Super-Gaussian directions of random vectors

Bo'az Klartag

Abstract

We establish the following universality property in high dimensions: Let X be a random vector with density in $\mathbb{R}^n$. The density function can be arbitrary. We show that there exists a fixed unit vector $\theta \in \mathbb{R}^n$ such that the random variable $Y = \langle X, \theta \rangle$ satisfies

$$\min\left\{\mathbb{P}(Y \geq tM), \mathbb{P}(Y \leq -tM)\right\} \geq ce^{-Ct^2} \qquad \text{for all } 0 \leq t \leq \tilde{c}\sqrt{n},$$

where $M > 0$ is any median of $|Y|$, i.e., $\min\{\mathbb{P}(|Y| \geq M), \mathbb{P}(|Y| \leq M)\} \geq 1/2$. Here, $c, \tilde{c}, C > 0$ are universal constants. The dependence on the dimension n is optimal, up to universal constants, improving upon our previous work.

1 Introduction

Consider a random vector X that is distributed uniformly in some Euclidean ball centered at the origin in $\mathbb{R}^n$. For any fixed vector $0 \neq \theta \in \mathbb{R}^n$, the density of the random variable $\langle X, \theta \rangle = \sum_i \theta_i X_i$ may be found explicitly, and in fact it is proportional to the function

$$t \mapsto \left(1 - \frac{t^2}{A^2 n}\right)_+^{(n-1)/2} \qquad (t \in \mathbb{R}) \tag{1}$$

where $x_+ = \max\{x, 0\}$ and $A > 0$ is a parameter depending on the length of θ and the radius of the Euclidean ball. It follows that when the dimension n is large, the density in (1) is close to a Gaussian density, and the random variable $Y = \langle X, \theta \rangle$ has a tail of considerable size:

$$\mathbb{P}(Y \geq tM) \geq c\exp(-Ct^2) \qquad \text{for all } 0 \leq t \leq \tilde{c}\sqrt{n}. \tag{2}$$

Here, $M = \mathrm{Median}(|Y|)$ is any median of $|Y|$, i.e., $\min\{\mathbb{P}(|Y| \geq M), \mathbb{P}(|Y| \leq M)\} \geq 1/2$, and $c, \tilde{c}, C > 0$ are universal constants. Both the median and the expectation of $|Y|$ differ from A by a factor which is at most a universal constant. We prefer to work with a median since in the cases we will consider shortly, the expectation of $|Y|$ is not guaranteed to be finite. The inequality in (2) expresses the property that the tail distribution of Y/M is at least as heavy as the standard

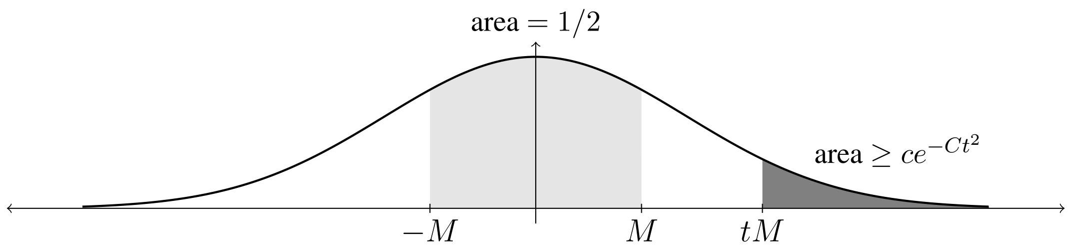

Figure 1: An example of a density of a Super-Gaussian random variable

Gaussian tail distribution, for $\sqrt{n}$ standard deviations. The dependence on the dimension n is optimal, since for $t > \tilde{C}\sqrt{n}$, the probability on the left-hand side of (2) vanishes.

Our goal in this paper is to show that a similar phenomenon occurs for essentially any random vector in $\mathbb{R}^n$, and not only for the uniform distribution on the high-dimensional Euclidean ball. Recall that when n is large and the random vector $X = (X_1, \ldots, X_n)$ has independent coordinates, the classical central limit theorem implies that under mild assumptions, there exists $0 \neq \theta \in \mathbb{R}^n$ for which $\langle X, \theta \rangle$ is approximately Gaussian. It is curious to note that a Gaussian lower bound on the tail persists, even when the independence assumption is completely dropped.

Let Y be a real-valued random variable and let $L > 0$. We say that Y is *Super-Gaussian of length* L with parameters $\alpha, \beta > 0$ if $\mathbb{P}(Y = 0) = 0$ and for any $0 \leq t \leq L$,

$$\min\left\{\mathbb{P}(Y \geq tM), \mathbb{P}(Y \leq -tM)\right\} \geq \alpha e^{-t^2/\beta},$$

where $M = \mathrm{Median}(|Y|)$ is any median of $|Y|$. The requirement that $\mathbb{P}(Y = 0) = 0$ is necessary only to avoid trivialities. A Gaussian random variable is certainly super-Gaussian of infinite length, as well as a symmetric exponential random variable. Write $|x| = \sqrt{\langle x, x \rangle}$ for the standard Euclidean norm of $x \in \mathbb{R}^n$, and denote $S^{n-1} = \{x \in \mathbb{R}^n \, ; \, |x| = 1\}$.

Theorem 1.1. *Let X be a random vector with density in $\mathbb{R}^n$. Then there exists a fixed vector $\theta \in S^{n-1}$ such that $\langle X, \theta \rangle$ is Super-Gaussian of length $c_1\sqrt{n}$ with parameters $c_2, c_3 > 0$, where $c_1, c_2, c_3 > 0$ are universal constants.*

Theorem 1.1 improves upon Corollary 1.4 from [5], in which the dependence on the dimension n was logarithmic. In the case where X is distributed uniformly in a 1-unconditional convex body in $\mathbb{R}^n$, Theorem 1.1 goes back to Pivovarov [9] up to logarithmic factors. In the case where X is distributed uniformly in a convex body satisfying the hyperplane conjecture with a uniform constant, Theorem 1.1 is due to Paouris [8]. Theorem 1.1 provides a universal lower bound on the tail distribution, which is tight up to constants in the case where X is uniformly distributed in a Euclidean ball centered at the origin. In particular, the dependence on the dimension in Theorem 1.1 is optimal, up to the value of the universal constants.

The assumption that the random vector X has a density in $\mathbb{R}^n$ may be somewhat relaxed. The following definition appears in [2, 5] with minor modifications:

Definition 1.2. *Let X be a random vector in a finite-dimensional vector space $\mathcal{B}$ and let $d > 0$. We say that "the effective rank of X is at least d", or in short that X is of class $\mathrm{eff.rank}_{\geq d}$ if for any linear subspace $E \subseteq \mathcal{B}$,*

$$\mathbb{P}(X \in E) \leq \dim(E)/d, \tag{3}$$

with equality if and only if there is a subspace $F \subseteq B$ with $E \oplus F = \mathcal{B}$ and $\mathbb{P}(X \in E \cup F) = 1$.

Intuitively, when X is of class $\mathrm{eff.rank}_{\geq d}$ we think of the support of X as effectively spanning a subspace whose dimension is at least d. Note, however, that d is not necessarily an integer. By substituting $E = \mathcal{B}$ in (3), we see that there are no random vectors in $\mathbb{R}^n$ of class $\mathrm{eff.rank}_{\geq d}$ with $d > n$. We say that the effective rank of X is d when X is of class $\mathrm{eff.rank}_{\geq d}$, but for any $\varepsilon > 0$ the random vector X is not of class $\mathrm{eff.rank}_{\geq d+\varepsilon}$. The effective rank of X is d^- if X is of class $\mathrm{eff.rank}_{\geq d-\varepsilon}$ for all $0 < \varepsilon < d$ but X is not of class $\mathrm{eff.rank}_{\geq d}$. In the terminology of [5], the random vector X has an effective rank greater than d if and only if it is ε-decent for some $\varepsilon < 1/d$.

There are many random vectors in $\mathbb{R}^n$ whose effective rank is precisely n. For example, any random vector with density in $\mathbb{R}^n$, or any random vector X that is distributed uniformly on a finite set that spans $\mathbb{R}^n$ and does not contain the origin. It was shown by Böröczky, Lutwak, Yang, and Zhang [1] and by Henk and Linke [4] that the cone volume measure of any convex body in $\mathbb{R}^n$ with barycenter at the origin is of class $\mathrm{eff.rank}_{\geq n}$ as well. Note that a random variable Y is Super-Gaussian of length L with parameters $\alpha, \beta > 0$ if and only if for any number $0 \neq r \in \mathbb{R}$, also rY is Super-Gaussian of length L with the same parameters $\alpha, \beta > 0$. Theorem 1.1 is thus a particular case of the following:

Theorem 1.3. *Let $d \geq 1$ and let $\mathcal{B}$ be a finite-dimensional linear space. Let X be a random vector in $\mathcal{B}$ whose effective rank is at least d. Then there exists a non-zero, fixed, linear functional $\ell : \mathcal{B} \to \mathbb{R}$ such that the random variable $\ell(X)$ is Super-Gaussian of length $c_1\sqrt{d}$ with parameters $c_2, c_3 > 0$, where $c_1, c_2, c_3 > 0$ are universal constants.*

Theorem 1.3 admits the following corollary, pertaining to infinite-dimensional spaces:

Corollary 1.4. *Let $\mathcal{B}$ be a topological vector space with a countable family of continuous linear functionals that separates points in $\mathcal{B}$. Let X be a random vector, distributed according to a Borel probability measure in $\mathcal{B}$. Assume that $d \geq 1$ is such that $\mathbb{P}(X \in E) \leq \dim(E)/d$ for any finite-dimensional subspace $E \subseteq \mathcal{B}$.*

Then there exists a non-zero, fixed, continuous linear functional $\ell : \mathcal{B} \to \mathbb{R}$ such that the random variable $\ell(X)$ is Super-Gaussian of length $c_1\sqrt{d}$ with parameters $c_2, c_3 > 0$, where $c_1, c_2, c_3 > 0$ are universal constants.

The remainder of this paper is devoted to the proof of Theorem 1.3 and Corollary 1.4. We use the letters $c, C, \tilde{C}, c_1, C_2$ etc. to denote various positive universal constants, whose value may change from one line to the next. We use upper-case C to denote universal constants that we think of as "sufficiently large", and lower-case c to denote universal constants that are "sufficiently small". We write $\#(A)$ for the cardinality of a set A. When we write that a certain set or a certain number are fixed, we intend to emphasize that they are non-random.

We denote by σ_{n-1} the uniform probability measure on the sphere S^{n-1}, which is the unique rotationally-invariant probability measure on S^{n-1}. When we say that a random vector θ is distributed uniformly on S^{n-1}, we refer to the probability measure σ_{n-1}. Similarly, when we write that a random subspace E is distributed uniformly over the Grassmannian $G_{n,k}$ of k-dimensional subspaces of $\mathbb{R}^n$, we refer to the unique rotationally-invariant probability measure on $G_{n,k}$.

