Berry–Esseen bounds
for typical weighted sums

S. G. Bobkov[*†‡]   G. P. Chistyakov[§†]   F. Götze[¶†]

Abstract

Under correlation-type conditions, we derive upper bounds of order $\frac{1}{\sqrt{n}}$ for the Kolmogorov distance between the distributions of weighted sums of dependent summands and the normal law.

Keywords: Sudakov's typical distributions; normal approximation.
AMS MSC 2010: 60E; 60F.
Submitted to EJP on October 12, 2017, final version accepted on July 4, 2018.
Supersedes arXiv:1709.06848.

1 Introduction

Given a random vector $X = (X_1, \ldots, X_n)$ in the Euclidean space $\mathbb{R}^n$ ($n \geq 2$), we consider the weighted sums

$$S_\theta = \theta_1 X_1 + \cdots + \theta_n X_n, \qquad \theta = (\theta_1, \ldots, \theta_n) \in S^{n-1},$$

parameterized by points of the unit sphere $S^{n-1} = \{\theta \in \mathbb{R}^n : \theta_1^2 + \cdots + \theta_n^2 = 1\}$. According to the celebrated result by Sudakov [23], if n is large, and if the covariance matrix of X has bounded spectral radius, the distribution functions $F_\theta(x) = \mathbb{P}\{S_\theta \leq x\}$ concentrate around a certain typical distribution function given by the mean

$$F(x) = \mathbb{E}_\theta F_\theta(x) \equiv \int_{S^{n-1}} F_\theta(x) \, d\mu_{n-1}(\theta), \qquad x \in \mathbb{R}, \tag{1.1}$$

over the uniform probability measure μ_{n-1} on S^{n-1}. In contrast to the classical scheme of independent summands, this theorem has a much wider range of applicability. However,

the problem of possible rates of concentration, including the rates for the μ_{n-1}-mean of the Kolmogorov distance

$$\rho(F_\theta, F) = \sup_x |F_\theta(x) - F(x)|,$$

is rather delicate, and the answers depend upon correlation-type characteristics of the distribution of X. A natural characteristic is for example the maximal L^p-norm

$$M_p = \sup_\theta \left(\mathbb{E}|S_\theta|^p\right)^{1/p}, \qquad p \geq 1.$$

Moreover, if we want to study the approximation for most of F_θ's by the standard normal distribution function

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-y^2/2}\, dy, \qquad x \in \mathbb{R},$$

one is led to study another concentration problem – namely rates for the distance $\rho(F, \Phi)$. To this aim, let us rewrite the definition (1.1) as $F(x) = \mathbb{P}\{rZ_n \leq x\}$ with

$$r^2 = \frac{|X|^2}{n} = \frac{X_1^2 + \cdots + X_n^2}{n} \qquad (r \geq 0),$$

where the random variable Z_n is independent of r and has the same distribution as $\sqrt{n}\theta_1$ under μ_{n-1}. Since Z_n is close to being standard normal, F itself is approximately normal, if and only if r^2 is nearly a constant, which translates into a weak law of large numbers for the sequence X_k^2. This property – that the distribution of r^2 is concentrated around a point – may be quantified by the variance-type functionals

$$\sigma_{2p} = \sqrt{n} \left(\mathbb{E}\,|r^2 - 1|^p\right)^{1/p}.$$

In reasonable situations, these functionals are expected to be bounded by quantities that are independent of n (at least, they are finite, as long as $M_{2p} < \infty$). For example, if $|X|^2 = n$ a.s., we have $\sigma_{2p} = 0$. If the components X_k are pairwise independent, identically distributed, and $\mathbb{E}X_1^2 = 1$, then

$$\sigma_4^2 = \frac{1}{n}\operatorname{Var}(|X|^2) = \operatorname{Var}(X_1^2).$$

It turns out that control of the two functionals, M_3 and σ_3 is sufficient to guarantee a Berry-Esseen type rate of normal approximation for F_θ on average, in analogy with the Berry-Esseen theorem for independent identically distributed random variables. Since the second moment for the typical distribution F is equal to $\mathbb{E}r^2$, a normalization condition for this moment is desirable.

Theorem 1.1. *If $\mathbb{E}\,|X|^2 = n$, then with some absolute constant c*

$$\mathbb{E}_\theta\, \rho(F_\theta, \Phi) \leq c\,(M_3^3 + \sigma_3^{3/2})\, \frac{1}{\sqrt{n}}. \tag{1.2}$$

In the case of non-correlated random variables X_k with mean zero and variance one, all S_θ have also mean zero and variance one, so that $M_2 = 1$. In many interesting examples, M_3 is known to be of the same order as M_2 (in particular, when Khinchine-type inequalities are available for linear functionals of X_k). In some other examples, the magnitude of M_3 is however much larger, and here control via M_2, as in the next assertion, is preferable.

Theorem 1.2. *If $\mathbb{E}\,|X|^2 = n$, then for some absolute constant c*

$$\mathbb{E}_\theta\, \rho(F_\theta, \Phi) \leq c\,(M_2^2 + \sigma_2)\, \frac{\log n}{\sqrt{n}}. \tag{1.3}$$

Thus, modulo an additional logarithmic factor, a Berry-Esseen type rate holds for this average under a second moment assumption, only.

For an illustration, consider the trigonometric system $X = (X_1, \ldots, X_n)$ with components

$$
\begin{aligned}
X_{2k-1}(\omega) &= \sqrt{2}\,\cos(k\omega), \\
X_{2k}(\omega) &= \sqrt{2}\,\sin(k\omega), \qquad -\pi < \omega < \pi, \ \ k = 1, \ldots, n/2,
\end{aligned}
$$

assuming that n is even. They may be treated as random variables on the probability space $\Omega = (-\pi, \pi)$ equipped with the normalized Lebesgue measure $\mathbb{P}$, such that the linear forms

$$
S_\theta = \sqrt{2} \sum_{k=1}^{n/2} \big(\theta_{2k-1} \cos(k\omega) + \theta_{2k} \sin(k\omega)\big)
$$

represent trigonometric polynomials of degree at most $\frac{n}{2}$. The normalization $\sqrt{2}$ is chosen for convenience only, since then X is isotropic, so that $M_2 = 1$. Since also $\sigma_2 = 0$, by Theorem 1.2, most of the distributions F_θ of S_θ are approximately standard normal, and we have an upper bound

$$
\mathbb{E}_\theta\, \rho(F_\theta, \Phi) \le c\, \frac{\log n}{\sqrt{n}}. \tag{1.4}
$$

The study of asymptotic normality for trigonometric polynomials has a long history, starting with results on lacunary systems due to Kac [13], Salem and Zygmund [21]-[22], Gaposhkin [12]; see also [9]-[10], [1], [11], [2]. As we see, normality with an almost Berry-Esseen type rate remains valid for most choices of coefficients even without an assumption of lacunarity. One can show that the inequality (1.4) still holds for many other functional orthogonal systems as well, including, for instance, Chebyshev's polynomials on the interval $\Omega = (-1, 1)$, the Walsh system on the Boolean cube $\{-1, 1\}^n$. It holds as well for any system of functions of the form $X_k(\omega_1, \omega_2) = f(k\omega_1 + \omega_2)$, $\omega_1, \omega_2 \in (0, 1)$, where f is 1-periodic and belongs to $L^4(0, 1)$ (this is a strictly stationary sequence of pairwise independent random variables). A common feature of all listed examples is that (1.4) may actually be reversed modulo a logarithmic factor, in the sense that

$$
\mathbb{E}_\theta\, \rho(F_\theta, \Phi) \ge c\, \frac{1}{\sqrt{n}\, (\log n)^s}
$$

with some $s > 0$. (However, we do not derive lower bounds here referring the interested reader to [8]).

The conditions of Theorem 1.2 may be further relaxed in order to eliminate dependence on σ_2^2. This can be achieved by replacing it by the requirement that the probabilities $\mathbb{P}\{|X - Y|^2 \le n/4\}$ are small enough, where Y is an independent copy of X, cf. Theorem 6.3 below. This extends the applicability of our results to further groups of examples, while replacing Φ by a certain mixture of centered Gaussian measures. More precisely, define G to be the law of rZ, where $Z \sim N(0, 1)$ is independent of $r = \frac{1}{n}\, |X|$. In particular, we have:

Theorem 1.3. *If the components X_k of the random vector X in $\mathbb{R}^n$ are independent, identically distributed, have mean zero and finite second moment, then*

$$
\mathbb{E}_\theta\, \rho(F_\theta, G) \le c\, \sqrt{\frac{\log n}{n}},
$$

where the constant c depends on the distribution of X_1 only.

At first sight it seems surprising that an approximate Berry-Esseen type rate holds under no additional assumption beyond the finiteness of the second moment. Indeed, in the classical situation of equal coefficients, and when $\mathbb{E}X_1 = 0$, $\mathbb{E}X_1^2 = 1$, the distributions F_n of the normalized sums $S_n = (X_1 + \cdots + X_n)/\sqrt{n}$ may approach the standard normal law at an arbitrary slow rate: For any sequence $\varepsilon_n \to 0^+$, one may choose the distribution of X_1 such that

$$\rho(F_n, \Phi) \geq \varepsilon_n$$

for all n large enough (cf. [17]). This shows that for typical coefficients, the distributions F_θ behave in a more stable way in comparison to F_n. This interesting phenomenon has been studied before. For example, Klartag and Sodin [15] have shown in the i.i.d. case and under the 4-th moment assumption, that

$$\mathbb{E}_\theta\, \rho(F_\theta, \Phi) \leq c\, \frac{\beta_4}{n}, \qquad \beta_4 = \mathbb{E}X_1^4,$$

thus essentially improving the standard rate in the Berry-Esseen theorem. A similar observation was also made by Klartag under a certain dependency hypothesis. It was shown that, if the random vector X has an isotropic, coordinatewise symmetric distribution with a logarithmically concave density, then, for all $\theta \in S^{n-1}$,

$$\rho(F_\theta, \Phi) \leq c \sum_{k=1}^{n} \theta_k^4$$

with some absolute constant c ([14], Theorem 1). In particular, $\mathbb{E}_\theta\, \rho(F_\theta, \Phi) \leq \frac{c}{n}$.

