

Variations on the Berry-Esseen theorem

Bo'az Klartag¹ and Sasha Sodin²

Abstract

Suppose that X_1, \dots, X_n are independent, identically-distributed random variables of mean zero and variance one. Assume that $\mathbb{E}|X_1|^4 \leq \delta^4$. We observe that there exist many choices of coefficients $\theta_1, \dots, \theta_n \in \mathbb{R}$ with $\sum_j \theta_j^2 = 1$ for which

$$\sup_{\substack{\alpha, \beta \in \mathbb{R} \\ \alpha < \beta}} \left| \mathbb{P} \left(\alpha \leq \sum_{j=1}^n \theta_j X_j \leq \beta \right) - \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-t^2/2} dt \right| \leq \frac{C\delta^4}{n}, \quad (1)$$

where $C > 0$ is a universal constant. Inequality (1) should be compared with the classical Berry-Esseen theorem, according to which the left-hand side of (1) may decay with n at the slower rate of $O(1/\sqrt{n})$, for the unit vector $\theta = (1, \dots, 1)/\sqrt{n}$. An explicit, universal example for coefficients $\theta = (\theta_1, \dots, \theta_n)$ for which (1) holds is

$$\theta = (1, \sqrt{2}, -1, -\sqrt{2}, 1, \sqrt{2}, -1, -\sqrt{2}, \dots) / \sqrt{3n/2}$$

when n is divisible by four. Parts of the argument are applicable also in the more general case, in which X_1, \dots, X_n are independent random variables of mean zero and variance one, yet they are not necessarily identically distributed. In this general setting, the bound (1) holds with $\delta^4 = n^{-1} \sum_{j=1}^n \mathbb{E}|X_j|^4$ for most selections of a unit vector $\theta = (\theta_1, \dots, \theta_n) \in \mathbb{R}^n$. Here “most” refers to the uniform probability measure on the unit sphere.

1 Introduction

This note brings further evidence for the fundamental rôle played by the geometry of the high-dimensional sphere in the analysis of the central limit theorem, in the spirit of works by Sudakov [14], Diaconis and Freedman [5] and others. Suppose that X_1, \dots, X_n are

¹Supported in part by the Israel Science Foundation and by a Marie Curie Reintegration Grant from the Commission of the European Communities.

²Supported in part by the Adams Fellowship Program of the Israel Academy of Sciences and Humanities and by the Israel Science Foundation.

independent random variables with finite third moments such that $\mathbb{E}X_j = 0$ and $\mathbb{E}X_j^2 = 1$ for all j . The classical Berry-Esseen theorem (see, e.g., Feller [7, Vol. II, Chapter XVI]) states that

$$\sup_{\substack{\alpha, \beta \in \mathbb{R} \\ \alpha < \beta}} \left| \mathbb{P} \left(\alpha \leq \frac{1}{\sqrt{n}} \sum_{j=1}^n X_j \leq \beta \right) - \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-t^2/2} dt \right| \leq \frac{C\gamma^3}{\sqrt{n}} \quad (2)$$

where $\gamma = \left(\sum_j \mathbb{E}|X_j|^3/n \right)^{1/3} \leq \max_j (\mathbb{E}|X_j|^3)^{1/3}$ and $C > 0$ is a universal constant. In the general case, where X_1, \dots, X_n are non-symmetric random variables, the bound (2) is sharp. Even when X_1, \dots, X_n are symmetric random variables, the bound (2) may not be improved in general: If the random variables are symmetric Bernoulli variables, for instance, then the probability $\mathbb{P}(\sum_j X_j = 0)$ is approximately $(\pi n/2)^{-1/2}$ for large even n . Therefore (2) is an asymptotically optimal bound in this case, up to the value of the constant C .

Quite unexpectedly, we find that there exists a linear combination of the random variables X_1, \dots, X_n that is much closer to the standard gaussian distribution. As it turns out, selecting the coefficients of the linear combination in a probabilistic fashion may significantly improve the rate of convergence to the gaussian distribution. We denote $S^{n-1} = \{(x_1, \dots, x_n) \in \mathbb{R}^n; \sum_i x_i^2 = 1\}$, the unit sphere in \mathbb{R}^n . Let σ_{n-1} be the unique rotationally-invariant probability measure on S^{n-1} , referred to as the uniform distribution on the sphere. Whenever we say that a random vector is distributed uniformly on the sphere, we mean that it is distributed according to σ_{n-1} . The coefficients of the linear combination will be selected randomly, uniformly over the sphere.

Theorem 1.1. *Let $n \geq 1$ be an integer, $0 < \rho < 1$. Suppose that X_1, \dots, X_n are independent random variables with finite fourth moments, such that $\mathbb{E}X_j = 0$ and $\mathbb{E}X_j^2 = 1$ for $j = 1, \dots, n$. Denote*

$$\delta = \left(\frac{1}{n} \sum_{j=1}^n \mathbb{E}X_j^4 \right)^{1/4}.$$

Then, there exists a subset $\mathcal{F} \subseteq S^{n-1}$ with $\sigma_{n-1}(\mathcal{F}) \geq 1 - \rho$ for which the following holds: For any $\theta = (\theta_1, \dots, \theta_n) \in \mathcal{F}$,

$$\sup_{\substack{\alpha, \beta \in \mathbb{R} \\ \alpha < \beta}} \left| \mathbb{P} \left(\alpha \leq \sum_{j=1}^n \theta_j X_j \leq \beta \right) - \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-t^2/2} dt \right| \leq \frac{C(\rho)\delta^4}{n} \quad (3)$$

where $C(\rho)$ is a constant depending solely on ρ . In fact, $C(\rho) \leq C \log^2(1/\rho)$, where $C > 0$ is a universal constant.

A case of interest is when X_1, X_2, \dots is an infinite sequence of independent, identically-distributed random variables of mean zero and variance one, with finite fourth moment.

In this case, Theorem 1.1 provides a convergence to the gaussian distribution, with rate of convergence of the order $O(1/n)$, for appropriate (or random) choice of linear combinations.

The case where X_1, \dots, X_n are identically-distributed, independent random variables is quite remarkable here. In this case the subset $\mathcal{F} \subseteq S^{n-1}$ from Theorem 1.1 can be described explicitly, and it does not depend on the distribution of the random variables. Following Rudelson and Vershynin [13], we make use of arithmetic properties of the vector θ . Define

$$d(x, \mathbb{Z}) = \min_{p \in \mathbb{Z}} |p - x|, \quad (x \in \mathbb{R}),$$

and for $\theta = (\theta_1, \dots, \theta_n) \in \mathbb{R}^n$ set,

$$d(\theta, \mathbb{Z}^n) = \sqrt{\sum_{j=1}^n d^2(\theta_j, \mathbb{Z})}.$$

Given $\theta \in S^{n-1}$ we denote by $\mathcal{N}(\theta)$ the minimal $R \geq 1$ for which the following three conditions hold:

(i) $\left| \sum_{j=1}^n \theta_j^3 \right| \leq R/n.$

(ii) $\sum_{j=1}^n \theta_j^4 \leq R/n.$

(iii) For any $|\xi| \leq n$,

$$d(\xi\theta, \mathbb{Z}^n) \geq \frac{1}{10} \min \left\{ |\xi|, \frac{n/R}{|\xi|} \right\}.$$

(the number 10 does not play any special rôle)

The three conditions above are satisfied by many unit vectors in the unit sphere, and are not very difficult to verify in certain examples. In order to appreciate the third condition, observe that for a typical unit vector $\theta \in S^{n-1}$ we have

$$d(\xi\theta, \mathbb{Z}^n) \geq \min\{|\xi|, c\sqrt{n}\} \quad \text{for any } |\xi| \leq e^{cn},$$

where $c > 0$ is a universal constant. For a concrete example, consider the unit vector

$$\theta^0 = (1, \sqrt{2}, -1, -\sqrt{2}, 1, \sqrt{2}, -1, -\sqrt{2}, \dots) / \sqrt{3n/2} \quad (4)$$

for n divisible by four. For this unit vector, the sum in (i) is zero, whereas (ii) clearly holds for any $R \geq 2$. The third condition is verified in Lemma 5.4, hence

$$\mathcal{N}(\theta^0) \leq C$$

for a universal constant $C > 0$.

