=1}^m \varepsilon_i \alpha_i x_i \right\|,$$

for all scalars $(\alpha_i)_{i=1}^m$. Note that $\mathrm{rud}(X) \leq \mathrm{unc}(X)$ (see [FJ80, Section 6] and [LAT16] for further details). With this terminology we have the following.

**Theorem 3.4.** *Let $X = (\mathbb{R}^m, \|\cdot\|)$ be a normed space and let $L = \mathrm{rud}(X)$. Then, there exists $T \in GL(m)$ such that*

$$\mathbb{P}\left( \big| \|TG\| - \mathbb{E}\|TG\| \big| > \varepsilon \mathbb{E}\|TG\| \right) \leq C \exp\left( -c \max\{\varepsilon, \varepsilon^2\} \log\left( e + \frac{m}{L^2} \right) \right),$$

*for all $\varepsilon > 0$, where $G$ is the standard $m$-dimensional Gaussian vector and $c, C > 0$ are universal constants.*

*Proof.* After applying an invertible linear transformation we may assume that

$$\|y\| \leq L \mathbb{E}_\varepsilon \left\| \sum_{i=1}^m \varepsilon_i y_i e_i \right\|, \quad y \in \mathbb{R}^m,$$

where $(e_i)_{i \leq m}$ is the standard basis on $\mathbb{R}^m$. Equivalently, we have

$$(3.3) \qquad \sup_{\varepsilon_i = \pm 1} \left\| \sum_{i=1}^m \varepsilon_i y_i e_i \right\| \leq L \mathbb{E}_\varepsilon \left\| \sum_{i=1}^m \varepsilon_i y_i e_i \right\|, \quad y \in \mathbb{R}^m.$$

First we consider the case that the given norm $\|\cdot\|$ is smooth. Let $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_m)$ be the diagonal matrix from Lemma 3.3. We set

$$a_\Lambda := \|\Lambda : \ell_\infty^m \to X\|, \quad b_\Lambda := \|\Lambda : \ell_2^m \to X\|, \quad k_\Lambda := \frac{(\mathbb{E}\|\Lambda G\|)^2}{b_\Lambda^2}.$$

Then, by the distributional inequality (2.6) in conjunction with Lemma 3.3 we obtain

$$(3.4) \qquad \mathbb{P}\left( \big| \|\Lambda G\| - \mathbb{E}\|\Lambda G\| \big| > \varepsilon \mathbb{E}\|\Lambda G\| \right) \leq 4 \exp\left( -c\varepsilon \sqrt{k_\Lambda \log\left( e + \frac{m b_\Lambda^2}{a_\Lambda^2} \right)} \right), \quad \varepsilon > 0.$$

Employing (3.3) we arrive at the following estimate:

$$a_\Lambda = \max_{\varepsilon_i = \pm 1} \left\| \sum_{i=1}^m \varepsilon_i \lambda_i e_i \right\| \leq L \mathbb{E}_\varepsilon \left\| \sum_{i=1}^m \varepsilon_i \lambda_i e_i \right\| \leq L \sqrt{\frac{\pi}{2}} \mathbb{E}\|\Lambda G\|,$$

where in the last passage we have used the contraction principle (see e.g. [LT91, Chapter 4] or [Pis74, Proposition 1]). Plugging the above estimate into (3.4) we obtain

$$\mathbb{P}\left( \big| \|\Lambda G\| - \mathbb{E}\|\Lambda G\| \big| > \varepsilon \mathbb{E}\|\Lambda G\| \right) \leq 4 \exp\left( -c\varepsilon \sqrt{k_\Lambda \log\left( e + \frac{m}{L^2 k_\Lambda} \right)} \right), \quad \varepsilon > 0.$$

Applying the latter for $\varepsilon \simeq 1$ and taking into account (1.3) we readily see that $k_\Lambda \geq c \log(e + m/L^2)$. This proves the result in the smooth case.

For the general case, recall that for the given norm $\|\cdot\|$ and for any $\delta \in (0,1)$ there exists a smooth norm $\|\cdot\|_\delta$ such that

$$(3.5) \qquad\qquad (1-\delta)\|x\| \leq \|x\|_\delta \leq (1+\delta)\|x\|,$$

for all $x \in \mathbb{R}^m$, see e.g. [Sch14]. We fix $0 < \delta \leq (7 + \log m)^{-1}$ and we apply the result for $\|\cdot\|_\delta$, thus we get $T = T_\delta \in GL(m)$ such that

$$\mathbb{P}\left( \left| \|TG\|_\delta - \mathbb{E}\|TG\|_\delta \right| > \varepsilon \mathbb{E}\|TG\|_\delta \right) \leq 4 e^{-c\varepsilon \log(e + m/L_\delta^2)},$$

for all $\varepsilon > 0$, where $L_\delta = \mathrm{rud}(X_\delta)$ for $X_\delta = (\mathbb{R}^m, \|\cdot\|_\delta)$. One may check, using (3.5), that for $\varepsilon > 8\delta$ we have

$$\mathbb{P}\left( \left| \|TG\| - \mathbb{E}\|TG\| \right| > \varepsilon \mathbb{E}\|TG\| \right) \leq \mathbb{P}\left( \left| \|TG\|_\delta - \mathbb{E}\|TG\|_\delta \right| > \frac{\varepsilon}{2} \mathbb{E}\|TG\|_\delta \right).$$