Acknowledgements. I would like to thank Bo Berndtsson and Emanuel Milman for interesting discussions and for encouraging me to write this paper. Supported by a grant from the European Research Council.

2 Proof strategy

The main ingredient in the proof of Theorem 1.3 is the following proposition:

Proposition 2.1. *Let X be a random vector in $\mathbb{R}^n$ with $\mathbb{P}(X = 0) = 0$ such that*

$$\mathbb{E} \left\langle \frac{X}{|X|}, \theta \right\rangle^2 \leq \frac{5}{n} \qquad \textit{for all } \theta \in S^{n-1}. \tag{4}$$

Then there exists a fixed vector $\theta \in S^{n-1}$ such that the random variable $\langle X, \theta \rangle$ is Super-Gaussian of length $c_1\sqrt{n}$ with parameters $c_2, c_3 > 0$, where $c_1, c_2, c_3 > 0$ are universal constants.

The number 5 in Proposition 2.1 does not play any particular role, and may be replaced by any other universal constant, at the expense of modifying the values of c_1, c_2 and c_3. Let us explain the key ideas in the proof of Proposition 2.1. In our previous work [5], the unit vector $\theta \in S^{n-1}$ was chosen randomly, uniformly on S^{n-1}. In order to improve the dependence on the dimension, here we select θ a bit differently. We shall define θ_1 and θ_2 via the following procedure:

(i) Let $M > 0$ be a $1/3$-quantile of $|X|$, i.e., $\mathbb{P}(|X| \geq M) \geq 1/3$ and $\mathbb{P}(|X| \leq M) \geq 2/3$. We fix a vector $\theta_1 \in S^{n-1}$ such that

$$\mathbb{P}\left(|X| \geq M \text{ and } \left|\frac{X}{|X|} - \theta_1\right| \leq \frac{1}{5}\right) \geq \frac{1}{2} \cdot \sup_{\eta \in S^{n-1}} \mathbb{P}\left(|X| \geq M \text{ and } \left|\frac{X}{|X|} - \eta\right| \leq \frac{1}{5}\right).$$

(ii) Next, we fix a vector $\theta_2 \in S^{n-1}$ with $|\langle \theta_1, \theta_2 \rangle| \leq 1/10$ such that

$$\mathbb{P}\left(|X| \geq M \text{ and } \left|\frac{X}{|X|} - \theta_2\right| \leq \frac{1}{5}\right) \geq \frac{1}{2} \cdot \sup_{\substack{\eta \in S^{n-1} \\ |\langle \eta, \theta_1 \rangle| \leq 1/10}} \mathbb{P}\left(|X| \geq M \text{ and } \left|\frac{X}{|X|} - \eta\right| \leq \frac{1}{5}\right).$$

In the following pages we will describe a certain subset $\mathcal{F}_3 \subseteq S^{n-1}$ which satisfies $\sigma_{n-1}(\mathcal{F}_3) \geq 1 - C/n^c$ and $\theta_2 - \theta_1 \notin \mathcal{F}_3$. We will show that for any $\theta_3 \in \mathcal{F}_3$, the random variable $\langle X, \theta \rangle$ is Super-Gaussian of length $c\sqrt{n}$ with parameters $c_1, c_2 > 0$, where θ is defined as follows:

$$\theta = \frac{\theta_1 - \theta_2 + \theta_3}{|\theta_1 - \theta_2 + \theta_3|}. \tag{5}$$

Thus, θ_1 and θ_2 are fixed vectors, while most choices of θ_3 will work for us, where by "most" we refer to the uniform measure on S^{n-1}. The first step the proof below is to show that for any unit vector $\theta \in S^{n-1}$,

$$\text{Median}\,(|\langle X, \theta \rangle|) \leq CM/\sqrt{n}, \tag{6}$$

that is, any median of $|\langle X, \theta \rangle|$ is at most $CM/\sqrt{n}$. Then we need to show that when $\theta_3 \in \mathcal{F}_3$ and θ is defined as in (5), for all $0 \leq t \leq c\sqrt{n}$,

$$\min \left\{ \mathbb{P}\left(Y \geq \frac{tM}{\sqrt{n}} \right), \mathbb{P}\left(Y \leq -\frac{tM}{\sqrt{n}} \right) \right\} \geq \tilde{c} e^{-\tilde{C} t^2}. \tag{7}$$

The proof of (7) is divided into three sections. The case where $t \in [0, \sqrt{\log n}]$ may essentially be handled by using the methods of [5], see Section 3. Let $t_0 > 0$ be defined via

$$e^{-t_0^2} = \mathbb{P}\left(|X| \geq M \text{ and } \left| \frac{X}{|X|} - \theta_2 \right| \leq \frac{1}{5} \right). \tag{8}$$

In order to prove (7) in the range $t \in [\sqrt{\log n}, t_0]$, we will use tools from the local theory of Banach spaces, such as Sudakov's inequality as well as the concentration of measure on the sphere. Details in Section 4 below. The remaining interval $t \in [t_0, c\sqrt{n}]$ is analyzed in Section 5. In Section 6 we deduce Theorem 1.3 and Corollary 1.4 from Proposition 2.1 by using the angularly-isotropic position, along the lines of [5].

3 Central limit regime

This section is the first in a sequence of three sections that are dedicated to the proof of Proposition 2.1. Thus, we are given a random vector X in $\mathbb{R}^n$ with $\mathbb{P}(X = 0) = 0$ such that (4) holds true. We fix a number $M > 0$ with the property that

$$\mathbb{P}(|X| \geq M) \geq 1/3, \qquad \mathbb{P}(|X| \leq M) \geq 2/3. \tag{9}$$

That is, M is a $1/3$-quantile of $|X|$. Our first lemma verifies (6), as it states that for any choice of a unit vector θ, any median of the random variable $|\langle X, \theta \rangle|$ is at most $CM/\sqrt{n}$.

Lemma 3.1. *For any $\theta \in S^{n-1}$,*

$$\mathbb{P}\left(|\langle X, \theta \rangle| \geq CM/\sqrt{n} \right) < 1/2,$$

where $C > 0$ is a universal constant.

Proof. It follows from (4) that for any $\theta \in S^{n-1}$,

$$\mathbb{E}\left[\langle X,\theta\rangle^2 1_{\{|X|\le M\}}\right] \le \mathbb{E}\left[\langle X,\theta\rangle^2 \cdot \frac{M^2}{|X|^2}\right] = M^2 \cdot \mathbb{E}\left\langle \frac{X}{|X|},\theta\right\rangle^2 \le \frac{5M^2}{n}.$$

By the Markov-Chebyshev inequality,

$$\mathbb{P}\left(\langle X,\theta\rangle^2 1_{\{|X|\le M\}} \ge 35M^2/n\right) \le 1/7.$$

Since $\mathbb{P}(|X|>M) \le 1/3$, we obtain

$$\mathbb{P}\left(|\langle X,\theta\rangle| \ge \frac{6M}{\sqrt{n}}\right) \le \mathbb{P}(|X|>M) + \mathbb{P}\left(|\langle X,\theta\rangle| \ge \frac{6M}{\sqrt{n}} \text{ and } |X|\le M\right) \le \frac{1}{3}+\frac{1}{7} < \frac{1}{2}.$$

The lemma follows with $C=6$. □

The rest of this section is devoted to the proof of (7) in the range $t \in [0,\sqrt{\log n}]$. The defining properties of $\theta_1,\theta_2 \in S^{n-1}$ from the previous section will not be used here, the entire analysis in this section applies for arbitrary unit vectors θ_1 and θ_2.

Lemma 3.2. *Let $\theta_1,\theta_2 \in S^{n-1}$ be any two fixed vectors. Then,*

$$\mathbb{P}\left(|X|\ge M,\ |\langle X,\theta_1\rangle| \le \frac{10|X|}{\sqrt{n}} \ \textit{and}\ |\langle X,\theta_2\rangle| \le \frac{10|X|}{\sqrt{n}}\right) > \frac{1}{5}.$$

Proof. By (4) and the Markov-Chebyshev inequality, for $j=1,2$,

$$\mathbb{P}\left(|\langle X,\theta_j\rangle| \ge \frac{10|X|}{\sqrt{n}}\right) \le \frac{n}{100}\cdot \mathbb{E}\left\langle \frac{X}{|X|},\theta_j\right\rangle^2 \le \frac{n}{100}\cdot\frac{5}{n} = \frac{1}{20}.$$

Thanks to (9), we conclude that

$$\mathbb{P}\left(|X|\ge M,\ |\langle X,\theta_1\rangle| \le \frac{10|X|}{\sqrt{n}},\ |\langle X,\theta_2\rangle| \le \frac{10|X|}{\sqrt{n}}\right) \ge 1-\left(\frac{2}{3}+\frac{1}{20}+\frac{1}{20}\right) > \frac{1}{5}. \quad \square$$

Let $1\le k\le n$. Following [5], we write $\mathcal{O}_k \subseteq (\mathbb{R}^n)^k$ for the collection of all k-tuples $(v_1,\dots,v_k)$ with the following property: There exist orthonormal vectors $w_1,\dots,w_k \in \mathbb{R}^n$ and real numbers $(a_{ij})_{i,j=0,\dots,k}$ such that $|a_{ij}| < a_{ii}/k^2$ for $j<i$, and

$$v_i = \sum_{j=1}^{i} a_{ij}w_j \qquad \text{for } i=1,\dots,k. \tag{10}$$

In other words, $\mathcal{O}_k$ consists of k-tuples of vectors that are almost orthogonal. By recalling the Gram-Schmidt process from linear algebra, we see that $(v_1,\dots,v_k)\in\mathcal{O}_k$ assuming that

$$|\mathrm{Proj}_{E_{i-1}} v_i| < |v_i|/k^2 \qquad \text{for } i=1,\dots,k, \tag{11}$$

where E_i is the subspace spanned by the vectors $v_1,\dots,v_i \in \mathbb{R}^n$ and Proj_{E_i} is the orthogonal projection operator onto E_i in $\mathbb{R}^n$. Here, $E_0=\{0\}$.