The paper is organized as follows. We start with comments on general properties of the moment and variance-type functionals. Then we turn to the normal approximation for distributions of the first coordinate on the sphere (with rate of order $1/n$), which is used in Section 4 to describe proper bounds on the distance from the typical distributions to the standard normal law. Proofs of both Theorems 1.1 and 1.2 rely upon the spherical Poincaré inequality and Berry-Esseen-type estimates in terms of characteristic functions. The characteristic functions of the weighted sums are discussed separately in Section 5. Their properties are used in Section 6 to complete the proof of Theorem 1.2 (in a more general form). Theorem 1.1 is proved in Section 8, and in the last section we add some remarks concerning Theorem 1.3.

2 Moment and variance-type functionals

First let us describe some basic properties of the functionals $M_p = M_p(X)$ and $\sigma_{2p} = \sigma_{2p}(X)$. We shall as well introduce a few additional functionals. Define

$$m_p = m_p(X) = \frac{1}{\sqrt{n}} \left(\mathbb{E}\, |\langle X, Y \rangle|^p \right)^{1/p}, \qquad p \geq 1, \tag{2.1}$$

where Y is an independent copy of X. All these quantities do not depend on the systems of coordinates, that is,

$$m_p(UX) = m_p(X), \qquad M_p(UX) = M_p(X)$$

for any orthogonal linear map $U : \mathbb{R}^n \to \mathbb{R}^n$.

We call M_p the p-th moment of X. In case M_2 is finite, one may consider the covariance operator (matrix) of X which is defined by the equality

$$\mathbb{E}\, \langle X, a \rangle^2 = \langle Ra, a \rangle, \qquad a \in \mathbb{R}^n.$$

It is symmetric, positive definite, and has non-negative eigenvalues λ_k ($1 \leq k \leq n$). Choosing a system of coordinates such that R is diagonal with entries λ_k, we see that

$$M_2^2 = \max_k \lambda_k, \qquad m_2^2 = \frac{1}{n} \mathbb{E} \langle X, Y \rangle^2 = \frac{1}{n} \sum_{k=1}^{n} \lambda_k^2, \qquad \mathbb{E} |X|^2 = \sum_{k=1}^{n} \lambda_k. \qquad (2.2)$$

The random vector X is called isotropic (or having an isotropic distribution), if the covariance matrix of X is an identity, i.e.,

$$\mathbb{E} \langle X, a \rangle^2 = |a|^2 \quad \text{for all} \ \ a \in \mathbb{R}^n.$$

In this case, $m_2 = M_2 = 1$, and $\mathbb{E} |X|^2 = n$. Isotropic distributions are invariant under orthogonal transformations of the space.

Applying Cauchy's inequality, from (2.2) we immediately obtain:

Proposition 2.1. *For any random vector X in $\mathbb{R}^n$ with $\mathbb{E} |X|^2 = n$, we have $m_2 \geq 1$, where equality is attained, if and only if X is isotropic.*

The p-th moments of X may easily be related to the moments of $|X|$.

Proposition 2.2. *Given $p \geq 2$, for any random vector X in $\mathbb{R}^n$,*

$$(\mathbb{E} |X|^p)^{1/p} \leq M_p \sqrt{n}.$$

If X is isotropic, there is an opposite inequality $(\mathbb{E} |X|^p)^{1/p} \geq (\mathbb{E} |X|^2)^{1/2} = \sqrt{n}$.

Proof. By the rotational invariance of the uniform distribution on S^{n-1}, we have

$$\mathbb{E}_\theta | \langle \theta, a \rangle |^p = |a|^p \, \mathbb{E}_\theta |\theta_1|^p, \qquad a \in \mathbb{R}^n,$$

where $\mathbb{E}_\theta$ denotes the integral over the uniform measure μ_{n-1}. Inserting here $a = X$, we get

$$|X|^p \, \mathbb{E}_\theta |\theta_1|^p = \mathbb{E}_\theta | \langle X, \theta \rangle |^p.$$

Next, take the expectation with respect to X and use $\mathbb{E} | \langle X, \theta \rangle |^p \leq M_p$ to arrive at the upper bound

$$\mathbb{E} |X|^p \leq \frac{M_p^p(X)}{\mathbb{E}_\theta |\theta_1|^p}.$$

Here, since $\mathbb{E}_\theta \theta_1^2 = \frac{1}{n}$, we have

$$(\mathbb{E}_\theta |\theta_1|^p)^{1/p} \geq (\mathbb{E}_\theta |\theta_1|^2)^{1/2} = \frac{1}{\sqrt{n}}. \quad \square$$

Corollary 2.3. $m_p \leq M_p^2$ *for any $p \geq 2$.*

Indeed, let Y be an independent copy of the random vector X. By the very definition, for any particular value of Y, we have $\mathbb{E}_X | \langle X, Y \rangle |^p \leq M_p^p |Y|^p$. It remains to take the expectation with respect to Y.

In particular, $m_2 \leq M_2^2$, as can also be seen from (2.2). The identities in (2.2) also show that, in the general non-isotropic case, M_2^2 may be larger than m_2.

Let us now turn to the functionals

$$\sigma_{2p} = \sigma_{2p}(X) = \sqrt{n} \left(\mathbb{E} \left| \frac{|X|^2}{n} - 1 \right|^p \right)^{1/p}, \qquad p \geq 1,$$

where it is natural to assume that $\mathbb{E} |X|^2 = n$. Note that σ_{2p} represents a non-decreasing function of p, which attains its minimum at $p = 1$ with value

$$\sigma_2 = \sigma_2(X) = \frac{1}{\sqrt{n}} \mathbb{E} \big| |X|^2 - n \big|.$$

Another important value is $\sigma_4 = \frac{1}{n} \mathrm{Var}(|X|^2)$. They may be related to the variance of the Euclidean norm.

Proposition 2.4. *If* $\mathbb{E}\,|X|^2 = n$, *then* $\mathrm{Var}(|X|) \leq \sigma_4^2$. *In addition,*

$$\frac{1}{4}\,\sigma_2^2 \leq \mathrm{Var}(|X|) \leq \sigma_2\sqrt{n}.$$

Proof. Put $\xi = \frac{1}{\sqrt{n}}\,|X|$ and $a = \sqrt{\mathbb{E}\xi^2}$. Then, since $\xi \geq 0$,

$$
\begin{aligned}
\mathrm{Var}(\xi^2) &= \mathbb{E}\,(\xi^2 - a^2)^2 \\
&= \mathbb{E}\,(\xi - a)^2(\xi + a)^2 \;\geq\; \mathbb{E}\,(\xi - a)^2 \cdot a^2 \;\geq\; \mathrm{Var}(\xi) \cdot a^2.
\end{aligned}
$$

That is, $\mathbb{E}\xi^2\,\mathrm{Var}(\xi) \leq \mathrm{Var}(\xi^2)$, which is exactly the first required relation.

Now, in terms of ξ, one may write

$$\mathrm{Var}(|X|) = n\,\mathrm{Var}(\xi) = n\,(1 - (\mathbb{E}\xi)^2) = n\,(1 - \mathbb{E}\xi)\,(1 + \mathbb{E}\xi),$$

while $\sigma_2 = \sqrt{n}\,\mathbb{E}\,|1 - \xi^2|$. By Cauchy's inequality,

$$(\mathbb{E}\,|1 - \xi^2|)^2 \;\leq\; \mathbb{E}\,(1 - \xi)^2\,\mathbb{E}\,(1 + \xi)^2 \;=\; 4\,\mathbb{E}\,(1 - \xi)\,\mathbb{E}\,(1 + \xi),$$

implying that $\sigma_2^2 \leq 4\,\mathrm{Var}(|X|)$.

The last inequality of the proposition may be rewritten as $1 - (\mathbb{E}\xi)^2 \leq \mathbb{E}\,|1 - \xi^2|$. If $(\Omega, \mathbb{P})$ is the underlying probability space, define the probability measure

$$dQ = (1 + \xi)\,d\mathbb{P}/\mathbb{E}\,(1 + \xi)$$

and write $\mathbb{E}_Q$ for the expectation with respect to it. The required inequality then takes the form $\mathbb{E}_Q\,|1 - \xi| \geq \mathbb{E}_Q(1 - \xi)$, which is obvious. $\qquad\square$

The functionals σ_{2p}^2 and m_p are useful in the problem of estimation of "small" ball probabilities.

Proposition 2.5. *Let Y be an independent copy of a random vector X in $\mathbb{R}^n$ such that* $\mathbb{E}\,|X|^2 = n$. *For all $p, q \geq 1$,*

$$\mathbb{P}\Big\{|X - Y|^2 \leq \frac{1}{4}\,n\Big\} \;\leq\; \frac{4^q}{n^{q/2}}\,m_q^q + \frac{4^{2p}}{n^p}\,\sigma_{2p}^{2p}.$$

In particular,

$$\mathbb{P}\Big\{|X - Y|^2 \leq \frac{1}{4}\,n\Big\} \;\leq\; \frac{C}{n^p}$$

with $C = 4^{2p}\,(m_{2p}^{2p} + \sigma_{2p}^{2p})$.