Theorem 1.2. *Suppose that X_1, \dots, X_n are independent, identically-distributed random variables with finite fourth moments, such that*

$$\mathbb{E}X_1 = 0, \quad \mathbb{E}X_1^2 = 1 \quad \text{and} \quad \mathbb{E}X_1^4 \leq \delta^4.$$

Then, for any $\theta \in S^{n-1}$,

$$\sup_{\substack{\alpha, \beta \in \mathbb{R} \\ \alpha < \beta}} \left| \mathbb{P} \left(\alpha \leq \sum_{j=1}^n \theta_j X_j \leq \beta \right) - \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-t^2/2} dt \right| \leq \frac{C\mathcal{N}(\theta)\delta^4}{n}, \quad (5)$$

where $C > 0$ is a universal constant.

Once formulated, Theorem 1.1 and Theorem 1.2 require nothing but an adaptation of the proofs of the classical quantitative bounds in the central limit theorem. The following pages contain the details of the argument. Section 2 serves mostly as a remainder for the proof of the Berry-Esseen bound using the Fourier transform. In Section 3 and Section 4 we exploit the randomness involved in the selection of $\theta_1, \dots, \theta_n$ in Theorem 1.1. Section 5 is devoted to the proof of Theorem 1.2.

Throughout this text, the letters $c, \tilde{c}, c', C, \tilde{C}, \bar{C}$ etc. stand for various positive universal constants, whose value may change from one line to the next. We usually 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”. The notation $O(x)$, for some expression x , is an abbreviation for some complicated quantity y with the property that $|y| \leq Cx$ for some universal constant $C > 0$. A standard Gaussian random variable is a random variable whose density is $t \mapsto (2\pi)^{-1/2} \exp(-t^2/2)$ on the real line.

Acknowledgement. We would like to thank Shahar Mendelson for his help on the subject of Bernstein-type inequalities in the absence of exponential moments.

2 The Fourier inversion formula

Throughout this note, X_1, \dots, X_n are independent random variables with finite fourth moments such that $\mathbb{E}X_j = 0$ and $\mathbb{E}X_j^2 = 1$ for all j . Denote

$$\gamma_j = (\mathbb{E}X_j^3)^{1/3}, \quad \bar{\gamma}_j = (\mathbb{E}|X_j|^3)^{1/3}, \quad \delta_j = (\mathbb{E}X_j^4)^{1/4} \quad (1 \leq j \leq n).$$

Note that γ_j may be negative, since $\mathbb{E}X_j^3$ does not have a definite sign, and the third root of a negative number is negative. To help the reader remember the rôles of the Greek letters, we confess right away that γ , the third letter in the Greek alphabet, represents third moments, while δ , the fourth letter in the Greek alphabet, represents fourth moments. We also set

$$\delta = \left(\frac{1}{n} \sum_{j=1}^n \delta_j^4 \right)^{1/4}.$$

According to the Cauchy-Schwartz inequality, $\bar{\gamma}_j^3 = \mathbb{E}|X_j|^3 \leq \sqrt{\mathbb{E}X_j^4 \mathbb{E}X_j^2} = \delta_j^2$. Hence,

$$|\gamma_j| \leq \bar{\gamma}_j \leq \delta_j^{2/3}, \quad \bar{\gamma}_j \geq 1, \quad \delta_j \geq 1 \quad \text{for } j = 1, \dots, n. \quad (6)$$

Consider the Fourier transform

$$\varphi_j(\xi) = \mathbb{E} \exp(-i\xi X_j), \quad (\xi \in \mathbb{R}, 1 \leq j \leq n)$$

where $i^2 = -1$. Clearly $|\varphi_j(\xi)| \leq 1$ for any $\xi \in \mathbb{R}$. The k^{th} derivative of φ_j is

$$\varphi_j^{(k)}(\xi) = (-i)^k \mathbb{E} X_j^k \exp(-i\xi X_j)$$

for any $\xi \in \mathbb{R}, 1 \leq j \leq n$ and $0 \leq k \leq 4$. Consequently,

$$\varphi_j(0) = 1, \quad \varphi_j'(0) = 0, \quad \varphi_j''(0) = -1, \quad \varphi_j^{(3)}(0) = i\gamma_j^3 \quad (7)$$

for all j , and

$$|\varphi_j^{(3)}(\xi)| \leq \bar{\gamma}_j^3, \quad |\varphi_j^{(4)}(\xi)| \leq \delta_j^4 \quad \text{for all } \xi \in \mathbb{R}, 1 \leq j \leq n. \quad (8)$$

For a unit vector $\theta = (\theta_1, \dots, \theta_n) \in S^{n-1}$, by independence,

$$\varphi_\theta(\xi) := \mathbb{E} \exp\left(-i\xi \sum_{j=1}^n \theta_j X_j\right) = \prod_{j=1}^n \mathbb{E} \exp(-i\xi \theta_j X_j) = \prod_{j=1}^n \varphi_j(\theta_j \xi).$$

Denote, for $\theta \in S^{n-1}$,

$$F_\theta(t) = \mathbb{P}\left(\sum_{j=1}^n \theta_j X_j \leq t\right), \quad \Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-s^2/2} ds \quad (t \in \mathbb{R}).$$

Recall that when Γ is a standard Gaussian random variable, $\mathbb{E} \exp(-i\xi \Gamma) = \exp(-\xi^2/2)$. In order to control $\sup_t |F_\theta(t) - \Phi(t)|$ it is customary to try and bound the difference of the Fourier transforms $|\varphi_\theta(\xi) - \exp(-\xi^2/2)|$ for ξ in a large enough interval. According to Lemma 2 in [7, Vol. II, Section XVI.3], whose proof is based on a simple smoothing technique,

$$\sup_{t \in \mathbb{R}} |F_\theta(t) - \Phi(t)| \leq C \int_{-T}^T \frac{|\varphi_\theta(\xi) - \exp(-\xi^2/2)|}{|\xi|} d\xi + \frac{C}{T}, \quad (9)$$

for any $T > 0$, where $C > 0$ is a universal constant.

It is important to mention that for large classes of probability distributions, the error term C/T in (9) is non-optimal, and may be improved upon to C/T^2 in some cases. See, for

instance, Lemma 9 in [12] for an improvement of this nature pertaining to even, log-concave distributions. We are dealing, however, with arbitrary random variables, hence we must rely on the bound (9). Thus, in order to prove Theorem 1.1, we need to establish

$$\int_{-n/\delta^4}^{n/\delta^4} \frac{|\varphi_\theta(\xi) - \exp(-\xi^2/2)|}{|\xi|} d\xi \leq \frac{C(\rho)\delta^4}{n} \quad (10)$$

for all $\theta \in \mathcal{F}$ where \mathcal{F} is a certain subset of the sphere with $\sigma_{n-1}(\mathcal{F}) \geq 1 - \rho$. The rest of this paper is devoted to the proof of (10) and of the analogous inequality in the context of Theorem 1.2. We divide the domain of integration in (10) into three parts. The contribution of two of these domains is analyzed in the following two lemmas.

Lemma 2.1. *Let $\theta = (\theta_1, \dots, \theta_n) \in S^{n-1}$. Denote $\varepsilon = \left(\sum_{j=1}^n \theta_j^4 \delta_j^4\right)^{1/4}$ and $R_1 = \left|\sum_{j=1}^n \gamma_j^3 \theta_j^3\right|$. Suppose that $\varepsilon \leq 1$. Then,*

$$\int_{-\varepsilon^{-2/3}}^{\varepsilon^{-2/3}} \left| \varphi_\theta(\xi) - e^{-\xi^2/2} \right| \frac{d\xi}{|\xi|} \leq C [R_1 + \varepsilon^4]$$

where $C > 0$ is a universal constant.

