Les Cahiers du GERAD ISSN: 0711–2440

Densities of sums and small ball probability

J. K. Dzahini

G–2017–68

August 2017

La publication de ces rapports de recherche est rendue possible grâce au soutien de HEC Montréal, Polytechnique Montréal, Université McGill, Université du Québec à Montréal, ainsi que du Fonds de recherche du Québec – Nature et technologies.

Dépôt légal – Bibliothèque et Archives nationales du Québec, 2017
– Bibliothèque et Archives Canada, 2017

The publication of these research reports is made possible thanks to the support of HEC Montréal, Polytechnique Montréal, McGill University, Université du Québec à Montréal, as well as the Fonds de recherche du Québec – Nature et technologies.

Legal deposit – Bibliothèque et Archives nationales du Québec, 2017
– Library and Archives Canada, 2017

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

Densities of sums and small ball probabilities

Kwassi Joseph Dzahini [a]

[a] *Department of Mathematics and Industrial Engineering, Polytechnique Montréal (Québec) Canada, H3C 3A7*

kwassi-joseph.dzahini@polymtl.ca

August 2017
Les Cahiers du GERAD
G–2017–68

Abstract: We propose a lemma that clarifies the proof of Theorem 4.1 on densities of sums in [3]. More precisely, by denoting by f_{S+Y} the density of an absolutely continuous real-valued random variable S augmented by an independent real-valued Gaussian random variable Y with mean zero and an arbitrarily small variance, we prove that if f_{S+Y} is bounded almost everywhere by a strictly positive constant C, then almost everywhere, the density f_S is also bounded by the same constant C. Then, using these results, we show how small ball probability estimates such as

$$\mathbb{P}\left(\left|\sum_{k=1}^{n} a_k \xi_k\right| \leq \varepsilon\right) \leq C\varepsilon \quad \text{for all} \quad \varepsilon > 0,$$

with a_k's real numbers still hold when a_k's are arbitrary random variables.

1 Notations

Given a vector $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, we denote by $\|*\|_2$ the ℓ_2-norm defined as:

$$\|x\|_2 = \left(\sum_{j=1}^{n} |x_j|^2 \right)^{\frac{1}{2}},$$

and $\mathbb{S}^{n-1}$ stands for the unit sphere of this norm, that is

$$\mathbb{S}^{n-1} = \{x = (x_1, \dots, x_n) \in \mathbb{R}^n : \|x\|_2 = 1\}.$$

Let $d \in \mathbb{N}^*$ and $p \in [1, +\infty]$:
We denote by $L^p(\mathbb{R}^d)$ the Lebesgue space of (classes of) measurable functions

$$u : \mathbb{R}^d \longrightarrow \mathbb{C}$$

such that:
$\star$ If $1 \le p < +\infty$, $\int_{\mathbb{R}^d} |u(x)|^p \, dx < +\infty$.
$\star$ If $p = +\infty$, $\operatorname{supess}_{x \in \mathbb{R}^d} |u(x)| := \inf\{M \ge 0, |u(x)| \le M \text{ a.e}\} < +\infty$,
where, "a.e": almost everywhere, means that $|u(x)| \le M$ except on a subset of $\mathbb{R}^d$ with Lebesgue measure zero.
Let $1 \le p < +\infty$, then

$$\|u\|_{L^p(\mathbb{R}^d)} = \left(\int_{\mathbb{R}^d} |u(x)|^p \, dx \right)^{\frac{1}{p}},$$

and

$$\|u\|_{L^\infty(\mathbb{R}^d)} = \operatorname*{supess}_{x \in \mathbb{R}^d} |u(x)|.$$

