Asymptotic Expansions and Two-Sided Bounds in Randomized Central Limit Theorems

Sergey G. Bobkov, Gennadiy P. Chistyakov, and Friedrich Götze

Abstract Lower and upper bounds are explored for the uniform (Kolmogorov) and L^2-distances between the distributions of weighted sums of dependent summands and the normal law. The results are illustrated for several classes of random variables whose joint distributions are supported on Euclidean spheres. We also survey several results on improved rates of normal approximation in randomized central limit theorems.

Keywords Typical distributions · Normal approximation · Central limit theorem

1 Introduction

A random vector $X = (X_1, \dots, X_n)$ in $\mathbb{R}^n$ ($n \geq 2$) defined on the probability space $(\Omega, \mathfrak{F}, \mathbb{P})$ is called isotropic, if

$$\mathbb{E} X_i X_j = \delta_{ij} \quad \text{for all } i, j \leq n,$$

where δ_{ij} is the Kronecker symbol. Equivalently, all weighted sums

$$S_\theta = \theta_1 X_1 + \dots + \theta_n X_n, \qquad \theta = (\theta_1, \dots, \theta_n), \ \ \theta_1^2 + \dots + \theta_n^2 = 1,$$

with coefficients from the unit sphere $\mathbb{S}^{n-1}$ in $\mathbb{R}^n$ have a second moment $\mathbb{E} S_\theta^2 = 1$. In this case, provided that the Euclidean norm $|X|$ is almost constant, and if n is large, a theorem due to Sudakov [27] asserts that the distribution functions

$$F_\theta(x) = \mathbb{P}\{S_\theta \leq x\}, \quad x \in \mathbb{R},$$

S. G. Bobkov (✉)
School of Mathematics, University of Minnesota, Minneapolis, MN, USA
e-mail: bobko001@umn.edu

G. P. Chistyakov · F. Götze
Faculty of Mathematics, University of Bielefeld, Bielefeld, Germany

R. Eldan et al. (eds.), *Geometric Aspects of Functional Analysis*, Lecture Notes in Mathematics 2327, https://doi.org/10.1007/978-3-031-26300-2_3

are well approximated for most of $\theta \in \mathbb{S}^{n-1}$ by the standard normal distribution function

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

Here, "most" should refer to the normalized Lebesgue measure $\mathfrak{s}_{n-1}$ on the sphere. This property may be quantified, for example, in terms of the Kolmogorov distance

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

Being rather universal (since no independence of the components X_k is required), randomized central limit theorems of such type have received considerable interest in recent years. For the history, bibliography, and interesting connections with other concentration problems we refer an interested reader to [8, 9, 12]. Let us mention one general upper bound

$$\mathbb{E}_\theta\, \rho(F_\theta, \Phi) \ \le\ c\,(1 + \sigma_4)\, \frac{\log n}{\sqrt{n}}, \tag{1.1}$$

which holds true with an absolute constant $c > 0$ for any isotropic random vector X (cf. Theorem 1.2 in [8]). Here and elsewhere, $\mathbb{E}_\theta$ denotes an integral over $\mathbb{S}^{n-1}$ with respect to the measure $\mathfrak{s}_{n-1}$, and the bound involves the variance-type functional

$$\sigma_4^2 = \sigma_4^2(X) = \frac{1}{n}\,\mathrm{Var}(|X|^2) \quad (\sigma_4 \ge 0).$$

Modulo a logarithmic factor, the bound (1.1) exhibits a standard rate of normal approximation for F_θ, in analogy with the classical case of independent identically distributed (iid) summands with equal coefficients. It turns out, however, that in the model with arbitrary $\theta \in \mathbb{S}^{n-1}$ and independent components X_k, the standard rate for $\rho(F_\theta, \Phi)$ is dramatically improved to the order $1/n$ on average and actually for most of θ. Motivated by the seminal paper of Klartag and Sodin [20], this interesting phenomenon was recently studied in [9, 10] for dependent data under certain correlation-type conditions. The last chapters of this paper provide a short account of these improved rates of normal approximation.

One of the main aims of this work is to develop lower bounds with a similar standard rate as in (1.1) (modulo logarithmic factors) and to illustrate them with a number of examples of random variables X_k often appearing in Functional Analysis. These results rely on a careful examination of the closely related L^2-distance

$$\omega(F_\theta, \Phi) = \Big(\int_{-\infty}^{\infty} (F_\theta(x) - \Phi(x))^2\, dx\Big)^{1/2}.$$

Similarly to (1.1), it can be shown that for the class of isotropic random vectors the inequality

$$\mathbb{E}_\theta\, \omega^2(F_\theta, \Phi) \ \le\ c\,(1 + \sigma_4^2)\,\frac{1}{n} \tag{1.2}$$

holds without an unnecessary logarithmic term. However, in order to explore the real behavior of the average L^2-distance, some other characteristics of the distribution of X are required. For example, assuming that the distribution is supported on the sphere $\sqrt{n}\,\mathbb{S}^{n-1}$, the L^2-distance admits an asymptotic expansion in terms of the moment functionals (normalized L^p-norms)

$$m_p = m_p(X) = \frac{1}{\sqrt{n}}\,\big(\mathbb{E}\,\langle X, Y\rangle^p\big)^{1/p} = \frac{1}{\sqrt{n}}\,\Big(\sum (\mathbb{E} X_{i_1} \dots X_{i_p})^2\Big)^{1/p}.$$

Here, Y is an independent copy of X, and the summation is performed over all indices $1 \le i_1, \dots, i_p \le n$. The second representation shows that these functionals are non-negative for any integer $p \ge 1$. Note that $m_1 = 0$ if X has mean zero, $m_2 = 1$ if X is isotropic, and $m_p = 0$ with odd p when the distribution of X is symmetric about the origin. The following expansion involves the moments m_p up to order 4.

Theorem 1.1 *Let X be an isotropic random vector in $\mathbb{R}^n$ with mean zero and such that $|X|^2 = n$ a.s. We have*

$$\mathbb{E}_\theta\, \omega^2(F_\theta, \Phi) \ =\ \frac{c}{n^{3/2}}\, m_3^3 + O\Big(\frac{1}{n^2}\, m_4^4\Big) \tag{1.3}$$

with $c = \frac{1}{16\sqrt{\pi}}$. Similarly, with some absolute constants $c_1, c_2 > 0$,

$$\mathbb{E}_\theta\, \rho^2(F_\theta, \Phi) \ \le\ \frac{c_1 \log n}{n^{3/2}}\, m_3^3 + \frac{c_2 (\log n)^2}{n^2}\, m_4^4. \tag{1.4}$$

As we will see, in the general isotropic case without the support assumption, but with bounded σ_4, the average L^2-distance is described by a more complicated formula

$$\begin{aligned}\mathbb{E}_\theta\, \omega^2(F_\theta, \Phi) = {} & \frac{1}{\sqrt{2\pi n}}\,\Big(1 + \frac{1}{8n}\Big)\,\mathbb{E}\sqrt{|X|^2 + |Y|^2} \\ & - \frac{1}{\sqrt{2\pi n}}\,\Big(1 + \frac{1}{4n}\Big)\,\mathbb{E}\,|X - Y| + O\Big(\frac{1 + \sigma_4^2}{n^2}\Big),\end{aligned} \tag{1.5}$$

which holds whenever $\mathbb{E}\,|X|^2 = n$.

In the setting of Theorem 1.1, using the pointwise bound $|\,\langle X, Y\rangle\,| \le n$ together with the isotropy assumption, we have $\mathbb{E}\,\langle X, Y\rangle^3 \ \le\ n^2$ and $\mathbb{E}\,\langle X, Y\rangle^4 \ \le\ n^3$.

Therefore, the inequalities (1.3) and (1.4) yield with some absolute constant $c > 0$

$$\mathbb{E}_\theta\, \omega^2(F_\theta, \Phi) \le \frac{c}{n}, \quad \mathbb{E}_\theta\, \rho^2(F_\theta, \Phi) \le \frac{c\,(\log n)^2}{n}, \tag{1.6}$$

thus recovering the upper bounds (1.1) and (1.2) for this particular case (since $\sigma_4 = 0$). On the other hand, for a large variety of examples, such bounds turn out to be optimal and may be reversed modulo a logarithmic factor (for large n). To see this, one may use the following lower bound which will be derived from a slightly modified variant of (1.5).