Lemma 3.3. *Assume that* $1 \le k \le n$ *and fix* $(v_1, \ldots, v_k) \in \mathcal{O}_k$. *Then there exists* $\mathcal{F} \subseteq S^{n-1}$ *with* $\sigma_{n-1}(\mathcal{F}) \ge 1 - C\exp(-c\sqrt{k})$ *such that for any* $\theta \in \mathcal{F}$ *and* $0 \le t \le \sqrt{\log k}$,

$$\#\left\{1 \le i \le k\,;\, \langle v_i, \theta\rangle \ge c_1 \frac{|v_i|}{\sqrt{n}} \cdot t\right\} \ge c_2 e^{-C_3 t^2} \cdot k,$$

where $c_1, c_2, C_3, c, C > 0$ *are universal constants.*

Proof. Let $w_1, \ldots, w_k$ and (a_{ij}) be as in (10). By applying an orthogonal transformation in $\mathbb{R}^n$, we may assume that $w_i = e_i$, the standard i^{th} unit vector. Let $\Gamma = (\Gamma_1, \ldots, \Gamma_n) \in \mathbb{R}^n$ be a standard Gaussian random vector in $\mathbb{R}^n$. For $i = 1, \ldots, n$ and $t > 0$, it is well-known that

$$\mathbb{P}(\Gamma_i \ge t) = \frac{1}{\sqrt{2\pi}} \int_t^\infty e^{-s^2/2} ds \in [ce^{-t^2}, Ce^{-t^2/2}].$$

Therefore, by the Chernoff large deviations bound (e.g., [3, Chapter 2]), for any $t > 0$,

$$\mathbb{P}\left(\#\left\{1 \le i \le k\,;\, \Gamma_i \ge t\right\} \ge \frac{c}{2} \cdot e^{-t^2} \cdot k\right) \ge 1 - \tilde{C}\exp\left(-\tilde{c}e^{-t^2} k\right). \tag{12}$$

From the Bernstein large deviation inequality (e.g., [3, Chapter 2]),

$$\mathbb{P}\left(|\Gamma| \le 2\sqrt{n}\right) \ge 1 - Ce^{-cn}, \qquad \mathbb{P}\left(\sum_{i=1}^k |\Gamma_i| \le 2k\right) \ge 1 - \hat{C}e^{-\hat{c}k}. \tag{13}$$

Note that when $\sum_{i=1}^k |\Gamma_i| \le 2k$, for any $i = 1, \ldots, k$,

$$\langle \Gamma, v_i\rangle = a_{ii} \cdot \left\langle \Gamma, e_i + \sum_{j=2}^i \frac{a_{ij}}{a_{ii}} e_j \right\rangle \ge a_{ii}\left(\Gamma_i - \frac{\sum_{j=1}^k |\Gamma_j|}{k^2}\right) \ge a_{ii}\left(\Gamma_i - \frac{2}{k}\right). \tag{14}$$

Moreover, $a_{ii} = |v_i - \sum_{j \le 2} a_{ij} e_j| \ge |v_i| - a_{ii}/k$ for all $i = 1, \ldots, k$. Therefore $a_{ii} \ge |v_i|/2$ for all i. It thus follows from (14) that when $\sum_{i=1}^k |\Gamma_i| \le 2k$, for any i,

$$\Gamma_i \ge t \qquad \Longrightarrow \qquad \langle \Gamma, v_i\rangle \ge a_{ii}\Gamma_i/2 \ge |v_i|t/4 \qquad\qquad \text{for all } t \ge 4/k.$$

Hence we deduce from (12) and (13) that for all $t \ge 4/k$,

$$\mathbb{P}\left(\#\left\{i\,;\, \langle \Gamma, v_i\rangle \ge \frac{t|v_i|}{4}\right\} \ge \frac{c}{2} \cdot e^{-t^2} \cdot k\right) \ge 1 - \tilde{C}\exp\left(-\tilde{c}e^{-t^2} k\right). \tag{15}$$

Write $I = \{\ell \in \mathbb{Z}\,;\, \ell \ge 2,\ 2^\ell \le \sqrt{\log k}/5\}$. By substituting $t = 2^\ell$ into (15) we see that

$$\mathbb{P}\left(\forall \ell \in I, \#\left\{i\,;\, \langle \Gamma, v_i\rangle \ge 2^{\ell-2}|v_i|\right\} \ge \frac{c}{2} \cdot e^{-(2^\ell)^2} \cdot k\right) \ge 1 - \tilde{C}\sum_{\ell \in I} \exp\left(-\tilde{c}e^{-(2^\ell)^2} k\right).$$

The latter sum is at most $\overline{C}\exp(-\bar{c}\sqrt{k})$. Moreover, suppose that $x \in \mathbb{R}^n$ is a fixed vector such that $\#\{i\,;\,\langle x, v_i\rangle \geq t|v_i|/4\} \geq (c/2)e^{-t^2}k$ for all $1 \leq t \leq \sqrt{\log k}/5$ of the form $t = 2^\ell$ for an integer $\ell \geq 2$. By adjusting the constants, we see that for any real number t with $0 \leq t \leq \sqrt{\log k}$,

$$\#\{i\,;\,\langle x, v_i\rangle \geq c_1 t|v_i|\} \geq \tilde{c}e^{-\tilde{C}t^2}k.$$

Consequently,

$$\mathbb{P}\left(\forall t \in [0, \sqrt{\log k}],\ \#\{i\,;\,\langle \Gamma, v_i\rangle \geq c_1 t|v_i|\} \geq \tilde{c}e^{-\tilde{C}t^2}\cdot k\right) \geq 1 - \overline{C}e^{-\bar{c}\sqrt{k}}.$$

Recall that $|\Gamma| \leq 2\sqrt{n}$ with a probability of at least $1 - Ce^{-cn}$. Therefore, as $k \leq n$,

$$\mathbb{P}\left(\forall t \in [0, \sqrt{\log k}],\ \#\left\{i\,;\,\left\langle \frac{\Gamma}{|\Gamma|}, v_i\right\rangle \geq c_1\frac{t|v_i|}{2\sqrt{n}}\right\} \geq \tilde{c}e^{-\tilde{C}t^2}\cdot k\right) \geq 1 - \hat{C}e^{-\hat{c}\sqrt{k}}. \tag{16}$$

Since $\Gamma/|\Gamma|$ is distributed uniformly on S^{n-1}, the lemma follows from (16). □

Let $E \subseteq \mathbb{R}^n$ be an arbitrary subspace. It follows from (4) that

$$\mathbb{E}\left|\mathrm{Proj}_E\frac{X}{|X|}\right|^2 = \mathbb{E}\sum_{i=1}^{\dim(E)}\left\langle\frac{X}{|X|}, u_i\right\rangle^2 \leq 5\frac{\dim(E)}{n}, \tag{17}$$

where $u_1, \ldots, u_m$ is an orthonormal basis of the subspace E for $m = \dim(E)$.

Lemma 3.4. *Set $\ell = \lfloor n^{1/8}\rfloor$ and let $\theta_1, \theta_2 \in S^{n-1}$ be any fixed vectors. Let $X_1, \ldots, X_\ell$ be independent copies of the random vector X. Then with a probability of at least $1 - C/\ell$ of selecting $X_1, \ldots, X_\ell$, there exists a subset $I \subseteq \{1, \ldots, \ell\}$ with the following three properties:*

(i) $k := \#(I) \geq \ell/10$.

(ii) We may write $I = \{i_1, \ldots, i_k\}$ such that $(X_{i_1}, \ldots, X_{i_k}) \in \mathcal{O}_k$.

(iii) For $j = 1, \ldots, k$,

$$|X_{i_j}| \geq M, \quad |\langle X_{i_j}, \theta_1\rangle| \leq 10|X_{i_j}|/\sqrt{n} \quad \text{and} \quad |\langle X_{i_j}, \theta_2\rangle| \leq 10|X_{i_j}|/\sqrt{n}.$$

Here, $C > 0$ is a universal constant.

Proof. We may assume that $\ell \geq 10$, as otherwise the lemma trivially holds with any $C \geq 10$. Define

$$I = \left\{1 \leq i \leq \ell\,;\,|X_i| \geq M,\ |\langle X_i, \theta_1\rangle| \leq 10|X_i|/\sqrt{n},\ |\langle X_i, \theta_2\rangle| \leq 10|X_i|/\sqrt{n}\right\}.$$

Denote $k = \#(I)$ and let $i_1 < i_2 < \ldots < i_k$ be the elements of I. We conclude from Lemma 3.2 and the Chernoff large deviation bound that

$$\mathbb{P}(\#(I) \geq \ell/10) \geq 1 - C\exp(-c\ell). \tag{18}$$

Thus (i) holds with a probability of at least $1 - C\exp(-c\ell)$. Clearly (iii) holds true with probability one, by the definition of I. All that remains is to show that (ii) holds true with a probability of at least $1 - 1/\ell$. Write F_i for the subspace spanned by $X_1, \ldots, X_i$, with $F_0 = \{0\}$. It follows from (17) that for $i = 1, \ldots, \ell$,

$$\mathbb{E}\left|Proj_{F_{i-1}}\frac{X_i}{|X_i|}\right|^2 \leq \frac{5\cdot\dim(F_{i-1})}{n} \leq \frac{5(i-1)}{n} \leq \frac{5\ell}{n} < \frac{1}{\ell^6},$$

as $10 \leq \ell \leq n^{1/8}$. It follows from the Markov-Chebyshev inequality that with a probability of at least $1 - 1/\ell$,

$$\left|Proj_{F_{i-1}}\frac{X_i}{|X_i|}\right| < \frac{1}{\ell^2} \qquad \text{for all } i = 1, \ldots, \ell.$$

Write E_j for the subspace spanned by $X_{i_1}, \ldots, X_{i_j}$. Then $E_{j-1} \subseteq F_{i_j-1}$. Therefore, with a probability of at least $1 - 1/\ell$,

$$\left|Proj_{E_{j-1}}\frac{X_{i_j}}{|X_{i_j}|}\right| \leq \left|\mathrm{Proj}_{F_{i_j-1}}\frac{X_{i_j}}{|X_{i_j}|}\right| < \frac{1}{\ell^2} \leq \frac{1}{k^2} \qquad \text{for all } j = 1, \ldots, k.$$

In view of (11), we see that (ii) holds true with a probability of at least $1 - 1/\ell$, thus completing the proof of the lemma. □

By combining Lemma 3.3 and Lemma 3.4 we arrive at the following:

Lemma 3.5. *Let ℓ, θ_1, θ_2 be as in Lemma 3.4. Then there exists a fixed subset $\mathcal{F} \subseteq S^{n-1}$ with $\sigma_{n-1}(\mathcal{F}) \geq 1 - C/\sqrt{\ell}$ such that for any $\theta_3 \in \mathcal{F}$ the following holds: Define θ via (5). Let $X_1, \ldots, X_\ell$ be independent copies of the random vector X. Then with a probability of at least $1 - C/\sqrt{\ell}$ of selecting $X_1, \ldots, X_\ell$,*

$$\#\left\{1 \leq i \leq \ell\,;\, \langle X_i, \theta\rangle \geq c_1\frac{M}{\sqrt{n}}\cdot t\right\} \geq c_2 e^{-C_3t^2}\cdot\ell, \qquad \textit{for all } 0 \leq t \leq \sqrt{\log \ell}, \tag{19}$$

and

$$\#\left\{1 \leq i \leq \ell\,;\, \langle X_i, \theta\rangle \leq -c_1\frac{M}{\sqrt{n}}\cdot t\right\} \geq c_2 e^{-C_3t^2}\cdot\ell, \qquad \textit{for all } 0 \leq t \leq \sqrt{\log \ell}. \tag{20}$$

Here, $c_1, c_2, C_3, c, C > 0$ are universal constants.