Proof. Recall (7) and (8). Taylor's theorem implies that for any $j = 1, \dots, n$ and $s \in \mathbb{R}$,

$$\left| \varphi_j(s) - \left[1 - \frac{1}{2}s^2 + \frac{i\gamma_j^3}{6}s^3 \right] \right| \leq \frac{\delta_j^4}{24}s^4.$$

Recall that $\max\{1, |\gamma_j|\} \leq \delta_j^{2/3} \leq \delta_j$ according to (6). In particular,

$$|\varphi_j(s) - 1| \leq 3/4 \quad \text{for } |s| \leq \delta_j^{-1}, j = 1, \dots, n.$$

Thus $\log \varphi_j(s)$ is well-defined for $|s| \leq \delta_j^{-1}$, and for any $|s| \leq \delta_j^{-1}$ and $j = 1, \dots, n$,

$$\log \varphi_j(s) = -\frac{1}{2}s^2 + \frac{i\gamma_j^3}{6}s^3 + O(\delta_j^4 s^4) \quad (11)$$

since $|\log(1+z) - z| \leq 8|z|^2$ whenever $|z| \leq 3/4$. Note that for any $\xi \in \mathbb{R}$ with $|\xi| \leq \varepsilon^{-2/3}$,

$$|\theta_j \xi| \leq |\theta_j| \varepsilon^{-2/3} \leq |\theta_j| \varepsilon^{-1} \leq \delta_j^{-1}.$$

Summing (11) over $j = 1, \dots, n$, we conclude that for any $|\xi| \leq \varepsilon^{-2/3}$,

$$\sum_{j=1}^n \log \varphi_j(\theta_j \xi) = -\frac{\xi^2}{2} + \frac{i \sum_{j=1}^n \gamma_j^3 \theta_j^3}{6} \xi^3 + O(\varepsilon^4 \xi^4) \quad (12)$$

as $\sum_j \theta_j^2 = 1$ and $\varepsilon^4 = \sum_{j=1}^n \delta_j^4 \theta_j^4$. Recall that $|\gamma_j|^3 \leq \delta_j^2$. By the Cauchy-Schwartz inequality,

$$R_1 = \left| \sum_{j=1}^n \gamma_j^3 \theta_j^3 \right| \leq \sum_{j=1}^n \delta_j^2 |\theta_j| \theta_j^2 \leq \left(\sum_{j=1}^n \delta_j^4 |\theta_j|^2 \theta_j^2 \right)^{1/2} = \varepsilon^2. \quad (13)$$

Hence $R_1 |\xi|^3 + \varepsilon^4 \xi^4 \leq 2$ for all $|\xi| \leq \varepsilon^{-2/3}$. From (12) we learn that for any $|\xi| \leq \varepsilon^{-2/3}$,

$$e^{\xi^2/2} \left| \varphi_\theta(\xi) - e^{-\xi^2/2} \right| = \left| e^{\xi^2/2} \prod_{j=1}^n \varphi_j(\theta_j \xi) - 1 \right| = \left| e^{O(R_1 |\xi|^3 + \varepsilon^4 \xi^4)} - 1 \right| \leq C' [R_1 |\xi|^3 + \varepsilon^4 \xi^4].$$

We integrate the above, to conclude that

$$\int_{-\varepsilon^{-2/3}}^{\varepsilon^{-2/3}} \left| \varphi_\theta(\xi) - e^{-\xi^2/2} \right| \frac{d\xi}{|\xi|} \leq \int_{-\infty}^{\infty} C' [R_1 |\xi|^3 + \varepsilon^4 \xi^4] e^{-\xi^2/2} \frac{d\xi}{|\xi|} \leq \tilde{C} [R_1 + \varepsilon^4].$$

□

Lemma 2.2. Let $\theta = (\theta_1, \dots, \theta_n) \in S^{n-1}$. Denote, as before, $\varepsilon = \left(\sum_{j=1}^n \theta_j^4 \delta_j^4 \right)^{1/4}$, and suppose that $R_2 > 0$ satisfies

$$\sum_{j \in \mathcal{S}} \theta_j^2 \geq 1/8 \quad \text{where} \quad \mathcal{S} = \{1 \leq j \leq n; |\theta_j| \leq R_2 / \bar{\gamma}_j^3\}. \quad (14)$$

Then, whenever $\varepsilon^{-2/3} \leq cR_2^{-1}$,

$$\int_{\varepsilon^{-2/3}}^{cR_2^{-1}} \left| \varphi_\theta(\xi) - e^{-\xi^2/2} \right| \frac{d\xi}{|\xi|} \leq C\varepsilon^4.$$

The right-hand side is also an upper bound for the integral from $-cR_2^{-1}$ to $-\varepsilon^{-2/3}$. Here $C, c > 0$ are universal constants.

Proof. As in the beginning of the proof of Lemma 2.1, we use Taylor's theorem. We conclude that for $j = 1, \dots, n$,

$$\left| \log \varphi_j(s) + \frac{1}{2}s^2 \right| \leq C\bar{\gamma}_j^3 |s|^3 \quad \text{when} \quad |s| \leq 1/\bar{\gamma}_j.$$

Hence, for $j = 1, \dots, n$ and $s \in \mathbb{R}$,

$$|\varphi_j(s)| \leq \exp(-s^2/4) \quad \text{when} \quad |s| \leq c/\bar{\gamma}_j^3. \quad (15)$$

Let $\xi \in \mathbb{R}$ be such that $|\xi| \leq cR_2^{-1}$ where c is the constant from (15). For any $j \in \mathcal{S}$, we have $|\theta_j \xi| \leq c/\bar{\gamma}_j^3$. Therefore,

$$|\varphi_\theta(\xi)| = \prod_{j=1}^n |\varphi_j(\theta_j \xi)| \leq \prod_{j \in \mathcal{S}} |\varphi_j(\theta_j \xi)| \leq \exp \left(- \sum_{j \in \mathcal{S}} \theta_j^2 \xi^2 / 4 \right) \leq \exp(-\tilde{c}\xi^2)$$

where the last inequality follows from (14). Consequently,

$$\int_{\varepsilon^{-2/3}}^{cR_2^{-1}} \left| \varphi_\theta(\xi) - e^{-\xi^2/2} \right| \frac{d\xi}{|\xi|} \leq \int_{\varepsilon^{-2/3}}^{cR_2^{-1}} \left[e^{-\tilde{c}\xi^2} + e^{-\xi^2/2} \right] \frac{d\xi}{|\xi|} \leq C\varepsilon^{2/3} e^{-c/\varepsilon^{4/3}} \leq \bar{C}\varepsilon^4.$$

□

3 Properties of a random direction

We retain the notation of the previous section, and our first goal is to estimate R_2 from Lemma 2.2. The following lemma serves that purpose. For a random variable Y and $a \in \mathbb{R}$ we write $1_{\{Y > a\}}$ for the random variable that equals one when $Y > a$ and vanishes otherwise.

Lemma 3.1. *Let $M \geq 1$ and suppose that Y is a non-negative random variable with $\mathbb{E}Y = 1$ and $\mathbb{E}Y^2 \leq M$. Then,*

- (i) $\mathbb{P}(Y \geq 1/2) \geq 1/(4M)$,
- (ii) $\mathbb{E}Y 1_{\{Y \leq 5M\}} \geq 4/5$.

Proof. The first inequality is due to Paley and Zygmund (see, e.g., Kahane [11, Section 1.6]). To prove (ii), observe that

$$\mathbb{E}1_{\{Y \leq 5M\}}Y = 1 - \mathbb{E}1_{\{Y > 5M\}}Y \geq 1 - \frac{1}{5M}\mathbb{E}Y^2 \geq 4/5.$$

□

The rest of this section and the next section are devoted to the proof of Theorem 1.1. Readers interested only in the proof of Theorem 1.2 may proceed to Section 5. Suppose that $\Theta = (\Theta_1, \dots, \Theta_n)$ is a random vector, distributed uniformly on the unit sphere S^{n-1} .

Lemma 3.2. *Let $\mathcal{J} \subseteq \{1, \dots, n\}$ be a subset, denote its cardinality by $k = \#(\mathcal{J})$, and assume that $k \geq 4n/5$. Then with probability greater than $1 - C \exp(-cn)$ of selecting the random vector $\Theta \in S^{n-1}$,*

$$\sum_{j \in \mathcal{S}} \Theta_j^2 \geq 1/8 \quad \text{where} \quad \mathcal{S} = \{j \in \mathcal{J}; |\Theta_j| \leq 40/\sqrt{n}\}.$$

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

Proof. Let us introduce independent, standard gaussian random variables Γ_j ($j \in \mathcal{J}$), that are independent of the Θ_j 's. Let Z be a chi-square random variable with $k = \#(\mathcal{J})$ degrees of freedom, independent of the Γ_j 's and the Θ_j 's. Then Z has the same distribution as $\sum_{j \in \mathcal{J}} \Gamma_j^2$. Bernstein's inequality (see, e.g., Ibragimov and Linnik [10, Chapter 7]) yields

$$\mathbb{P}\left(\frac{k}{2} \leq Z \leq 2k\right) \geq 1 - C \exp(-ck) \geq 1 - \tilde{C} \exp(-\tilde{c}n). \quad (16)$$