Theorem 1.2 *Let X be a random vector in $\mathbb{R}^n$ satisfying $\mathbb{E}\,|X|^2 = n$, and let Y be its independent copy. For some absolute constants $c_1, c_2 > 0$, we have*

$$\mathbb{E}_\theta\, \omega^2(F_\theta, \Phi) \ \ge\ c_1\, \mathbb{P}\Big\{|X - Y| \le \frac{1}{2}\sqrt{n}\Big\} - c_2\, \frac{1 + \sigma_4^4}{n^2}. \tag{1.7}$$

Thus, if the probability in (1.7) is of order at least $1/n$, and σ_4 is bounded, the right-hand side of this bound will be of the same order. If, for example, $|X| = \sqrt{n}$ a.s., we then obtain that $\mathbb{E}_\theta\, \omega^2(F_\theta, \Phi) \sim 1/n$. In order to derive a similar conclusion for the Kolmogorov distance, one may refer to the next statement.

Theorem 1.3 *Let X be an isotropic random vector in $\mathbb{R}^n$ such that $|X| \le b\sqrt{n}$ a.s. Suppose that we have a lower bound at the standard rate*

$$\mathbb{E}_\theta\, \omega^2(F_\theta, \Phi) \ \ge\ \frac{D}{n}$$

with some $D > 0$. Then with some absolute constants $c_0, c_1 > 0$

$$\mathbb{E}_\theta\, \rho(F_\theta, F) \ \ge\ \frac{c_0}{(1 + \sigma_4)^3\, b^2}\, \frac{D^2}{(\log n)^4\, \sqrt{n}} - \frac{c_1\,(1 + \sigma_4^2)}{n}.$$

These estimates may be employed to arrive at the two-sided bounds of the form

$$\frac{c_0}{n} \ \le\ \mathbb{E}_\theta\, \omega^2(F_\theta, \Phi) \ \le\ \frac{c_1}{n}, \qquad \frac{c_0}{(\log n)^4 \sqrt{n}} \le \mathbb{E}_\theta\, \rho(F_\theta, \Phi) \le \frac{c_1 \log n}{\sqrt{n}} \tag{1.8}$$

with some absolute constants $c_0 > 0$ and $c_1 > 0$. Examples where both inequalities in (1.8) are fulfilled include the following uniformly bounded orthonormal systems in $L^2(\Omega, \mathfrak{F}, \mathbb{P})$:

(i) The trigonometric system $X = (X_1, \dots, X_n)$ with components

$$X_{2k-1}(t) = \sqrt{2}\, \cos(kt),$$
$$X_{2k}(t) = \sqrt{2}\, \sin(kt) \quad (-\pi < t < \pi,\ k = 1, \dots, n/2,\ n \text{ even})$$

on the interval $\Omega = (-\pi, \pi)$ equipped with the normalized Lebesgue measure $\mathbb{P}$.

(ii) The cosine trigonometric system $X = (X_1, \dots, X_n)$ with

$$X_k(t) = \sqrt{2}\, \cos(kt)$$

on the interval $\Omega = (0, \pi)$ equipped with the normalized Lebesgue measure $\mathbb{P}$.

(iii) The normalized Chebyshev polynomials $X_1, \dots, X_n$ defined by

$$\begin{aligned} X_k(t) &= \sqrt{2} \cos(k \arccos t) \\ &= \sqrt{2} \Big[t^n - \binom{n}{2} t^{n-2}(1-t^2) + \binom{n}{4} t^{n-4}(1-t^2)^2 - \dots \Big] \end{aligned}$$

on $\Omega = (-1, 1)$ equipped with the probability measure $d\mathbb{P}(t) = \frac{1}{\pi\sqrt{1-t^2}}\, dt$, $|t| < 1$.

(iv) The systems of functions of the form

$$X_k(t, s) = \Psi(kt + s), \quad k = 1, \dots, n \ \ (0 < t, s < 1)$$

on the square $\Omega = (0, 1) \times (0, 1)$ equipped with the Lebesgue measure $\mathbb{P}$. In this case, (1.8) holds true for any 1-periodic Lipschitz function Ψ on the real line such that $\int_0^1 \Psi(x)\, dx = 0$ and $\int_0^1 \Psi(x)^2\, dx = 1$ with constants c_0 and c_1 depending on Ψ only.

(v) The Walsh system

$$X = \{X_\tau\}_{\tau \neq \emptyset}, \quad \tau \subset \{1, \dots, d\},$$

of dimension $n = 2^d - 1$ on the discrete cube $\Omega = \{-1, 1\}^d$ (the ordering of the components does not play any role). Here, $\mathbb{P}$ denotes the normalized counting measure, and

$$X_\tau(t) = \prod_{k \in \tau} t_k \quad \text{for } t = (t_1, \dots, t_d) \in \Omega.$$

(vi) Random vectors X with associated empirical distribution functions F_θ based on the "observations" $X_k = \sqrt{n}\, \theta_k$ $(1 \le k \le n)$.

The paper is organized as follows. We start in Sect. 2 with a review of several results on the so-called typical distributions F which serve as main approximations for F_θ (in general, they do not need to be normal, or even nearly normal). Sections 3–7 deal with the L^2-distances $\omega(F_\theta, F)$ only, while Sects. 8–12 are mostly focused on the Kolmogorov distances $\rho(F_\theta, F)$. In Sect. 13, the examples described in items (i)–(vi) illustrate the applicability of Theorems 1.1–1.3, thus with a standard rate of normal approximation. In Sect. 14 we consider lacunary trigonometric systems and show that the typical rate is improved to the order $1/n$. Similar improved rates are also reviewed in the last section in presence of certain correlation-type conditions. Thus an outline of all sections reads as:

1. Introduction
2. Typical distributions
3. Upper bound for the L^2-distance at standard rate
4. General approximations for the L^2-distance with error of order at most $1/n$
5. Proof of Theorem 1.1 for the L^2-distance
6. General lower bounds for the L^2-distance. Proof of Theorem 1.2
7. Lipschitz systems
8. Berry-Esseen-type bounds
9. Quantitative forms of Sudakov's theorem for the Kolmogorov distance
10. Proof of Theorem 1.1 for the Kolmogorov Distance
11. Relations between L^1, L^2 and Kolmogorov distances
12. Lower bounds. Proof of Theorem 1.3
13. Functional examples
14. The Walsh system; Empirical measures
15. Improved rates for lacunary systems
16. Improved rates for independent and log-concave summands
17. Improved rates under correlation-type conditions

As usual, the Euclidean space $\mathbb{R}^n$ is endowed with the canonical norm $|\cdot|$ and the inner product $\langle\cdot,\cdot\rangle$. In the sequel, we denote by $\mathbb{E}_\theta$ an integral over $\mathbb{S}^{n-1}$ with respect to the measure $\mathfrak{s}_{n-1}$. By $c, c_1, c_2, \dots$, we denote positive absolute constants which may vary from place to place (if not stated explicitly that c depends on some parameter). Similarly C will denote a quantity bounded by an absolute constant. Throughout, we assume that X is a given random vector in $\mathbb{R}^n$ ($n \geq 2$) and Y is its independent copy.

2 Typical Distributions

In the sequel, we denote by

$$F(x) = \mathbb{E}_\theta F_\theta(x) = \mathbb{E}_\theta\, \mathbb{P}\{S_\theta \leq x\}, \quad x \in \mathbb{R},$$

the mean distribution function of the weighted sums $S_\theta = \langle X, \theta\rangle$ with respect to the uniform measure $\mathfrak{s}_{n-1}$. It is also called a typical distribution function using the terminology of [27]. Indeed, according to Sudakov's theorem, if X is isotropic, then most of F_θ are concentrated about F in a weak sense (cf. [1, 2, 8] for quantitative statements).

However, whether or not F itself is close to the normal distribution function Φ is determined by the concentration properties of the distribution of $|X|$. Note that, due to the rotational invariance of $\mathfrak{s}_{n-1}$, the typical distribution can be described as the distribution of the product $\theta_1 \, |X|$, assuming that $\theta = (\theta_1, \dots, \theta_n)$ is a random vector which is independent of X and has distribution $\mathfrak{s}_{n-1}$. In this product, $\theta_1 \sqrt{n}$ is almost standard normal, so that F is almost standard normal, if and only if $\frac{1}{\sqrt{n}} \, |X|$ is almost 1 (like in the weak law of large numbers). This assertion can be quantified in terms of the weighted total variation distance by virtue of the following upper bound derived in [7].