Proof. According to the definition,

$$\sigma_{2p}^p = n^{-p/2}\,\mathbb{E}\,\big|\,|X|^2 - n\,\big|^p.$$

Hence, for any $\lambda \in (0, 1)$, by Chebyshev's inequality,

$$
\begin{aligned}
\mathbb{P}\big\{|X|^2 \leq \lambda n\big\} &= \mathbb{P}\Big\{\mathbb{E}\,|X|^2 - |X|^2 \geq (1 - \lambda)\,\mathbb{E}\,|X|^2\Big\} \\
&\leq \frac{\sigma_{2p}^p\, n^{p/2}}{(1 - \lambda)^p\,(\mathbb{E}\,|X|^2)^p} \;=\; \frac{\sigma_{2p}^p}{(1 - \lambda)^p\,n^{p/2}}.
\end{aligned}
$$

In particular, choosing $\lambda = 3/4$, we get

$$\mathbb{P}\Big\{|X|^2 + |Y|^2 \leq \frac{3}{4}\,n\Big\} \;\leq\; \mathbb{P}\Big\{|X|^2 \leq \frac{3}{4}\,n\Big\}\,\mathbb{P}\Big\{|Y|^2 \leq \frac{3}{4}\,n\Big\} \;\leq\; \frac{4^{2p}\,\sigma_{2p}^{2p}}{n^p}.$$

On the other hand, by Markov's inequality,

$$\mathbb{P}\left\{\left|\langle X, Y\rangle\right| \geq \frac{1}{4}\, n\right\} \leq \frac{4^q\, \mathbb{E}\left|\langle X, Y\rangle\right|^q}{n^q} = \frac{4^q\, m_q^q}{n^{q/2}}.$$

One may now write

$$|X - Y|^2 = |X|^2 + |Y|^2 - 2\langle X, Y\rangle$$

and split the event $|X - Y|^2 \leq \frac{1}{4}\,n$ into the case $\left|\langle X, Y\rangle\right| \geq \frac{1}{4}\,n$ and the case of the opposite inequality. In view of the set inclusion

$$\left\{|X - Y|^2 \leq \frac{1}{4}\,n\right\} \subset \left\{\left|\langle X, Y\rangle\right| \geq \frac{1}{4}\,n\right\} \cup \left\{|X|^2 + |Y|^2 \leq \frac{3}{4}\,n\right\},$$

the proposition follows. $\qquad\square$

3 Linear functionals on the sphere

The aim of this section is to quantify the asymptotic normality of distributions of linear functionals with respect to the normalized Lebesgue measure μ_{n-1} on the unit sphere $S^{n-1} \subset \mathbb{R}^n$ ($n \geq 2$). By the rotational invariance of this measure, all linear functionals $f(\theta) = \langle \theta, v\rangle$ with $|v| = 1$ have equal distributions, and it is sufficient to focus just on the first coordinate θ_1 of the vector $\theta \in S^{n-1}$. As a random variable on the probability space (S^{n-1}, μ_{n-1}), it has density

$$c_n \left(1 - x^2\right)_+^{\frac{n-3}{2}}, \qquad x \in \mathbb{R},$$

where $c_n = \frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi}\,\Gamma\left(\frac{n-1}{2}\right)}$ is a normalizing constant.

Let us denote by φ_n the density of the normalized first coordinate $Z_n = \sqrt{n}\,\theta_1$ under the measure μ_{n-1}, i.e.,

$$\varphi_n(x) = c_n' \left(1 - \frac{x^2}{n}\right)_+^{\frac{n-3}{2}}, \qquad c_n' = \frac{c_n}{\sqrt{n}} = \frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi n}\,\Gamma\left(\frac{n-1}{2}\right)}.$$

Clearly, as $n \to \infty$,

$$\varphi_n(x) \to \varphi(x) = \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2}, \qquad c_n' \to \frac{1}{\sqrt{2\pi}},$$

and one can show that $c_n' < \frac{1}{\sqrt{2\pi}}$ for all $n \geq 2$.

We are interested in non-uniform deviation bounds of $\varphi_n(x)$ from $\varphi(x)$.

Proposition 3.1. *If $n \geq 3$, then for all $x \in \mathbb{R}$, with some universal constant C*

$$\left|\varphi_n(x) - \varphi(x)\right| \leq \frac{C}{n}\, e^{-x^2/8}. \tag{3.1}$$

Proof. Since the random variable Z_3 has a uniform distribution on $[-\sqrt{3}, \sqrt{3}]$, inequality (3.1) obviously holds for $n = 3$. Hence, let $n \geq 4$.

First we consider the asymptotic behavior of the functions

$$p_n(x) = \left(1 - \frac{x^2}{n}\right)_+^{\frac{n-3}{2}}, \qquad x \in \mathbb{R}.$$

Clearly, $p_n(x) \to e^{-x^2/2}$ for all x. Moreover, for $|x| < \sqrt{n}$, we have

$$-\log p_n(x) = -\frac{n-3}{2} \log\left(1 - \frac{x^2}{n}\right) \geq \frac{n-3}{2}\,\frac{x^2}{n} \geq \frac{x^2}{8},$$

so that there is a uniform bound

$$p_n(x) \le e^{-x^2/8}, \qquad x \in \mathbb{R}. \tag{3.2}$$

To study the rate of convergence of $p_n(x)$, assume that $|x| \le \frac{1}{2}\sqrt{n}$. By Taylor's expansion, with some $0 \le \varepsilon \le 1$

$$
\begin{aligned}
-\log p_n(x) &= \frac{n-3}{2}\left[\frac{x^2}{n} + \left(\frac{x^2}{n}\right)^2 \sum_{k=2}^{\infty} \frac{1}{k}\left(\frac{x^2}{n}\right)^{k-2}\right] \\
&= \frac{n-3}{2}\left(\frac{x^2}{n} + \frac{x^4}{n^2}\varepsilon\right) = \frac{x^2}{2} + \frac{x^2}{2n}\left(-3 + \frac{n-3}{n}x^2\varepsilon\right),
\end{aligned}
$$

that is,

$$p_n(x) = e^{-x^2/2}\, e^{-\delta} \quad \text{with} \quad \delta = \frac{x^2}{2n}\left(-3 + \frac{n-3}{n}x^2\varepsilon\right).$$

Since $\delta \ge -\frac{3x^2}{2n} \ge -\frac{3}{8n} \ge -\frac{3}{32}$, we have

$$|e^{-\delta} - 1| \le |\delta|\, e^{3/32} \le 1.1\,|\delta|.$$

On the other hand,

$$\delta \le \frac{x^2}{2n}\left(-3 + \frac{n-3}{n}x^2\right) \le \frac{x^4}{2n},$$

which together with the lower bound on δ yields

$$1.1\,|\delta| \le 1.1\left(\frac{3x^2}{2n} + \frac{x^4}{2n}\right) \le \frac{1}{n}(3x^2 + x^4).$$

Thus,

$$\left|p_n(x) - e^{-x^2/2}\right| \le \frac{1}{n}(3x^2 + x^4)\, e^{-x^2/2}, \qquad |x| \le \frac{1}{2}\sqrt{n}.$$

Combining this inequality with (3.2), we also get a non-uniform bound on the whole real line, namely

$$\left|p_n(x) - e^{-x^2/2}\right| \le \frac{C}{n}\, e^{-x^2/8}, \qquad x \in \mathbb{R},$$

where C is an absolute constant. Let us integrate this inequality over x. Since

$$\int_{-\infty}^{\infty} p_n(x)\, dx = \frac{1}{c_n'}, \qquad \int_{-\infty}^{\infty} e^{-x^2/2}\, dx = \sqrt{2\pi},$$

we get that $\left|\frac{1}{c_n'} - \sqrt{2\pi}\right| \le \frac{C}{n}$ with some absolute constant C. Hence, we arrive at the conclusion (3.1) for the densities φ_n for $n \ge 4$ as well. $\qquad\square$

In the sequel we denote by J_n the characteristic function of the first coordinate θ_1 of a random vector θ which is uniformly distributed on the unit sphere S^{n-1}. In a more explicit form, for any $t \in \mathbb{R}$,

$$
\begin{aligned}
J_n(t) &= c_n \int_{-\infty}^{\infty} e^{itx}\,(1-x^2)_+^{\frac{n-3}{2}}\, dx \\
&= c_n' \int_{-\infty}^{\infty} e^{itx/\sqrt{n}}\left(1 - \frac{x^2}{n}\right)_+^{\frac{n-3}{2}}\, dx.
\end{aligned}
$$

Note that the equality

$$\widetilde{J}_\nu(t) = \frac{1}{\sqrt{\pi}\,\Gamma(\nu + \frac{1}{2})}\left(\frac{t}{2}\right)^\nu \int_{-1}^{1} e^{itx}\,(1-x^2)^{\nu-\frac{1}{2}}\, dx$$

defines the classical Bessel function of the first kind with index ν ([3], p. 81). Therefore,

$$J_n(t) = \frac{1}{c_n} \sqrt{\pi}\, \Gamma\left(\nu + \frac{1}{2}\right) \left(\frac{t}{2}\right)^{-\nu} \widetilde{J}_\nu(t), \qquad \nu = \frac{n}{2} - 1.$$

However, this relationship will not be used in the sequel.

Thus, the characteristic function of $Z_n = \theta_1 \sqrt{n}$ is given by

$$\hat{\varphi}_n(t) = J_n\big(t\sqrt{n}\big) = \int_{-\infty}^{\infty} e^{itx} \varphi_n(x)\, dx,$$

which is the Fourier transform of the probability density φ_n. One immediate consequence from Proposition 3.1 is the following:

Corollary 3.2. *For all $t \in \mathbb{R}$, we have*

$$\left| J_n\big(t\sqrt{n}\big) - e^{-t^2/2} \right| \leq \frac{C}{n},$$

where C is an absolute constant.

For large t, this bound may be improved by virtue of the following upper bound.

Proposition 3.3. *For all $t \in \mathbb{R}$,*

$$\left| J_n(t\sqrt{n}) \right| \,\leq\, 4.1\, e^{-t^2/2} + 4\, e^{-n/12}. \tag{3.3}$$

Proof. One may assume that $n \geq 4$ (since $4\, e^{-n/12} > 1$ for $n = 2$ and $n = 3$, while $|J_n| \leq 1$). For the approximation we shall use an approach based on contour integration in complex analyis.