Observe that the random variables $(\Gamma_j)_{j \in \mathcal{J}}$ have exactly the same joint distribution as the random variables $(\sqrt{Z}\Theta_j)_{j \in \mathcal{J}}$. Therefore, in order to prove the lemma, it suffices to show that with probability greater than $1 - C \exp(-cn)$,

$$\sum_{j \in \mathcal{S}} \Gamma_j^2 \geq n/2 \quad \text{where} \quad \mathcal{S} = \{j \in \mathcal{J}; |\Gamma_j| \leq 20\}.$$

Denote $Y_j = \Gamma_j^2 1_{\{|\Gamma_j| \leq 20\}}$. Then $(Y_j)_{j \in \mathcal{J}}$ are independent, identically-distributed random variables, and our goal is to prove that

$$\mathbb{P}\left(\sum_{j \in \mathcal{J}} Y_j \geq n/2\right) \geq 1 - C \exp(-cn). \quad (17)$$

Since $\mathbb{E}\Gamma_j^4 = 3$, then Lemma 3.1(ii) yields that

$$\mathbb{E}Y_j \geq 4/5, \quad \text{and clearly} \quad \text{Var}(Y_j) \leq \mathbb{E}Y_j^2 \leq \mathbb{E}\Gamma_j^2 = 3$$

for $j \in \mathcal{J}$. According to Bernstein's inequality,

$$\mathbb{P}\left(\sum_{j \in \mathcal{J}} Y_j \leq \frac{4k}{5} - t\sqrt{3k}\right) \leq C \exp(-ct^2) \quad \text{for any } t \geq 0. \quad (18)$$

Recall that $k/n \geq 4/5$. Inequality (17) follows by setting $t = \sqrt{n/200}$ in (18). \square

Corollary 3.3. *Set $R = 200\delta^2/\sqrt{n}$. Then with probability greater than $1 - C \exp(-cn)$ of selecting the random vector $\Theta \in S^{n-1}$,*

$$\sum_{j \in \mathcal{S}} \Theta_j^2 \geq 1/8 \quad \text{where} \quad \mathcal{S} = \{1 \leq j \leq n; |\Theta_j| \leq R/\bar{\gamma}_j^3\}.$$

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

Proof. Denote $\mathcal{J} = \{1 \leq j \leq n; \bar{\gamma}_j^3 \leq 5\delta^2\}$. Then,

$$\delta^4 = \frac{1}{n} \sum_{j=1}^n \delta_j^4 \geq \frac{1}{n} \sum_{j=1}^n \bar{\gamma}_j^6 \geq \frac{1}{n} \sum_{j \notin \mathcal{J}} \bar{\gamma}_j^6 > \frac{n - \#(\mathcal{J})}{n} (5\delta^2)^2.$$

Denoting $k = \#(\mathcal{J})$, we thus see that $k/n \geq 24/25 \geq 4/5$. For $j \in \mathcal{J}$, we have $40/\sqrt{n} \leq R/\bar{\gamma}_j^3$. In order to prove the lemma, it therefore suffices to show that with probability greater than $1 - C \exp(-cn)$,

$$\sum_{j \in \mathcal{S}} \Theta_j^2 \geq 1/8 \quad \text{where} \quad \mathcal{S} = \{j \in \mathcal{J}; |\Theta_j| \leq 40/\sqrt{n}\}.$$

This is precisely the content of Lemma 3.2. \square

Our goal is to bound the integral in (10). Lemma 2.1 and Lemma 2.2 (with the help of Corollary 3.3) control the contribution of the interval $[-c\sqrt{n}/\delta^2, c\sqrt{n}/\delta^2]$. Next we aim at bounding the contributions of $\xi \in \mathbb{R}$ with $c\sqrt{n}/\delta^2 \leq |\xi| \leq n/\delta^4$. Denote

$$J_n(\xi) = \mathbb{E} \exp(-i\xi\Theta_1) \quad (\xi \in \mathbb{R}).$$

The function J_n is even and real-valued, and is related to the Bessel function of order $n/2-1$.

Lemma 3.4. *We have*

$$J_n(\xi) \leq 1 - c \min \{ \xi^2/n, 1 \} \quad \text{for all } \xi \in \mathbb{R}, \quad (19)$$

where $c > 0$ is a universal constant.

Proof. Since $\mathbb{E}(\sqrt{n}\Theta_1)^2 = 1$ and $\mathbb{E}(\sqrt{n}\Theta_1)^4 \leq C$, then Taylor's theorem yields

$$J_n(\sqrt{n}\tau) = 1 - \frac{\tau^2}{2} + O(\tau^4)$$

for $|\tau| \leq 1$, as the odd moments vanish. This implies (19) for $|\xi| \leq c\sqrt{n}$. The density f_n of the random variable Θ_1 vanishes outside $[-1, 1]$, and is proportional to $t \mapsto (1-t^2)^{(n-3)/2}$ on $[-1, 1]$. Denote $g_n(t) = n^{-1/2}f_n(n^{-1/2}t)$. Then

$$\int_{-\infty}^{\infty} \left| g_n(t) - \frac{1}{\sqrt{2\pi}} \exp(-t^2/2) \right| dt \xrightarrow{n \rightarrow \infty} 0,$$

as be may verified routinely (see, e.g., Diaconis and Freedman [6] for quantitative bounds). Therefore the Fourier transform satisfies

$$\sup_{\xi \in \mathbb{R}} |J_n(\sqrt{n}\xi) - \exp(-\xi^2/2)| \xrightarrow{n \rightarrow \infty} 0$$

which implies (19) in the range $|\xi| \geq c\sqrt{n}$. \square

Lemma 3.5. *Let $j = 1, \dots, n$. Then, for any $\tau \in \mathbb{R}$,*

$$\mathbb{E}|\varphi_j(\tau\Theta_j)|^2 \leq 1 - c \min \{ \tau^2/n, \delta_j^{-4} \},$$

where $c > 0$ is a universal constant.

Proof. As before, denote by f_n the density of the random variable Θ_1 . Then,

$$\mathbb{E}|\varphi_j(\tau\Theta_j)|^2 = \int_{-\infty}^{\infty} |\varphi_j(\tau\xi)|^2 f_n(\xi) d\xi.$$

Let \tilde{X}_j be an independent copy of X_j . Define $Y = X_j - \tilde{X}_j$, a symmetric random variable. Then the Fourier transform of Y is

$$\mathbb{E} \exp(-i\xi Y) = \mathbb{E} \exp(-i\xi X_j) \overline{\mathbb{E} \exp(-i\xi \tilde{X}_j)} = \varphi_j(\xi) \overline{\varphi_j(\xi)} = |\varphi_j(\xi)|^2.$$

Hence the function $|\varphi_j(\tau\xi)|^2$ is the Fourier transform of the random variable τY . Recall that $J_n(\xi)$ is the Fourier transform of the density f_n . The central observation is that according to the Plancherel theorem,

$$\mathbb{E}|\varphi_j(\tau\Theta_j)|^2 = \int_{-\infty}^{\infty} |\varphi_j(\tau\xi)|^2 f_n(\xi) d\xi = \mathbb{E} J_n(\tau Y) \leq 1 - c \mathbb{E} \min\{\tau^2 Y^2/n, 1\},$$

where the last inequality is the content of Lemma 3.4. Denote $r = \tau^2/n$ and $Z = Y^2/2$. In order to complete the proof of the lemma, it suffices to show that for any $r \geq 0$,

$$\mathbb{E} \min\{rZ, 1\} \geq c \min\{r, \delta_j^{-4}\}. \quad (20)$$

The left-hand side of (20) is non-decreasing in r , hence it is enough to prove (20) when $r \leq \delta_j^{-4}/10$. Since $Y = X_j - \tilde{X}_j$ and $Z = Y^2/2$, then $\mathbb{E}Z = 1$ and $\mathbb{E}Z^2 = (3 + \delta_j^4)/2 \leq 2\delta_j^4$. According to Lemma 3.1(ii), $\mathbb{E}1_{\{Z \leq 10\delta_j^4\}} Z \geq 4/5$. Therefore, for $0 \leq r \leq \delta_j^{-4}/10$,

$$\mathbb{E} \min\{rZ, 1\} \geq \mathbb{E}1_{\{Z \leq 10\delta_j^4\}} \min\{rZ, 1\} = r \mathbb{E}1_{\{Z \leq 10\delta_j^4\}} Z \geq r/2,$$

and (20) follows. The lemma is thus proven. \square

When the dimension n is large, the random variables $\Theta_1, \dots, \Theta_n$ are ‘‘approximately independent’’. One would thus expect that usually, for functions $f_1, \dots, f_n : \mathbb{R} \rightarrow \mathbb{C}$,

$$\mathbb{E} \prod_{j=1}^n f_j(\Theta_j) \approx \prod_{j=1}^n \mathbb{E} f_j(\Theta_j). \quad (21)$$

The most straightforward way to obtain estimates in the spirit of (21) is to compare the distribution of Θ with that of a gaussian random vector of the same expectation and covariance as in the proof of Lemma 3.2 above. Even though this approach works well in our present context, we prefer to invoke below a recent inequality due to Carlen, Lieb and Loss [3]. This inequality provides a particularly elegant way to exploit the ‘‘approximate

independence” of $\Theta_1, \dots, \Theta_n$. It states that for any non-negative, measurable functions $f_1, \dots, f_n : [-1, 1] \rightarrow \mathbb{R}$,

$$\mathbb{E} \prod_{j=1}^n f_j(\Theta_j) \leq \prod_{j=1}^n (\mathbb{E} f_j(\Theta_j)^2)^{1/2}. \quad (22)$$

See Barthe, Cordero-Erausquin, Ledoux and Maurey [2] for ramifications of the Brascamp-Lieb type inequality (22). Recall that $\varphi_1, \dots, \varphi_n$ are the Fourier transforms of the independent random variables X_1, \dots, X_n .