Thus, we obtain

$$\mathbb{P}\left( \left| \|TG\| - \mathbb{E}\|TG\| \right| > \varepsilon \mathbb{E}\|TG\| \right) \leq 4 e^{-\frac{c}{2}\varepsilon \log(e + m/L_\delta^2)},$$

for all $\varepsilon > 8\delta$. By adjusting the universal constants the previous estimate holds true for all $\varepsilon > 0$. On the other hand we may easily check that

$$L_\delta \leq \frac{1+\delta}{1-\delta} L \leq 2L.$$

The proof is complete. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

**Remarks 3.5.** 1. It is somewhat unexpected that the almost optimal concentration is established with an isomorphic parameter. This suggests that Question 1.1 seems plausible to have an affirmative answer and the above partial result consists of a deficiency of the approach, which uses the $L_1 - L_2$ estimate for the quantification of the problem. In fact, in the next paragraph we show how one can eliminate $\mathrm{rud}(X)$ and overcome this obstacle.

2. The *random unconditional convergence constant* $\mathrm{ruc}(X)$ is defined similarly as the least $R > 0$ for which there exists a basis $(x_i)$ of $X = (\mathbb{R}^n, \|\cdot\|)$ such that

$$\mathbb{E}\left\| \sum_{i=1}^n \varepsilon_i \alpha_i x_i \right\| \leq R \left\| \sum_{i=1}^n \alpha_i x_i \right\|,$$

for all scalars $(\alpha_i)_{i \leq n}$. While RUC bases have been previously studied (see e.g. [BKPS86], [LAT16]) and their extremal asymptotic behavior was established in [Bal91], it was not until recently that its left analogue, i.e. $\mathrm{rud}(X)$, was put forward in systematic study; see e.g. [LAT16]. Consulting specialists we couldn't locate a precise reference for the study of the latter in the context of high-dimensional normed spaces. These notions and more will be part of a detailed study which will appear elsewhere.

# 4  Probabilistic dichotomy and Dvoretzky's theorem

In this Section we prove Theorem 1.4 and its corollaries mentioned in the Introduction. Our first main ingredient is the classical Dvoretzky-Rogers lemma from [DR50]. Another crucial tool in our approach is the Alon-Milman theorem. The idea to use this dichotomy in this problem can be traced back to the work of Schechtman [Sch06].

**Lemma 4.1** (Dvoretzky-Rogers). *Let $X = (\mathbb{R}^n, \|\cdot\|)$ be a normed space for which $B_X$ is in John's position. Then, there exists an orthonormal basis $v_1, \ldots, v_n$ such that*

$$1 = \|v_k\|_2 \geq \|v_k\| \geq \sqrt{1 - \frac{k-1}{n}},$$

*for $k = 1, 2, \ldots, n$. In particular, $\|v_j\| \geq 1/\sqrt{2}$ for $j = 1, \ldots, \lfloor n/2 \rfloor$.*

**Remark 4.2** (W.B. Johnson). Starting with the above orthonormal basis one may redefine the vectors to get a new orthonormal basis $(w_i)$ with $\|w_i\| \geq 1/4$ for all $i \leq n$. This remark is due to Bill Johnson [Joh13]. We would like to thank him for allowing us to include his elegant argument here. Assume for simplicity that $n = 2s$. For each $i = 1, \ldots, s$, if $\|u_{s+i}\| \geq 1/4$ we set $w_{s-i+1} = u_{s-i+1}$ and $w_{s+i} = u_{s+i}$, while if $\|u_{s+i}\| < 1/4$ we replace $w_{s-i+1} = \frac{u_{s-i+1} + u_{s+i}}{\sqrt{2}}$ and $w_{s+i} = \frac{u_{s-i+1} - u_{s+i}}{\sqrt{2}}$. Note that the vectors $(w_i)_{i \leq n}$ are still orthonormal and $\|w_{s-i+1}\|, \|w_{s+i}\| \geq 1/4$. Indeed; by construction we have

$$\min\{\|w_{s-i+1}\|, \|w_{s+i}\|\} \geq \frac{1}{\sqrt{2}}(\|u_{s-i+1}\| - \|u_{s+i}\|) > \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}} - \frac{1}{4}\right) > \frac{1}{4},$$

where we have used that $\|u_{s-i+1}\| \geq 1/\sqrt{2}$ for all $i \leq s$ by Lemma 4.1. $\qquad\square$

We will also need the following theorem of Alon and Milman [AM83]; see also [Tal95] for an alternative simpler proof.