The function $z \to (1 - z^2)^{\frac{n-3}{2}}$ is analytic in the whole complex plane when n is odd and in the strip $z = x + iy$, $|x| < 1$, when n is even. Therefore, integrating along the boundary of the rectangle $\mathfrak{C} = [-1, 1] \times [0, y]$ with $y > 0$ (slightly modifying the contour in a standard way near the points -1 and 1), we have

$$\int_{\mathfrak{C}} e^{itz} \left(1 - z^2\right)^{\frac{n-3}{2}} dz = 0.$$

Then we obtain a natural decomposition $J_n(t\sqrt{n}) = c_n \big(I_1(t) + I_2(t) + I_3(t)\big)$ for $t > 0$, where

$$I_1(t) \;=\; e^{-ty\sqrt{n}} \left(1 + y^2\right)^{\frac{n-3}{2}} \int_{-1}^{1} e^{itx\sqrt{n}} \left(\frac{1 - (x + iy)^2}{1 + y^2}\right)^{\frac{n-3}{2}} dx,$$

$$I_2(t) \;=\; -e^{it\sqrt{n}} \int_0^y e^{-ts\sqrt{n}} \left(1 - (1 + is)^2\right)^{\frac{n-3}{2}} ds,$$

$$I_3(t) \;=\; e^{-it\sqrt{n}} \int_0^y e^{-ts\sqrt{n}} \left(1 - (1 - is)^2\right)^{\frac{n-3}{2}} ds.$$

For $0 \leq s \leq y \leq \alpha$,

$$\left| 1 - (1 + is)^2 \right| \;=\; s\sqrt{s^2 + 4} \;\leq\; \alpha\, \sqrt{\alpha^2 + 4} \equiv \beta.$$

Choosing $\alpha = \frac{1}{\sqrt{6}}$, we have $\beta = \frac{5}{6}$. Hence, for all $t > 0$,

$$|I_2(t)| \;\leq\; \beta^{\frac{n-3}{2}} \int_0^y e^{-ts\sqrt{n}}\, ds \;\leq\; \frac{1}{t\sqrt{n}}\, \beta^{\frac{n-3}{2}}. \tag{3.4}$$

The same estimates hold for $I_3(t)$.

In order to estimate $I_1(t)$, we use an elementary identity

$$\left|1 - (x + iy)^2\right|^2 = (1 - 2x^2)(1 + y^2)^2 + x^2\,(x^2 + 6y^2 + 2y^4) \qquad (x, y \in \mathbb{R}),$$

which for the region $|x| \leq 1$ yields

$$\left|\frac{1 - (x + iy)^2}{1 + y^2}\right|^2 \leq 1 - 2x^2 + x^2\,v(y^2), \qquad v(z) = \frac{1 + 6z + 2z^2}{(1 + z)^2}.$$

Since $v'(z) = \frac{4 - 2z}{(1 + z)^3} > 0$, this function increases in $0 \leq z \leq 2$, and since $z = y^2 \leq \frac{1}{6}$, we have $v(y^2) \leq v(1/6) = \frac{74}{49}$. Hence

$$\left|\frac{1 - (x + iy)^2}{1 + y^2}\right|^2 \leq 1 - \frac{24}{49}\,x^2 \leq e^{-\frac{24}{49}\,x^2}.$$

Using this estimate together with $n - 3 \geq \frac{1}{4}\,n$, we have

$$
\begin{aligned}
\int_{-1}^{1} \left(\frac{|1 - (x + iy)^2|}{1 + y^2}\right)^{\frac{n-3}{2}} dx \quad &\leq \quad \int_{1}^{1} e^{-\frac{12}{49}\,\frac{n-3}{2}\,x^2} dx \\
&= \quad \sqrt{\frac{49}{12}}\,\sqrt{\frac{2\pi}{n - 3}} \leq \frac{14}{\sqrt{12}}\,\sqrt{\frac{2\pi}{n}} \leq 4.1\,\sqrt{\frac{2\pi}{n}}.
\end{aligned}
$$

This upper bound allows us to conclude that

$$
\begin{aligned}
|I_1(t)| \quad &\leq \quad 4.1\,\sqrt{\frac{2\pi}{n}}\,e^{-ty\sqrt{n}}\left(1 + y^2\right)^{\frac{n-3}{2}} \\
&\leq \quad 4.1\,\sqrt{\frac{2\pi}{n}}\,\exp\left\{-ty\sqrt{n} + \frac{n - 3}{2}\,y^2\right\}.
\end{aligned}
$$

Choosing here $y = \frac{t}{\sqrt{n}}$, the expression in the exponent will be smaller than $t^2/2$, hence

$$|I_1(t)| \leq 4.1\,e^{-t^2/2}, \qquad 0 \leq t \leq \sqrt{n/6}. \tag{3.5}$$

In the case $t > \sqrt{n/6}$, we choose $y = \frac{1}{\sqrt{6}}$ and then $-ty\sqrt{n} + \frac{n-3}{2}\,y^2 < -\frac{n}{12} - \frac{1}{4}$, so that

$$|I_1(t)| \leq 4.1\,\sqrt{\frac{2\pi}{n}}\,e^{-\frac{n}{12} - \frac{1}{4}}, \qquad t \geq \sqrt{n/6}. \tag{3.6}$$

Let us collect these estimates. For $2 \leq t \leq \sqrt{n/6}$, we combine (3.5) with (3.4) and a similar bound for $I_3(t)$, and use $c_n < \frac{1}{\sqrt{2\pi}}\,\sqrt{n}$ with $\beta^{\frac{n-3}{2}} < 0.77\,e^{-n/12}$. This leads to

$$
\begin{aligned}
c_n\left(|I_1(t)| + |I_2(t)| + |I_3(t)|\right) \quad &\leq \quad \frac{c_n}{\sqrt{n}}\,\beta^{\frac{n-3}{2}} + 4.1\,c_n\,\sqrt{\frac{2\pi}{n}}\,e^{-t^2/2} \\
&\leq \quad 0.31\,e^{-n/12} + 4.1\,c_n\,e^{-t^2/2}.
\end{aligned}
$$

Similarly, in case $t > \sqrt{n/6}$, we use (3.6) leading to

$$c_n\left(|I_1(t)| + |I_2(t)| + |I_3(t)|\right) \leq 0.31\,e^{-n/12} + 4.1\,e^{-1/4}\,e^{-n/12} < 4e^{-n/12}.$$

Finally, if $0 \leq t \leq 2$, then $|J_n(t\sqrt{n})| \leq 1 \leq 4\,e^{-t^2/2}$. $\qquad \square$

4 Typical distributions and mixtures of Gaussian measures

The asymptotic normality of the typical distributions F in Sudakov's theorem, defined in (1.1), may be described in the next assertion proved in [7].

Proposition 4.1. *Given a random vector X in $\mathbb{R}^n$, suppose that $\mathbb{E}\,|X|^2 = n$. With some absolute constant $c > 0$ we have*

$$\int_{-\infty}^{\infty} (1 + x^2)\,|F(dx) - \Phi(dx)| \;\leq\; c\left(\frac{1}{n} + \mathrm{Var}(r)\right), \tag{4.1}$$

where $r = \frac{1}{\sqrt{n}}\,|X|$.

Here the positive measure $|F - \Phi|$ denotes the variation in the sense of measure theory, and the left integral represents the weighted total variation of $F - \Phi$. In particular, we have a similar bound for the usual total variation distance between F and Φ, as well as for the Kolmogorov distance $\rho(F, \Phi)$. Applying Proposition 2.4, the latter may be related to the variance-type functionals σ_{2p} (cf. also [18]).

Corollary 4.2. *In particular (under the same conditions),*

$$\rho(F, \Phi) \;\leq\; c\,\frac{1 + \sigma_4^2}{n}, \qquad \rho(F, \Phi) \;\leq\; c\,\frac{1 + \sigma_2}{\sqrt{n}}.$$

The proof of Proposition 4.1 is based on the following observation about general mixtures of centered Gaussian measure on the real line. Given a random variable $r \geq 0$, let us denote by Φ_r the distribution function of the random variable rZ, where $Z \sim N(0,1)$ is independent of r. That is,

$$\Phi_r(x) = \mathbb{P}\{rZ \leq x\} = \mathbb{E}\,\Phi(x/r), \qquad x \in \mathbb{R}.$$

As shown in [7], if $\mathbb{E}\,r^2 = 1$, then with some absolute constant c we have

$$\int_{-\infty}^{\infty} (1 + x^2)\,|\Phi_r - \Phi|(dx) \;\leq\; c\,\mathrm{Var}(r). \tag{4.2}$$

To explain the transition from (4.2) to (4.1), assume that $n \geq 3$. Let Φ_n and φ_n denote respectively the distribution function and the density of $Z_n = \theta_1 \sqrt{n}$, where θ_1 is the first coordinate of a random point θ uniformly distributed in S^{n-1}. If $r^2 = \frac{1}{n}\,|X|^2$ is independent of Z_n ($r \geq 0$), then, by the definition of the typical distribution,

$$F(x) = \mathbb{P}\{rZ_n \leq x\} = \mathbb{E}\,\Phi_n(x/r), \qquad x \in \mathbb{R},$$

so that

$$\int_{-\infty}^{\infty} (1 + x^2)\,|F(dx) - \Phi_r(dx)| = \int_{-\infty}^{\infty} (1 + x^2)\,|\mathbb{E}\,\Phi_n(dx/r) - \mathbb{E}\,\Phi(dx/r)|. \tag{4.3}$$

But, for any fixed value of r,

$$\int_{-\infty}^{\infty} (1 + x^2)\,|\Phi_n(dx/r) - \Phi(dx/r)| = \int_{-\infty}^{\infty} (1 + r^2 x^2)\,|\Phi_n(dx) - \Phi(dx)|,$$

hence, by (4.3), taking the expectation with respect to r and using Jensen's inequality, we get

$$\begin{aligned}
\int_{-\infty}^{\infty} (1 + x^2)\,|F(dx) - \Phi_r(dx)| \;&\leq\; \mathbb{E}\int_{-\infty}^{\infty} (1 + x^2)\,|\Phi_n(dx/r) - \Phi(dx/r)| \\
&=\; \mathbb{E}\int_{-\infty}^{\infty} (1 + r^2 x^2)\,|\Phi_n(dx) - \Phi(dx)| \\
&=\; \int_{-\infty}^{\infty} (1 + x^2)\,|\Phi_n(dx) - \Phi(dx)|.
\end{aligned}$$

It remains to apply (3.1), which yields

$$\int_{-\infty}^{\infty} (1+x^2) \, |\Phi_n(dx) - \Phi(dx)| \, = \, \int_{-\infty}^{\infty} (1+x^2) \, |\varphi_n(x) - \varphi(x)| \, dx \, \leq \, \frac{C}{n}$$

with some universal constant C. □

5 Characteristic functions of weighted sums

As before, let $X = (X_1, \ldots, X_n)$ denote a random vector in $\mathbb{R}^n$, $n \geq 2$. The concentration problems for distributions of weighted sums $S_\theta = \langle X, \theta \rangle$ may be studied by means of their characteristic functions

$$f_\theta(t) = \mathbb{E} \, e^{it\langle X, \theta \rangle}, \qquad t \in \mathbb{R}. \tag{5.1}$$

In particular, we intend to quantify the concentration of f_θ around the characteristic function f of the typical distribution F on average over the directions θ in terms of correlation-type functionals. Note that the characteristic function of F is given by

$$f(t) = \mathbb{E}_\theta f_\theta(t) = \mathbb{E}_\theta \, \mathbb{E} \, e^{it\langle X, \theta \rangle} = \mathbb{E} \, J_n(t|X|), \qquad t \in \mathbb{R},$$

where J_n is the characteristic function of the first coordinate θ_1 under the uniform measure μ_{n-1} on the unit sphere S^{n-1}.