Lemma 3.6. *Let $\alpha > 0$ and assume that $\alpha\sqrt{n}/\delta^2 \leq n/\delta^4$. Then, with probability greater than $1 - C(\alpha) \exp(-c(\alpha)n/\delta^4)$ of selecting $(\Theta_1, \dots, \Theta_n) \in S^{n-1}$,*

$$\int_{\alpha\sqrt{n}/\delta^2}^{n/\delta^4} \left| \prod_{j=1}^n \varphi_j(\Theta_j \xi) - e^{-\xi^2/2} \right| \frac{d\xi}{|\xi|} \leq C(\alpha) \exp(-c(\alpha)n/\delta^4) \leq \frac{\bar{C}(\alpha)\delta^4}{n}.$$

The right-hand side is also an upper bound for the integral from $-n/\delta^4$ to $-\alpha\sqrt{n}/\delta^2$. Here $C(\alpha), \bar{C}(\alpha), c(\alpha) > 0$ are constants depending solely on α .

Proof. Lemma 3.5 and (22) imply that for any $\xi \in \mathbb{R}$,

$$\mathbb{E} \left| \prod_{j=1}^n \varphi_j(\Theta_j \xi) \right| \leq \prod_{j=1}^n \sqrt{\mathbb{E} |\varphi_j(\Theta_j \xi)|^2} \leq \prod_{j=1}^n \left(1 - c \min \{ \xi^2/n, \delta_j^{-4} \} \right).$$

Denote $\mathcal{J} = \{1 \leq j \leq n; \delta_j \leq 2\delta\}$. Repeating a simple argument, we have

$$\delta^4 = \frac{1}{n} \sum_{j=1}^n \delta_j^4 \geq \frac{1}{n} \sum_{j \notin \mathcal{J}} \delta_j^4 \geq \frac{n - \#\mathcal{J}}{n} 16\delta^4,$$

hence $\#\mathcal{J} \geq n/2$. For any $\xi \in \mathbb{R}$,

$$\begin{aligned} \mathbb{E} \left| \prod_{j=1}^n \varphi_j(\Theta_j \xi) \right| &\leq \prod_{j=1}^n \left(1 - \tilde{c} \min \{ \xi^2/n, \delta_j^{-4} \} \right) \leq \prod_{j \in \mathcal{J}} \left(1 - \tilde{c} \min \{ \xi^2/n, \delta_j^{-4} \} \right) \\ &\leq \left(1 - c \min \{ \xi^2/n, \delta^{-4} \} \right)^{n/2} \leq \exp(-c' \min \{ \xi^2, n/\delta^4 \}). \end{aligned}$$

Therefore,

$$\begin{aligned} &\mathbb{E} \int_{\alpha\sqrt{n}/\delta^2}^{n/\delta^4} \left| \prod_{j=1}^n \varphi_j(\Theta_j \xi) - e^{-\xi^2/2} \right| \frac{d\xi}{|\xi|} \\ &\leq \frac{\delta^2}{\alpha\sqrt{n}} \int_{\alpha\sqrt{n}/\delta^2}^{n/\delta^4} e^{-\tilde{c} \min \{ \xi^2, n/\delta^4 \}} + e^{-\xi^2/2} d\xi \leq C(\alpha) e^{-c(\alpha)n/\delta^4}. \end{aligned}$$

From the Chebyshev inequality,

$$\mathbb{P} \left(\int_{\alpha\sqrt{n}/\delta^2}^{n/\delta^4} \left| \prod_{j=1}^n \varphi_j(\Theta_j \xi) - e^{-\xi^2/2} \right| \frac{d\xi}{|\xi|} \geq \sqrt{C(\alpha)e^{-c(\alpha)n/\delta^4}} \right) \leq \sqrt{C(\alpha)e^{-c(\alpha)n/\delta^4}}.$$

□

The results obtained so far may be summarized as follows:

Lemma 3.7. *There exists a subset $\mathcal{F} \subseteq S^{n-1}$ with $\sigma_{n-1}(\mathcal{F}) \geq 1 - C \exp(-cn/\delta^4)$ such that for any $\theta = (\theta_1, \dots, \theta_n) \in \mathcal{F}$ with $\sum_{j=1}^n \theta_j^4 \delta_j^4 \leq 1$,*

$$\int_{-n/\delta^4}^{n/\delta^4} \left| \prod_{j=1}^n \varphi_j(\theta_j \xi) - e^{-\xi^2/2} \right| \frac{d\xi}{|\xi|} \leq C \left(\sum_{j=1}^n \theta_j^4 \delta_j^4 + \left| \sum_{j=1}^n \gamma_j^3 \theta_j^3 \right| + \frac{\delta^4}{n} \right), \quad (23)$$

where $C > 0$ is a universal constant.

Proof. Let $\mathcal{F}_1 \subseteq S^{n-1}$ be the set of directions with $\sigma_{n-1}(\mathcal{F}_1) \geq 1 - C \exp(-cn)$ whose existence is guaranteed by Corollary 3.3. Assume that $\theta \in \mathcal{F}_1$ is such that $\sum_j \theta_j^4 \delta_j^4 \leq 1$. The bound (23) is the culmination of three arguments: Lemma 2.1 controls the contribution of $|\xi| \leq \varepsilon^{-2/3}$, for $\varepsilon = \left(\sum_{j=1}^n \theta_j^4 \delta_j^4 \right)^{1/4} \leq 1$. Thanks to the definition of \mathcal{F}_1 , Lemma 2.2 with $R_2 = 200\delta^2/\sqrt{n}$ provides an upper bound for the contribution up to $|\xi| \leq c\sqrt{n}/\delta^2$. We conclude with an application of Lemma 3.6, with α being a universal constant. Denote by $\mathcal{F}_2 \subseteq S^{n-1}$ the set with $\sigma_{n-1}(\mathcal{F}_2) \geq 1 - C \exp(-cn/\delta^4)$ whose existence is guaranteed by Lemma 3.6. Setting $\mathcal{F} = \mathcal{F}_1 \cap \mathcal{F}_2$, we see that (23) holds for any $\theta \in \mathcal{F}$ with $\sum_j \theta_j^4 \delta_j^4 \leq 1$. □

Corollary 3.8. *There exists a subset $\mathcal{F}_1 \subseteq S^{n-1}$ with $\sigma_{n-1}(\mathcal{F}_1) \geq 1 - C \exp(-cn/\delta^4)$ with the following property: For any $\theta = (\theta_1, \dots, \theta_n) \in \mathcal{F}_1$ and $t \in \mathbb{R}$,*

$$\left| \mathbb{P} \left(\sum_{j=1}^n \theta_j X_j \leq t \right) - \Phi(t) \right| \leq C \left[\sum_{j=1}^n \delta_j^4 \theta_j^4 + \left| \sum_{j=1}^n \gamma_j^3 \theta_j^3 \right| + \frac{\delta^4}{n} \right]. \quad (24)$$

Here $\Phi(t) = (2\pi)^{-1/2} \int_{-\infty}^t e^{-t^2/2} dt$ and $C, c > 0$ are universal constants.