If $x \in \mathbb{R}^d$, we denote by x^T its transpose.
The scalar product of two column vectors $x = (x_1, \dots, x_d)$ and $y = (y_1, \dots, y_d)$ of $\mathbb{R}^d$ is denoted by $x \cdot y$ or $\langle x, y \rangle$, and defined as

$$x \cdot y = \langle x, y \rangle = x^T y = \sum_{k=1}^{d} x_k y_k.$$

The Fourier transform of a function $f \in L^1(\mathbb{R}^d)$ is denoted by $\hat{f}$, and defined for $\xi \in \mathbb{R}^d$ as:

$$\hat{f}(\xi) = \int_{\mathbb{R}^d} f(x) e^{-i\xi \cdot x} dx.$$

2 Introduction

In [3], to prove Theorem 4.1 about "Densities of sums" which is as follows,

Theorem 1 (Densities of sums) *Let $X_1, \dots, X_n$ be real-valued independent random variables whose densities are bounded by $K > 0$ almost everywhere. Then there exists a strictly positive constant C independent of n such that for all sequence of (deterministic) real numbers $a_1, \dots, a_n$ satisfying $\sum_{j=1}^n a_j^2 = 1$, the density of $S = \sum_{j=1}^n a_j X_j$ is bounded by CK almost everywhere.*

Mark Rudelson and Roman Vershynin had stated that one "may assume that $\phi_{X_j} \in L^1(\mathbb{R})$ by adding to X_j an independent normal random variable with an arbitrarily small variance", and that "Fourier inversion formula associated with the Fourier transform yields that the density of S at the origin can be reconstructed from its Fourier transform", where ϕ_{X_j} denotes the characteristic function of X_j. They then bounded the density $f_S(0)$ of S at the origin (under many other assumptions) to complete their proof. However, one way of making this proof simpler and clearer is to define the real-valued random variable $S_m = S + a_1 Y_m = a_1(X_1 + Y_m) + \sum_{j=2}^{n} a_j X_j$, where Y_m is a centered real-valued Gaussian random variable, independent of X_j for all $1 \leq j \leq n$, with variance ϵ_m^2 satisfying $\lim\limits_{m \to +\infty} \epsilon_m = 0$, then prove that $f_{S_m}(0)$ is bounded and then show how this last result implies that $f_S(0)$ is bounded. This leads us to state the following lemma which is our main result.

Lemma 1 *Let S be a real-valued random variable with density f_S and $(Y_n)_{n\geq 1}$ a sequence of Gaussian random variables independent of S, with mean zero and variance ϵ_n^2 for all $n \geq 1$, where $(\epsilon_n)_{n\geq 1}$ is a sequence of positive real numbers tending to zero at infinity. We assume that almost everywhere the random variable $S_n = S + Y_n$ has a density f_{S_n} bounded by a strictly positive constant C independent of n. Then almost everywhere, f_S is bounded by C.*

In this paper, we first give a complete proof of Lemma 1 in Section 3; then we show in Section 4 how it implies the main result in Theorem 1 using the same suggested assumptions and approaches of M. Rudelson and R. Vershynin in [3]. Futhermore, using an approach based on a basic property of conditional expectation, we also decided to give in Section 5 a rigorous proof of the fact that, small ball probability estimates such as

$$\mathbb{P}\left(\left|\sum_{k=1}^{n} a_k \xi_k\right| \leq \varepsilon\right) \leq C\varepsilon \quad \text{for all} \quad \varepsilon > 0, \tag{1}$$

with $C > 0$ independent of n, $a = (a_1, \ldots, a_n) \in \mathbb{S}^{n-1}$ a deterministic vector and $X_n = (\xi_1, \ldots, \xi_n)$ a random vector, still hold with the same constant C when $a = (a_1, \ldots, a_n)$ is an arbitrary random vector belonging to the unit sphere of $\mathbb{R}^n$. In fact, such an estimate follows obviously from Theorem 1 when $a \in \mathbb{S}^{n-1}$ is a deterministic vector and was one of the most important ingredients that helped to solve (when $a \in \mathbb{S}^{n-1}$ is a random vector) square matrices (with subGaussian entries and bounded density) invertibility problem in [4].

3 Proof of Lemma 1

In this section we present the proof of Lemma 1 that helped us to make the proof of Theorem 1 clearer.