Proposition 2.1 *If* $\mathbb{E} \, |X|^2 = n$ *(in particular, when* X *is isotropic), then*

$$\int_{-\infty}^{\infty} (1 + x^2) \, |F(dx) - \Phi(dx)| \, \leq \, \frac{c}{n} \, \big(1 + \mathrm{Var}(|X|)\big).$$

In particular, this gives a non-uniform bound for the normal approximation, namely

$$|F(x) - \Phi(x)| \, \leq \, \frac{c}{n \, (1 + x^2)} \, \big(1 + \mathrm{Var}(|X|)\big), \quad x \in \mathbb{R}. \tag{2.1}$$

In these bounds we shall rely on the following monotone functionals (of p)

$$\sigma_{2p} = \sqrt{n} \, \bigg(\mathbb{E} \, \Big|\frac{|X|^2}{n} - 1\Big|^p\bigg)^{1/p}, \quad p \geq 1, \tag{2.2}$$

where the particular cases $p = 1$ and $p = 2$ will be most important. If $\mathbb{E} \, |X|^2 = n$, we thus deal with a more tractable quantity

$$\sigma_4^2 = \frac{1}{n} \, \mathrm{Var}\big(|X|^2\big).$$

Using an elementary inequality $\mathrm{Var}(\xi) \, \mathbb{E}\xi^2 \leq \mathrm{Var}(\xi^2)$ (which is true for any random random variable $\xi \geq 0$), we have $\mathrm{Var}(|X|) \leq \sigma_4^2$. Another similar relation

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

can be found in [8]. From (2.1), we therefore obtain the following bounds for the normal approximation in all L^p-norms

$$\|F - \Phi\|_p = \Big(\int_{-\infty}^{\infty} |F(x) - \Phi(x)|^p \, dx\Big)^{1/p},$$

including the limit case

$$\|F - \Phi\|_\infty = \rho(F, \Phi) = \sup_x |F(x) - \Phi(x)|.$$

Corollary 2.2 *If* $\mathbb{E}\,|X|^2 = n$*, then, for all* $p \geq 1$*,*

$$\|F - \Phi\|_p \,\leq\, c\,\frac{1+\sigma_2}{\sqrt{n}}, \qquad \|F - \Phi\|_p \,\leq\, c\,\frac{1+\sigma_4^2}{n}. \tag{2.3}$$

Note that the characteristic function associated to F is given by

$$f(t) \,=\, \mathbb{E}_\theta\,\mathbb{E}\,e^{it\langle X,\theta\rangle} \,=\, \mathbb{E}_\theta\,\mathbb{E}\,e^{it|X|\,\theta_1} \,=\, \mathbb{E}\,J_n(t|X|), \qquad t \in \mathbb{R}, \tag{2.4}$$

where J_n denotes the characteristic function of the first coordinate θ_1 of θ under $\mathfrak{s}_{n-1}$. Hence, by the Plancherel theorem,

$$\omega^2(F,\Phi) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\big(\mathbb{E}\,J_n(t|X|) - e^{-t^2/2}\big)^2\,\frac{dt}{t^2}. \tag{2.5}$$

For $p = 2$, the relations in (2.3) can also be derived by means of (2.5) and by virtue of the following Edgeworth-type approximations derived in [8] and [10].

Lemma 2.3 *For all* $t \in \mathbb{R}$*,*

$$\big|J_n\big(t\sqrt{n}\big) - e^{-t^2/2}\big| \,\leq\, \frac{c}{n}\,\min\{1,t^2\}. \tag{2.6}$$

Moreover,

$$\Big|J_n\big(t\sqrt{n}\big) - \Big(1 - \frac{t^4}{4n}\Big)\,e^{-t^2/2}\Big| \,\leq\, \frac{c}{n^2}\,\min\{1,t^4\}. \tag{2.7}$$

The functions J_n have a subgaussian (although oscillatory) decay on a long interval of the real line. In particular, as was shown in [8],

$$\big|J_n\big(t\sqrt{n}\big)\big| \,\leq\, 5\,e^{-t^2/2} + 4\,e^{-n/12}, \quad t \in \mathbb{R}. \tag{2.8}$$

This bound can be used for the estimation of the characteristic function of the typical distribution, by involving the variance-type functionals σ_{2p}.

Lemma 2.4 *The characteristic function of the typical distribution satisfies, for all* $t \in \mathbb{R}$*,*

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

with constants $c_p > 0$ *depending on* $p \geq 1$ *only. Consequently, for all* $T > 0$*,*

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

Proof One may split the expectation in (2.4) to the event $A = \{|X|^2 \le \lambda n\}$ and its complement $B = \{|X|^2 > \lambda n\}$, $0 < \lambda < 1$. By (2.8),

$$\begin{aligned}\mathbb{E}\,|J_n(t|X|)|\,1_B &\le \mathbb{E}\left(5\,e^{-t^2|X|^2/2n} + 4\,e^{-n/12}\right)1_B \\ &\le 5\,e^{-\lambda t^2/2} + 4\,e^{-n/12}.\end{aligned}$$

On the other hand, recalling the definition (2.2), we have

$$\begin{aligned}\mathbb{P}(A) &= \mathbb{P}\big\{n - |X|^2 \ge (1-\lambda)n\big\} \\ &\le \frac{1}{((1-\lambda)n)^p}\,\mathbb{E}\,|n - |X|^2|^p \;=\; \frac{\sigma_{2p}^p}{(1-\lambda)^p\,n^{p/2}}.\end{aligned} \tag{2.9}$$

Choosing $\lambda = \frac{1}{2}$, and since $|J_n(s)| \le 1$ for all $s \in \mathbb{R}$, we get

$$\mathbb{E}\,|J_n(t|X|)|\,1_A \;\le\; (2\sigma_{2p})^p\,n^{-p/2},$$

thus implying that

$$|f(t)| \;\le\; 5\,e^{-t^2/4} + 4\,e^{-n/12} + (2\sigma_{2p})^p\,n^{-p/2}.$$

This readily yields the desired pointwise and integral bounds of the lemma. □

If $|X| = \sqrt{n}$ a.s., the typical distribution F is just the distribution of $\sqrt{n}\,\theta_1$, the normalized first coordinate of a point on the unit sphere under $\mathfrak{s}_{n-1}$, whose characteristic function is $J_n(t\sqrt{n})$. In this case, the subgaussian character of F manifests itself in corresponding deviation and moment inequalities such as the following.

Lemma 2.5 *For all $p > 0$,*

$$\mathbb{E}_\theta\,|\theta_1|^p \;\le\; 2\left(\frac{p}{n}\right)^{p/2}. \tag{2.10}$$

This inequality can be derived from the well-known bound on the Laplace transform

$$\mathbb{E}_\theta\,e^{t\theta_1} \;\le\; \exp\left\{\frac{t^2}{2(n-1)}\right\}, \quad t \in \mathbb{R},$$

which follows from the fact that the logarithmic Sobolev constant for the unit sphere is equal to $n - 1$ (cf. [21]). Using $x^p \le (\frac{p}{e})^p e^x$, $x \ge 0$, we have $|x|^p \le 2\,(\frac{p}{e})^p \cosh(x)$, $x \in \mathbb{R}$, and the above bound implies

$$t^p\, \mathbb{E}_\theta\, |\theta_1|^p \le 2\left(\frac{p}{e}\right)^p e^{\frac{t^2}{2(n-1)}} \quad \text{for all } t \ge 0.$$

The latter can be optimized over t, which leads to (2.10), even in a sharper form.

In this connection, let us emphasize that rates for the normal approximation for F that are better than $1/n$ cannot be obtained under the support assumption as above.