Proof. It is enough to consider θ for which $\sum_j \delta_j^4 \theta_j^4 \leq 1$, as otherwise (24) holds trivially. The bound (24) is thus an immediate consequence of the smoothing inequality (9) and Lemma 3.7. □

4 Deviation inequalities

It remains to deduce Theorem 1.1 from Corollary 3.8. To that end, we need to analyze the terms $\sum_{j=1}^n \gamma_j^3 \theta_j^3$ and $\sum_{j=1}^n \delta_j^4 \theta_j^4$ appearing in Corollary 3.8. We would like to get a bound in (3) of the form $C(\rho)\delta^4/n$, where $C(\rho)$ depends on ρ solely. This is the reason we use the following crude lemma.

Lemma 4.1. *Suppose that $(\Theta_1, \dots, \Theta_n) \in S^{n-1}$ is a random vector, distributed uniformly on S^{n-1} . Then, for any $t \geq 0$,*

$$\mathbb{P} \left(\left| \sum_{j=1}^n \gamma_j^3 \Theta_j^3 \right| \geq t \frac{\delta^4}{n} \right) \leq C \exp \left(-ct^{2/3} \right), \quad (25)$$

and additionally,

$$\mathbb{P} \left(\sum_{j=1}^n \delta_j^4 \Theta_j^4 \geq t \frac{\delta^4}{n} \right) \leq C \exp \left(-c\sqrt{t} \right). \quad (26)$$

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

Proof. Introduce independent, standard gaussian random variables $\Gamma_1, \dots, \Gamma_n$ that are independent of the Θ_j 's. Let Z be a chi-square random variable with n degrees of freedom, independent of the Γ_j 's and Θ_j 's. As in (16), we know that $n/2 \leq Z \leq 2n$ with probability greater than $1 - C \exp(-cn)$. Thus,

$$\mathbb{P} \left(\left| \sum_{j=1}^n \gamma_j^3 \Theta_j^3 \right| \geq t \right) = \mathbb{P} \left(\left| \frac{\sum_{j=1}^n \gamma_j^3 \Gamma_j^3}{Z^{3/2}} \right| \geq t \right) \leq \mathbb{P} \left(\left| \sum_{j=1}^n \gamma_j^3 \Gamma_j^3 \right| \geq \frac{n^{3/2}t}{4} \right) + C e^{-cn}. \quad (27)$$

The random variable $Y = \sum_{j=1}^n \gamma_j^3 \Gamma_j^3$ is the sum of independent, mean zero random variables. We will apply a moment inequality we learned from Adamczak, Litvak, Pajor and Tomczak-Jaegermann [1, Section 3], which builds upon previous work by Hitczenko, Montgomery-Smith, and Oleszkiewicz [9]. Recall that $\mathbb{E} \exp(c\Gamma_j^2) \leq 2$ for a universal constant $c > 0$. In the terminology of [1], the random variables $\Gamma_1^3, \dots, \Gamma_n^3$ are random variables of class $\psi_{2/3}$, hence for any $p \geq 2$,

$$(\mathbb{E}|Y|^p)^{1/p} \leq C \left(p^{1/2} \sqrt{\sum_{j=1}^n \gamma_j^6} + p^{3/2} \left(\sum_{j=1}^n |\gamma_j|^{3p} \right)^{1/p} \right) \leq \tilde{C} p^{3/2} \sqrt{\sum_{j=1}^n \gamma_j^6} \leq \bar{C} p^{3/2} \sqrt{n} \delta^2,$$

as $\gamma_j^6 \leq \delta_j^4$ for all j . According to the Chebyshev inequality, for any $t \geq C\delta^2\sqrt{n}$,

$$\mathbb{P} \left(\left| \sum_{j=1}^n \gamma_j^3 \Gamma_j^3 \right| \geq t \right) \leq \frac{\mathbb{E}|Y|^p}{t^p} \leq \frac{(\bar{C} p^{3/2} \sqrt{n} \delta^2)^p}{t^p} \leq e^{-p} \leq \exp \left(-\tilde{c} \frac{t^{2/3}}{(\delta^4 n)^{1/3}} \right)$$

where $p = ct^{2/3}/(\delta^4 n)^{1/3}$ for an appropriate small universal constant $c > 0$. From (27) and the last inequality,

$$\mathbb{P} \left(\left| \sum_{j=1}^n \gamma_j^3 \Theta_j^3 \right| \geq t \frac{\delta^2}{n} \right) \leq C \exp(-ct^{2/3}) + C \exp(-cn) \quad \text{for all } t > C. \quad (28)$$

According to the Cauchy-Schwartz inequality in (13) we always have $\left| \sum_{j=1}^n \gamma_j^3 \Theta_j^3 \right| \leq \sqrt{\sum_{j=1}^n \delta_j^4 \Theta_j^4} \leq \sqrt{n} \delta^2$ and hence the probability on the left-hand side of (28) vanishes for $t \geq n^{3/2}$. We may thus deduce (25) from (28). Inequality (26) is proven in a similar vein: Denote $W = \sum_{j=1}^n \delta_j^4 [\Gamma_j^4 - 3]$. Then, $\mathbb{E}W = 0$ and for any $t \geq 0$,

$$\mathbb{P} \left(\sum_{j=1}^n \delta_j^4 \Theta_j^4 \geq 12 \sum_{j=1}^n \frac{\delta_j^4}{n^2} + t \right) \leq \mathbb{P}(W \geq n^2 t/4) + C \exp(-cn). \quad (29)$$

The random variables $\Gamma_1^4 - 3, \dots, \Gamma_n^4 - 3$ are independent random variables of class $\psi_{1/2}$. Again, using the inequality from [1, Section 3] we see that for $p \geq 2$,

$$(\mathbb{E}|W|^p)^{1/p} \leq C \left(p^{1/2} \sqrt{\sum_{j=1}^n \delta_j^8} + p^2 \left(\sum_{j=1}^n \delta_j^{4p} \right)^{1/p} \right) \leq \tilde{C} p^2 \sum_{j=1}^n \delta_j^4 = \tilde{C} p^2 n \delta^4.$$

Using the Chebyshev inequality, as before, we deduce that for any $t > C$,

$$\mathbb{P}(W \geq tn\delta^4) \leq C \exp(-c\sqrt{t}). \quad (30)$$

Inequalities (29) and (32) lead to the bound

$$\mathbb{P} \left(\sum_{j=1}^n \delta_j^4 \Theta_j^4 \geq t \frac{\delta^4}{n} \right) \leq C \exp(-c\sqrt{t}) + C \exp(-cn) \quad \text{for all } t \geq 15.$$

Since with probability one $\sum_{j=1}^n \delta_j^4 \Theta_j^4 \leq n\delta^4$, the bound (26) follows. \square

Proof of Theorem 1.1. We may assume that $(\log 1/\rho)^2 \leq \tilde{c}n/\delta^4$, for a small universal constant $\tilde{c} > 0$, since otherwise the conclusion (3) of the theorem is trivial, for an appropriate choice of a universal constant C in Theorem 1.1. Therefore,

$$\rho \geq \exp\left(-\sqrt{\tilde{c}n/\delta^4}\right) \geq \exp(-cn/\delta^4)$$

where $c > 0$ is the constant from Corollary 3.8. Let $\mathcal{F}_1 \subseteq S^{n-1}$ be the subset of the sphere with $\sigma_{n-1}(\mathcal{F}_1) \geq 1 - C \exp(-cn/\delta^4) \geq 1 - C\rho$ whose existence is guaranteed by Corollary

3.8. According to Lemma 4.1, there exists a subset $\mathcal{F}_2 \subseteq S^{n-1}$ with $\sigma_{n-1}(\mathcal{F}_2) \geq 1 - \rho$ such that for any $\theta = (\theta_1, \dots, \theta_n) \in \mathcal{F}_2$,

$$\left| \sum_{j=1}^n \gamma_j^3 \theta_j^3 \right| + \sum_{j=1}^n \delta_j^4 \theta_j^4 \leq \tilde{C} \left(\log \frac{1}{\rho} \right)^{3/2} \frac{\delta^4}{n} + \tilde{C} \left(\log \frac{1}{\rho} \right)^2 \frac{\delta^4}{n} \leq \bar{C} \left(\log \frac{1}{\rho} \right)^2 \frac{\delta^4}{n}. \quad (31)$$

Denote $\mathcal{F} = \mathcal{F}_1 \cap \mathcal{F}_2$. Then $\sigma_{n-1}(\mathcal{F}) \geq 1 - C'\rho$. Furthermore, according to Corollary 3.8 and to (31), the desired bound (3) holds for any $\theta \in \mathcal{F}$, with $C(\rho) \leq \hat{C}(\log 1/\rho)^2$. \square

Remark. There is some wiggle room in the bound for $C(\rho)$ in Theorem 1.1. One easily notices that the bounds stated in Lemma 4.1 are, in many cases, quite weak: When all the δ_j are comparable, a better analysis of the moment inequality from [1, Section 3] leads to a sub-gaussian tail, at least in some range. If one is interested in a version of Theorem 1.1 where $\delta^4 = \sum_j \mathbb{E}X_j^4/n$ is replaced by the larger quantity $\max_j \mathbb{E}X_j^4$, finer analogs of Lemma 4.1 may be employed. For such a version of Theorem 1.1, the power of the logarithm in the bound for $C(\rho)$ may essentially be improved, from 2 to 1/2, at least for ρ in some range.