Proof. Let Θ be a random vector, distributed uniformly on S^{n-1}. According to Lemma 3.4, with a probability of at least $1 - C/\ell$ of selecting $X_1, \ldots, X_\ell$, there exists a subset

$$I = \{i_1, \ldots, i_k\} \subseteq \{1, \ldots, \ell\}$$

such that properties (i), (ii) and (iii) of Lemma 3.4 hold true. Let us apply Lemma 3.3. Then under the event where properties (i), (ii) and (iii) hold true, with a probability of at least $1 - \tilde{C}\exp(-\tilde{c}\sqrt{\ell})$ of selecting $\Theta \in S^{n-1}$,

$$\#\left\{1 \leq j \leq k\,;\, \langle X_{i_j}, \Theta\rangle \geq c_1\frac{|X_{i_j}|}{\sqrt{n}}\cdot t\right\} \geq c_2 e^{-C_3t^2}\cdot k \qquad \text{for all } 0 \leq t \leq \sqrt{\log k},$$

and moreover $k \geq \ell/10$ with

$$\max\left\{\left|\left\langle \frac{X_{i_j}}{|X_{i_j}|}, \theta_1 \right\rangle\right|, \left|\left\langle \frac{X_{i_j}}{|X_{i_j}|}, \theta_2 \right\rangle\right|\right\} \leq \frac{10}{\sqrt{n}} \qquad \text{for } j = 1, \ldots, k.$$

Consequently, under the event where properties (i), (ii) and (iii) hold true, with a probability of at least $1 - \tilde{C}\exp(-\tilde{c}\sqrt{\ell})$ of selecting $\Theta \in S^{n-1}$,

$$\#\left\{1 \leq j \leq k\,;\, \left\langle \frac{X_{i_j}}{|X_{i_j}|}, \theta_1 - \theta_2 + \Theta \right\rangle \geq \frac{c_1}{2}\frac{t}{\sqrt{n}}\right\} \geq c_2 e^{-C_3 t^2} \cdot k \qquad \text{for } t \in [80/c_1, \sqrt{\log k}].$$

Since $k \geq \ell/10$, the condition $t \in [80/c_1, \sqrt{\log k}]$ can be upgraded to $t \in [0, \sqrt{\log \ell}]$ at the cost of modifying the universal constants. Recall that by Lemma 3.3(iii), we have that $|X_{i_j}| \geq M$ for all j. By the triangle inequality, with probability one, $0 < |\theta_1 - \theta_2 + \Theta| \leq 3$. Hence,

$$|X_{i_j}|/|\theta_1 - \theta_2 + \Theta| \geq M/3.$$

Therefore, under the event where properties (i), (ii) and (iii) hold true, with a probability of at least $1 - \tilde{C}\exp(-\tilde{c}\sqrt{\ell})$ of selecting $\Theta \in S^{n-1}$,

$$\forall t \in [0, \sqrt{\log \ell}], \quad \#\left\{1 \leq i \leq \ell\,;\, \left\langle X_i, \frac{\theta_1 - \theta_2 + \Theta}{|\theta_1 - \theta_2 + \Theta|} \right\rangle \geq \bar{c}_1 \frac{M}{\sqrt{n}} \cdot t\right\} \geq \bar{c}_2 e^{-\bar{C}_3 t^2} \cdot \ell. \tag{21}$$

Write $\mathcal{A}$ for the event that the statement in (21) holds true. Denoting $\vec{X} = (X_1, \ldots, X_\ell)$, we have shown that

$$\mathbb{P}((\Theta, \vec{X}) \in \mathcal{A}) \geq 1 - \tilde{C}\exp(-\tilde{c}\sqrt{\ell}) - C/\ell \geq 1 - \bar{C}/\ell.$$

Denote

$$\mathcal{F} = \left\{\theta \in S^{n-1}\,;\, \mathbb{P}_{\vec{X}}((\theta, \vec{X}) \in \mathcal{A}) \geq 1 - \bar{C}/\sqrt{\ell}\right\}.$$

Then,

$$1 - \frac{\bar{C}}{\ell} \leq \mathbb{P}((\Theta, \vec{X}) \in \mathcal{A}) \leq \mathbb{P}(\Theta \in \mathcal{F}) + \left(1 - \frac{\bar{C}}{\sqrt{\ell}}\right)\mathbb{P}(\Theta \notin \mathcal{F}). \tag{22}$$

It follows from (22) that $\sigma_{n-1}(\mathcal{F}) = \mathbb{P}(\Theta \in \mathcal{F}) \geq 1 - 1/\sqrt{\ell}$. By the definition of $\mathcal{F} \subseteq S^{n-1}$, for any $\theta_3 \in \mathcal{F}$, with a probability of at least $1 - \bar{C}\sqrt{\ell}$ of selecting $X_1, \ldots, X_\ell$,

$$\forall t \in [0, \sqrt{\log \ell}], \quad \#\left\{1 \leq i \leq \ell\,;\, \left\langle X_i, \frac{\theta_1 - \theta_2 + \theta_3}{|\theta_1 - \theta_2 + \theta_3|} \right\rangle \geq \bar{c}_1 \frac{M}{\sqrt{n}} \cdot t\right\} \geq \bar{c}_2 e^{-\bar{C}_3 t^2} \cdot \ell.$$

This completes the proof of (19). The argument for (20) requires only the most trivial modifications, and we leave it for the reader to complete. $\square$

We will use the well-known fact that for any random variable Y and measurable sets $A_1, \ldots, A_\ell$, by the Markov-Chebyshev inequality,

$$\frac{1}{s} \cdot \sum_{i=1}^{\ell} \mathbb{P}(Y \in A_i) = \frac{1}{s} \cdot \mathbb{E} \sum_{i=1}^{\ell} 1_{\{Y \in A_i\}} \geq \mathbb{P}\big(\# \{i \,;\, Y \in A_i\} \geq s\big) \qquad (s > 0).$$

Corollary 3.6. *Let $\theta_1, \theta_2 \in S^{n-1}$ be any fixed vectors. Then there exists a fixed subset $\mathcal{F} \subseteq S^{n-1}$ with $\sigma_{n-1}(\mathcal{F}) \geq 1 - C/n^c$ such that for any $\theta_3 \in \mathcal{F}$, defining θ via (5),*

$$\forall t \in [0, 5\sqrt{\log n}], \quad \min\left\{ \mathbb{P}\left(\langle X, \theta \rangle \geq c_1 \frac{M}{\sqrt{n}} \cdot t \right), \mathbb{P}\left(\langle X, \theta \rangle \leq -c_1 \frac{M}{\sqrt{n}} \cdot t \right) \right\} \geq c_2 e^{-C_3 t^2},$$

where $c, C, c_1, c_2, C_3 > 0$ are universal constants.

Proof. We may assume that n exceeds a certain fixed universal constant, as otherwise the conclusion of the lemma trivially holds for $\mathcal{F} = \emptyset$. Set $\ell = \lfloor n^{1/8} \rfloor$ and let $\mathcal{F}$ be the set from Lemma 3.5. Let $\theta_3 \in \mathcal{F}$ and define θ via (5). Suppose that $X_1, \ldots, X_\ell$ are independent copies of the random vector X. Then for any $0 \leq t \leq \sqrt{\log \ell}$,

$$\mathbb{P}\left(\langle X, \theta \rangle \geq c_1 \frac{M}{\sqrt{n}} \cdot t \right) = c_2 e^{-C_3 t^2} \frac{1}{c_2 e^{-C_3 t^2} \cdot \ell} \sum_{i=1}^{\ell} \mathbb{P}\left(\langle X_i, \theta \rangle \geq c_1 \frac{M}{\sqrt{n}} \cdot t \right)$$
$$\geq c_2 e^{-C_3 t^2} \cdot \mathbb{P}\left(\# \left\{ i; \langle X_i, \theta \rangle \geq c_1 \frac{M}{\sqrt{n}} \cdot t \right\} \geq c_2 e^{-C_3 t^2} \cdot \ell \right) \geq c_2 e^{-C_3 t^2} \cdot (1 - C/\sqrt{\ell}),$$

where the last passage is the content of Lemma 3.5. We may similarly obtain a corresponding lower bound for $\mathbb{P}\left(\langle X, \theta \rangle \leq -c_1 t M / \sqrt{n} \right)$. Since $\ell = \lfloor n^{1/8} \rfloor$, the desired conclusion follows by adjusting the constants. ☐

4 Geometry of the high-dimensional sphere

This is the second section dedicated to the proof of Proposition 2.1. A few geometric properties of the high-dimensional sphere will be used here. For example, the sphere S^{n-1} does not contain more than n mutually orthogonal vectors, yet it contains $e^{\varepsilon n}$ mutually almost-orthogonal vectors. Moreover, for the purpose of computing the expectation of the supremum, a family of $e^{\varepsilon n}$ standard Gaussians which are almost-orthogonal in pairs behaves approximately like a collection of independent Gaussians.

While Corollary 3.6 takes care of the interval $t \in [0, 5\sqrt{\log n}]$, in this section we deal with the range $t \in [5\sqrt{\log n}, t_0]$ where t_0 is defined in (8). We begin with some background on Sudakov's minoration theorem and the concentration of measure inequality on the sphere. Given a bounded, non-empty subset $S \subseteq \mathbb{R}^n$, its *supporting functional* is defined via

$$h_S(\theta) = \sup_{x \in S} \langle x, \theta \rangle \qquad (\theta \in \mathbb{R}^n).$$

The supporting functional h_S is a convex function on $\mathbb{R}^n$ whose Lipschitz constant is bounded by $R(S) = \sup_{x \in S} |x|$. The *mean width* of S is $2M^*(S)$ where

$$M^*(S) = \int_{S^{n-1}} h_S(\theta) d\sigma_{n-1}(\theta).$$

The concentration inequality for Lipschitz functions on the sphere (see, e.g., [7, Appendix V]) states that for any $r > 0$,

$$\sigma_{n-1}\left(\left\{v \in S^{n-1}\,;\, |h_S(v) - M^*(S)| \geq r \cdot R(S)\right\}\right) \leq Ce^{-cr^2 n}. \tag{23}$$

A lower bound for $M^*(S)$ is provided by the following Sudakov's minoration theorem (see, e.g., [6, Section 3.3]):

Theorem 4.1 (Sudakov)**.** *Let $N \geq 1, \alpha > 0$ and let $x_1, \ldots, x_N \in \mathbb{R}^n$. Set $S = \{x_1, \ldots, x_N\}$ and assume that $|x_i - x_j| \geq \alpha$ for any $i \neq j$. Then,*

$$M^*(S) \geq c\alpha\sqrt{\frac{\log N}{n}},$$

where $c > 0$ is a universal constant.