Proposition 2.6 *For any random vector X in $\mathbb{R}^n$ such that $|X|^2 = n$ a.s., we have*

$$\mathbb{E}_\theta\, \rho(F, \Phi) \ \ge\ \frac{c}{n}.$$

Proof One may apply the following lower bound

$$\rho(F, \Phi) \ \ge\ \frac{1}{3T}\, \Big| \int_0^T (f(t) - e^{-t^2/2}) \Big(1 - \frac{t}{T}\Big)\, dt \Big|, \tag{2.11}$$

which holds for any $T > 0$ (cf. [3]). Since $|X|^2 = n$ a.s., we have $f(t) = J_n(t\sqrt{n})$. Choosing $T = 1$ and applying (2.7), it follows from (2.11) that $\rho(F, \Phi) \ge \frac{c}{n}$ for all $n \ge n_0$ where n_0 is determined by c only. But, a similar bound also holds for $n < n_0$ since F is supported on the interval $[-\sqrt{n}, \sqrt{n}]$. □

3 Upper Bound for the L^2-Distance at Standard Rate

Like in the problem of normal approximation for the typical distribution function $F = \mathbb{E}_\theta F_\theta$, the closeness of distribution functions F_θ of the weighted sums $S_\theta = \langle X, \theta \rangle$ ($\theta \in \mathbb{S}^{n-1}$) to F in the metric ω can also be explored in terms of the associated characteristic functions (the Fourier-Stieltjes transforms)

$$f_\theta(t) \ =\ \mathbb{E}\, e^{it\langle X,\theta\rangle} \ =\ \int_{-\infty}^{\infty} e^{it\langle x,\theta\rangle}\, dF_\theta(x), \quad t \in \mathbb{R}. \tag{3.1}$$

Again, let us start with the identity

$$\omega^2(F_\theta, F) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{|f_\theta(t) - f(t)|^2}{t^2}\, dt. \tag{3.2}$$

Here, the mean value of the numerator represents the variance $\mathbb{E}_\theta\, |f_\theta(t)|^2 - |f(t)|^2$ with respect to $\mathfrak{s}_{n-1}$. Moreover, using an independent copy Y of X, we have

$$\mathbb{E}_\theta\, |f_\theta(t)|^2 \ =\ \mathbb{E}_\theta\, \mathbb{E}\, e^{it\langle X-Y,\theta\rangle} \ =\ \mathbb{E} J_n(t|X - Y|). \tag{3.3}$$

Hence, the Plancherel formula (3.2) together with (2.4) yields

$$\mathbb{E}_\theta\, \omega^2(F_\theta, F) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Big(\mathbb{E} J_n(t|X-Y|) - \big(\mathbb{E} J_n(t|X|)\big)^2\Big)\, \frac{dt}{t^2}. \tag{3.4}$$

In this section our aim is to show that the above expression is of order at most $O(1/n)$ provided that the mean $a = \mathbb{E}X$, $m_2 = m_2(X)$ and $\sigma_4^2 = \sigma_4^2(X)$ are of order 1. The next statement contains the upper bound (1.2) as a partial case.

Proposition 3.1 *Given a random vector X in $\mathbb{R}^n$ with $\mathbb{E}X = a$ and $\mathbb{E}\,|X|^2 = n$, we have*

$$\mathbb{E}_\theta\, \omega^2(F_\theta, F) \,\le\, \frac{cA}{n} \tag{3.5}$$

with $A = 1 + |a|^2 + m_2^2 + \sigma_4^2$. A similar inequality continues to hold with the normal distribution function Φ in place of F.

If X is isotropic, then $m_2 = 1$, while $|a| \le 1$ (by Bessel's inequality). Hence, both characteristics m_2 and a may be removed from the parameter A in this case. However, in the general case, it may happen that m_2 and σ_4 are bounded, while $|a|$ is large. The example in Remark 3.2 shows that this parameter can not be removed.

Proof Note that, for any $\eta > 0$,

$$\int_{-\infty}^{\infty} \frac{\min\{1, t^2\eta^2\}}{t^2}\, dt \,=\, 4\eta, \tag{3.6}$$

Hence, in the formula (3.4), the expectation $\mathbb{E} J_n(t|X-Y|)$ can be replaced using the normal approximation (2.6) at the expense of an error not exceeding

$$\frac{c}{n}\, \mathbb{E} \int_{-\infty}^{\infty} \min\Big\{1, \frac{t^2|X-Y|^2}{n}\Big\}\, \frac{dt}{t^2} \,=\, \frac{4c}{n}\, \mathbb{E}\, \frac{|X-Y|}{\sqrt{n}} \,\le\, \frac{8c}{n},$$

where we used that $\mathbb{E}\,|X| \le \sqrt{n}$. Similarly, by (2.6) and (3.6),

$$\begin{aligned}
\int_{-\infty}^{\infty} \Big| \big(\mathbb{E} J_n(t|X|)\big)^2 - \big(\mathbb{E}\, e^{-t^2|X|^2/2n}\big)^2\Big|\, \frac{dt}{t^2} &\le 2\,\mathbb{E} \int_{-\infty}^{\infty} \big|J_n(t|X|) - e^{-t^2|X|^2/2n}\big|\, \frac{dt}{t^2} \\
&\le \frac{2c}{n}\, \mathbb{E} \int_{-\infty}^{\infty} \min\Big\{1, \frac{t^2|X|^2}{n}\Big\}\, \frac{dt}{t^2} \\
&= \frac{8c}{n}\, \mathbb{E}\, \frac{|X|}{\sqrt{n}} \,\le\, \frac{8c}{n}.
\end{aligned}$$

Hence, using these bounds in (3.4), we arrive at the general approximation

$$\mathbb{E}_\theta\, \omega^2(F_\theta, F) \;=\; \frac{1}{2\pi} \int_{-\infty}^{\infty} \Big(\mathbb{E}\, e^{-t^2|X-Y|^2/2n} - \big(\mathbb{E}\, e^{-t^2|X|^2/2n}\big)^2\Big)\, \frac{dt}{t^2} + \frac{C}{n}, \qquad (3.7)$$

where we recall that C denotes a quantity bounded by an absolute constant.

Introduce the random variable

$$\rho^2 = \frac{|X-Y|^2}{2n} \quad (\rho \geq 0).$$

By Jensen's inequality, $\mathbb{E}\, e^{-t^2|X|^2/2n} \geq e^{-t^2/2}$, so that, by (3.7),

$$\mathbb{E}_\theta\, \omega^2(F_\theta, F) \;\leq\; \frac{1}{2\pi}\, \mathbb{E} \int_{-\infty}^{\infty} \frac{e^{-\rho^2 t^2} - e^{-t^2}}{t^2}\, dt + \frac{c}{n}.$$

The above integral is easily evaluated (by differentiating with respect to the variable "ρ^2"), and we arrive at the bound

$$\mathbb{E}_\theta\, \omega^2(F_\theta, F) \leq \frac{1}{\sqrt{\pi}}\, (1 - \mathbb{E}\rho) + \frac{c}{n}. \qquad (3.8)$$

To further simplify, one may apply an elementary inequality $1 - x \;\leq\; \frac{1}{2}\,(1 - x^2) + (1-x^2)^2$ $(x \geq 0)$, which gives

$$\mathbb{E}_\theta\, \omega^2(F_\theta, F) \leq \frac{1}{2\sqrt{\pi}}\, \mathbb{E}\,(1-\rho^2) + \frac{1}{\sqrt{\pi}}\, \mathbb{E}\,(1-\rho^2)^2 + \frac{c}{n}.$$

Since

$$1 - \rho^2 = \frac{n - |X|^2}{2n} + \frac{n - |Y|^2}{2n} + \frac{\langle X, Y\rangle}{n},$$

we have

$$1 - \mathbb{E}\rho^2 = \frac{1}{n}\, \mathbb{E}\,\langle X, Y\rangle = \frac{1}{n}\, |\mathbb{E}X|^2 = \frac{1}{n}\, |a|^2.$$

In addition,

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

which implies

$$\mathbb{E}\,(1-\rho^2)^2 \;\le\; \frac{\mathrm{Var}(|X|^2)}{n^2} + 2\,\frac{\mathbb{E}\,\langle X,Y\rangle^2}{n^2} \;=\; \frac{\sigma_4^2 + 2m_2^2}{n}.$$

Using this estimate in (3.8), the inequality (3.5) follows immediately.