Proof. Let's suppose that there exists $0 < \delta < \frac{C}{4}$ and an interval $I = [a, b]$ with non-empty interior such that almost everywhere,

$$\text{for all } y \in I, \;\; f_S(y) \geq C + \delta. \tag{2}$$

Let $a < t < b$ and $F_n = \left[\frac{t-b}{\epsilon_n}, \frac{t-a}{\epsilon_n}\right]$. Using the independence of S and Y_n, we have:

$$f_{S_n}(t) = \frac{1}{\sqrt{2\pi}\epsilon_n} \int_{\mathbb{R}} f_S(t-x) e^{-\frac{x^2}{2\epsilon_n^2}} dx. \tag{3}$$

By the change of variable $y = \frac{x}{\epsilon_n}$, we get almost everywhere,

$$f_{S_n}(t) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f_S(t - \epsilon_n y) e^{-\frac{y^2}{2}} dy \quad > \quad \frac{1}{\sqrt{2\pi}} \int_{F_n} f_S(t - \epsilon_n y) e^{-\frac{y^2}{2}} dy \geq (C + \delta) J_n, \tag{4}$$

where for all $n \geq 1$,

$$J_n = \frac{1}{\sqrt{2\pi}} \int_{F_n} e^{-\frac{y^2}{2}} dy. \tag{5}$$

Since the sequence of intervals $(F_n)_{n\geq 1}$ converges to $\mathbb{R}$ when n tends to $+\infty$, then according to Lebesgue Dominated Convergence Theorem,

$$\lim_{n \to +\infty} J_n = 1. \tag{6}$$

Hence there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, $J_n > 1 - \frac{\delta}{2C}$. Thus, almost everywhere,

$$f_{S_n}(t) > (C+\delta)\left(1 - \frac{\delta}{2C}\right) = C + \frac{\delta}{2} - \frac{\delta^2}{2C} > C + \frac{3\delta}{8}, \tag{7}$$

which is absurd since $f_{S_n}(t) \leq C$ almost everywhere. □

4 Proof of Theorem 1

Now we show how Lemma 1 implies the main result in Theorem 1 using the same suggested assumptions and approaches of M. Rudelson and R. Vershynin in [3]. For this purpose, we need in addition the following results:

Definition 1 (Distribution function and non-increasing rearrangement [8]) *Let E be a measurable subspace of $\mathbb{R}^n$ and $f : E \longrightarrow \mathbb{R}$ a measurable function. Let m be the Lebesgue measure of $\mathbb{R}^n$. The distribution function μ_f of f is given by:*

$$\mu_f(\lambda) := m(\{x \in E : |f(x)| > \lambda\}), \quad \textit{for all } \lambda \geq 0, \tag{8}$$

and its non-increasing rearrangement f^* is defined by:

$$f^\star(t) := \inf\{\lambda \geq 0, \mu_f(\lambda) \leq t\}, \quad \textit{for all } \ t \in (0, m(E)). \tag{9}$$

We have the following properties:

Lemma 2 *[6] If $f \in L^p(E)$, $p \geq 1$, then:*

i)

$$\int_E |f|^p dm = p \int_0^{m(E)} \lambda^{p-1} \mu_f(\lambda) d\lambda. \tag{10}$$

ii)

$$\|f\|_{L^p(E)} = \left(\int_E |f|^p \, dm\right)^{\frac{1}{p}} = \left(\int_0^{m(E)} [f^\star(t)]^p dt\right)^{\frac{1}{p}} = \|f^\star\|_{L^p((0,m(E)))} \, . \tag{11}$$

Lemma 3 (Decay of characteristic functions [3]) *Let X be a random variable whose density is bounded by $K > 0$. Then the non-increasing rearrangement of the characteristic function $\phi_X(t) = \mathbb{E}\left(e^{itX}\right)$ of X satisfies:*

$$|\phi_X|^\star(t) \leq \begin{cases} 1 - c\left(\frac{t}{K}\right)^2 & \textit{if } 0 < t < 2\pi K \\ \sqrt{\frac{2\pi K}{t}} & \textit{if } t \geq 2\pi K, \end{cases} \tag{12}$$

where c is a strictly positive constant.