**Theorem 4.3** (Alon-Milman, Talagrand). *Let $X$ be a normed space and let $T : \ell_\infty^n \to X$. We set*

$$a = \|T : \ell_\infty^n \to X\| \quad and \quad M_n = \mathbb{E}_\varepsilon \left\| \sum_{i=1}^n \varepsilon_i T e_i \right\|.$$

*Assuming that $\|Te_i\| \geq 1$ for all $i$, there exists $\sigma \subset [n]$ with $|\sigma| \geq cn/a$ such that*

$$\frac{1}{2} \max_{i \in \sigma} |\alpha_i| \leq \left\| \sum_{i \in \sigma} \alpha_i T e_i \right\| \leq 4M_n \max_{i \in \sigma} |\alpha_i|,$$

*for all scalars $(\alpha_i) \subset \mathbb{R}$.*

Alon-Milman's proof yields $\sigma \subset [n]$ with $|\sigma| \geq cn^{1/2}/M_n$. The improved estimate stated above is due to Talagrand. We are now ready to prove the key result of this section. Note that Theorem 1.4 will follow from the next result.

**Theorem 4.4.** *Let $X = (\mathbb{R}^n, \|\cdot\|)$ be a normed space for which $B_X$ is in John's position and let $0 < \delta < 1/2$. Then, at least one of the following conditions holds:*

- *Either $k(X) \geq n^{1/2-\delta}$,*

- *Or there exists a subspace $F$ with $\dim F = m \geq cn^{1/2}$ and a linear isomorphism $T : F \to F$ such that for all $t > 0$,*

$$\mathbb{P}\left( \left| \|TZ\| - \mathbb{E}\|TZ\| \right| > t\mathbb{E}\|TZ\| \right) \leq Ce^{-c\delta \max\{t^2, t\}\log m}, \quad Z \sim N(\mathbf{0}, I_F),$$

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

*Proof.* Fix $0 < \delta < 1/2$. Let $(w_j)$ be an orthonormal basis with $\|w_j\| \geq 1/4$ for all $j$ and let $k(X) \leq n^{1/2-\delta}$. Then, we may write

$$\sqrt{\frac{2}{\pi}}\mathbb{E}\left\| \sum_{i=1}^{n} \varepsilon_i w_i \right\| \leq \mathbb{E}\left\| \sum_{i=1}^{n} g_i w_i \right\| = \mathbb{E}\|G\| = \sqrt{k(X)},$$

where in the first inequality we have used the contraction principle, see [LT91, Chapter 4] or [Pis74, Proposition 1]. Using Theorem 4.3 we obtain a subset $\sigma \subset [n]$ with $|\sigma| \geq cn/a \geq c\sqrt{n}$, where $a = \|i : \ell_\infty^n \to X\|$ and

$$\frac{1}{8}\max_{i \in \sigma}|\alpha_i| \leq \left\| \sum_{i \in \sigma} \alpha_i w_i \right\| \leq 4M_n \max_{i \in \sigma}|\alpha_i|,$$

for all $(\alpha_i)_{i \in \sigma} \subset \mathbb{R}$, where $M_n = \mathbb{E}\|\sum_i \varepsilon_i w_i\|$. Note that the subpace $(F, \|\cdot\|)$ with $F = \text{span}\{w_i : i \in \sigma\}$ satisfies $d(F, \ell_\infty^\sigma) \leq 32M_n \leq C\sqrt{k(X)}$. Thus, by Theorem 3.4 there exists a linear isomorphism $T : F \to F$ such that

$$\mathbb{P}\left( \left| \|TZ\| - \mathbb{E}\|TZ\| \right| > t\mathbb{E}\|TZ\| \right) \leq Ce^{-c\max\{t^2, t\}\log(e+c|\sigma|/k(X))}$$

$$\leq Ce^{-c'\delta \max\{t^2, t\}\log|\sigma|},$$

for all $t > 0$. The proof is complete. $\qquad\qquad\qquad\qquad\qquad\qquad\square$

In the following theorem, which is immediate consequence of Theorem 4.4, we summarize Theorem 1.5 and Corollary 1.6. We would like to thank K. Tikhomirov who kindly pointed out to us that the linear map in Theorem 4.4 can be lifted up to an invertible linear map, whence obtaining an affirmative answer to Question 1.1 as it is stated.