For the second assertion, it remains to apply Corollary 2.2. □

Remark 3.2 Let us illustrate the inequality (3.5) in the example where the random vector X has a normal distribution with a large mean value. Given a standard normal random vector $Z = (Z_1, \dots, Z_{n-1})$ in $\mathbb{R}^{n-1}$ (which we identify with the space of all points in $\mathbb{R}^n$ with zero last coordinate), define

$$X = \alpha Z + \lambda e_n \quad \text{with } 1 \le \lambda \le n^{1/4},\;\; \alpha^2(n-1) + \lambda^2 = n,$$

where $e_n = (0, \dots, 0, 1)$ is the last unit vector in the canonical basis of $\mathbb{R}^n$. Since Z is orthogonal to e_n, so that $|X|^2 = \alpha^2\,|Z|^2 + \lambda^2$, we have $\mathbb{E}\,|X|^2 = n$, and

$$\sigma_4^2 \;=\; \frac{\alpha^4}{n}\,\mathrm{Var}(|Z|^2) \;=\; \frac{2\alpha^4\,(n-1)}{n} \;=\; 2\,\frac{(n-\lambda^2)^2}{n(n-1)} \;<\; 2.$$

Let Z' be an independent copy of Z. Then $Y = \alpha Z' + \lambda e_n$ is an independent copy of X, so that

$$m_2^2 \;=\; \frac{1}{n}\,\mathbb{E}\,\langle X,Y\rangle^2 \;=\; \frac{1}{n}\,(\alpha^4\,(n-1) + \lambda^4) \;<\; 2.$$

Thus, both m_2 and σ_4 are bounded, while the mean $a = \mathbb{E}X = \lambda e_n$ has the Euclidean length $|a| = \lambda \ge 1$. Hence, the inequality (3.5) being stated for the normal distribution function in place of F simplifies to

$$\mathbb{E}_\theta\,\omega^2(F_\theta, \Phi) \;\le\; \frac{c\lambda^2}{n}.$$

Let us show that this bound may be reversed up to an absolute factor (which would imply that $|a|^2$ may not be removed from A). For any unit vector $\theta = (\theta_1, \dots, \theta_n)$, the linear form

$$S_\theta = \langle X, \theta\rangle = \alpha\theta_1 Z_1 + \dots + \alpha\theta_{n-1} Z_{n-1} + \lambda\theta_n$$

has a normal distribution on the line with mean $\mathbb{E}S_\theta = \lambda\theta_n$ and variance $\mathrm{Var}(S_\theta) = \alpha^2(1-\theta_n^2)$. Consider the normal distribution function $\Phi_{\mu,\sigma^2}(x) = \Phi(\frac{x-\mu}{\sigma})$ with parameters $0 \le \mu \le 1$ and $\frac{1}{2} \le \sigma^2 \le 1$ ($\sigma > 0$). If $x \le \frac{\mu}{1+\sigma}$, then $\frac{x-\mu}{\sigma} \le x$, and on the interval with these endpoints the standard normal density $\varphi(y)$ attains minimum at the left endpoint. Hence

$$|\Phi_{\mu,\sigma^2}(x) - \Phi(x)| \, = \, \int_{\frac{x-\mu}{\sigma}}^{x} \varphi(y)\,dy \, \geq \, \Big(x - \frac{x-\mu}{\sigma}\Big)\,\varphi\Big(\frac{x-\mu}{\sigma}\Big),$$

so that

$$\begin{aligned}
\omega^2(\Phi_{\mu,\sigma^2}, \Phi) &\geq \int_{-\infty}^{\frac{\mu}{1+\sigma}} \Big(x - \frac{x-\mu}{\sigma}\Big)^2 \varphi\Big(\frac{x-\mu}{\sigma}\Big)^2 dx \\
&= \frac{\sigma}{2\pi} \int_{-\infty}^{-\frac{\mu}{1+\sigma}} (\mu - (1-\sigma)y)^2\, e^{-y^2/2}\,dy \\
&\geq \frac{\sigma\mu^2}{2\pi} \int_{-\infty}^{-\frac{\mu}{1+\sigma}} e^{-y^2/2}\,dy \, \geq \, c\mu^2.
\end{aligned}$$

In our case, since $\lambda \leq n^{1/4}$ and

$$\alpha^2 \, = \, \frac{n-\lambda^2}{n-1} \, \geq \, \frac{n-\sqrt{n}}{n-1} \, \geq \, 1 - \frac{1}{\sqrt{n}},$$

we have $|\mathbb{E} S_\theta| \leq 1$ and $\mathrm{Var}(S_\theta) \geq \frac{1}{2}$ on the set $\Omega_n = \{\theta \in \mathbb{S}^{n-1} : |\theta_n| < \frac{\log n}{\sqrt{n}}\}$ with n large enough. It follows that

$$\mathbb{E}_\theta\, \omega^2(F_\theta, \Phi) \, \geq \, c\lambda^2\, \mathbb{E}\,\theta_n^2\, 1_{\{\theta\in\Omega_n\}} \, \geq \, \frac{c'\lambda^2}{n}.$$

4 General Approximations for the L^2-Distance with Error of Order at Most $1/n$

We now turn to general representations for the average L^2-distance between F_θ and the typical distribution function F with error of order at most $1/n$.

Proposition 4.1 *Suppose that $\mathbb{E}\,|X| \leq b\sqrt{n}$ for some $b \geq 0$. Then*

$$\mathbb{E}_\theta\, \omega^2(F_\theta, F) = \frac{1}{\sqrt{2\pi}}\,\mathbb{E} R + \frac{Cb}{n^2}, \tag{4.1}$$

where

$$R \, = \, \frac{(|X|^2 + |Y|^2)^{1/2}}{\sqrt{n}} \Big(1 + \frac{1}{4n}\,\frac{|X|^4 + |Y|^4}{(|X|^2 + |Y|^2)^2}\Big) - \frac{|X-Y|}{\sqrt{n}}\Big(1 + \frac{1}{4n}\Big). \tag{4.2}$$

We use the convention that $R = 0$ if $X = Y = 0$. Note that $|R| \leq 3\, \frac{|X|+|Y|}{\sqrt{n}}$, so $\mathbb{E}R \leq 3b$.

Let us give a simpler expression by involving the functional $\sigma_4^2 = \frac{1}{n}\,\mathrm{Var}(|X|^2)$ and assuming that $\mathbb{E}\,|X|^2 = n$. Since

$$\frac{|X|^4 + |Y|^4}{(|X|^2 + |Y|^2)^2} - \frac{1}{2} = \frac{(|X|^2 - |Y|^2)^2}{2\,(|X|^2 + |Y|^2)^2},$$

we may write

$$R = \frac{1}{8n^{3/2}}\,\frac{(|X|^2 - |Y|^2)^2}{(|X|^2 + |Y|^2)^{3/2}} + \frac{(|X|^2 + |Y|^2)^{1/2}}{\sqrt{n}}\Big(1 + \frac{1}{8n}\Big) - \frac{|X - Y|}{\sqrt{n}}\Big(1 + \frac{1}{4n}\Big). \tag{4.3}$$

As we will see, the first term here is actually of order at most σ_4^2/n^2. As a result, we arrive at the relation (1.5).

Proposition 4.2 *If* $\mathbb{E}\,|X|^2 = n$*, then*

$$\mathbb{E}_\theta\, \omega^2(F_\theta, F) = \frac{1}{\sqrt{2\pi}}\,\mathbb{E}R + C\,\frac{1 + \sigma_4^2}{n^2}, \tag{4.4}$$

where

$$R = \frac{(|X|^2 + |Y|^2)^{1/2}}{\sqrt{n}}\Big(1 + \frac{1}{8n}\Big) - \frac{|X - Y|}{\sqrt{n}}\Big(1 + \frac{1}{4n}\Big). \tag{4.5}$$

Proof of Proposition 4.1 Let us return to the Plancherel formula (3.4). To simplify the integrand therein, we apply the inequality (2.7) in Lemma 2.3, by replacing t^4 with t^2 in the remainder term. Using the equality (3.6), the expectation $\mathbb{E}J_n(t|X - Y|)$ in the formula (3.4) can be therefore replaced according to (2.7) at the expense of an error not exceeding

$$\frac{c}{n^2}\,\mathbb{E}\int_{-\infty}^{\infty} \min\Big\{1, \frac{t^2|X - Y|^2}{n}\Big\}\,\frac{dt}{t^2} = \frac{4c}{n^2}\,\mathbb{E}\,\frac{|X - Y|}{\sqrt{n}} \leq \frac{8cb}{n^2}.$$