Lemma 4 (The Fourier inversion formula [1]) *Let $d \in \mathbb{N}^*$ and $f \in L^1(\mathbb{R}^d)$ with $\hat{f}$ its Fourier transform. We assume that $\hat{f} \in L^1(\mathbb{R}^d)$. Then almost everywhere:*

$$f(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \hat{f}(\xi) e^{i\xi\cdot x} d\xi \qquad i.e \qquad \check{f} = \frac{1}{(2\pi)^d} \hat{\hat{f}}, \tag{13}$$

where $\check{f}(\cdot) = f(-\cdot)$.

Lemma 5 (Convolution and inequalities [1]) *Let $p, q \in [1, +\infty]$ satisfying $\frac{1}{p} + \frac{1}{q} = 1$ and $d \in \mathbb{N}^*$.*

i) If $f \in L^1(\mathbb{R}^d)$ and $g \in L^p(\mathbb{R}^d)$, then $f * g(x)$ exists for almost all $x \in \mathbb{R}^d$, $f * g \in L^p(\mathbb{R}^d)$ and:

$$\|f * g\|_{L^p(\mathbb{R}^d)} \leq \|f\|_{L^1(\mathbb{R}^d)} \|g\|_{L^p(\mathbb{R}^d)} \, . \tag{14}$$

ii) If $f \in L^p(\mathbb{R}^d)$ and $g \in L^q(\mathbb{R}^d)$, then $f * g(x)$ exists for almost all $x \in \mathbb{R}^d$, $f * g \in L^\infty(\mathbb{R}^d)$ and:

$$\|f * g\|_{L^\infty(\mathbb{R}^d)} \leq \|f\|_{L^p(\mathbb{R}^d)} \|g\|_{L^q(\mathbb{R}^d)} . \tag{15}$$

Moreover, $f * g$ is uniformly continuous.

Proof. (Theorem 1) Let X be an absolutely continuous real-valued random variable. We denote by f_X its density with respect to the Lebesgue measure on $\mathbb{R}$. Let $\beta \in \mathbb{R}^*$ and $g : \mathbb{R} \longrightarrow \mathbb{R}$ be a continuous and bounded function. We have

$$\mathbb{E}\left(g(\beta X)\right) = \int_{\mathbb{R}} g(y) \times \frac{1}{\beta} f_X\left(\frac{y}{\beta}\right) dy,$$

then the density of βX is $f_{\beta X}(x) = \frac{1}{\beta} f_X\left(\frac{x}{\beta}\right)$. Thus, "by replacing X_j with KX_j, we can assume that $K = 1$. By replacing X_j with $(-X_j)$ when necessary, we can assume that for all $1 \leq j \leq n$, $a_j \geq 0$. We can further assume that $a_j > 0$ by dropping all zero terms from the sum."

Let's suppose that there exists $1 \leq j_0 \leq n$ such that $a_{j_0} > \frac{1}{2}$. By reodering the numbers a_j, we can assume that $j_0 = n$. We denote by h_n the density of $\sum_{j=1}^n a_j X_j$. We have:

$$\begin{aligned} h_n(x) = h_{n-1} * f_{a_n X_n}(x) &= \int_{\mathbb{R}} h_{n-1}(y) \times \frac{1}{a_n} f_{X_n}\left(\frac{x-y}{a_n}\right) dy \\ &\leq 2 \int_{\mathbb{R}} h_{n-1}(y) dy = 2. \end{aligned}$$

Now we assume in the rest of the proof that a_j is fixed, and $0 < a_j \leq \frac{1}{2}$ for all $1 \leq j \leq n$.