**Theorem 4.5.** *Let $X = (\mathbb{R}^n, \|\cdot\|)$ be a normed space. Then, there exists a linear transformation $S \in GL(n)$ such that for all $t > 0$,*

$$\mathbb{P}\left( \left| \|SG\| - \mathbb{E}\|SG\| \right| > t\mathbb{E}\|SG\| \right) \leq Ce^{-c\max\{t^2, t\}\log n}, \quad G \sim N(\mathbf{0}, I_n).$$

*In particular, for any $\varepsilon \in (0, 1)$ the random $k$-dimensional section of $S^{-1}B_X$ with $k \leq c\varepsilon \log n / \log(1/\varepsilon)$ is $(1 + \varepsilon)$-spherical with probability greater than $1 - n^{-c\varepsilon}$.*

17

*Proof.* We may assume that $B_X$ is in John's position and $k(X) \le n^{1/3}$, otherwise there is nothing to prove. Then, Theorem 4.4 yields the existence of a subspace $F$ with $\dim F = m \simeq \sqrt{n}$ and $T : F \to F$ with $T \in GL(F)$ such that for every $\varepsilon > 0$,

$$\mathbb{P}\left( \big| \|TZ\| - \mathbb{E}\|TZ\| \big| > \varepsilon \mathbb{E}\|TZ\| \right) \le C \exp\left(-c\varepsilon \log n\right), \quad Z \sim N(\mathbf{0}, I_F).$$

Let $S : \mathbb{R}^n \to \mathbb{R}^n$ be the operator defined by

$$S(x, y) = Tx + \lambda y, \quad x \in F, \ y \in F^\perp, \quad \lambda = \frac{\mathbb{E}\|TZ\|}{\mathbb{E}\|W\|_2 \log n}, \quad W \sim N(\mathbf{0}, I_{F^\perp}).$$

Note that by the contraction principle, the triangle inequality and the fact that $\|\cdot\| \le \|\cdot\|_2$ we have

$$(4.1) \qquad \mathbb{E}\|TZ\| \le \mathbb{E}\|SG\| \le \left(1 + \frac{1}{\log n}\right) \mathbb{E}\|TZ\|, \quad G \sim N(\mathbf{0}, I_n), \ Z \sim N(\mathbf{0}, I_F).$$

Next, we have the following:

*Claim.* For all $\varepsilon > 0$ we have

$$(4.2) \qquad \mathbb{P}\left( \big| \|SG\| - \mathbb{E}\|SG\| \big| > \varepsilon \mathbb{E}\|SG\| \right) \le C e^{-c\varepsilon \log n}.$$

*Proof of Claim.* Let $Z \sim N(\mathbf{0}, I_F)$, $W \sim N(\mathbf{0}, I_{F^\perp})$ and $G = Z + W$. Then, we may write

$$
\begin{aligned}
\mathbb{P}\left(\|SG\| > (1+\varepsilon)\mathbb{E}\|SG\|\right) &\le \mathbb{P}\left(\|TZ\| > (1+\varepsilon)\mathbb{E}\|SG\| - \lambda\|W\|\right) \\
&\le \mathbb{P}\left(\|TZ\| > (1+\varepsilon)\mathbb{E}\|TZ\| - 10\lambda\mathbb{E}\|W\|_2\right) + \mathbb{P}\left(\|W\| > 10\mathbb{E}\|W\|_2\right) \\
&\le \mathbb{P}\left(\|TZ\| > (1+\varepsilon)\mathbb{E}\|TZ\| - \frac{10\mathbb{E}\|TZ\|}{\log n}\right) + \mathbb{P}\left(\|W\|_2 > 10\mathbb{E}\|W\|_2\right) \\
&\le \mathbb{P}\left(\|TZ\| > \left(1 + \frac{\varepsilon}{2}\right)\mathbb{E}\|TZ\|\right) + \mathbb{P}\left(\|W\|_2 > 10\mathbb{E}\|W\|_2\right) \\
&\le C e^{-c\varepsilon \log n},
\end{aligned}
$$

for all $\frac{20}{\log n} < \varepsilon < 1$, where we have also used the fact that

$$\mathbb{P}(\|W\|_2 > 10\mathbb{E}\|W\|_2) \le e^{-cn}, \quad W \sim N(\mathbf{0}, I_{F^\perp}).$$

For the deviation below the mean we may write

$$
\begin{aligned}
\mathbb{P}(\|SG\| < (1-\varepsilon)\mathbb{E}\|SG\|) &\le \mathbb{P}\left(\|TG\| < (1-\varepsilon)\left(1 + \frac{1}{\log n}\right)\mathbb{E}\|TG\|\right) \\
&\le \mathbb{P}\left(\|TG\| < \left(1 - \frac{\varepsilon}{2}\right)\mathbb{E}\|TG\|\right) \\
&\le C e^{-c\varepsilon \log n},
\end{aligned}
$$

for all $\frac{2}{\log n} < \varepsilon < 1$, where we have used (4.1). The claim follows for all $\varepsilon \in (0, 1)$ by adjusting the universal constants. Note that the estimate (4.2), combined with (1.3), yields that $k(S^{-1}(B_X)) \ge c \log n$, hence the two-level tail estimate readily follows. $\square$