As for the main term $(1 - \frac{t^4}{4n})\,e^{-t^2/2}$ in (2.7), it is bounded by an absolute constant, which implies that

$$\begin{aligned} J_n\big(t\sqrt{n}\big)J_n\big(s\sqrt{n}\big) &= \Big(1 - \frac{t^4}{4n}\Big)\Big(1 - \frac{s^4}{4n}\Big)\,e^{-(t^2+s^2)/2} + O\big(n^{-2}\min\{1, t^2 + s^2\}\big) \\ &= \Big(1 - \frac{t^4 + s^4}{4n}\Big)\,e^{-(t^2+s^2)/2} + O\big(n^{-2}\min\{1, t^2 + s^2\}\big). \end{aligned}$$

Hence

$$|\mathbb{E} J_n(t|X|)|^2 \,=\, \mathbb{E}\, J_n(t|X|)\, J_n(t|Y|) = \mathbb{E}\Big(1 - \frac{t^4\,(|X|^4 + |Y|^4)}{4n^3}\Big)\, e^{-\frac{t^2\,(|X|^2+|Y|^2)}{2n}} + O\Big(n^{-2} \min\Big\{1, \frac{t^2\,(|X|^2 + |Y|^2)}{n}\Big\}\Big).$$

As before, after integration in (3.4) the latter remainder term will produce a quantity not exceeding a multiple of b/n^2. As a preliminary step, we therefore obtain the representation

$$\mathbb{E}_\theta\, \omega^2(F_\theta, F) = \frac{1}{2\pi}\, I + \frac{Cb}{n^2} \tag{4.6}$$

with

$$I \,=\, \mathbb{E}\int_{-\infty}^{\infty} \Big[\Big(1 - \frac{t^4|X-Y|^4}{4n^3}\Big)\, e^{-\frac{t^2|X-Y|^2}{2n}} - \Big(1 - \frac{t^4\,(|X|^4+|Y|^4)}{4n^3}\Big)\, e^{-\frac{t^2\,(|X|^2+|Y|^2)}{2n}}\Big]\, \frac{dt}{t^2}.$$

To evaluate the integrals of this type, consider the functions

$$\psi_r(\alpha) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \Big((1 - rt^4)\, e^{-\alpha t^2/2} - e^{-t^2/2}\Big)\, \frac{dt}{t^2} \qquad (\alpha > 0,\ r \in \mathbb{R}).$$

Clearly,

$$\psi_r(1) = -\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} rt^2\, e^{-t^2/2}\, dt = -r$$

and

$$\psi_r'(\alpha) = -\frac{1}{2\sqrt{2\pi}} \int_{-\infty}^{\infty} (1 - rt^4)\, e^{-\alpha t^2/2}\, dt$$
$$= -\frac{1}{2\sqrt{\alpha}}\, \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \Big(1 - \frac{r}{\alpha^2}\, s^4\Big)\, e^{-s^2/2}\, ds \,=\, -\frac{1}{2\sqrt{\alpha}}\, \Big(1 - \frac{3r}{\alpha^2}\Big).$$

Hence

$$\psi_r(\alpha) - \psi_r(1) = \int_1^{\alpha} \Big(-\frac{1}{2}\, z^{-1/2} + \frac{3r}{2}\, z^{-5/2}\Big)\, dz$$
$$= (1 + r) - (\alpha^{1/2} + r\alpha^{-3/2}),$$

and we get

$$\psi_r(\alpha) = 1 - (\alpha^{1/2} + r\alpha^{-3/2}). \tag{4.7}$$

Here, when α and r both approach zero subject to the relation $r = O(\alpha^2)$, we get in the limit $\psi_0(0) = 1$. From this,

$$\frac{1}{\sqrt{2\pi}}\, I = \mathbb{E}\,(\psi_{r_1}(\alpha_1) - \psi_{r_2}(\alpha_2))$$
$$= \mathbb{E}\,(\alpha_2^{1/2} + r_2\alpha_2^{-3/2}) - \mathbb{E}\,(\alpha_1^{1/2} + r_1\alpha_1^{-3/2}),$$

which we need with

$$\alpha_1 = \frac{|X-Y|^2}{n}, \quad r_1 = \frac{|X-Y|^4}{4n^3},$$
$$\alpha_2 = \frac{|X|^2+|Y|^2}{n}, \quad r_2 = \frac{|X|^4+|Y|^4}{4n^3}.$$

It follows that

$$\alpha_2^{1/2} + r_2\alpha_2^{-3/2} = \Big(\frac{|X|^2+|Y|^2}{n}\Big)^{1/2}\Big(1 + \frac{1}{4n}\,\frac{|X|^4+|Y|^4}{(|X|^2+|Y|^2)^2}\Big),$$
$$\alpha_1^{1/2} + r_1\alpha_1^{-3/2} = \Big(\frac{|X-Y|^2}{n}\Big)^{1/2}\Big(1 + \frac{1}{4n}\Big)$$

with the assumption that both expressions are equal to zero in the case $X = Y = 0$. As a result, (4.6) yields the desired representation (4.1) with quantity R described in (4.2). □

In order to modify (4.1) and (4.2) to the form (4.4) and (4.5), first let us verify the following general relation.

Lemma 4.3 *Let ξ be a non-negative random variable with finite second moment (not identically zero), and let η be its independent copy. Then*

$$\mathbb{E}\,\frac{(\xi-\eta)^2}{(\xi+\eta)^{3/2}}\,1_{\{\xi+\eta>0\}} \,\le\, 12\,\frac{\mathrm{Var}(\xi)}{(\mathbb{E}\,\xi)^{3/2}}.$$

Applying the lemma with $\xi = |X|^2$, $\eta = |Y|^2$ and assuming that $\mathbb{E}\,|X|^2 = n$, we get that

$$\mathbb{E}\,\frac{(|X|^2-|Y|^2)^2}{(|X|^2+|Y|^2)^{3/2}} \,\le\, 12\,\frac{\mathrm{Var}(|X|^2)}{(\mathbb{E}\,|X|^2)^{3/2}} \,=\, 12\,\frac{\mathrm{Var}(|X|^2)}{n^{3/2}} \,=\, 12\,\frac{\sigma_4^2}{n^{1/2}}.$$

In view of (4.3), this proves Proposition 4.2.

Proof of Lemma 4.3 By homogeneity, we may assume that $\mathbb{E}\xi = 1$. In particular, $\mathbb{E}\,|\xi - \eta| \le 2$. We have

$$\mathbb{E}\,\frac{(\xi-\eta)^2}{(\xi+\eta)^{3/2}}\,1_{\{\xi+\eta>1/2\}} \le 2^{3/2}\,\mathbb{E}\,(\xi-\eta)^2\,1_{\{\xi+\eta>1/2\}}$$
$$\le 2^{3/2}\,\mathbb{E}\,(\xi-\eta)^2 \;=\; 4\sqrt{2}\,\mathrm{Var}(\xi).$$

Also note that, by Chebyshev's inequality,

$$\mathbb{P}\,\{\xi \le 1/2\} \;=\; \mathbb{P}\,\{1-\xi \ge 1/2\} \;\le\; 4\,\mathrm{Var}(\xi)^2,$$

so

$$\mathbb{P}\,\{\xi+\eta \le 1/2\} \;\le\; \mathbb{P}\,\{\xi \le 1/2\}\,\mathbb{P}\,\{\eta \le 1/2\} \;\le\; 16\,\mathrm{Var}(\xi)^2.$$

Hence, since $\frac{|\xi-\eta|}{\xi+\eta} \le 1$ for $\xi+\eta > 0$, we have, by Cauchy's inequality,

$$\mathbb{E}\,\frac{(\xi-\eta)^2}{(\xi+\eta)^{3/2}}\,1_{\{0<\xi+\eta\le 1/2\}} \le \mathbb{E}\,\sqrt{|\xi-\eta|}\,1_{\{0<\xi+\eta\le 1/2\}}$$
$$\le \sqrt{\mathbb{E}\,|\xi-\eta|}\,\sqrt{\mathbb{P}\,\{\xi+\eta \le 1/2\}} \;\le\; 4\sqrt{2}\,\mathrm{Var}(\xi).$$

It remains to combine both inequalities, which yield

$$\mathbb{E}\,\frac{(\xi-\eta)^2}{(\xi+\eta)^{3/2}}\,1_{\{\xi+\eta>0\}} \;\le\; 8\sqrt{2}\,\mathrm{Var}(\xi) \;\le\; 12\,\mathrm{Var}(\xi).$$

□

5 Proof of Theorem 1.1 for the L^2-Distance

The expression (4.5) may be further simplified in the particular case where the distribution of X is supported on the sphere $\sqrt{n}\,\mathbb{S}^{n-1}$. Introduce the random variable

$$\xi = \frac{\langle X, Y\rangle}{n},$$

where Y is an independent copy of X. Since $|X-Y|^2 = 2n\,(1-\xi)$, Proposition 4.2 yields:

Corollary 5.1 *If* $|X|^2 = n$ *a.s., then*

$$\sqrt{\pi}\,\mathbb{E}_\theta\,\omega^2(F_\theta, F) \;=\; \Big(1+\frac{1}{4n}\Big)\,\mathbb{E}\,\Big(1-(1-\xi)^{1/2}\Big) - \frac{1}{8n} + O\Big(\frac{1}{n^2}\Big). \tag{5.1}$$

Note that $|\xi| \leq 1$. Therefore, the relation (5.1) suggests to develop an expansion in powers of ε for the function $w(\varepsilon) = 1 - \sqrt{1-\varepsilon}$ near zero, which will be needed up to the term ε^4.