The density of S is

$$f_S(x) = h_n(x) = f_{a_1 X_1} * f_{a_2 X_2} * f_{a_3 X_3} * \cdots * f_{a_{n-1} X_{n-1}} * f_{a_n X_n}(x)$$

and for all $1 \leq i \leq n$,

$$\int_{\mathbb{R}} f^2_{a_i X_i}(x) dx \leq \frac{1}{a_i} \int_{\mathbb{R}} f_{a_i X_i}(x) dx = \frac{1}{a_i} < +\infty$$

whence $f_{a_i X_i} \in L^2(\mathbb{R}) \cap L^1(\mathbb{R})$. Thus, it follows from Lemma 5 that $f_{a_1 X_1} * f_{a_2 X_2} \in L^2(\mathbb{R})$ which implies $(f_{a_1 X_1} * f_{a_2 X_2}) * f_{a_3 X_3} \in L^2(\mathbb{R})$ and step by step we finally get $h_{n-1} \in L^2(\mathbb{R})$. It follows again from Lemma 5 that

$$f_S = h_{n-1} * f_{a_n X_n} \in L^\infty(\mathbb{R})$$

and moreover f_S is uniformly continuous.

For all $t \in \mathbb{R}$, let

$$\tilde{X}_j = X_j - a_j^2 t \;\text{ and }\; \tilde{S}_t = \sum_{j=1}^n \tilde{X}_j = S - t.$$

Since f_S is continuous, then

$$f_S(t) = f_{\tilde{S}_t}(0) \quad \text{for all } t \in \mathbb{R},$$

thus

$$\|f_S\|_{L^\infty(\mathbb{R})} \leq C_0 \iff \forall t \in \mathbb{R} \left|f_{\tilde{S}_t}(0)\right| \leq C_0, \quad \text{with } C_0 > 0,$$

in other words, by translating X_j when necessary, our problem is reduced to a problem of boundary of the density of S at the origin.

Let Y_m be a real-valued Gaussian random variable with mean zero, independent of X_j for all $1 \leq j \leq n$ and with variance ϵ_m^2 satisfying $\lim_{m \to +\infty} \epsilon_m = 0$. For all $2 \leq j \leq n$, Let's denote by ϕ_j the characteristic function of X_j and by ϕ_1^m the characteristic function of X_1 augmented by the Gaussian random variable Y_m, that is $\phi_1^m(t) = \mathbb{E}\left(e^{it(X_1 + Y_m)}\right)$. Let $S_m = S + a_1 Y_m = a_1(X_1 + Y_m) + \sum_{j=2}^n a_j X_j$.

$$\int_{\mathbb{R}} |\phi_1^m(t)|dt = \int_{\mathbb{R}} \left|\mathbb{E}\left(e^{itX_1}e^{itY_m}\right)\right| dt \leq \int_{\mathbb{R}} e^{-\frac{1}{2}t^2\epsilon_m^2}dt < \infty,$$

then $\phi_1^m \in L^1(\mathbb{R})$. Hence,

$$\int_{\mathbb{R}} |\phi_{S_m}(t)|dt \leq \int_{\mathbb{R}} |\phi_1^m(a_1 t)|dt < \infty.$$

Since

$$\phi_{S_m}(-t) = \mathbb{E}\left(e^{-itS_m}\right) = \int_{\mathbb{R}} e^{-it\xi} f_{S_m}(\xi)d\xi = \hat{f}_{S_m}(t),$$

it follows that $\hat{f}_{S_m} \in L^1(\mathbb{R})$ whence by Fourier inversion formula (Lemma 4), almost everywhere,

$$f_{S_m}(0) = \frac{1}{2\pi}\int_{\mathbb{R}} \hat{f}_{S_m}(t)dt = \frac{1}{2\pi}\int_{\mathbb{R}} \phi_{S_m}(t)dt \leq \int_{\mathbb{R}} |\phi_{S_m}(t)|dt := I. \tag{16}$$