5 Explicit, universal coefficients

This section is devoted to the proof of Theorem 1.2 and related statements. We assume that the independent random variables X_1, \dots, X_n are identically distributed, and that they have the same distribution as a certain random variable X . This random variable X has mean zero, variance one, and we denote $\delta = (\mathbb{E}X^4)^{1/4}$. Its Fourier transform is

$$\varphi(\xi) = \mathbb{E} \exp(-i\xi X) \quad (\xi \in \mathbb{R}).$$

As before we fix $\theta \in S^{n-1}$ and set

$$\varphi_\theta(\xi) = \prod_{j=1}^n \varphi(\theta_j \xi) \quad (\xi \in \mathbb{R}).$$

Let \tilde{X} be an independent copy of X , and define $Y = X - \tilde{X}$. The next three lemmas bound an integral of $|\varphi_\theta|$ in terms of a certain arithmetic property of θ . The property is quite similar to the one introduced by Rudelson and Vershynin [13]; we closely follow their presentation, taking into account several simplifications proposed in [8].

Lemma 5.1. *For any $\xi \in \mathbb{R}$,*

$$|\varphi_\theta(\xi)| \leq \exp \left\{ -4\mathbb{E}d^2 \left(\frac{\xi Y}{2\pi} \theta, \mathbb{Z}^n \right) \right\}.$$

Proof. It is easily verified that

$$\cos \theta \leq 1 - \frac{2}{\pi^2} d^2(\theta, 2\pi\mathbb{Z}) = 1 - 8d^2(\theta/(2\pi), \mathbb{Z}) \quad (\theta \in \mathbb{R}).$$

As in the proof of Lemma 3.5, for any $\xi \in \mathbb{R}$,

$$\begin{aligned} |\varphi(\xi)|^2 &= \mathbb{E} \exp(-i\xi Y) = \mathbb{E} \cos(\xi Y) \\ &\leq 1 - 8\mathbb{E}d^2(\xi Y/(2\pi), \mathbb{Z}) \leq \exp\{-8\mathbb{E}d^2(\xi Y/(2\pi), \mathbb{Z})\}. \end{aligned}$$

Therefore,

$$|\varphi_\theta(\xi)| = \prod_{j=1}^n |\varphi(\xi\theta_j)| \leq \exp\left\{-4\mathbb{E} \sum_{j=1}^n d^2(\xi\theta_j Y/(2\pi), \mathbb{Z})\right\}.$$

□

The following lemma summarizes a few properties of the even function

$$S(\xi) = \sqrt{\mathbb{E}d^2\left(\frac{\xi Y}{2\pi}\theta, \mathbb{Z}^n\right)} \quad (\xi \in \mathbb{R}).$$

Recall the definition of $\mathcal{N}(\theta)$, for a unit vector $\theta \in S^{n-1}$.

Lemma 5.2. For any $\xi_1, \xi_2 \in \mathbb{R}$,

$$S(\xi_1 + \xi_2) \leq S(\xi_1) + S(\xi_2). \quad (32)$$

Furthermore, denote $R = \mathcal{N}(\theta) \geq 1$. Then, for any $|\xi| \leq n/(R\delta^4)$,

$$S(\xi) \geq c \min\left\{|\xi|, \frac{n/(R\delta^4)}{|\xi|}\right\}, \quad (33)$$

where $c > 0$ is a universal constant.

Proof. Note that for any $x, y \in \mathbb{R}^n$,

$$d(x + y, \mathbb{Z}^n) \leq d(x, \mathbb{Z}^n) + d(y, \mathbb{Z}^n).$$

The inequality (32) thus follows from the Cauchy-Schwartz inequality. Let us move to the proof of (33). From the definition of $\mathcal{N}(\theta)$,

$$d(\xi\theta, \mathbb{Z}^n) \geq \frac{1}{10} \min\left\{|\xi|, \frac{n/R}{|\xi|}\right\} = \frac{|\xi|}{10} \min\left\{1, \frac{n/R}{\xi^2}\right\} \quad \text{for all } |\xi| \leq n. \quad (34)$$

Since $Y = X - \tilde{X}$ then $\mathbb{E}Y^2 = 2$ and $\mathbb{E}Y^4 = 2\delta^4 + 6 \leq 8\delta^4$. Denote $\tilde{Y} = Y1_{|Y| \leq 5\delta^2}$. According to Lemma 3.1(ii),

$$\mathbb{E}\tilde{Y}^2 \geq 4/5.$$

For any $|\xi| \leq n/\delta^4$ we have $|\tilde{Y}\xi|/(2\pi) \leq \delta^2\xi \leq n$ and (34) yields

$$\begin{aligned} \mathbb{E}d^2\left(\frac{\xi Y}{2\pi}\theta, \mathbb{Z}^n\right) &\geq \mathbb{E}d^2\left(\frac{\xi \tilde{Y}}{2\pi}\theta, \mathbb{Z}^n\right) \geq \frac{1}{40\pi^2} \mathbb{E}\left(\xi \tilde{Y} \min\left\{1, \frac{n/R}{|\tilde{Y}\xi|^2/(4\pi^2)}\right\}\right)^2 \\ &\geq \frac{1}{40\pi^2} \xi^2 \min\left\{1, \frac{n^2/R^2}{\delta^8\xi^4}\right\} \mathbb{E}\tilde{Y}^2 \geq \frac{1}{800} \min\left\{\xi^2, \frac{n^2/R^2}{\delta^8\xi^2}\right\}. \end{aligned}$$

Therefore (33) holds for all $|\xi| \leq n/\delta^4$. \square

Lemma 5.3. *Let $0 < \alpha < 1, T \geq 1$ and suppose that $f : \mathbb{R} \rightarrow [0, \infty)$ is an even, measurable function which satisfies*

$$f(\xi_1 + \xi_2) \leq f(\xi_1) + f(\xi_2) \quad (\xi_1, \xi_2 \in \mathbb{R}) \quad (35)$$

and

$$f(\xi) \geq \alpha \min\left\{|\xi|, \frac{T}{|\xi|}\right\} \quad \text{for any } |\xi| \leq T. \quad (36)$$

Then,

$$\int_{T^{1/6} \leq |\xi| \leq T} \exp\{-f^2(\xi)\} \frac{d\xi}{|\xi|} \leq \frac{C}{\alpha^6 T}$$

where $C > 0$ is a universal constant.

Proof. Fix $r > 0$. Denote

$$A_r = \{T^{1/2} \leq \xi \leq T; r \leq f(\xi) < 2r\}.$$

Then $A_r \subset [\alpha T/(2r), T]$, thanks to (36). Furthermore, let $\xi_1, \xi_2 \in A_r$. We learn from (35) and from the fact that f is even that $f(\xi_1 - \xi_2) \leq f(\xi_1) + f(\xi_2) \leq 4r$. According to (36), either $|\xi_1 - \xi_2| \leq 4r/\alpha$, or else

$$|\xi_1 - \xi_2| \geq \alpha T/f(\xi_1 - \xi_2) \geq \alpha T/(4r).$$

Therefore, the set A_r can be covered by closed intervals of length at most $4r/\alpha$, and the distance between two such intervals is at least $\alpha T/(4r)$. For this specific purpose, the distance between two closed intervals means the distance between their left-most points. Since $A_r \subset [\alpha T/(2r), T]$ then the number of such intervals is at most $4r/\alpha + 1$. Consequently, for any $r > 0$,

$$\int_{A_r} \exp(-f^2(\xi)) \frac{d\xi}{\xi} \leq \left(\frac{4r}{\alpha} + 1\right) \cdot \frac{4r}{\alpha} \cdot \exp(-r^2) \cdot \frac{2r}{\alpha T} = \frac{8r^2(4r + \alpha)}{\alpha^3 T} \exp(-r^2). \quad (37)$$