# 5 Further remarks and questions

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

**§1. Concentration for norms close to $\ell_2$.** Here we recall the fact that norms close to the $\ell_2$ norm share analogous concentration, as was claimed in the Introduction. Let $\alpha \geq 1$ and let $\|\cdot\|$ be a norm on $\mathbb{R}^n$ such that

$$\|x\|_2 \leq \|x\| \leq \alpha\|x\|_2, \quad x \in \mathbb{R}^n.$$

Note that $x \mapsto \|x\|$ is $\alpha$-Lipschitz and $\mathbb{E}\|G\| \geq \mathbb{E}\|G\|_2 \simeq \sqrt{n}$, hence the standard Gaussian concentration inequality (1.1) implies

$$\mathbb{P}\left(\left|\|G\| - \mathbb{E}\|G\|\right| > \varepsilon\mathbb{E}\|G\|\right) \leq C\exp(-c\varepsilon^2 n/\alpha^2), \quad \varepsilon > 0.$$

Moreover, one can show that this estimate is essentially optimal, i.e.

$$\mathbb{P}\left(\left|\|G\| - \mathbb{E}\|G\|\right| > \varepsilon\mathbb{E}\|G\|\right) \geq c\exp(-C\varepsilon^2 n\alpha^2), \quad \varepsilon > 0.$$

For a proof of the latter estimate the reader is referred to [PVZ17, Lemma 6.1].

**§2. Concentration in John's position.** We provide an explicit construction of norms which shows that in John's position the concentration estimate (1.2) cannot be improved.

**Proposition 5.1.** *Let $X = (\mathbb{R}^n, \|\cdot\|)$ be a normed space and let $Y := (X \oplus \ell_2^m)_\infty$, i.e.*

$$\|y\|_Y = \max\{\|x\|, \|z\|_2\}, \quad y = (x;z) \in \mathbb{R}^n \times \mathbb{R}^m.$$

*Suppose that $B_X$ is in John's position. Then, we have the following:*

  i. *$B_Y$ is also in John's position and for $m \geq Ck(X)$ we have $\mathrm{Var}\|G\|_Y \geq cb(Y)^2$.*

  ii. *For $m \simeq k(X)$, the norm $\|\cdot\|_Y$ exhibits the following concentration*

$$ce^{-Ct^2 k(X)} \leq \mathbb{P}\left(\left|\|G\|_Y - \mathbb{E}\|G\|_Y\right| > t\mathbb{E}\|G\|_Y\right) \leq Ce^{-ct^2 k(X)}, \quad t > 0.$$

  iii. *For any $\varepsilon \in (0,1)$ we have $k_r(Y,\varepsilon) \simeq \varepsilon^2 k(X)$.*

*Proof.* i. It is easy to verify that $B_Y$ is in John's position. Indeed; for all $y = (x;z)$ we clearly have

$$\|y\|_Y = \|(x;z)\|_Y \leq \max\{\|x\|_2, \|z\|_2\} \leq \sqrt{\|x\|_2^2 + \|z\|_2^2} = \|(x;z)\|_2.$$

Furthermore, there exist $u_1, \ldots, u_s \in S^{n-1}$ contact points, i.e. $\|u_i\| = \|u_i\|_* = 1$ and $c_1, \ldots c_s > 0$ such that $I_{\mathbb{R}^n} = \sum_{j=1}^s c_j u_j \otimes u_j$. Hence, we have

$$I_{\mathbb{R}^n \times \mathbb{R}^m} = \sum_{i=1}^s c_i (u_i; \mathbf{0}_{\mathbb{R}^m}) \otimes (u_i; \mathbf{0}_{\mathbb{R}^m}) + \sum_{j=1}^m e_{n+j} \otimes e_{n+j}.$$

19

By the converse of John's theorem (see [Bal92] for a proof) we conclude that $B_Y$ is in John's position. Let $Z, W$ be independent Gaussian vectors with $Z \sim N(\mathbf{0}, I_m)$ and $W \sim N(\mathbf{0}, I_n)$ and let $G = (W, Z) \sim N(\mathbf{0}, I_{n+m})$. If $A = \{y = (x; z) \in \mathbb{R}^{n+m} : \|x\| \leq \|z\|_2\}$, then we may write

$$\mathrm{Var}[\|G\|_Y] \geq \frac{1}{2} \iint_{A \times A} (\|z\|_2 - \|z'\|_2)^2 \, d\gamma_{n+m}(y) d\gamma_{n+m}(y')$$

$$\geq \mathrm{Var}[\|Z\|_2] - \frac{1}{2} \iint_{(A \times A)^c} (\|z\|_2 - \|z'\|_2)^2 \, d\gamma_{n+m}(y) \, d\gamma_{n+m}(y')$$

$$\geq c_0 - C_0 \sqrt{\mathbb{P}((A \times A)^c)},$$

where we have used the Cauchy-Schwarz inequality and the fact that

$$c_0 \leq \mathrm{Var}(\|Z\|_2) \leq \left( \mathbb{E} \big| \|Z\|_2 - \|Z'\|_2 \big|^4 \right)^{1/2} \leq C_0,$$