It follows from the independence of $X_1, \ldots, X_n, Y_m$ that:

$$\begin{aligned}\phi_{S_m}(t) &= \mathbb{E}\left(e^{it\left(a_1(X_1+Y_m)+\sum_{j=2}^n a_jX_j\right)}\right) = \mathbb{E}\left(e^{ia_1t(X_1+Y_m)}\prod_{j=2}^n e^{ia_jtX_j}\right)\\ &= \mathbb{E}\left(e^{ia_1t(X_1+Y_m)}\right)\prod_{j=2}^n \mathbb{E}\left(e^{ia_jtX_j}\right) = \phi_1^m(a_1t)\prod_{j=2}^n \phi_j(a_jt),\end{aligned}$$

then

$$I = \int_{\mathbb{R}} |\phi_1^m(a_1t)| \prod_{j=2}^n |\phi_j(a_jt)|dt.$$

In the rest of the proof, we set

$$\phi_1 := \phi_1^m.$$

Using the equality $\sum_{j=1}^n a_j^2 = 1$, and the generalized Hölder inequality [2], we get:

$$I \leq \prod_{j=1}^n \left(\int_{\mathbb{R}} |\phi_j(a_jt)|^{\frac{1}{a_j^2}} dt\right)^{a_j^2}.$$

$|\phi_j|$ is then replaced by $|\phi_j|^\star$ in the above integrals whose values remain the same according to Lemma 2, followed by the change of variable $a_jt = u$. Hence,

$$I \leq \prod_{j=1}^n \left(\frac{1}{a_j}\int_0^{+\infty} \left[|\phi_j|^\star(t)\right]^{\frac{1}{a_j^2}} dt\right)^{a_j^2}. \tag{17}$$

Let

$$I_j := \int_0^{+\infty} \left[|\phi_j|^\star(t)\right]^{\frac{1}{a_j^2}} dt.$$

Using Lemma 3, we get:

$$I_j = \int_0^{2\pi} \left[|\phi_j|^\star(t)\right]^{\frac{1}{a_j^2}} dt + \int_{2\pi}^{+\infty} \left[|\phi_j|^\star(t)\right]^{\frac{1}{a_j^2}} dt \leq I_j^1 + I_j^2$$

where

$$I_j^1 = \int_0^{2\pi} (1-ct^2)^{\frac{1}{a_j^2}} dt \quad \text{and} \quad I_j^2 = \int_{2\pi}^{+\infty} \left[\sqrt{\frac{2\pi}{t}}\right]^{\frac{1}{a_j^2}} dt.$$

Using the inequality $1 - t \leq e^{-t}$ for all $t \in \mathbb{R}$ we get

$$I_j^1 \leq \int_0^{2\pi} e^{-\frac{ct^2}{a_j^2}} dt \leq \int_0^{+\infty} \exp\left(-\frac{t^2}{2\left(\frac{a_j}{\sqrt{2c}}\right)^2}\right) dt = \sqrt{2\pi} \times \frac{a_j}{\sqrt{c}} = Ca_j.$$

Since $\frac{1}{2a_j^2} \geq 2$, the integral

$$I_j^2 = (2\pi)^{\frac{1}{2a_j^2}} \int_{2\pi}^{+\infty} \frac{1}{t^{\frac{1}{2a_j^2}}} dt$$

is a convergent Riemann integral satisfying:

$$\begin{aligned} I_j^2 &= (2\pi)^{\frac{1}{2a_j^2}} \left(\frac{1}{1 - \frac{1}{2a_j^2}} \right) \left[\frac{1}{t^{\frac{1}{2a_j^2}-1}} \right]_{2\pi}^{+\infty} = (2\pi)^{\frac{1}{2a_j^2}} \times (2\pi)^{1-\frac{1}{2a_j^2}} \times \frac{1}{\frac{1}{2a_j^2} - 1} \\ &= \frac{2\pi}{\frac{1}{2a_j^2} - 1} = \frac{4\pi a_j^2}{1 - 2a_j^2} \leq 8\pi a_j^2. \end{aligned}$$

Thus,

$$I_j \leq C a_j + 8\pi a_j^2 \leq (C + 8\pi) a_j \quad \text{since } a_j \in (0,1) \Rightarrow a_j^2 < a_j$$

whence $I_j \leq 2Ca_j$ for C large enough.