We shall need the following elementary lemma:

Lemma 4.2. *Let $Z_1, \ldots, Z_N$ be random variables attaining values in $\{0, 1\}$. Let $1 \leq k \leq N, 0 \leq \varepsilon \leq 1$, and assume that for any $A \subseteq \{1, \ldots, N\}$ with $\#(A) = k$,*

$$\mathbb{P}\left(\exists i \in A,\ Z_i = 1\right) \geq 1 - \varepsilon. \tag{24}$$

Then,

$$\mathbb{P}\left(\sum_{i=1}^{N} Z_i \geq \frac{N}{3k}\right) \geq 1 - 2\varepsilon. \tag{25}$$

Proof. If $k \geq N/3$ then (25) holds true, since it follows from (24) that with a probability of at least $1 - \varepsilon$, there is a non-zero element among $Z_1, \ldots, Z_N$. Suppose now that $k < N/3$. The number of k-elements subsets $A \subseteq \{1, \ldots, N\}$ with $\max_{i \in A} Z_i = 0$ equals

$$\binom{N - \sum_{i=1}^{N} Z_i}{k}.$$

Write $\mathcal{E}$ for the event that $\sum_{i=1}^{N} Z_i \leq N/(3k)$. Conditioning on the event $\mathcal{E}$,

$$\frac{1}{\binom{N}{k}} \sum_{\#(A)=k} \mathbb{P}\left(\forall i \in A,\ Z_i = 0 \,|\, \mathcal{E}\right) \geq \frac{\binom{N - \lfloor N/(3k) \rfloor}{k}}{\binom{N}{k}} \geq \left(1 - \frac{N/(3k)}{N-k}\right)^k > \left(1 - \frac{1}{2k}\right)^k \geq \frac{1}{2}.$$

However, by (24),

$$\varepsilon \geq \frac{1}{\binom{N}{k}} \sum_{\#(A)=k} \mathbb{P}\left(\forall i \in A,\ Z_i = 0\right) \geq \frac{1}{\binom{N}{k}} \sum_{\#(A)=k} \mathbb{P}(\mathcal{E}) \cdot \mathbb{P}\left(\forall i \in A,\ Z_i = 0 \,|\, \mathcal{E}\right) \geq \mathbb{P}(\mathcal{E})/2.$$

Hence $\mathbb{P}(\mathcal{E}) \leq 2\varepsilon$ and the lemma is proven. □

Sudakov's theorem is used in the following lemma:

Lemma 4.3. *Let $N \geq n$ and let $x_1, \ldots, x_N \in S^{n-1}$ be such that $\langle x_i, x_j \rangle \leq 49/50$ for any $i \neq j$. Then there exists $\mathcal{F} \subseteq S^{n-1}$ with $\sigma_{n-1}(\mathcal{F}) \geq 1 - C/n^c$ such that for any $\theta \in \mathcal{F}$,*

$$\frac{\#\left\{1 \leq i \leq N\,;\ \langle x_i, \theta \rangle \geq c_1 t/\sqrt{n}\right\}}{N} \geq c_2 e^{-C_3 t^2}, \qquad \text{for all } t \in [\sqrt{\log n}, \sqrt{\log N}], \tag{26}$$

where $c_1, c_2, C_3, c, C > 0$ are universal constants.

Proof. Denote $S = \{x_1, \ldots, x_N\} \subset S^{n-1}$ and note that $|x_i - x_j| \geq \sqrt{2 - 49/25} = 1/5$ for all $i \neq j$. Fix a number $t \in [\sqrt{\log n}, \sqrt{\log N}]$. Let $A \subseteq \{x_1, \ldots, x_N\}$ be any subset with $\#(A) \geq \exp(t^2)$. By Theorem 4.1,

$$M^*(A) \geq ct/\sqrt{n}. \tag{27}$$

Next we will apply the concentration inequality (23) with $r = M^*(A)/(2R(A))$. Since $R(A) = 1$, it follows from (23) and (27) that

$$\sigma_{n-1}\left(\left\{\theta \in S^{n-1}\,;\ h_A(\theta) \geq M^*(A)/2\right\}\right) \geq 1 - C\exp\left(-cn\left(\frac{M^*(A)}{R(A)}\right)^2\right) \geq 1 - \tilde{C}e^{-\tilde{c}t^2}.$$

Let Θ be a random vector, distributed uniformly over S^{n-1}. By combining the last inequality with (27), we see that for any fixed subset $\tilde{A} \subseteq \{1, \ldots, N\}$ with $\#(\tilde{A}) = \lceil \exp(t^2) \rceil$,

$$\mathbb{P}\left(\exists i \in \tilde{A}\,; \langle x_i, \Theta \rangle \geq ct/\sqrt{n}\right) \geq 1 - \tilde{C}e^{-\tilde{c}t^2}.$$

Let us now apply Lemma 4.2 for $Z_i = 1_{\{\langle x_i, \Theta \rangle \geq ct/\sqrt{n}\}}$. Lemma 4.2 now implies that with a probability of at least $1 - 2\tilde{C}e^{-\tilde{c}t^2}$ of selecting $\Theta \in S^{n-1}$,

$$\#\left\{1 \leq i \leq N\,;\ \langle x_i, \Theta \rangle \geq ct/\sqrt{n}\right\} \geq \frac{N}{3\lceil \exp(t^2) \rceil} \geq \frac{N}{6} \cdot e^{-t^2}.$$

We now let the parameter t vary. Let I be the collection of all integer powers of two that lie in the interval $[\sqrt{\log n}, \sqrt{\log N}]$. Then,

$$\mathbb{P}\left(\forall t \in I,\ \frac{\#\left\{1 \leq i \leq N\,;\ \langle x_i, \Theta \rangle \geq ct/\sqrt{n}\right\}}{N} \geq \frac{e^{-t^2}}{6}\right) \geq 1 - \sum_{t \in I} 2\tilde{C}e^{-\tilde{c}t^2} \geq 1 - \frac{\hat{C}}{n^{\hat{c}}}.$$

The restriction $t \in I$ may be upgraded to the condition $t \in [\sqrt{\log n}, \sqrt{\log N}]$ by adjusting the constants. The lemma is thus proven. □

Recall the construction of θ_1 and θ_2 from Section 2, and also the definition (8) of the parameter t_0. From the construction we see that for any $v \in S^{n-1}$ with $|\langle v, \theta_1 \rangle| \leq 1/10$,

$$\mathbb{P}\left(|X| \geq M \text{ and } \left| \frac{X}{|X|} - v \right| \leq \frac{1}{5} \right) \leq 2e^{-t_0^2}, \tag{28}$$

where $M > 0$ satisfies $\mathbb{P}(|X| \geq M) \geq 1/3$ and $\mathbb{P}(|X| \leq M) \geq 2/3$.

Lemma 4.4. *Assume that $t_0 \geq 5\sqrt{\log n}$ and set $N = \lfloor e^{t_0^2/4} \rfloor$. Let $X_1, \ldots, X_N$ be independent copies of X. Then with a probability of at least $1 - C/n$ of selecting $X_1, \ldots, X_N$, there exists $I \subseteq \{1, \ldots, N\}$ with the following three properties:*

(i) $\#(I) \geq N/10$.

(ii) For any $i, j \in I$ with $i \neq j$ we have $\langle X_i, X_j \rangle \leq (49/50) \cdot |X_i| \cdot |X_j|$.

(iii) For any $i \in I$,

$$|X_i| \geq M, \quad |\langle X_i, \theta_1 \rangle| \leq 10|X_i|/\sqrt{n} \quad \textit{and} \quad |\langle X_i, \theta_2 \rangle| \leq 10|X_i|/\sqrt{n}.$$

Here, $C > 0$ is a universal constant.

Proof. We may assume that $n \geq 10^4$, as otherwise for an appropriate choice of the constant C, all we claim is that a certain event holds with a non-negative probability. Write

$$\mathcal{A} = \{ v \in \mathbb{R}^n \, ; \, |v| \geq M, \, \max_{j=1,2} |\langle v/|v|, \theta_j \rangle| \leq 10/\sqrt{n} \}.$$

According to Lemma 3.2, for $i = 1, \ldots, N$,

$$\mathbb{P}(X_i \in \mathcal{A}) > 1/5.$$

Denote $I = \{ i = 1, \ldots, N \, ; \, X_i \in \mathcal{A} \}$. By the Chernoff large deviation bound,

$$\mathbb{P}(\#(I) \geq N/10) \geq 1 - C\exp(-cN).$$

Note that $10/\sqrt{n} \leq 1/10$ and that if $v \in \mathcal{A}$ then $|\langle v/|v|, \theta_1 \rangle| \leq 1/10$. It thus follows from (28) that for any $i, j \in \{1, \ldots, N\}$ with $i \neq j$,

$$\begin{aligned} &\mathbb{P}\left(i, j \in I \text{ and } \left| \frac{X_j}{|X_j|} - \frac{X_i}{|X_i|} \right| \leq \frac{1}{5} \right) \\ &\leq \mathbb{P}\left(X_j \in \mathcal{A} \text{ and } \left| \frac{X_j}{|X_j|} - \frac{X_i}{|X_i|} \right| \leq \frac{1}{5} \;\middle|\; X_i \in \mathcal{A} \right) \leq 2e^{-t_0^2} \leq \frac{2}{N^4}. \end{aligned}$$

Consequently,

$$\mathbb{P}\left(\exists i, j \in I \text{ with } i \neq j \text{ and } \left| \frac{X_i}{|X_i|} - \frac{X_j}{|X_j|} \right| \leq \frac{1}{5} \right) \leq \frac{N(N-1)}{2} \cdot \frac{2}{N^4} \leq \frac{1}{N^2}.$$

We conclude that with a probability of at least $1 - C\exp(-cN) - 1/N^2 \geq 1 - \tilde{C}/n$,

$$\#(I) \geq N/10 \qquad \text{and} \qquad \forall i,j \in I, i \neq j \implies \left| \frac{X_i}{|X_i|} - \frac{X_j}{|X_j|} \right| > \frac{1}{5}.$$

Note that $\langle X_i, X_j \rangle \leq (49/50) \cdot |X_i| \cdot |X_j|$ if and only if $|X_i/|X_i| - X_j/|X_j|| \geq 1/5$. Thus conclusions (i), (ii) and (iii) hold true with a probability of at least $1 - \tilde{C}/n$, thereby completing the proof. $\square$

By combining Lemma 4.3 and Lemma 4.4 we arrive at the following:

Lemma 4.5. *Assume that $t_0 \geq 5\sqrt{\log n}$ and set $N = \lfloor e^{t_0^2/4} \rfloor$. Then there exists a fixed subset $\mathcal{F} \subseteq S^{n-1}$ with $\sigma_{n-1}(\mathcal{F}) \geq 1 - C/n^c$ such that for any $\theta_3 \in \mathcal{F}$ the following holds: Define θ via (5). Let $X_1, \ldots, X_N$ be independent copies of the random vector X. Then with a probability of at least $1 - \tilde{C}/n^{\tilde{c}}$ of selecting $X_1, \ldots, X_N$,*

$$\frac{\#\left\{1 \leq i \leq N\,;\, \langle X_i, \theta \rangle \geq c_1 \frac{M}{\sqrt{n}} \cdot t\right\}}{N} \geq c_2 e^{-C_3 t^2}, \qquad \textit{for all } t \in [\sqrt{\log n}, t_0], \tag{29}$$

and

$$\frac{\#\left\{1 \leq i \leq N\,;\, \langle X_i, \theta \rangle \leq -c_1 \frac{M}{\sqrt{n}} \cdot t\right\}}{N} \geq c_2 e^{-C_3 t^2}, \qquad \textit{for all } t \in [\sqrt{\log n}, t_0]. \tag{30}$$

Here, $c_1, c_2, C_3, c, C, \tilde{c}, \tilde{C} > 0$ are universal constants.