where $Z'$ is an independent copy of $Z$ (see e.g. [PVZ17, Proposition 4.4]). On the other hand we have

$$\mathbb{P}((A \times A)^c) \leq 2\mathbb{P}(\|Z\|_2 < \|W\|)$$

$$\leq 2 \left[ \mathbb{P}\left( \|Z\|_2 \leq \delta \sqrt{m} \right) + \mathbb{P}\left( \|W\| > \delta \sqrt{m} \right) \right]$$

$$\leq (c_1 \delta)^m + c_2 e^{-c_3 \delta^2 m},$$

provided that $\delta \sqrt{m} \geq 2\mathbb{E}\|W\| = 2\sqrt{k(X)}$. Choosing $\delta$ sufficiently small universal constant we obtain $\mathbb{P}((A \times A)^c) \leq e^{-cm}$, hence for $m \geq C\delta^{-2} k(X)$ we get $\mathrm{Var}(\|G\|_Y) \geq c_0'$, as required.

ii. Note that $k(Y) \simeq \max\{k(X), m\}$. Taking into account (i) and employing the main result of [Val17] we get the assertion.

iii. Recall that $k_r(Y, \varepsilon)$ is the maximal $k$ for which the random $k$-dimensional subspace of $Y$ is $(1 + \varepsilon)$-Euclidean with probability at least $2/3$. Once we have established the sharp concentration for the norm, it is routine to check that $k_r(Y, \varepsilon) \simeq \varepsilon^2 k(X)$. For the details see [Val17, Section 3]. □

**Remarks 5.2.** 1. The above construction, when $X = \ell_\infty^n$, yields a 1-unconditional norm for which the concentration estimate (1.4) in John's position cannot be improved. Essentially this example is due to K. Tikhomirov, who proves part (iii) in [Tik18] since his focus is on the dependence on $\varepsilon$ in Dvoretzky's theorem. His approach is completely different from the one presented here and lies in delicate estimates for singular values of Gaussian matrices.

2. Applying the above construction for $X = \ell_q^n$, $2 \leq q \leq \infty$ and $m \simeq k(\ell_q^n)$, we get spaces which are in John's position, exhibit optimal concentration in terms of the Lipschitz constant and have Dvoretzky number of all possible range, i.e. $k(Y) \simeq k(\ell_q^n) \simeq k(\ell_q^N) \in (\log N, N)$.

3. The same spaces $Y$ as above yield examples of 1-unconditional normed spaces which are in John's position, are of cotype $q$ with constant $C_q(Y) \simeq C_q(\ell_q^n)$ and satisfy

optimal concentration as in (1.2). This shows that the consideration of [FLM77] on the random version of Dvoretzky's theorem for spaces with cotype in John's position cannot be improved. This also shows that their approach to study the corresponding question for subspaces of $L_q, 2 < q < \infty$, by viewing them as spaces with cotype $q$, is insufficient and other tools are required to obtain the optimal estimates, see [PV15] for the details.

§3. **Hypercontractive measures.** It is worth mentioning that Proposition 2.3 holds true for any hypercontractive measure, since such measures satisfy (1.6); see [CEL12]. Recall that a measure $\mu$ is said to be hypercontractive with constant $\rho$ if it satisfies a log-Sobolev inequality with constant $\rho > 0$, i.e.

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

for any smooth function $f$. In particular, we have the following: Let $\mu$ be a hypercontractive and symmetric Borel probability measure on $\mathbb{R}^n$ with constant $\rho > 0$, and let $\|\cdot\|$ be an arbitrary norm on $\mathbb{R}^n$. Then, for any smooth Lipschitz map $f : \mathbb{R}^n \to \mathbb{R}$ there exists $\lambda = (\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n$ with $\|\lambda\| = 1$ such that

$$\mu\left(z \in \mathbb{R}^n : \left|f(\Lambda z) - \mathbb{E}_\mu f(\Lambda z)\right| > t\right) \le 4\exp\left(-\frac{ct}{b_\Lambda}\sqrt{\rho \log\left(e + \frac{nb_\Lambda^2}{a_\Lambda^2}\right)}\right), \quad t > 0,$$

where $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ and $b_\Lambda = b(f \circ \Lambda)$, $a_\Lambda = a(f \circ \Lambda)$.

In addition, if $f$ is a norm, one may get the following variant of Theorem 3.4: For any normed space $X = (\mathbb{R}^n, \|\cdot\|)$ there exists a non-singular matrix $T$ such that

$$\mu\left(z : \left|\|Tz\| - \mathbb{E}_\mu\|Tz\|\right| > t\mathbb{E}_\mu\|Tz\|\right) \le C\exp\left(-ct\sqrt{\rho k \log\left(\frac{en}{d(X, \ell_\infty^n)^2}\right)}\right), \quad t > 0,$$

where $k = (\mathbb{E}_\mu\|Tz\|/\max_{\|z\|_2 \le 1}\|Tz\|)^2$. To this end, one needs to invoke the following elementary fact.