Lemma 5.2 *For all* $|\varepsilon| \leq 1$,

$$1 - \sqrt{1-\varepsilon} \ \leq \ \frac{1}{2}\,\varepsilon + \frac{1}{8}\,\varepsilon^2 + \frac{1}{16}\,\varepsilon^3 + 3\varepsilon^4.$$

In addition,

$$1 - \sqrt{1-\varepsilon} \ \geq \ \frac{1}{2}\,\varepsilon + \frac{1}{8}\,\varepsilon^2 + \frac{1}{16}\,\varepsilon^3 + 0.01\,\varepsilon^4.$$

Proof By Taylor's formula for the function $w(\varepsilon)$ around zero on the half-axis $\varepsilon < 1$,

$$1 - \sqrt{1-\varepsilon} \ = \ \frac{1}{2}\,\varepsilon + \frac{1}{8}\,\varepsilon^2 + \frac{1}{16}\,\varepsilon^3 + \frac{5}{128}\,\varepsilon^4 + \frac{w^{(5)}(\varepsilon_1)}{120}\,\varepsilon^5$$

for some ε_1 between zero and ε. Since $w^{(5)}(\varepsilon) = \frac{105}{32}\,(1-\varepsilon)^{-9/2} \geq 0$, we have an upper bound

$$1 - \sqrt{1-\varepsilon} \leq \frac{1}{2}\,\varepsilon + \frac{1}{8}\,\varepsilon^2 + \frac{1}{16}\,\varepsilon^3 + \frac{5}{128}\,\varepsilon^4, \qquad \varepsilon \leq 0.$$

Also, $w^{(5)}(\varepsilon) \leq \frac{105}{32}\,3^{9/2} < 461$ for $0 \leq \varepsilon \leq \frac{2}{3}$, so, in this interval

$$\frac{5}{128}\,\varepsilon^4 + \frac{w^{(5)}(\varepsilon_1)}{120}\,\varepsilon^5 \leq 3\varepsilon^4.$$

Thus, in both cases,

$$1 - \sqrt{1-\varepsilon} \ \leq \ \frac{1}{2}\,\varepsilon + \frac{1}{8}\,\varepsilon^2 + \frac{1}{16}\,\varepsilon^3 + 3\varepsilon^4, \qquad \varepsilon \leq \frac{2}{3}.$$

To treat the remaining values $\frac{2}{3} \leq \varepsilon \leq 1$, it is sufficient to select a positive constant b such that the polynomial

$$Q(\varepsilon) = \frac{1}{2}\,\varepsilon + \frac{1}{8}\,\varepsilon^2 + \frac{1}{16}\,\varepsilon^3 + b\varepsilon^4$$

is greater than or equal to 1 for $\varepsilon \geq \frac{2}{3}$. On this half-axis, $Q(\varepsilon) \geq \frac{11}{27} + b\,\frac{16}{81} \geq 1$ for $b \geq 3$. Thus, the upper bound of the lemma is proved.

Now, from Taylor's formula we also get that

$$1 - \sqrt{1-\varepsilon} \geq \frac{1}{2}\,\varepsilon + \frac{1}{8}\,\varepsilon^2 + \frac{1}{16}\,\varepsilon^3 + \frac{5}{128}\,\varepsilon^4, \qquad \varepsilon \geq 0.$$

In addition, if $-1 \le \varepsilon \le 0$, then $w^{(5)}(\varepsilon) \le \frac{105}{32}$, so

$$\begin{aligned} 1 - \sqrt{1-\varepsilon} &= \frac{1}{2}\,\varepsilon + \frac{1}{8}\,\varepsilon^2 + \frac{1}{16}\,\varepsilon^3 + \frac{5}{128}\,\varepsilon^4 \left(1 + \frac{w^{(5)}(\varepsilon_1)}{120}\,\varepsilon\right) \\ &\ge \frac{1}{2}\,\varepsilon + \frac{1}{8}\,\varepsilon^2 + \frac{1}{16}\,\varepsilon^3 + \varepsilon^4 \left(\frac{5}{128} - \frac{\frac{105}{32}}{120}\right) \\ &\ge \frac{1}{2}\,\varepsilon + \frac{1}{8}\,\varepsilon^2 + \frac{1}{16}\,\varepsilon^3 + 0.01\,\varepsilon^4. \end{aligned}$$

□

Proof of Theorem 1.1 *(First Part)* Using Lemma 5.2 with $\varepsilon = \xi$ and applying Corollary 5.1, we get an asymptotic representation

$$\sqrt{\pi}\,\mathbb{E}_\theta\,\omega^2(F_\theta, F) \,=\, \Big(1 + \frac{1}{4n}\Big)\Big(\frac{1}{8}\,\mathbb{E}\xi^2 + \frac{1}{16}\,\mathbb{E}\xi^3 + c\,\mathbb{E}\xi^4\Big) - \frac{1}{8n} + O\Big(\frac{1}{n^2}\Big)$$

for some quantity c such that $0.01 \le c \le 3$. If additionally X is isotropic, then $\mathbb{E}\,\langle X, Y\rangle^2 = n$, i.e. $\mathbb{E}\xi^2 = \frac{1}{n}$, and the representation is simplified to

$$\sqrt{\pi}\,\mathbb{E}_\theta\,\omega^2(F_\theta, F) \,=\, \Big(1 + \frac{1}{4n}\Big)\Big(\frac{1}{16}\,\mathbb{E}\xi^3 + c\,\mathbb{E}\xi^4\Big) + O\Big(\frac{1}{n^2}\Big),$$

thus removing the term of order $1/n$. Moreover, since $\mathbb{E}\xi^4 \le \mathbb{E}\,|\xi|^3 \le \mathbb{E}\xi^2 = \frac{1}{n}$, the fraction $\frac{1}{4n}$ may be removed from the brackets at the expense of the remainder term. Thus

$$\sqrt{\pi}\,\mathbb{E}_\theta\,\omega^2(F_\theta, F) \,=\, \frac{1}{16}\,\mathbb{E}\xi^3 + c\,\mathbb{E}\xi^4 + O\Big(\frac{1}{n^2}\Big),$$

which is exactly the expansion (1.3). □

Remark 5.3 In the isotropic case with $|X|^2 = n$ a.s., but without the mean zero assumption, the above expansion takes the form

$$\sqrt{\pi}\,\mathbb{E}_\theta\,\omega^2(F_\theta, F) \,=\, \frac{1}{2}\,\mathbb{E}\xi + \frac{1}{16}\,\mathbb{E}\xi^3 + c\,\mathbb{E}\xi^4 + O\Big(\frac{1}{n^2}\Big). \tag{5.2}$$

Since the last two expectations are non-negative, this implies in particular that

$$\mathbb{E}_\theta\,\omega^2(F_\theta, F) \,\ge\, \frac{1}{2\sqrt{\pi}}\,\mathbb{E}\xi + O\Big(\frac{1}{n^2}\Big). \tag{5.3}$$

6 General Lower Bounds for the L^2-Distance: Proof of Theorem 1.2

Proposition 4.1 may be used to establish the following general lower bound which will be the first step in the proof of Theorem 1.2. Recall that Y denotes an independent copy of a random vector X in $\mathbb{R}^n$.