First let us describe the decay of $t \to |f_\theta(t)|$ at infinity on average with respect to θ. Starting from (5.1), write

$$\mathbb{E}_\theta \, |f_\theta(t)|^2 = \mathbb{E}_\theta \, e^{it\langle X - Y, \theta \rangle} = \mathbb{E} \, J_n(t|X - Y|),$$

where Y is an independent copy of X. To proceed, let us rewrite the Gaussian-type bound (3.3) of Proposition 3.3 as

$$|J_n(t)| \, \leq \, 4.1 \, e^{-t^2/2n} + 4 \, e^{-n/12} \tag{5.2}$$

which gives

$$\mathbb{E}_\theta \, |f_\theta(t)|^2 \, \leq \, 4.1 \, \mathbb{E} \, e^{-t^2|X-Y|^2/2n} + 4 \, e^{-n/12}.$$

Splitting the latter expectation into the event $A = \{|X - Y|^2 \leq \lambda n\}$ and its complement, we get the following general bound.

Lemma 5.1. *The characteristic functions f_θ satisfy, for all $t \in \mathbb{R}$ and $\lambda > 0$,*

$$\frac{1}{2.1} \, \mathbb{E}_\theta \, |f_\theta(t)| \leq e^{-\lambda t^2/4} + e^{-n/24} + \sqrt{\mathbb{P}\{|X - Y|^2 \leq \lambda n\}},$$

where Y is an independent copy of X.

In case $\mathbb{E} \, |X|^2 = n$, the right-hand side of these bounds can be further quantified by using the moment and variance-type functionals, which we have discussed before, namely

$$m_p = \frac{1}{\sqrt{n}} \left(\mathbb{E} \, |\langle X, Y \rangle|^p \right)^{1/p}, \qquad \sigma_{2p} = \sqrt{n} \left(\mathbb{E} \left| \frac{|X|^2}{n} - 1 \right|^p \right)^{1/p}.$$

Note that both m_p and σ_p are non-decreasing functions in $p \geq 1$. In order to estimate the probability of the event A, we shall use Proposition 2.5, which gives

$$\mathbb{P}(A) \leq \frac{C}{n^p}$$

with a constant $C = 4^{2p} \, (m_{2p}^{2p} + \sigma_{2p}^{2p})$. Hence, from Lemma 5.1 and using $m_{2p} \geq m_2 \geq 1$ (cf. Proposition 2.1), we deduce:

Lemma 5.2. *Suppose that* $\mathbb{E}\,|X|^2 = n$. *If the moment* m_{2p} *is finite for* $p \geq 1$, *then with some constant* $c_p > 0$ *depending on* p,

$$c_p\,\mathbb{E}_\theta\,|f_\theta(t)| \;\leq\; \frac{m_{2p}^p + \sigma_{2p}^p}{n^{p/2}} + e^{-t^2/16}.$$

By the triangle inequality, $|f(t)| \leq \mathbb{E}_\theta\,|f_\theta(t)|$. Hence, the characteristic function of the typical distribution shares the same bounds. In fact, here the parameter m_{2p} is not needed. Indeed, as was shown in the proof of Proposition 2.5 with $\lambda = \frac{1}{2}$, we have

$$\mathbb{P}\Big\{|X|^2 \leq \frac{1}{2}\,n\Big\} \;\leq\; 2^p\,\frac{\sigma_{2p}^p}{n^{p/2}}.$$

Hence, by (5.2),

$$
\begin{aligned}
|f(t)| \;&\leq\; \mathbb{E}\,|J_n(t|X|)|\,1_{\{|X| \leq \sqrt{n/2}\}} + \mathbb{E}\,|J_n(t|X|)|\,1_{\{|X| > \sqrt{n/2}\}} \\
&\leq\; 2^p\,\frac{\sigma_{2p}^p}{n^{p/2}} + C\left(e^{-t^2/4} + e^{-n/12}\right).
\end{aligned}
$$

Thus, we get:

Lemma 5.3. *Suppose that* $\mathbb{E}\,|X|^2 = n$. *Then with some constant* $c_p > 0$ *depending on* $p \geq 1$, *for all* $t \in \mathbb{R}$,

$$c_p\,|f(t)| \leq \frac{1 + \sigma_{2p}^p}{n^{p/2}} + e^{-t^2/4},$$

and therefore, for all $T > 0$,

$$\frac{c_p}{T}\int_0^T |f(t)|\,dt \;\leq\; \frac{1 + \sigma_{2p}^p}{n^{p/2}} + \frac{1}{T}.$$

We first study the concentration properties of $f_\theta(t)$ as functions of θ on the sphere with fixed $t \in \mathbb{R}$ (rather than directly for the distributions F_θ). This can be done in terms of the moment functionals

$$M_p = M_p(X) \;=\; \sup_{\theta \in S^{n-1}} \Big(\mathbb{E}\,|\langle X, \theta\rangle|^p\Big)^{1/p}.$$

Our basic tool is a well-known spherical Poincaré inequality

$$\int_{S^{n-1}} |u(\theta) - a|^2\,d\mu_{n-1}(\theta) \leq \frac{1}{n-1}\int_{S^{n-1}} |\nabla u(\theta)|^2\,d\mu_{n-1}(\theta). \tag{5.3}$$

It holds true for any complex-valued function u which is defined and smooth in a neighborhood of the sphere, and has gradient ∇u and the mean $a = \int u\,d\mu_{n-1}$ (cf. [16]).

According to (5.1), the function $\theta \to f_\theta(t)$ is smooth on the whole space $\mathbb{R}^n$ and has partial derivatives

$$\frac{\partial_j f_\theta(t)}{\partial \theta_j} = it\,\mathbb{E}X_j\,e^{it\langle X,\theta\rangle}$$

or in the vector form

$$\langle \nabla f_\theta(t), v\rangle = it\,\mathbb{E}\,\langle X, v\rangle\,e^{it\langle X,\theta\rangle}, \qquad v \in \mathbb{R}^n.$$

Hence

$$|\langle \nabla f_\theta(t), v\rangle| \leq |t|\,\mathbb{E}\,|\langle X, v\rangle|.$$

Taking the sup over all $v \in S^{n-1}$, we obtain a uniform bound on the modulus of the gradient, namely $|\nabla f_\theta(t)| \leq M_1|t|$.

A similar bound holds as well in average. To this aim, let us square the vector representation and write

$$\langle \nabla f_\theta(t), v \rangle^2 = t^2 \, \mathbb{E} \, \langle X, v \rangle \, \langle Y, v \rangle \, e^{it\langle X-Y, \theta \rangle},$$

where Y is an independent copy of X. Integrating over v with respect to μ_{n-1}, we get the representation

$$|\nabla f_\theta(t)|^2 \;=\; t^2 \, \mathbb{E} \, \langle X, Y \rangle \, e^{it\langle X-Y, \theta \rangle},$$

so that

$$\mathbb{E}_\theta \, |\nabla f_\theta(t)|^2 \;=\; t^2 \, \mathbb{E} \, \langle X, Y \rangle \, J_n(t(X-Y)).$$

(where $\mathbb{E}_\theta$ refers to integration over μ_{n-1}). Applying (5.3), one can summarize.

Lemma 5.4. *Given a random vector X in $\mathbb{R}^n$ with finite moment M_1, for all $t \in \mathbb{R}$,*

$$\mathbb{E}_\theta \, |f_\theta(t) - f(t)|^2 \leq \frac{t^2}{n-1} \, M_1^2.$$

In addition,

$$\mathbb{E}_\theta \, |f_\theta(t) - f(t)|^2 \leq \frac{t^2}{n-1} \, \mathbb{E} \, \langle X, Y \rangle \, J_n(t(X-Y)),$$

where Y is an independent copy of X.

6 Berry-Esseen bounds. Theorem 1.2 and its generalization

Fourier Analysis provides a well-established tool to prove Berry-Esseen-type bounds for the Kolmogorov distance

$$\rho(F_\theta, F) = \sup_x |F_\theta(x) - F(x)|.$$

To study the average behavior of this distance with respect to θ using the uniform measure μ_{n-1} on the unit sphere, as a preliminary step, let us first introduce two auxiliary bounds.