**Fact 5.3.** *If the norm $\|\cdot\|$ on $\mathbb{R}^n$ satisfies $\|x\|_\infty \le \|x\| \le a\|x\|_\infty$ for all $x$, then for any diagonal matrix $\Lambda$ we have*

$$\max_{\|x\|_\infty \le 1}\|\Lambda x\| \le a \max_{\|x\|_2 \le 1}\|\Lambda x\|.$$

The details are left to the interested reader.

§4. **On the parameter $\beta$.** The following parameter, referred to as the *normalized variance* is introduced in [PV18] (see also [PV17]) for the study of sharp Gaussian small deviation inequalities and small ball probabilities for norms. For any normed space $X = (\mathbb{R}^n, \|\cdot\|)$ we define

$$\beta(X) = \beta(B_X) = \frac{\mathrm{Var}(\|G\|)}{(\mathbb{E}\|G\|)^2}, \quad G \sim N(\mathbf{0}, I_n).$$

It is also known (see e.g. [PV17]) that $\beta(X) \geq \beta(\ell_2^n) \simeq 1/n$. We define further

$$\mathfrak{B}(X) := \min_{T \in GL(n)} \beta(T B_X).$$

In [PV15] we prove that for any $n$-dimensional subspace $X$ of $L_p$, $1 \leq p < \infty$ one has

$$\mathfrak{B}(X) \leq \frac{e^{cp}}{n},$$

which is clearly of minimal possible order (up to constants of $p$). In [PVZ17] and [LT17] the parameter $\beta(\ell_p^n)$ is estimated asymptotically with respect to $n$ and $p$ (when $p$ grows along with $n$).

In the light of Proposition 2.5 we get for any 1-unconditional normed space $X$ in position of minimal $M$, that

$$\beta(X) \leq \frac{C}{k \log(en/k)}, \quad k = k(X) \geq c \log n,$$

which is clearly optimal for $X = \ell_\infty^n$. In particular,

$$\mathfrak{B}(X) \leq \frac{C}{(\log n)^2},$$

for any 1-unconditional normed space $X$. The latter is also derived by K. Tikhomirov in [Tik18].

The main result of Section 4 shows moreover, that for any normed space $X = (\mathbb{R}^n, \|\cdot\|)$ one has

$$\mathfrak{B}(X) \leq \frac{C}{(\log n)^2}.$$

# References

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

[AM83]      N. Alon and V. D. Milman. Embedding of $l_\infty^k$ in finite-dimensional Banach spaces. *Israel J. Math.*, 45(4):265–280, 1983.

[Bal91]     K. Ball. Normed spaces with a weak-Gordon-Lewis property. In *Functional analysis (Austin, TX, 1987/1989)*, volume 1470 of *Lecture Notes in Math.*, pages 36–47. Springer, Berlin, 1991.

[Bal92]     K. Ball. Ellipsoids of maximal volume in convex bodies. *Geom. Dedicata*, 41(2):241–250, 1992.

[BKPS86]    P. Billard, S. Kwapień, A. Pełczyński, and Ch. Samuel. Biorthogonal systems of random unconditional convergence in Banach spaces. In *Texas Functional Analysis Seminar 1985–1986 (Austin, TX, 1985–1986)*, Longhorn Notes, pages 13–35. Univ. Texas, Austin, TX, 1986.

[BL97]      S. Bobkov and M. Ledoux. Poincaré's inequalities and Talagrand's concentration phenomenon for the exponential distribution. *Probab. Theory Related Fields*, 107(3):383–400, 1997.

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

[Bor33]     K. Borsuk. Drei sätze über die $n$-dimensionale euklidische Sphäre. *Fund. Math.*, 20:177–190, 1933.

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

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

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

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

[EG92]      L. C. Evans and R. F. Gariepy. *Measure theory and fine properties of functions*. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.

[FJ80]      T. Figiel and W. B. Johnson. Large subspaces of $l_\infty^n$ and estimates of the Gordon-Lewis constant. *Israel J. Math.*, 37(1-2):92–112, 1980.

[FKP77]     T. Figiel, S. Kwapień, and A. Pełczyński. Sharp estimates for the constants of local unconditional structure of Minkowski spaces. *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.*, 25(12):1221–1226, 1977.

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

[GM97]      A. A. Giannopoulos and V. D. Milman. On the diameter of proportional sections of a symmetric convex body. *Internat. Math. Res. Notices*, (1):5–19, 1997.

[GM00]      A. A. Giannopoulos and V. D. Milman. Extremal problems and isotropic positions of convex bodies. *Israel J. Math.*, 117:29–60, 2000.

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

[Gro53]   A. Grothendieck. Sur certaines classes de suites dans les espaces de Banach et le théorème de Dvoretzky-Rogers. *Bol. Soc. Mat. São Paulo*, 8:81–110 (1956), 1953.