Hence it follows from (17) and the previous inequality that:

$$I \leq \prod_{j=1}^{n} \left(\frac{1}{a_j} \times 2Ca_j \right)^{a_j^2} = (2C)^{\sum_{j=1}^n a_j^2} = 2C < +\infty.$$

Thus, it follows from (16) that

$$f_{S_m}(0) = f_{S+a_1Y_m}(0) \leq I \leq 2C,$$

whence, using Lemma 1 we finally get,

$$f_S(0) \leq 2C,$$

and the proof is complete. □

5 A small ball probability estimate

In [4], after having reduced via Lemma 5.6, the invertibility problem of square random matrices A with subgaussian entries and bounded density, to a lower bound on the distance between a random vector and a random subspace, M. Rudelson then reduced bounding the distance to a small ball probability estimate". More precisely, given a *"random normal"* $X^* = (a_1, \dots, a_n) \in \mathbb{S}^{n-1}$ and the $n-th$ column $X_n = (\xi_1, \dots, \xi_n)$ of A, he made use of the following estimate:

$$\mathbb{P}\left(|\langle X^*, X_n \rangle| < t\right) = \mathbb{P}\left(\left| \sum_{k=1}^{n} a_k \xi_k \right| < t \right) \leq Ct \quad \text{for all } t > 0, \tag{18}$$

where C is a strictly positive constant independent of n. In fact, if X^* had deterministic components a_k, then almost everywhere the density of the real-valued random variable $\sum_{k=1}^n a_k \xi_k$ would have been bounded according to Theorem 1 and (18) would have followed obviously from this same theorem but unfortunately, this is not the case so one needs to prove rigorously that such an estimate holds when $(a_1, \dots, a_n) \in \mathbb{S}^{n-1}$ is an arbitrary random vector. We therefore propose a complete proof. More precisely, we show how such an estimate can be deduced from Theorem 1.

Theorem 2 (Small ball probability) *Let $\xi_1, \dots, \xi_n$ be real-valued independent random variables whose densities are bounded by $K > 0$ almost everywhere. Then there exists a strictly positive (deterministic) constant C independent of n such that for all arbitrary* **random vector** *$a = (a_1, \dots, a_n) \in \mathbb{S}^{n-1}$ independent of $X_n = (\xi_1, \dots, \xi_n)$,*

$$\mathbb{P}\left(|\langle a, X_n \rangle| < t\right) = \mathbb{P}\left(\left| \sum_{k=1}^{n} a_k \xi_k \right| < t \right) \leq 2CKt \quad \textit{for all } t > 0. \tag{19}$$

Before we prove Theorem 2, we also need in addition to Theorem 1, the following result which constitutes our main tool.

Lemma 6 (Substitution property of conditional expectation [7]) *Let U, V be random maps into $(S_1, \mathcal{S}_1)$ and $(S_2, \mathcal{S}_2)$ respectively. Let ψ be a measurable real-valued function on $(S_1 \times S_2, \mathcal{S}_1 \bigotimes \mathcal{S}_2)$. If U is $\mathcal{G}$-mesurable, $\sigma(V)$ and $\mathcal{G}$ are independent and $\mathbb{E}\left(|\psi(U,V)|\right) < +\infty$, then one has that*

$$\mathbb{E}\left(\psi(U,V)|\mathcal{G}\right) = h(U) \quad \text{where} \ \ h(u) := \mathbb{E}\left(\psi(u,V)\right). \tag{20}$$

Proof. (Theorem 2) Let

$$S_n^{a,\xi} = \sum_{k=1}^{n} a_k \xi_k.$$

We want to estimate the probability

$$\mathbb{P}\left(|\langle a, X_n\rangle| < t\right) = \mathbb{P}\left(\left|S_n^{a,\xi}\right| < t\right) \quad \text{for all } t > 0.$$