Lemma 6.1. *Let X be a random vector in $\mathbb{R}^n$. With some absolute constant $c > 0$, for all $T \geq T_0 > 0$,*

$$
\begin{aligned}
c \, \mathbb{E}_\theta \, \rho(F_\theta, F) \quad \leq \quad & \int_0^{T_0} \frac{\mathbb{E}_\theta \, |f_\theta(t) - f(t)|}{t} \, dt \\
& + \int_{T_0}^{T} \frac{\mathbb{E}_\theta \, |f_\theta(t)|}{t} \, dt + \frac{1}{T} \int_0^T |f(t)| \, dt.
\end{aligned}
\tag{6.1}
$$

As before, here F_θ denote distribution functions of the weighted sums $S_\theta = \langle X, \theta \rangle$ with their characteristic functions

$$f_\theta(t) = \mathbb{E} \, e^{it\langle X, \theta \rangle} = \int_{-\infty}^{\infty} e^{itx} \, dF_\theta(x), \qquad t \in \mathbb{R}, \; \theta \in S^{n-1},$$

and $F(x) = \mathbb{E}_\theta F_\theta(x)$ is the typical distribution function with characteristic function

$$f(t) = \mathbb{E}_\theta f_\theta(t) = \int_{-\infty}^{\infty} e^{itx} \, dF(x).$$

For an estimation of the Kolmogorov distance, the following general Berry-Esseen bound will be convenient:

$$c \, \rho(U, V) \leq \int_0^T \frac{|u(t) - v(t)|}{t} \, dt + \frac{1}{T} \int_0^T |v(t)| \, dt \qquad (T > 0).
\tag{6.2}$$

Here U and V may be arbitrary distribution functions on the line with characteristic functions u and v, respectively, and $c > 0$ is an absolute constant (cf. e.g. [6], [19], [20]).

In our situation, we take $U = F_\theta$ and $V = F$. In order to estimate the first integral in (6.2), we shall split the integration into the two intervals, $[0, T_0]$ (the interval of moderate values of t), where it is easier to control the closeness of the two characteristic functions, and the long interval $[T_0, T]$, where both characteristic functions can be shown to be sufficiently small. Note that, by the triangle inequality, we have $|f(t)| \leq \mathbb{E}_\theta |f_\theta(t)|$, which implies $\mathbb{E}_\theta |f_\theta(t) - f(t)| \leq 2\,\mathbb{E}_\theta |f_\theta(t)|$. Using this on the long interval, we arrive at the more specific variant of (6.2), namely (6.1).

The estimation of the integrals in (6.1) will be done in terms of the functionals $m_p = m_p(X)$, $M_p = M_p(X)$ and $\sigma_{2p} = \sigma_{2p}(X)$.

Lemma 6.2. *Suppose that X has a finite moment of order $2p$ $(p \geq 1)$, and $\mathbb{E}|X|^2 = n$. Then with some constant c_p depending on p only, for all $T \geq T_0 > 0$,*

$$
\begin{aligned}
c_p \, \mathbb{E}_\theta\, \rho(F_\theta, F) \quad \leq \quad & \int_0^{T_0} \mathbb{E}_\theta\, |f_\theta(t) - f(t)|\, \frac{dt}{t} \\
& + \frac{m_{2p}^p + \sigma_{2p}^p}{n^{p/2}} \left(1 + \log \frac{T}{T_0} \right) + \frac{1}{T} + e^{-T_0^2/16}.
\end{aligned}
$$

Proof. By the second inequality of Lemma 5.3 (on this step we use the assumption $\mathbb{E}|X|^2 = n$), we have

$$
\frac{c_p}{T} \int_0^T |f(t)|\, dt \; \leq \; \frac{1 + \sigma_{2p}^p}{n^{p/2}} + \frac{1}{T},
$$

while Lemma 5.2 yields the bound

$$
c_p \int_{T_0}^T \frac{\mathbb{E}_\theta\, |f_\theta(t)|}{t}\, dt \; \leq \; \frac{m_{2p}^p + \sigma_{2p}^p}{n^{p/2}} \log \frac{T}{T_0} + e^{-T_0^2/16}.
$$

This allows us to estimate the second last and last integrals in (6.1). $\qquad\square$

We are now prepared to establish Theorem 1.2, in fact – in somewhat more general form which requires the first moment, only. Recall that

$$
\sigma_2 = \frac{1}{\sqrt{n}}\, \mathbb{E}\, \big|\, |X|^2 - n\, \big|.
$$

Theorem 6.3. *If the random vector X in $\mathbb{R}^n$ has finite first moment M_1, then*

$$
c\, \mathbb{E}_\theta\, \rho(F_\theta, F) \; \leq \; M_1 \sqrt{\frac{\log n}{n}} + \sqrt{\mathbb{P}\{|X - Y|^2 \leq n/4\}}\, \log n + \frac{1}{n}, \qquad (6.3)
$$

where $c > 0$ is an absolute constant, and Y is an independent copy of X. As a consequence, if X has finite 2-nd moment M_2 and $\mathbb{E}|X|^2 = n$, then

$$
c\, \mathbb{E}_\theta\, \rho(F_\theta, F) \; \leq \; (M_1 + m_2 + \sigma_2) \frac{\log n}{\sqrt{n}}. \qquad (6.4)
$$

A similar bound also holds for the normal distribution function Φ in place of F.

The coefficient in (6.4) may be simplified by using $m_2 \leq M_2^2$ and $M_1 \leq M_2$. Since necessarily $M_2 \geq 1$, (6.4) implies the inequality (1.3) of Theorem 1.2.

Proof. We apply Lemma 6.1 with $T_0 = 5\sqrt{\log n}$ and $T = 5n$. The first integral in (6.1) can be bounded by virtue of the spherical Poincaré-type inequality, i.e., using the first bound of Lemma 5.4. It gives

$$
\mathbb{E}_\theta\, |f_\theta(t) - f(t)| \leq \frac{M_1 t}{\sqrt{n-1}} \qquad (t \geq 0)
$$

and hence

$$\int_0^{T_0} \frac{\mathbb{E}_\theta \, |f_\theta(t) - f(t)|}{t} \, dt \le \frac{5M_1}{\sqrt{n-1}} \sqrt{\log n}.$$

Next, we apply Lemma 5.1 with $\lambda = 1/4$ which gives

$$\begin{aligned}
\frac{1}{T} \int_0^T |f(t)| \, dt &\le & \frac{1}{T} \int_0^T \mathbb{E}_\theta \, |f_\theta(t)| \, dt \\
&\le & \frac{2.1}{T} \int_0^T \left(e^{-t^2/16} + e^{-n/24} + \sqrt{\mathbb{P}\{|X-Y|^2 \le n/4\}} \right) dt \\
&\le & \frac{c}{T} + 2.1\sqrt{\mathbb{P}\{|X-Y|^2 \le n/4\}}
\end{aligned}$$

with some absolute constant $c > 0$. Similarly,

$$c \int_{T_0}^T \frac{\mathbb{E}_\theta \, |f_\theta(t)|}{t} \, dt \le \left(e^{-n/24} + \sqrt{\mathbb{P}\{|X-Y|^2 \le n/4\}} \right) \log \frac{T}{T_0} + e^{-T_0^2/16}.$$

These bounds prove the first assertion of the theorem.

For the second assertion, it remains to recall that, by Proposition 2.5,

$$\mathbb{P}\{|X-Y|^2 \le n/4\} \le 16 \, \frac{m_2^2 + \sigma_2^2}{n},$$

so that from (6.3) we get

$$c \, \mathbb{E}_\theta \, \rho(F_\theta, F) \le M_1 \sqrt{\frac{\log n}{n}} + 4 \, \frac{m_2 + \sigma_2}{\sqrt{n}} \, \log n + \frac{1}{n}. \tag{6.5}$$

Here, the last term $1/n$ is dominated by m_2/n. This leads to the bound (6.4), in which F may be replaced with the standard normal distribution function Φ due to the estimate $\rho(F, \Phi) \le \frac{C}{\sqrt{n}} \left(1 + \sigma_2\right)$, cf. Corollary 4.2. $\qquad\square$

Remark 6.4. Working with the Lévy distance L, which in general is weaker then the Kolmogorov distance ρ, one can get guaranteed rates with respect to n for $\mathbb{E}_\theta \, L(F_\theta, F)$ in terms of M_1 or M_2. In particular, if X isotropic, it is known that

$$\mu_{n-1}\{L(F_\theta, F) \ge \delta\} \le 4n^{3/8} \, e^{-n\delta^4/8}, \qquad \delta > 0.$$

This deviation bound yields

$$\mathbb{E}_\theta \, L(F_\theta, F) \le C \left(\frac{\log n}{n} \right)^{1/4}$$

with some absolute constant C ([4]). See also [5] for similar results about the Kantorovich distance.

7 Proof of Theorem 1.1

In order to get rid of the logarithmic term in the bounds of Theorems 1.2/6.3, one may involve the 3-rd moment assumptions in terms of the moment and variance-type functionals m_p and σ_p of index $p = 3$. They are defined by

$$m_3 = m_3(X) = \frac{1}{\sqrt{n}} \left(\mathbb{E} \, | \langle X, Y \rangle |^3 \right)^{1/3},$$

where Y is an independent copy of X, and

$$\sigma_3 = \sigma_3(X) = \sqrt{n} \left(\mathbb{E} \left| \frac{|X|^2}{n} - 1 \right|^{\frac{3}{2}} \right)^{\frac{2}{3}} = \frac{1}{\sqrt{n}} \left(\mathbb{E} \, \big| |X|^2 - n \big|^{\frac{3}{2}} \right)^{\frac{2}{3}}.$$

Let us recall that $m_3 \leq M_3^2$. Hence, Theorem 1.1 will follow from the following, slightly sharpened assertion.