Proof. This proof is almost identical to the deduction of Lemma 3.5 from Lemma 3.3 and Lemma 3.4. Let us spell out the details. Set $\vec{X} = (X_1, \ldots, X_N)$ and let Θ be a random vector, independent of $\vec{X}$, distributed uniformly on S^{n-1}. We say that $\vec{X} \in \mathcal{A}_1$ if the event described in Lemma 4.4 holds true. Thus,

$$\mathbb{P}(\vec{X} \in \mathcal{A}_1) \geq 1 - C/n.$$

Assuming that $\vec{X} \in \mathcal{A}_1$, we may apply Lemma 4.3 and obtain that with a probability of at least $1 - \tilde{C}/n^{\tilde{c}}$ of selecting $\Theta \in S^{n-1}$,

$$\#\left\{1 \leq i \leq N\,;\, \left\langle \frac{X_i}{|X_i|}, \Theta \right\rangle \geq c_1 t/\sqrt{n}\right\} \geq c_2 e^{-C_3 t^2} \cdot (N/10) \quad \text{for all } t \in [\sqrt{\log n}, \sqrt{\log N}].$$

Assuming that $\vec{X} \in \mathcal{A}_1$, we may use Lemma 4.4(iii) in order to conclude that with a probability of at least $1 - \tilde{C}/n^{\tilde{c}}$ of selecting $\Theta \in S^{n-1}$, for $t \in [\sqrt{\log n}, 4\sqrt{\log N}]$,

$$\#\left\{1 \leq i \leq N\,;\, \left\langle X_i, \frac{\theta_1 - \theta_2 + \Theta}{|\theta_1 - \theta_2 + \Theta|} \right\rangle \geq \bar{c}_1 \frac{M}{\sqrt{n}} \cdot t\right\} \geq \bar{c}_2 e^{-\bar{C}_3 t^2} \cdot N. \tag{31}$$

Write $\mathcal{A}_2$ for the event that (31) holds true for all $t \in [\sqrt{\log n}, 4\sqrt{\log N}]$. Thus,

$$\mathbb{P}((\Theta, \vec{X}) \in \mathcal{A}_2) \geq 1 - C/n - \tilde{C}/n^{\tilde{c}} \geq 1 - \bar{C}/n^{\bar{c}}.$$

Consequently, there exists $\mathcal{F} \subseteq S^{n-1}$ with

$$\sigma_{n-1}(\mathcal{F}) \geq 1 - \hat{C}/n^{\hat{c}}$$

with the following property: For any $\theta_3 \in \mathcal{F}$, with a probability of at least $1 - \hat{C}/n^{\hat{c}}$ of selecting $X_1, \ldots, X_N$, for all $t \in [\sqrt{\log n}, 4\sqrt{\log N}]$,

$$\#\left\{1 \leq i \leq N\,;\, \left\langle X_i, \frac{\theta_1 - \theta_2 + \theta_3}{|\theta_1 - \theta_2 + \theta_3|} \right\rangle \geq c_1 \frac{M}{\sqrt{n}} \cdot t \right\} \geq c_2 e^{-C_3 t^2} \cdot N.$$

Recalling that $4\sqrt{\log N} \geq t_0$, we have established (29). The proof of (30) is similar. $\square$

The short proof of the following corollary is analogous to that of Corollary 3.6.

Corollary 4.6. *There exists a fixed subset $\mathcal{F} \subseteq S^{n-1}$ with $\sigma_{n-1}(\mathcal{F}) \geq 1 - C/n^c$ such that for any $\theta_3 \in \mathcal{F}$, defining θ via (5),*

$$\forall t \in [\sqrt{\log n}, t_0], \quad \min\left\{\mathbb{P}\left(\langle X, \theta \rangle \geq c_1 \frac{M}{\sqrt{n}} \cdot t\right), \mathbb{P}\left(\langle X, \theta \rangle \leq -c_1 \frac{M}{\sqrt{n}} \cdot t\right)\right\} \geq c_2 e^{-C_3 t^2},$$

where $c, C, c_1, c_2, C_3 > 0$ are universal constants.

Proof. We may assume that n exceeds a certain fixed universal constant. Let $\mathcal{F}$ be the set from Lemma 4.5, denote $N = \lfloor \exp(t_0^2/4) \rfloor$, and let $X_1, \ldots, X_N$ be independent copies of X. Then for any $\theta_3 \in \mathcal{F}$, defining θ via (5) we have that for any $t \in [\sqrt{\log n}, t_0]$,

$$\mathbb{P}\left(\langle X, \theta \rangle \geq c_1 \frac{M}{\sqrt{n}} \cdot t\right) \geq c_2 e^{-C_3 t^2} \cdot \mathbb{P}\left(\frac{\#\left\{i; \langle X_i, \theta \rangle \geq c_1 \frac{M}{\sqrt{n}} \cdot t\right\}}{N} \geq c_2 e^{-C_3 t^2}\right) \geq \frac{c_2}{2} e^{-C_3 t^2},$$

where the last passage is the content of Lemma 4.5. The bound for $\mathbb{P}\left(\langle X, \theta \rangle \leq -c_1 t M/\sqrt{n}\right)$ is proven similarly. $\square$

5 Proof of the main proposition

In this section we complete the proof of Proposition 2.1. We begin with the following standard observation:

Lemma 5.1. *Suppose that X is a random vector in $\mathbb{R}^n$ with $\mathbb{P}(X = 0) = 0$. Then there exists a fixed subset $\mathcal{F} \subseteq S^{n-1}$ of full measure, such that $\mathbb{P}(\langle X, \theta \rangle = 0) = 0$ for all $\theta \in \mathcal{F}$.*

Proof. For $a > 0$, we say that a subspace $E \subseteq \mathbb{R}^n$ is a-basic if $\mathbb{P}(X \in E) \geq a$ while $\mathbb{P}(X \in F) < a$ for all subspaces $F \subsetneq E$. Lemma 7.1 in [5] states that there are only finitely many subspaces that are a-basic for any fixed $a > 0$. Write $\mathcal{S}$ for the collection of all subspaces that are a-basic for some rational number $a > 0$. Then $\mathcal{S}$ is a countable family which does not contain the subspace $\{0\}$. Consequently, the set

$$\mathcal{F} = \{\theta \in S^{n-1}\, ; \, \forall E \in \mathcal{S}, \ E \not\subset \theta^{\perp}\}$$

is a set of full measure in S^{n-1}, as its complement is the countable union of spheres of lower dimension. Here, $\theta^{\perp} = \{x \in \mathbb{R}^n\, ;\, \langle x, \theta \rangle = 0\}$. Suppose that $\theta \in \mathcal{F}$, and let us prove that $\mathbb{P}(\langle X, \theta\rangle = 0) = 0$. Otherwise, there exists a rational number $a > 0$ such that

$$\mathbb{P}(\langle X, \theta \rangle = 0) \geq a.$$

Thus $\theta^{\perp}$ contains an a-basic subspace, contradicting the definition of $\mathcal{F}$. □

Recall the definition of M, θ_1 and θ_2 from Section 2.

Lemma 5.2. *Let* $\mathcal{F}_3 \subseteq \left\{\theta_3 \in S^{n-1}\, ;\, |\langle \theta_3, \theta_1 \rangle| \leq \frac{1}{10} \text{ and } |\langle \theta_3, \theta_2\rangle| \leq \frac{1}{10}\right\}$. *Then for any* $\theta_3 \in \mathcal{F}_3$ *and* $v \in S^{n-1}$,

$$|v - \theta_1| \leq \frac{1}{5} \qquad \Longrightarrow \qquad \langle v, \theta_1 - \theta_2 + \theta_3 \rangle \geq \frac{1}{10}, \tag{32}$$

and

$$|v - \theta_2| \leq \frac{1}{5} \qquad \Longrightarrow \qquad \langle v, \theta_1 - \theta_2 + \theta_3 \rangle \leq -\frac{1}{10}. \tag{33}$$

Proof. Recall that $|\langle \theta_1, \theta_2 \rangle| \leq 1/10$. Note that for any $\theta_3 \in \mathcal{F}_3$ and $i, j \in \{1, 2, 3\}$ with $i \neq j$,

$$\sqrt{9/5} \leq |\theta_i - \theta_j| \leq \sqrt{11/5}.$$

Let $v \in S^{n-1}$ be any vector with $|v - \theta_1| \leq 1/5$. Then for any $\theta_3 \in \mathcal{F}_3$ and $j = 2, 3$ we have that

$$\sqrt{\frac{9}{5}} - \frac{1}{5} \leq |\theta_j - \theta_1| - |\theta_1 - v| \leq |v - \theta_j| \leq |\theta_j - \theta_1| + |\theta_1 - v| \leq \sqrt{\frac{11}{5}} + \frac{1}{5},$$

and hence for $j = 2, 3$,

$$\langle v, \theta_j \rangle = 1 - \frac{1}{2} \cdot |v - \theta_j|^2 \in \left[1 - \frac{1}{2} \cdot \left(\sqrt{\frac{11}{5}} + \frac{1}{5}\right)^2, 1 - \frac{1}{2} \cdot \left(\sqrt{\frac{9}{5}} - \frac{1}{5}\right)^2\right] \subseteq \left[-\frac{3}{7}, \frac{3}{7}\right]. \tag{34}$$

However, $\langle v, \theta_1 \rangle \geq 49/50$ for such v, and hence (32) follows from (34). By replacing the triplet $(\theta_1, \theta_2, \theta_3)$ by $(\theta_2, \theta_1, -\theta_3)$ and repeating the above argument, we obtain (33). □