Let ψ_t be the measurable real-valued function on $(\mathbb{R}^n \times \mathbb{R}^n, \mathcal{B}(\mathbb{R}^n) \bigotimes \mathcal{B}(\mathbb{R}^n))$ defined by

$$\psi_t(a, X_n) = \mathbf{1}_{|\langle a, X_n\rangle| < t}.$$

Since $a \perp\!\!\!\perp X_n$, then the σ-algebra $\sigma(a)$ is independent of $\sigma(X_n)$. ψ_t satisfies the condition

$$\mathbb{E}\left(|\psi_t(a, X_n)|\right) < +\infty$$

then it follows from Lemma 6 that:

$$\mathbb{E}\left(\psi_t(a, X_n)|\sigma(a)\right) = \mathbb{E}\left(\mathbf{1}_{|\langle a, X_n\rangle| < t}|\sigma(a)\right) = h_t(a) \tag{21}$$

Let $(b_1, \ldots, b_n) \in \mathbb{R}^n$ be a deterministic vector, one has

$$h_t((b_1, \ldots, b_n)^T) = \mathbb{E}\left(\mathbf{1}_{\left|\sum_{k=1}^n b_k \xi_k\right| < t}\right) = \mathbb{P}\left(\left|\sum_{k=1}^{n} b_k \xi_k\right| < t\right). \tag{22}$$

Hence, since $\forall \omega \in \Omega,\ a(\omega) = (a_1(\omega), \ldots, a_n(\omega)) \in \mathbb{S}^{n-1}$ is a deterministic vector, it follows from Theorem 1 that, for all $\omega \in \Omega$,

$$h_t(a(\omega)) = \mathbb{P}\left(\left|\sum_{k=1}^{n} a_k(\omega) \xi_k\right| < t\right) = \int_{-t}^{t} f_{S_n^{a(\omega),\xi}}(u) du \leq 2CKt. \tag{23}$$

where the term on the right-hand side doesn't depend on ω, thus it follows from a basic property of conditional expectation (see point (b) of Theorem 2.7 in [7]) that:

$$\mathbb{E}\left(\mathbf{1}_{|\langle a, X_n\rangle| < t}\right) = \mathbb{E}\left[\mathbb{E}\left(\mathbf{1}_{|\langle a, X_n\rangle| < t}|\sigma(a)\right)\right] = \mathbb{E}\left(h_t(a)\right) \leq 2CKt, \tag{24}$$

that is,

$$\mathbb{P}\left(|\langle a, X_n\rangle| < t\right) = \mathbb{P}\left(\left|\sum_{k=1}^{n} a_k \xi_k\right| < t\right) \leq 2CKt. \tag{25}$$

□

References

[1] Arthur Leclaire, Analyse de Fourier, Prparation l'agrgation, ENS Cachan, Mathmatiques, Anne 2015–2016.

[2] Helmut Finner, A Generalization of Hlder's Inequality And Some Probability Inequalities, The Annals of Probability, 1992, 20(4), 1893–1901

[3] M. Rudelson, R. Vershynin, Small ball probabilities for linear images of high dimensional distributions.

[4] M. Rudelson, Lecture notes on non-asymptotic theory of random matrices, AMS Short Course on Random Matrices, San Diego, January 2013

[5] M. Rudelson, R. Vershynin, The Littlewood-Offord Problem and invertibility of random matrices, Adv. Math. 2008, 218 (2), 600–633.

[6] R. O'Neil, G. WEISS, The Hilbert transform and rearrangement of functions, ICM Studia Mathematica, T. XXIII. 1963

[7] Rabi Bhattacharya, Edward C. Waymire, A basic course in Probability Theory

[8] Ren Erlin Castillo, Humberto Rafeiro, An Introductory Course in Lebesgue Space, CMS Books in Mathematics