Theorem 7.1. *Let X be a random vector in $\mathbb{R}^n$ with finite 3-rd moment, and such that $\mathbb{E}|X|^2 = n$. Then with some absolute constant c*

$$\mathbb{E}_\theta \, \rho(F_\theta, \Phi) \leq c \, (m_3^{3/2} + \sigma_3^{3/2}) \, \frac{1}{\sqrt{n}}. \tag{7.1}$$

Proof. We now apply Lemma 6.2, choosing there $p = 3/2$, $T = 4n$ and $T_0 = 4\sqrt{\log n}$. Since necessarily $m_3 \geq 1$, the last term $e^{-T_0^2/16}$ is negligible, and we get the bound

$$c \, \mathbb{E}_\theta \, \rho(F_\theta, F) \leq \int_0^{T_0} \frac{\mathbb{E}_\theta \, |f_\theta(t) - f(t)|}{t} \, dt + (m_3^{3/2} + \sigma_3^{3/2}) \frac{\log n}{n^{3/4}}$$

with some absolute constant $c > 0$. To analyze the last integral over the interval $[0, T_0]$, we apply Lemma 5.4, which gives

$$\mathbb{E}_\theta \, |f_\theta(t) - f(t)| \leq \frac{t}{\sqrt{n-1}} \, \sqrt{\mathbb{E} \langle X, Y \rangle J_n(t(X-Y))}, \qquad t \geq 0,$$

and hence

$$\begin{aligned} c \, \mathbb{E}_\theta \, \rho(F_\theta, F) &\leq \frac{1}{\sqrt{n}} \int_0^{T_0} \sqrt{\mathbb{E} \langle X, Y \rangle J_n(t(X-Y))} \, dt \\ &\quad + (m_3^{3/2} + \sigma_3^{3/2}) \frac{\log n}{n^{3/4}}. \end{aligned} \tag{7.2}$$

Next, let us apply the bound of Corollary 3.2, $|J_n(t\sqrt{n}) - e^{-t^2/2}| \leq \frac{C}{n}$, which allows one to replace the J_n-term with $e^{-t^2|X-Y|^2/2n}$ at the expense of an error of order

$$\frac{1}{n} T_0 \sqrt{\mathbb{E} \, | \langle X, Y \rangle |} \leq \frac{\sqrt{m_2}}{n^{3/4}} \, T_0 \leq \frac{m_3^{3/2}}{n^{3/4}} \, T_0,$$

where we used the inequality $m_3 \geq m_2 \geq 1$. As a result, the bound (7.2) may be simplified to

$$c \, \mathbb{E}_\theta \, \rho(F_\theta, F) \leq \frac{1}{\sqrt{n}} \int_0^{T_0} \sqrt{I(t)} \, dt + (m_3^{3/2} + \sigma_3^{3/2}) \frac{\log n}{n^{3/4}} \tag{7.3}$$

with

$$I(t) = \mathbb{E} \langle X, Y \rangle \, e^{-t^2|X-Y|^2/2n}.$$

Note that $I(t) \geq 0$ which follows from $I(t) = \big| \, \mathbb{E} \, e^{it\langle X, Z \rangle/\sqrt{n}} \, \big|^2$, where the random vector Z is independent of X and has a standard normal distribution on $\mathbb{R}^n$.

Now, focusing on $I(t)$, consider the events

$$A = \Big\{ |X-Y|^2 \leq \frac{1}{4} \, n \Big\}, \qquad B = \Big\{ |X-Y|^2 > \frac{1}{4} \, n \Big\}.$$

We split the expectation in the definition of $I(t)$ into the sets A and B, so that $I(t) = I_1(t) + I_2(t)$, where

$$I_1(t) = \mathbb{E} \langle X, Y \rangle \, e^{-t^2|X-Y|^2/2n} \, 1_A, \qquad I_2(t) = \mathbb{E} \langle X, Y \rangle \, e^{-t^2|X-Y|^2/2n} \, 1_B.$$

As we know (cf. Proposition 2.5),

$$\mathbb{P}(A) \leq 64 \, \frac{m_3^3 + \sigma_3^3}{n^{3/2}}.$$

Hence, applying Hölder's inequality, we have

$$
\begin{aligned}
|I_1(t)| &\leq \left(\mathbb{E}\,|\langle X, Y\rangle|^3\right)^{1/3}\left(\mathbb{P}(A)\right)^{2/3} \\
&= m_3\sqrt{n}\cdot 16\,\frac{m_3^2+\sigma_3^2}{n} \leq \frac{32}{\sqrt{n}}\left(m_3^3+\sigma_3^3\right),
\end{aligned}
$$

where we used that $m_3 \geq 1$.

Now, we represent the second expectation as

$$
\begin{aligned}
I_2(t) &= e^{-t^2}\,\mathbb{E}\,\langle X, Y\rangle\,e^{-t^2\left(\frac{|X-Y|^2}{2n}-1\right)}\mathbf{1}_B \\
&= e^{-t^2}\,\mathbb{E}\,\langle X, Y\rangle\left(e^{-t^2\left(\frac{|X-Y|^2}{2n}-1\right)}-1\right)\mathbf{1}_B - e^{-t^2}\,\mathbb{E}\,\langle X, Y\rangle\,\mathbf{1}_A.
\end{aligned}
$$

Here the last expectation has been already bounded by $\frac{32}{\sqrt{n}}\left(m_3^3+\sigma_3^3\right)$. To estimate the first one, we use an elementary inequality

$$
|e^{-x}-1| \leq |x|\,e^{x_0} \qquad (x_0\geq 0,\ \ x\geq -x_0).
$$

Since on the set B, there is a uniform bound $t^2(\frac{|X-Y|^2}{2n}-1)\geq -\frac{7}{8}\,t^2$, we conclude by virtue of Hölder's inequality that

$$
\begin{aligned}
\mathbb{E}\,|\langle X, Y\rangle|\,\left|\,e^{-t^2\left(\frac{|X-Y|^2}{2n}-1\right)}-1\right|\mathbf{1}_B &\leq t^2 e^{7t^2/8}\,\mathbb{E}\,|\langle X, Y\rangle|\,\left|\frac{|X-Y|^2}{2n}-1\right| \\
&\leq t^2 e^{7t^2/8}\left(\mathbb{E}\,|\langle X, Y\rangle|^3\right)^{\frac{1}{3}}\left(\mathbb{E}\,\left|\frac{|X-Y|^2}{2n}-1\right|^{\frac{3}{2}}\right)^{\frac{2}{3}}.
\end{aligned}
$$

The first expectation on the right-hand side is $\mathbb{E}\,|\langle X, Y\rangle|^3 = (m_3\sqrt{n})^3$. Writing

$$
\frac{|X-Y|^2}{2n}-1 = \frac{1}{2}\left(\frac{|X|^2}{n}-1\right)+\frac{1}{2}\left(\frac{|Y|^2}{n}-1\right)-\frac{1}{n}\,\langle X, Y\rangle,
$$

we also have, by Jensen's inequality,

$$
\left|\frac{|X-Y|^2}{2n}-1\right|^{\frac{3}{2}} \leq \left|\frac{|X|^2}{n}-1\right|^{\frac{3}{2}}+\left|\frac{|Y|^2}{n}-1\right|^{\frac{3}{2}}+\frac{2}{n^{3/2}}\,|\langle X, Y\rangle|^{\frac{3}{2}}.
$$

Therefore

$$
\begin{aligned}
\mathbb{E}\,\left|\frac{|X-Y|^2}{2n}-1\right|^{\frac{3}{2}} &\leq 2\left|\frac{|X|^2}{n}-1\right|^{\frac{3}{2}}+\frac{2}{n^{3/2}}\left(\mathbb{E}\,|\langle X, Y\rangle|^3\right)^{1/2} \\
&= \frac{2}{n^{3/4}}\left(\sigma_3^{3/2}+m_3^{3/2}\right),
\end{aligned}
$$

which gives

$$
\left(\mathbb{E}\,\left|\frac{|X-Y|^2}{2n}-1\right|^{\frac{3}{2}}\right)^{\frac{2}{3}} \leq \frac{2}{\sqrt{n}}\left(\sigma_3+m_3\right).
$$

Hence

$$
\begin{aligned}
\mathbb{E}\,|\langle X, Y\rangle|\,\left|\,e^{-t^2\left(\frac{|X-Y|^2}{2n}-1\right)}-1\right|\mathbf{1}_B &\leq 2t^2 e^{7t^2/8}\,m_3\left(m_3+\sigma_3\right) \\
&\leq 4t^2 e^{7t^2/8}\left(m_3^2+\sigma_3^2\right),
\end{aligned}
$$

and, as a result,

$$
I_2(t) \leq 32\,\frac{e^{-t^2}}{\sqrt{n}}\left(m_3^3+\sigma_3^3\right)+4t^2 e^{-t^2/8}\left(m_3^2+\sigma_3^2\right),
$$

where the factor e^{-t^2} in the first term can be removed without loss of strength.

Together with the estimate on $I_1(t)$, we get

$$I(t) \leq \frac{64}{\sqrt{n}}\left(m_3^3 + \sigma_3^3\right) + 4t^2 e^{-t^2/8}\left(m_3^2 + \sigma_3^2\right),$$

so

$$\sqrt{I(t)} \leq \frac{8}{n^{1/4}}\left(m_3^{3/2} + \sigma_3^{3/2}\right) + 2|t|\, e^{-t^2/16}\left(m_3 + \sigma_3\right)$$

and

$$\frac{1}{\sqrt{n}}\int_0^{T_0} \sqrt{I(t)}\, dt \leq \frac{4T_0}{n^{3/4}}\left(m_3^{3/2} + \sigma_3^{3/2}\right) + \frac{C}{\sqrt{n}}\left(m_3 + \sigma_3\right)$$

with some absolute constant C.

Returning to the bound (7.3), we thus obtain that

$$c\,\mathbb{E}_\theta\,\rho(F_\theta, F) \leq \frac{\log n}{n^{3/4}}\left(m_3^{3/2} + \sigma_3^{3/2}\right) + \frac{C}{\sqrt{n}}\left(m_3 + \sigma_3\right)$$

To simplify it, one may use again that $m_3 \geq 1$, which implies that $m_3 + \sigma_3 \leq 2(m_3^{3/2} + \sigma_3^{3/2})$ for all values of σ_3. Thus, with some absolute constant $c > 0$,

$$c\,\mathbb{E}_\theta\,\rho(F_\theta, F) \leq \frac{C}{\sqrt{n}}\left(m_3^{3/2} + \sigma_3^{3/2}\right).$$

To get a similar bound with Φ in place of F, i.e. (7.1), one may apply the estimate $\rho(F, \Phi) \leq C\,\frac{1+\sigma_2}{\sqrt{n}}$, where $1 + \sigma_2^2$ may further be bounded by $2(m_3^{3/2} + \sigma_3^{3/2})$. $\qquad\square$

8 The i.i.d. case

Theorem 1.3 follows from Theorem 6.3, by taking into account the following elementary statement (various variants of which under higher moment assumptions are well-known).