Proof of Proposition 2.1. From Corollary 3.6 and Corollary 4.6 we learn that there exists $\mathcal{F} \subseteq S^{n-1}$ with $\sigma_{n-1}(\mathcal{F}_3) \geq 1 - C/n^c$ such that for any $\theta_3 \in \mathcal{F}$, defining θ via (5),

$$\forall t \in [0, t_0], \ \min\left\{\mathbb{P}\left(\langle X, \theta\rangle \geq c_1 \frac{M}{\sqrt{n}} \cdot t\right), \mathbb{P}\left(\langle X, \theta\rangle \leq -c_1 \frac{M}{\sqrt{n}} \cdot t\right)\right\} \geq c_2 e^{-C_3 t^2}. \quad (35)$$

According to Lemma 5.1, we may remove a set of measure zero from $\mathcal{F}$ and additionally assume that $\mathbb{P}(\langle X, \theta\rangle = 0) = 0$. From Lemma 3.1 we learn that any median of $|\langle X, \theta\rangle|$ is at most $CM/\sqrt{n}$. Hence (35) shows that for any $\theta_3 \in \mathcal{F}$, defining θ via (5) we have that $\langle X, \theta\rangle$ is Super-Gaussian of length $c_1 t_0$, with parameters $c_2, c_3 > 0$. We still need to increase the length to $c_1\sqrt{n}$. To this end, denote

$$\mathcal{F}_3 = \left\{\theta_3 \in \mathcal{F} \ ; \ |\langle \theta_3, \theta_1\rangle| \leq \frac{1}{10} \quad \text{and} \quad |\langle \theta_3, \theta_2\rangle| \leq \frac{1}{10}\right\}.$$

Then $\sigma_{n-1}(\mathcal{F}_3) \geq \sigma_{n-1}(\mathcal{F}) - C\exp(-cn) \geq 1 - \tilde{C}/n^{\tilde{c}}$. Recall from Section 2 that for $j = 1, 2$,

$$\mathbb{P}\left(|X| \geq M \text{ and } \left|\frac{X}{|X|} - \theta_j\right| \leq \frac{1}{5}\right) \geq \frac{1}{2} \cdot e^{-t_0^2}. \quad (36)$$

Let us fix $t \in [t_0, \sqrt{n}], \theta_3 \in \mathcal{F}_3$ and define θ via (5). Since $0 < |\theta_1 - \theta_2 + \theta_3| \leq 3$, by (36) and Lemma 5.2,

$$\begin{aligned}\mathbb{P}\left(\langle X, \theta\rangle \geq \frac{Mt}{30\sqrt{n}}\right) &\geq \mathbb{P}\left(\langle X, \theta_1 - \theta_2 + \theta_3\rangle \geq \frac{Mt}{10\sqrt{n}}\right) \geq \mathbb{P}\left(\left\langle \frac{X}{|X|}, \theta_1 - \theta_2 + \theta_3\right\rangle \geq \frac{M}{10|X|}\right) \\ &\geq \mathbb{P}\left(|X| \geq M, \ \left|\frac{X}{|X|} - \theta_1\right| \leq \frac{1}{5}\right) \geq \frac{1}{2} \cdot e^{-t_0^2} \geq \frac{1}{2} \cdot e^{-t^2}.\end{aligned}$$

Similarly,

$$\begin{aligned}\mathbb{P}\left(\langle X, \theta\rangle \leq -\frac{Mt}{30\sqrt{n}}\right) &\geq \mathbb{P}\left(\left\langle \frac{X}{|X|}, \theta_1 - \theta_2 + \theta_3\right\rangle \leq -\frac{M}{10|X|}\right) \\ &\geq \mathbb{P}\left(|X| \geq M, \ \left|\frac{X}{|X|} - \theta_2\right| \leq \frac{1}{5}\right) = e^{-t_0^2} \geq e^{-t^2}.\end{aligned}$$

Therefore, we may upgrade (35) to the following statement: For any $\theta_3 \in \mathcal{F}$ and $t \in [0, \sqrt{n}]$, defining θ via (5),

$$\min\left\{\mathbb{P}\left(\langle X, \theta\rangle \geq c_1 \frac{M}{\sqrt{n}} \cdot t\right), \mathbb{P}\left(\langle X, \theta\rangle \leq -\hat{c}_1 \frac{M}{\sqrt{n}} \cdot t\right)\right\} \geq \hat{c}_2 e^{-\hat{C}_3 t^2}.$$

We have thus proven that $\langle X, \theta\rangle$ is Super-Gaussian of length $\overline{c}_1\sqrt{n}$ with parameters $\overline{c}_2, \overline{c}_3 > 0$. □

6 Angularly-isotropic position

In this section we deduce Theorem 1.3 from Proposition 2.1 by using the angularly-isotropic position which is discussed below. We begin with the following:

Lemma 6.1. *Let $d, X, \mathcal{B}$ be as in Theorem 1.3. Set $n = \lceil d \rceil$. Then there exists a fixed linear map $T : \mathcal{B} \to \mathbb{R}^n$ such that for any $\varepsilon > 0$, the random vector $T(X)$ is of class* $\mathrm{eff.rank}_{\geq d-\varepsilon}$.

Proof. We will show that a generic linear map T works. Denote $N = \dim(\mathcal{B})$ and identify $\mathcal{B} \cong \mathbb{R}^N$. Since the effective rank of X is at least d, necessarily $d \leq N$ and hence also $n = \lceil d \rceil \leq N$. Let $L \subseteq \mathbb{R}^N$ be a random n-dimensional subspace, distributed uniformly in the Grassmannian $G_{N,n}$. Denote $T = \mathrm{Proj}_L : \mathbb{R}^N \to L$, the orthogonal projection operator onto the subspace L.

For any fixed subspace $E \subseteq \mathbb{R}^N$, with probability one of selecting $L \in G_{N,n}$,

$$\dim(\ker(T) \cap E) = \max\{0, \dim(E) - n\},$$

or equivalently,

$$\dim(T(E)) = \dim(E) - \dim(\ker(T) \cap E) = \min\{n, \dim(E)\}. \tag{37}$$

Recall that for $a > 0$, a subspace $E \subseteq \mathbb{R}^N$ is a-basic if $\mathbb{P}(X \in E) \geq a$ while $\mathbb{P}(X \in F) < a$ for all subspaces $F \subsetneq E$. Lemma 7.1 in [5] states that there exist only countably many subspaces that are a-basic with a being a positive, rational number. Write $\mathcal{G}$ for the collection of all these basic subspaces. Then with probability one of selecting $L \in G_{N,n}$,

$$\forall E \in \mathcal{G}, \qquad \dim(T(E)) = \min\{n, \dim(E)\}. \tag{38}$$

We now fix a subspace $L \in G_{N,n}$ for which $T = \mathrm{Proj}_L$ satisfies (38). Let $S \subseteq L$ be any subspace and assume that $a \in \mathbb{Q} \cap (0, 1]$ satisfies

$$\mathbb{P}(T(X) \in S) \geq a.$$

Then $\mathbb{P}(X \in T^{-1}(S)) \geq a$. Therefore $T^{-1}(S)$ contains an a-basic subspace E. Thus $E \in \mathcal{G}$ while $E \subseteq T^{-1}(S)$ and $\mathbb{P}(X \in E) \geq a$. Since the effective rank of X is at least d, necessarily $\dim(E) \geq a \cdot d$. Since $T(E) \subseteq S$, from (38),

$$\dim(S) \geq \dim(T(E)) = \min\{n, \dim(E)\} \geq \min\{n, \lceil a \cdot d \rceil\} = \lceil a \cdot d \rceil.$$

We have thus proven that for any subspace $S \subseteq L$ and $a \in \mathbb{Q} \cap (0, 1]$,

$$\mathbb{P}(T(X) \in S) \geq a \qquad \Longrightarrow \qquad \dim(S) \geq \lceil a \cdot d \rceil. \tag{39}$$

It follows from (39) that for any subspace $S \subseteq L$,

$$\mathbb{P}(T(X) \in S) \leq \dim(S)/d.$$

This implies that for any $\varepsilon > 0$, the random vector $T(X)$ is of class $\mathrm{eff.rank}_{\geq d-\varepsilon}$. $\square$

Lemma 6.2. *Let $d, X, \mathcal{B}$ be as in Theorem 1.3. Assume that $d < \dim(\mathcal{B})$ and that for any subspace $\{0\} \neq E \subsetneq \mathcal{B}$,*

$$\mathbb{P}(X \in E) < \dim(E)/d. \tag{40}$$

Then there exists $\varepsilon > 0$ such that X is of class $\mathrm{eff.rank}_{\geq d+\varepsilon}$.

Proof. Since the effective rank of X is at least d, necessarily $\mathbb{P}(X = 0) = 0$. Assume by contradiction that for any $\varepsilon > 0$, the random vector X is not of class $\mathrm{eff.rank}_{\geq d+\varepsilon}$. Then for any $\varepsilon > 0$ there exists a subspace $\{0\} \neq E \subseteq \mathcal{B}$ with

$$\mathbb{P}(X \in E) \geq -\varepsilon + \dim(E)/d.$$

The Grassmannian of all k-dimensional subspaces of $\mathcal{B}$ is compact. Hence there is a dimension $1 \leq k \leq \dim(\mathcal{B})$ and a converging sequence of k-dimensional subspaces $E_1, E_2, \ldots \subseteq \mathcal{B}$ with

$$\mathbb{P}(X \in E_\ell) \geq -1/\ell + \dim(E_\ell)/d = -1/\ell + k/d \qquad \text{for all } \ell \geq 1. \tag{41}$$

Denote $E_0 = \lim_\ell E_\ell$, which is a k-dimensional subspace in $\mathcal{B}$. Let $U \subseteq \mathcal{B}$ be an open neighborhood of E_0 with the property that $tx \in U$ for all $x \in U, t \in \mathbb{R}$. Then $E_\ell \subseteq U$ for a sufficiently large ℓ, and we learn from (41) that

$$\mathbb{P}(X \in U) \geq k/d. \tag{42}$$

Since E_0 is the intersection of a decreasing sequence of such neighborhoods U, it follows from (42) that

$$\mathbb{P}(X \in E_0) \geq k/d = \dim(E_0)/d. \tag{43}$$

Since $d < \dim(\mathcal{B})$, the inequality in (43) shows that $E_0 \neq \mathcal{B}$. Hence $1 \leq \dim(E_0) \leq \dim(\mathcal{B})-1$, and (43) contradicts (40). The lemma is thus proven. □

The following lemma is a variant of Lemma 5.4 from [5].