From Fubini's theorem,

$$\int_{\sqrt{T} \leq |\xi| \leq T} e^{-f^2(\xi)} \frac{d\xi}{|\xi|} = 2 \sum_{i=-\infty}^{\infty} \int_{A_{2^i}} e^{-f^2(\xi)} \frac{d\xi}{|\xi|} \leq 4 \int_0^{\infty} \left(\int_{A_r} e^{-f^2(\xi)} \frac{d\xi}{|\xi|} \right) \frac{dr}{r}.$$

We plug in the information from (37) to obtain the bound

$$\int_{\sqrt{T} \leq |\xi| \leq T} \exp\{-f^2(\xi)\} \frac{d\xi}{|\xi|} \leq C \int_0^{\infty} \frac{r^2(r+\alpha)}{\alpha^3 T} \exp\{-r^2\} \frac{dr}{r} \leq \frac{\tilde{C}}{\alpha^3 T}. \quad (38)$$

Additionally, according to (36),

$$\int_{T^{1/6} \leq |\xi| \leq T^{1/2}} e^{-f^2(\xi)} \frac{d\xi}{|\xi|} \leq 2 \int_{T^{1/6}}^{\infty} e^{-\alpha^2 \xi^2} \frac{d\xi}{|\xi|} \leq \frac{C}{\alpha T^{1/6}} e^{-\alpha^2 T^{1/3}} \leq \frac{\bar{C}}{\alpha^6 T}. \quad (39)$$

The lemma follows from (38) and (39). \square

Proof of Theorem 1.2. Set $R = \mathcal{N}(\theta)$. We assume that $R\delta^4 \leq n$, since otherwise the conclusion of the theorem is vacuous. According to Lemma 2.1,

$$\int_{-[n/(R\delta^4)]^{1/6}}^{[n/(R\delta^4)]^{1/6}} \left| \varphi_{\theta}(\xi) - e^{-\xi^2/2} \right| \frac{d\xi}{|\xi|} \leq \frac{CR}{n} [\gamma^3 + \delta^4] \leq \frac{\tilde{C}R\delta^4}{n},$$

for $\gamma^3 = \mathbb{E}X^3 \leq \delta^2$. It still remains to bound the integral for $[n/(R\delta^4)]^{1/6} \leq |\xi| \leq n/(R\delta^4)$. To that end, we use Lemma 5.1, which states that,

$$|\varphi_{\theta}(\xi)| \leq \exp\{-4S^2(\xi)\} \quad (\xi \in \mathbb{R}).$$

According to Lemma 5.2, we may apply Lemma 5.3 for the function $S(\xi)$, with $T = n/(R\delta^4)$ and with α being a universal constant. We deduce that

$$\begin{aligned} & \int_{[n/(R\delta^4)]^{1/6} \leq |\xi| \leq n/(R\delta^4)} \left| \varphi_{\theta}(\xi) - e^{-\xi^2/2} \right| \frac{d\xi}{|\xi|} \\ & \leq C \exp\left\{-c[n/(R\delta^4)]^{1/3}\right\} + \int_{[n/(R\delta^4)]^{1/6} \leq |\xi| \leq n/(R\delta^4)} \exp\{-4S^2(\xi)\} \frac{d\xi}{|\xi|} \leq \frac{\tilde{C}\delta^4 R}{n}. \end{aligned}$$

The theorem now follows from (9). \square

Let us verify that $\mathcal{N}(\theta^0) \leq C$ for a universal constant $C > 0$, where θ^0 is the unit vector defined in (4).

Lemma 5.4. For any $\xi \in \mathbb{R}$,

$$d(\xi\theta^0, \mathbb{Z}^n) \geq \min\left\{|\xi|, \frac{cn}{|\xi|}\right\},$$

where $c > 0$ is a universal constant.

Proof. Liouville's theorem (see, e.g., [4, Section II.6]) states that for any integer $p, q \neq 0$,

$$\left| \sqrt{2} - \frac{q}{p} \right| \geq \frac{c_1}{p^2},$$

for some universal constant $c_1 > 0$. Let $|\xi| > 1/2$ and suppose that $p, q \in \mathbb{Z}$ are integers that satisfy $|\xi - p| = d(\xi, \mathbb{Z})$ and $|\xi\sqrt{2} - q| = d(\xi\sqrt{2}, \mathbb{Z})$. Then,

$$\frac{c}{|\xi|} \leq \frac{c_1}{|p|} \leq \left| p\sqrt{2} - q \right| \leq |p\sqrt{2} - \xi\sqrt{2}| + |\xi\sqrt{2} - q| = \sqrt{2}d(\xi, \mathbb{Z}) + d(\xi\sqrt{2}, \mathbb{Z}).$$

We deduce that for any $\xi \in \mathbb{R}$,

$$d^2(\xi, \mathbb{Z}) + d^2(\xi\sqrt{2}, \mathbb{Z}) \geq \min\{3\xi^2, \tilde{c}\xi^{-2}\}.$$

According to the definition of the unit vector θ^0 , and see that for $\xi \in \mathbb{R}$,

$$d^2(\xi\theta^0, \mathbb{Z}^n) = \frac{n}{2} \left[d^2\left(\sqrt{\frac{2}{3n}}\xi, \mathbb{Z}\right) + d^2\left(\sqrt{\frac{2}{3n}}\xi\sqrt{2}, \mathbb{Z}\right) \right] \geq \min\{\xi^2, cn^2/\xi^2\}.$$

□

References

- [1] Adamczak, R., Litvak, A. E., Pajor, A., Tomczak-Jaegermann, N., *Restricted isometry property of matrices with independent columns and neighborly polytopes by random sampling*. Preprint. Available under <http://arxiv.org/abs/0904.4723>
- [2] Barthe, F., Cordero-Erausquin, D., Ledoux, M., Maurey, M., *Correlation and Brascamp-Lieb inequalities for Markov semigroups*. Preprint. Available under <http://arxiv.org/abs/0907.2858>
- [3] Carlen, E., Lieb, E., Loss, M., *A sharp analog of Young's inequality on S^N and related entropy inequalities*. J. Geom. Anal., 14, (2004), 487–520.
- [4] Courant, R., Robbins, R., Stewart, I., *What Is Mathematics? An Elementary Approach to Ideas and Methods*. Oxford University Press, 1996.
- [5] Diaconis, P., Freedman, D., *Asymptotics of graphical projection pursuit*. Ann. Statist., 12, no. 3, (1984), 793–815.
- [6] Diaconis, P., Freedman, D., *A dozen de Finetti-style results in search of a theory*. Ann. Inst. H. Poincaré Probab. Statist., 23, no. 2, (1987), 397–423.
- [7] Feller, W., *An introduction to probability theory and its applications, volume I+II*. John Wiley & Sons, Inc., New York-London-Sydney, 1971.

- [8] Friedland, O., Sodin, S., *Bounds on the concentration function in terms of the Diophantine approximation*. C. R. Math. Acad. Sci. Paris, 345, no. 9, (2007), 513–518.
- [9] Hitczenko, P., Montgomery-Smith, S. J., Oleszkiewicz, K., *Moment inequalities for sums of certain independent symmetric random variables*. Studia Math., 123, no. 1, (1997), 15–42.
- [10] Ibragimov, I., Linnik, Yu. V., *Independent and stationary sequences of random variables*. Translation from the Russian, edited by J. F. C. Kingman. Wolters-Noordhoff Publishing, Groningen, 1971.
- [11] Kahane, J. P., *Some random series of functions*. Second edition. Cambridge Studies in Advanced Mathematics, 5. Cambridge University Press, Cambridge, 1985.
- [12] Klartag, B., *A Berry-Esseen type inequality for convex bodies with an unconditional basis*. Probab. Theory Related Fields, vol. 45, no. 1, (2009), 1–33.
- [13] Rudelson, M., Vershynin, R., *The Littlewood-Offord problem and invertibility of random matrices*. Adv. Math. 218, no. 2, (2008), 600–633.
- [14] Sudakov, V. N., *Typical distributions of linear functionals in finite-dimensional spaces of high-dimension*. (Russian) Dokl. Akad. Nauk. SSSR, 243, no. 6, (1978), 1402–1405. English translation in Soviet Math. Dokl., 19, (1978), 1578–1582.

School of Mathematical Sciences, Tel-Aviv University, Tel-Aviv 69978, Israel
e-mail address: [klartagb,sodinale]@tau.ac.il