Lemma 8.1. *Assume that the non-negative random variables $\xi_1, \ldots, \xi_n$ are independent and identically distributed, with $\mathbb{E}\xi_1 = 1$. Given $0 < \lambda < 1$, let a number $\kappa > 0$ is chosen to satisfy*

$$\mathbb{E}\,\xi_1\,1_{\{\xi_1 > \kappa\}} \leq \frac{1 - \lambda}{2}.$$

Then for the sum $S_n = \xi_1 + \cdots + \xi_n$, we have

$$\mathbb{P}\{S_n \leq \lambda n\} \leq \exp\left\{-\frac{(1-\lambda)^2}{8\kappa}\,n\right\}. \tag{8.1}$$

Proof. Let V denote the common distribution of ξ_k. The function

$$u(t) = \mathbb{E}\,e^{-t\xi_1} = \int_0^\infty e^{-tx}\, dV(x), \qquad t \geq 0,$$

is positive, convex, non-increasing, and has a continuous, non-decreasing derivative

$$u'(t) = -\mathbb{E}\,\xi_1 e^{-t\xi_1} = -\int_0^\infty x e^{-tx}\, dV(x),$$

with $u(0) = 1$, $u'(0) = -1$.

Let κ_p denote the maximal quantile for the probability measure $x\,dV(x)$ on $(0,\infty)$ of a given order $p \in (0,1)$, i.e., the minimal number such that

$$\int_{\kappa_p}^{\infty} x\,dV(x) \le 1 - p,$$

where the integration is performed over the open half-axis (κ_p, ∞). Using the elementary inequality $1 - e^{-y} \le y$ $(y \ge 0)$, we have, for all $s > 0$,

$$
\begin{aligned}
1 + u'(s) &= \int_0^{\infty} x(1 - e^{-sx})\,dV(x) \\
&= \int_{0 < x \le \kappa_p} x(1 - e^{-sx})\,dV(x) + \int_{x > \kappa_p} x(1 - e^{-tx})\,dV(x) \\
&\le s \int_{0 < x \le \kappa_p} x^2\,dV(x) + p \le p + \kappa_p s.
\end{aligned}
$$

This gives

$$
\begin{aligned}
u(t) &= 1 - t + \int_0^t (1 + u'(s))\,ds \\
&\le 1 - t + pt + \kappa_p \frac{t^2}{2} \le \exp\left\{ -t + pt + \kappa_p \frac{t^2}{2} \right\},
\end{aligned}
$$

and therefore

$$
\begin{aligned}
\mathbb{P}\{S_n \le \lambda n\} &\le e^{\lambda n t}\,\mathbb{E}\,e^{-tS_n} = e^{\lambda n t}\,u(t)^n \\
&\le \exp\left\{ -n\Big((1 - \lambda - p)t - \kappa_p \frac{t^2}{2}\Big) \right\}.
\end{aligned}
$$

If $p < 1 - \lambda$, the right-hand side is minimized at $t = \frac{1 - \lambda - p}{\kappa_p}$, and we get

$$\mathbb{P}\{S_n \le \lambda n\} \le e^{-n\,(1 - \lambda - p)^2 / 2\kappa_p}.$$

One may take, for example, $p = (1 - \lambda)/2$, and then we arrive at (8.1). $\qquad\square$

Proof of Theorem 1.3. First let us derive the inequality

$$\mathbb{E}_\theta\,\rho(F_\theta, F) \le c\,\sqrt{\frac{\log n}{n}} \tag{8.2}$$

with the typical distribution F in place of G. Let $Y = (Y_1, \ldots, Y_n)$ be an independent copy of X. Since the Kolmogorov distance is scale invariant, without loss of generality one may assume that $\mathbb{E}\,(X_1 - Y_1)^2 = 1$. But then, by Lemma 8.1, applied to the random variables $\xi_k = (X_k - Y_k)^2$, we have

$$\mathbb{P}\{|X - Y|^2 \le n/4\} \le e^{-cn}$$

with some constant $c > 0$ depending on the distribution of X_1 only. In addition, $M_1 \le M_2 = \frac{1}{2}$. As a result, Theorem 6.3 yields (8.2).

Now, in order to replace F with G in (8.2), one may apply Proposition 3.1. Indeed, F represents the distribution function of rZ_n, where $r = \frac{1}{\sqrt{n}}\,|X|$ and $Z_n = \sqrt{n}\,\theta_1$ is independent of r. Similarly, G is the distribution function of rZ where $Z \sim N(0,1)$ is independent of r. Let Φ_n denote the distribution function of Z_n and φ_n its density.

Since $F(x) = \mathbb{E}\,\Phi_n(x/r)$ and $G(x) = \mathbb{E}\,\Phi(x/r)$, we conclude, by the triangle inequality, that

$$
\begin{aligned}
\rho(F, G) &\leq \mathbb{E}\sup_x |\Phi_n(x/r) - \Phi(x/r)| \\
&= \sup_x |\Phi_n(x) - \Phi(x)| \leq \int_{-\infty}^{\infty} |\varphi_n(x) - \varphi(x)|\,dx \leq \frac{C}{n}. \quad \square
\end{aligned}
$$

References

[1] Aistleitner, C. and Berkes, I.: On the central limit theorem for $f(n_k x)$. *Probab. Theory Related Fields* **146**, (2010), no. 1-2, 267–289. MR-2550364

[2] Aistleitner, C. and Elsholtz, C.: The central limit theorem for subsequences in probabilistic number theory. *Canad. J. Math.* **64**, (2012), no. 6, 1201–1221. MR-2994662

[3] Bateman, H.: Higher transcendental functions, Vol. II. *McGraw-Hill Book Company, Inc.*, 1953, 396 pp. MR-0066496

[4] Bobkov, S. G.: On concentration of distributions of random weighted sums. *Ann. Probab.* **31**, (2003), no. 1, 195–215. MR-1959791

[5] Bobkov, S. G.: On a theorem of V. N. Sudakov on typical distributions. (Russian) *J. Math. Sciences (New York)* **167**, (2010), no. 4, 464–473. Translated from: *Zap. Nauchn. Semin. POMI*, vol. 368, (2009), 59–74. MR-2749183

[6] Bobkov, S. G.: Closeness of probability distributions in terms of Fourier-Stieltjes transforms. *Russian Math. Surveys*, vol. 71, issue 6, (2016), 1021–1079. Translated from: *Uspekhi Matem. Nauk*, vol. 71, issue 6 (432), (2016), 37–98. MR-3588939

[7] Bobkov, S. G., Chistyakov, G. P. and Götze, F.: Gaussian mixtures and normal approximation for V. N. Sudakov's typical distributions. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov.* (POMI) **457** (2017), Veroyatnost i Statistika. 25, 37–52. MR-3723575

[8] Bobkov, S. G., Chistyakov, G. P. and Götze, F.: Concentration and Gaussian approximation for randomized sums. In preparation.

[9] Bobkov, S. G. and Götze, F. On the central limit theorem along subsequences of noncorrelated observations. *Theory Probab. Appl.* **48**, (2004), no. 4, 604–621. MR-2142522

[10] Bobkov, S. G. and Götze, F.: Concentration inequalities and limit theorems for randomized sums. *Probab. Theory Related Fields* **137**, (2007), no. 1–2, 49–81. MR-2278452

[11] Fukuyama, K.: A central limit theorem to trigonometric series with bounded gaps. *Probab. Theory Related Fields* **149**, (2011), no. 1–2, 139–148. MR-2773027

[12] Gaposhkin, V. F.: The rate of approximation to the normal law of the distributions of weighted sums of lacunary series. (Russian) *Teor. Verojatnost. i Primenen.* **13**, (1968), 445–461. MR-0239644

[13] Kac, M.: On the distribution of values of sums of the type $\sum f(2^k t)$. *Ann. Math.* **47** (2), (1946), 33–49. MR-0015548

[14] Klartag, B.: A Berry-Esseen type inequality for convex bodies with an unconditional basis. *Probab. Theory Related Fields* **145**, (2009), no. 1-2, 1–33. MR-2520120

[15] Klartag, B. and Sodin, S.: Variations on the Berry-Esseen theorem. *Teor. Veroyatn. Primen.* **56**, (2011), no. 3, 514–533; reprinted in: *Theory Probab. Appl.* **56**, (2012), no. 3, 403–419. MR-3136463

[16] Ledoux, M.: Concentration of measure and logarithmic Sobolev inequalities. Séminaire de Probabilités XXXIII. Lect. Notes in Math. 1709 (1999), 120–216, Springer. MR-1767995

[17] Matskyavichyus, V. K.: A lower bound for the rate of convergence in the central limit theorem. (Russian) *Teor. Veroyatnost. i Primenen.* **28**, (1983), no. 3, 565–569. MR-0716313

[18] Meckes, E. S. and Meckes, M. W.: The central limit problem for random vectors with symmetries. *J. Theoret. Probab.* **20**, (2007), no. 4, 697–720. MR-2359052

[19] Petrov, V. V.: Sums of independent random variables. Translated from the Russian by A. A. Brown. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 82. *Springer–Verlag*, New York–Heidelberg, 1975. x+346 pp. MR-0388499

[20] Petrov, V. V.: Limit theorems for sums of independent random variables (Russian), *Nauka*, Moscow, 1987. 318 pp. MR-0896036

[21] Salem, R. and Zygmund, A.: On lacunary trigonometric systems. *Proc. Nat. Acad. Sci. USA* **33**, (1947), 333–338. MR-0022263

[22] Salem, R. and Zygmund, A.: On lacunary trigonometric series. II. *Proc. Nat. Acad. Sci. USA* **34**, (1948), 54–62. MR-0023936

[23] Sudakov, V. N.: Typical distributions of linear functionals in finite-dimensional spaces of high dimension. (Russian) *Soviet Math. Dokl.* **19**, (1978), 1578–1582; translation in: *Dokl. Akad. Nauk SSSR* **243**, (1978), no. 6, 1402–1405. MR-0517198