Lemma 6.3. *Let $d, X, \mathcal{B}$ be as in Theorem 1.3. Then there exists a fixed scalar product $\langle \cdot, \cdot \rangle$ on $\mathcal{B}$ such that denoting $|\theta| = \sqrt{\langle \theta, \theta \rangle}$, we have*

$$\mathbb{E} \left\langle \frac{X}{|X|}, \theta \right\rangle^2 \leq \frac{|\theta|^2}{d} \qquad \textit{for all } \theta \in \mathcal{B}. \tag{44}$$

Proof. By induction on the dimension $n = \dim(\mathcal{B})$. Assume first that there exists a subspace $\{0\} \neq E \subsetneq \mathcal{B}$, such that equality holds true in (3). In this case, there exists a subspace $F \subseteq \mathcal{B}$ with $E \oplus F = \mathcal{B}$ and $\mathbb{P}(X \in E \cup F) = 1$. We will construct a scalar product in $\mathcal{B}$ as follows: Declare that E and F are orthogonal subspaces, and use the induction hypothesis in order to find appropriate scalar products in the subspace E and in the subspace F. This induces a scalar product in $\mathcal{B}$ which satisfies

$$\mathbb{E} \left\langle \frac{X}{|X|}, \theta \right\rangle^2 \leq \frac{|\theta|^2}{d} \qquad \text{for all } \theta \in E \cup F.$$

For any $\theta \in \mathcal{B}$ we may decompose $\theta = \theta_E + \theta_F$ with $\theta_E \in E, \theta_F \in F$. Since $\mathbb{P}(X \in E \cup F) = 1$, we obtain

$$\mathbb{E}\left\langle \frac{X}{|X|}, \theta \right\rangle^2 = \mathbb{E}\left\langle \frac{X}{|X|}, \theta_E \right\rangle^2 + \mathbb{E}\left\langle \frac{X}{|X|}, \theta_F \right\rangle^2 \leq \frac{|\theta_E|^2 + |\theta_F|^2}{d} = \frac{|\theta|^2}{d},$$

proving (44).

Next, assume that for any subspace $\{0\} \neq E \subsetneq \mathcal{B}$, the inequality in (3) is strict. There are two distinct cases, either $d = n$ or $d < n$. Consider first the case where $d = n = \dim(\mathcal{B})$. Thus, for any subspace $E \subseteq \mathcal{B}$ with $E \neq \{0\}$ and $E \neq \mathcal{B}$,

$$\mathbb{P}(X \in E) < \dim(E)/n.$$

This is precisely the main assumption of Corollary 5.3 in [5]. By the conclusion of the corollary, there exists a scalar product in $\mathcal{B}$ such that (44) holds true. We move on to the case where $d < n$. Here, we apply Lemma 6.2 and conclude that X is of class $\mathrm{eff.rank}_{\geq d+\varepsilon}$ for some $\varepsilon > 0$. Therefore, for some $\varepsilon > 0$,

$$\mathbb{P}(X \in E) < \dim(E)/(d+\varepsilon) \qquad \forall E \subseteq \mathcal{B}. \tag{45}$$

Now we invoke Lemma 5.4 from [5]. Its assumptions are satisfies thanks to (45). From the conclusion of that lemma, there exists a scalar product in $\mathcal{B}$ for which (44) holds true. □

The condition that the effective rank of X is at least d is not only sufficient but is also necessary for the validity of conclusion (44) from Lemma 6.3. Indeed, it follows from (44) that for any subspace $E \subseteq \mathcal{B}$,

$$\mathbb{P}(X \in E) \leq \mathbb{E}\left|\mathrm{Proj}_E \frac{X}{|X|}\right|^2 = \sum_{i=1}^{\dim(E)} \mathbb{E}\left\langle \frac{X}{|X|}, u_i \right\rangle^2 \leq \frac{\dim(E)}{d}, \tag{46}$$

where $u_1, \ldots, u_m$ is an orthonormal basis of the subspace E with $m = \dim(E)$. Equality in (46) holds true if and only if $\mathbb{P}(X \in E \cup E^{\perp}) = 1$, where $E^{\perp}$ is the orthogonal complement to E. Consequently, the effective rank of X is at least d.

Definition 6.4. *Let X be a random vector in $\mathbb{R}^n$ with $\mathbb{P}(X = 0) = 0$. We say that X is angularly-isotropic if*

$$\mathbb{E}\left\langle \frac{X}{|X|}, \theta \right\rangle^2 = \frac{1}{n} \qquad \text{for all } \theta \in S^{n-1}. \tag{47}$$

For $0 < d \leq n$ we say that $X/|X|$ is sub-isotropic with parameter d if

$$\mathbb{E}\left\langle \frac{X}{|X|}, \theta \right\rangle^2 \leq \frac{1}{d} \qquad \text{for all } \theta \in S^{n-1}. \tag{48}$$

We observe that X is angularly-isotropic if and only if $X/|X|$ is sub-isotropic with parameter n. Indeed, suppose that (48) holds true with $d = n$. Given any $\theta \in S^{n-1}$ we may find an orthonormal basis $\theta_1, \ldots, \theta_n \in \mathbb{R}^n$ with $\theta_1 = \theta$. Hence

$$1 = \mathbb{E}\left|\frac{X}{|X|}\right|^2 = \mathbb{E}\sum_{i=1}^n \left\langle \frac{X}{|X|}, \theta_i \right\rangle^2 \leq \sum_{i=1}^n \frac{1}{n} = 1,$$

and (47) is proven.

Proof of Theorem 1.3. According to Lemma 6.1, we may project X to a lower-dimensional space, and assume that $\dim(\mathcal{B}) = n = \lceil d \rceil$ and that the effective rank of X is at least $n/2$. Lemma 6.3 now shows that there exists a scalar product in $\mathcal{B}$ with respect to which $X/|X|$ is sub-isotropic with parameter $n/2$. We may therefore identify $\mathcal{B}$ with $\mathbb{R}^n$ so that

$$\mathbb{E}\left\langle \frac{X}{|X|}, \theta \right\rangle^2 \leq \frac{2}{n} \qquad \text{for all } \theta \in S^{n-1}.$$

Thus condition (4) of Proposition 2.1 is verified. By the conclusion of Proposition 2.1, there exists a non-zero linear functional $\ell : \mathbb{R}^n \to \mathbb{R}$ such that $\ell(X)$ is Super-Gaussian of length $c_1\sqrt{n} \geq c\sqrt{d}$ with parameters $c_2, c_3 > 0$. □

Proof of Corollary 1.4. By assumption, $\mathbb{P}(X \in E) \leq \dim(E)/d$ for any finite-dimensional subspace $E \subseteq \mathcal{B}$. Lemma 7.2 from [5] states that there exists a continuous, linear map $T : \mathcal{B} \to \mathbb{R}^N$ such that $T(X)$ has an effective rank of at least $d/2$. We may now invoke Theorem 1.3 for the random vector $T(X)$, and conclude that for some non-zero, fixed, linear functional $\ell : \mathbb{R}^N \to \mathbb{R}$, the random variable $(\ell \circ T)(X)$ is Super-Gaussian of length $c_1\sqrt{d}$ with parameters $c_2, c_3 > 0$. □

Remark 6.5. We were asked by Yaron Oz about analogs of Theorem 1.1 in the hyperbolic space. We shall work with the standard hyperboloid model

$$\mathbb{H}^n = \left\{ (x_0, \ldots, x_n) \in \mathbb{R}^{n+1} \, ; \, -x_0^2 + \sum_{i=1}^n x_i^2 = -1, x_0 > 0 \right\}$$

where the Riemannian metric tensor is $g = -dx_0^2 + \sum_{i=1}^n dx_i^2$. For any linear subspace $L \subseteq \mathbb{R}^{n+1}$, the intersection $L \cap \mathbb{H}^n$ is a totally-geodesic submanifold of $\mathbb{H}^n$ which is called a hyperbolic subspace. When we discuss the dimension of a hyperbolic subspace, we refer to its dimension as a smooth manifold. Note that an $(n-1)$-dimensional hyperbolic subspace $E \subseteq \mathbb{H}^n$ divides $\mathbb{H}^n$ into two sides. A signed distance function $d_E : \mathbb{H}^n \to \mathbb{R}$ is a function that equals the hyperbolic distance to E on one of these sides, and minus the distance to E on the other side. Given a linear functional $\ell : \mathbb{R}^{n+1} \to \mathbb{R}$ such that $E = \mathbb{H}^n \cap \{x \in \mathbb{R}^{n+1} \, ; \, \ell(x) = 0\}$ we may write

$$d_E(x) = \operatorname{arcsinh}(\alpha \cdot \ell(x)) \qquad (x \in \mathbb{H}^n)$$

for some $0 \neq \alpha \in \mathbb{R}$. It follows from Theorem 1.3 that for any absolutely-continuous random vector X in $\mathbb{H}^n$, there exists an $(n-1)$-dimensional hyperbolic subspace $E \subseteq \mathbb{H}^n$ and an associated signed distance function d_E such that the random variable $\sinh(d_E(X))$ is Super-Gaussian of length $c_1\sqrt{n}$ with parameters $c_2, c_3 > 0$. In general, we cannot replace the random variable $\sinh(d_E(X))$ in the preceding statement by $d_E(X)$ itself. This is witnessed by the example of the random vector

$$X = \left(\sqrt{1 + R^2 \sum_{i=1}^n Z_i^2}, RZ_1, \ldots, RZ_n \right) \in \mathbb{R}^{n+1}$$

which is supported in $\mathbb{H}^n$. Here, $Z_1, \ldots, Z_n$ are independent standard Gaussian random variables, and $R > 1$ is a fixed, large parameter.

References

[1] Böröczky, K., Lutwak, E., Yang, D., Zhang, G., *The logarithmic Minkowski problem.* J. Amer. Math. Soc., Vol. 26, No. 3, (2013), 831-852.

[2] Böröczky, K., Lutwak, E., Yang, D., Zhang, G., *Affine images of isotropic measures.* J. Diff. Geom., Vol. 99, No. 3, (2015), 407-442.

[3] Boucheron, S., Lugosi, G., Massart, P., *Concentration inequalities. A nonasymptotic theory of independence.* Oxford University Press, Oxford, 2013.

[4] Henk, M., Linke., E., *Cone-volume measures of polytopes.* Adv. Math., Vol. 253, (2014), 50-62.

[5] Klartag, B., *On nearly radial marginals of high-dimensional probability measures.* J. Eur. Math. Soc., Vol. 12, No. 3, (2010), 723-754.

[6] Ledoux, M., Talagrand, M., *Probability in Banach spaces. Isoperimetry and processes.* Ergeb. Math. Grenzgeb. (3), Vol. 23, Springer-Verlag, Berlin, 1991.

[7] Milman, V. D., Schechtman, G., *Asymptotic theory of finite-dimensional normed spaces.* Lecture Notes in Math., Vol. 1200. Springer-Verlag, Berlin, 1986.

[8] Paouris, G., *On the existence of supergaussian directions on convex bodies.* Mathematika, Vol. 58, No. 2, (2012), 389-408.

[9] Pivovarov, P., *On the volume of caps and bounding the mean-width of an isotropic convex body.* Math. Proc. Cambridge Philos. Soc., Vol. 149, no. 2, (2010), 317-331.

School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel.
E-mail: klartagb@tau.ac.il