[Jam64]   R. C. James. Uniformly non-square Banach spaces. *Ann. of Math. (2)*, 80:542–550, 1964.

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

[Joh13]   W. B. Johnson. Private communication, Spring 2013.

[KL08]    N. J. Kalton and G. Lancien. Best constants for Lipschitz embeddings of metric spaces into $c_0$. *Fund. Math.*, 199(3):249–272, 2008.

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

[LAT16]   J. Lopez-Abad and P. Tradacete. Bases of random unconditional convergence in Banach spaces. *Trans. Amer. Math. Soc.*, 368(12):9001–9032, 2016.

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

[Lin64]   J. Lindenstrauss. Extension of compact operators. *Mem. Amer. Math. Soc. No.*, 48:112, 1964.

[LMS98]   A. E. Litvak, V. D. Milman, and G. Schechtman. Averages of norms and quasi-norms. *Math. Ann.*, 312(1):95–124, 1998.

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

[LT91]    M. Ledoux and M. Talagrand. *Probability in Banach spaces*, volume 23 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1991. Isoperimetry and processes.

[LT17]    A. Lytova and K. E. Tikhomirov. The variance of the $\ell_p^n$-norm of the gaussian vector, and Dvoretzky's theorem. `https://arxiv.org/abs/1705.05052`, (2017). preprint.

[Mat03]   J. Matoušek. *Using the Borsuk-Ulam theorem*. Universitext. Springer-Verlag, Berlin, 2003. Lectures on topological methods in combinatorics and geometry, Written in cooperation with Anders Björner and Günter M. Ziegler.

[Mau91]   B. Maurey. Some deviation inequalities. *Geom. Funct. Anal.*, 1(2):188–197, 1991.

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

[Mil88]   V. D. Milman. A few observations on the connections between local theory and some other fields. In *Geometric aspects of functional analysis (1986/87)*, volume 1317 of *Lecture Notes in Math.*, pages 283–289. Springer, Berlin, 1988.

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

[Pis74]   G. Pisier. "Type" des espaces normés. In *Séminaire Maurey-Schwartz 1973–1974: Espaces $L^p$, applications radonifiantes et géométrie des espaces de Banach, Exp. No. 3*, pages 12 pp. (errata, p. E.1). Centre de Math., école Polytech., Paris, 1974.

[Pis86]    G. Pisier. Probabilistic methods in the geometry of Banach spaces. In *Probability and analysis (Varenna, 1985)*, volume 1206 of *Lecture Notes in Math.*, pages 167–241. Springer, Berlin, 1986.

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

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

[PV17]    G. Paouris and P. Valettas. Variance estimates and almost Euclidean structure. *Advances in Geometry*, (to appear), 2017. https://arxiv.org/abs/1703.10244.

[PV18]    G. Paouris and P. Valettas. A Gaussian small deviation inequality for convex functions. *Ann. Probab.*, 46(3):1441–1454, 2018.

[PVZ17]    G. Paouris, P. Valettas, and J. Zinn. Random version of Dvoretzky's theorem in $\ell_p^n$. *Stochastic Process. Appl.*, 127(10):3187–3227, 2017.

[Ros83]    H. P. Rosenthal. Some remarks concerning unconditional basic sequences. In *Texas functional analysis seminar 1982–1983 (Austin, Tex.)*, Longhorn Notes, pages 15–47. Univ. Texas Press, Austin, TX, 1983.

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

[Sch06]    G. Schechtman. Two observations regarding embedding subsets of Euclidean spaces in normed spaces. *Adv. Math.*, 200(1):125–135, 2006.

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

[Sch13]    G. Schechtman. Euclidean sections of convex bodies. In *Asymptotic geometric analysis*, volume 68 of *Fields Inst. Commun.*, pages 271–288. Springer, New York, 2013.

[Sch14]    R. Schneider. *Convex bodies: the Brunn-Minkowski theory*, volume 151 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, expanded edition, 2014.

[Tal91]    M. Talagrand. A new isoperimetric inequality and the concentration of measure phenomenon. In *Geometric aspects of functional analysis (1989–90)*, volume 1469 of *Lecture Notes in Math.*, pages 94–124. Springer, Berlin, 1991.

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

[Tal95]    M. Talagrand. Embedding of $l_k^\infty$ and a theorem of Alon and Milman. In *Geometric aspects of functional analysis (Israel, 1992–1994)*, volume 77 of *Oper. Theory Adv. Appl.*, pages 289–293. Birkhäuser, Basel, 1995.

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

[Tik18]    K. E. Tikhomirov. Superconcentration, and randomized Dvoretzky's theorem for spaces with 1-unconditional bases. *J. Funct. Anal.*, 274(1):121–151, 2018.

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

[Val17]    P. Valettas. On the tightness of Gaussian concentration for convex functions. *Journal d' Analyse Mathématique*, (to appear), 2017.

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