Proposition 6.1 *If* $\mathbb{E}\,|X| \le b\sqrt{n}$, *then*

$$\mathbb{E}_\theta\, \omega^2(F_\theta, F) \ \ge\ c_1\, \mathbb{E}\, \rho\, \xi^4 - c_2 \frac{b}{n^2}, \tag{6.1}$$

where

$$\rho = \Big(\frac{|X|^2 + |Y|^2}{2n}\Big)^{1/2}, \qquad \xi = \frac{2\,\langle X, Y\rangle}{|X|^2 + |Y|^2}.$$

The argument employs two elementary lemmas.

Lemma 6.2 *If* $\mathbb{E}\,|X|^2$ *is finite, then*

$$\mathbb{E}\,\langle X, Y\rangle^2 \ \ge\ \frac{1}{n}\,\big(\mathbb{E}\,|X|^2\big)^2. \tag{6.2}$$

By the invariance of (6.2) under linear orthogonal transformations, we may assume that $\mathbb{E} X_i X_j = \lambda_i \delta_{ij}$ where λ_i's appear as eigenvalues of the covariance operator of X. Since

$$\mathbb{E}\,|X|^2 = \sum_{i=1}^n \lambda_i, \qquad \mathbb{E}\,\langle X, Y\rangle^2 = \sum_{i=1}^n \lambda_i^2,$$

the inequality (6.2) follows by applying Cauchy's inequality.

Lemma 6.3 *If* $\mathbb{E}\,|X|^p$ *is finite for an integer* $p \ge 1$, *then, for any real number* $0 \le \alpha \le p$,

$$\mathbb{E}\,\frac{\langle X, Y\rangle^p}{(|X|^2 + |Y|^2)^\alpha} \ge 0,$$

where the ratio is defined to be zero in case $X = Y = 0$. *In addition, for* $\alpha \in [0, 2]$,

$$\mathbb{E}\,\frac{\langle X, Y\rangle^2}{(|X|^2 + |Y|^2)^\alpha} \ \ge\ \frac{1}{n}\,\mathbb{E}\,\frac{|X|^2\,|Y|^2}{(|X|^2 + |Y|^2)^\alpha}.$$

Proof First, let us note that

$$\mathbb{E}\,\frac{|\,\langle X,Y\rangle\,|^p}{(|X|^2+|Y|^2)^\alpha} \,\le\, \mathbb{E}\,\frac{(|X|\,|Y|)^p}{(|X|\,|Y|)^\alpha} \,=\, (\mathbb{E}\,|X|^{p-\alpha})^2,$$

so, the expectation on the left is finite. Without loss of generality, we may assume that $0<\alpha\le p$ and $r=|X|^2+|Y|^2>0$ with probability 1. We use the identity

$$\int_0^\infty e^{-rt^{1/\alpha}}\,dt \,=\, c_\alpha\, r^{-\alpha} \quad \text{where } c_\alpha = \int_0^\infty e^{-s^{1/\alpha}}\,ds,$$

which gives

$$c_\alpha\,\mathbb{E}\,\langle X,Y\rangle^p\, r^{-\alpha} \,=\, \int_0^\infty \mathbb{E}\,\langle X,Y\rangle^p\ e^{-rt^{1/\alpha}}\,dt.$$

Writing $X=(X_1,\dots,X_n)$ and $Y=(Y_1,\dots,Y_n)$, we have

$$\begin{aligned}\mathbb{E}\,\langle X,Y\rangle^p\ e^{-rt^{1/\alpha}} &= \mathbb{E}\,\langle X,Y\rangle^p\ e^{-t^{1/\alpha}(|X|^2+|Y|^2)}\\ &= \sum_{i_1,\dots,i_p=1}^n \Big(\mathbb{E}\,X_{i_1}\dots X_{i_p}\ e^{-t^{1/\alpha}\,|X|^2}\Big)^2,\end{aligned}$$

which shows that the left expectation is always non-negative. Integrating over $t>0$, this proves the first assertion.

For the second assertion, write

$$c_\alpha\,\mathbb{E}\,\langle X,Y\rangle^2\, r^{-\alpha} \,=\, \int_0^\infty \mathbb{E}\,\langle X,Y\rangle^2\ e^{-t^{1/\alpha}(|X|^2+|Y|^2)}\,dt \,=\, \int_0^\infty \mathbb{E}\,\langle X_t,Y_t\rangle^2\ dt,$$

where

$$X_t = e^{-t^{1/\alpha}|X|^2/2}\,X, \qquad Y_t = e^{-t^{1/\alpha}|Y|^2/2}\,Y.$$

Since Y_t represents an independent copy of X_t, one may apply Lemma 6.2 which gives

$$\mathbb{E}\,\langle X_t,Y_t\rangle^2 \,\ge\, \frac{1}{n}\,\mathbb{E}\,|X_t|^2\,|Y_t|^2.$$

Hence,

$$\begin{aligned}\int_0^\infty \mathbb{E}\,\langle X_t,Y_t\rangle^2\ dt &\ge \frac{1}{n}\int_0^\infty \mathbb{E}|X_t|^2\,|Y_t|^2\,dt\\ &= \frac{1}{n}\int_0^\infty \mathbb{E}|X|^2\,|Y|^2\,e^{-t^{1/\alpha}(|X|^2+|Y|^2)}\,dt = \frac{c_\alpha}{n}\,\mathbb{E}\,|X|^2\,|Y|^2\,r^{-\alpha}.\end{aligned}$$

□

Proof of Proposition 6.1 Let us return to the representation (4.3) in Proposition 4.1 and write

$$\mathbb{E}_\theta\, \omega^2(F_\theta, F) = \frac{1}{\sqrt{2\pi}}\, \mathbb{E}\,(R_0 + R_1) + \frac{Cb}{n^2},$$

where

$$R_0 \;=\; \frac{1}{8n^{3/2}}\, \frac{(|X|^2 - |Y|^2)^2}{(|X|^2 + |Y|^2)^{3/2}}$$

and

$$\begin{aligned} R_1 &= \frac{(|X|^2 + |Y|^2)^{1/2}}{\sqrt{n}} \Big(1 + \frac{1}{8n}\Big) - \frac{|X - Y|}{\sqrt{n}} \Big(1 + \frac{1}{4n}\Big) \\ &= \frac{(|X|^2 + |Y|^2)^{1/2}}{\sqrt{n}} \bigg[\Big(1 + \frac{1}{4n}\Big)\big(1 - \sqrt{1 - \xi}\,\big) - \frac{1}{8n}\bigg] \end{aligned}$$

with the assumption that $R_0 = 0$ when $X = Y = 0$. Since $|\xi| \le 1$, one may apply Lemma 5.2 which gives

$$R_1 \;\ge\; \frac{(|X|^2 + |Y|^2)^{1/2}}{\sqrt{n}} \bigg[\Big(1 + \frac{1}{4n}\Big)\Big(\frac{1}{2}\,\xi + \frac{1}{8}\,\xi^2 + \frac{1}{16}\,\xi^3 + 0.01\,\xi^4\Big) - \frac{1}{8n}\bigg].$$

The expectation of the terms on the right-hand side containing ξ and ξ^3 is non-negative according to Lemma 6.3 with $\alpha = \frac{1}{2}$, $p = 1$, and with $\alpha = \frac{5}{2}$, $p = 3$, respectively. Hence, removing the unnecessary factor $1 + \frac{1}{4n}$, we get

$$\begin{aligned} \mathbb{E}_\theta\, \omega^2(F_\theta, F) \ge \frac{1}{\sqrt{2\pi}}\, \mathbb{E} R_0 + \frac{1}{\sqrt{2\pi}}\, \mathbb{E}\, \frac{(|X|^2 + |Y|^2)^{1/2}}{8\sqrt{n}} \Big(\xi^2 - \frac{1}{n}\Big) \\ + c_1\, \mathbb{E}\, \frac{(|X|^2 + |Y|^2)^{1/2}}{\sqrt{n}}\, \xi^4 - c_2 \frac{b}{n^2}. \end{aligned} \tag{6.3}$$

Now, by the second inequality of Lemma 6.3 applied with $\alpha = 3/2$, $p = 2$, we have

$$\begin{aligned} \mathbb{E}\,(|X|^2 + |Y|^2)^{1/2}\, \xi^2 &= 4\, \mathbb{E}\, \frac{\langle X, Y\rangle^2}{(|X|^2 + |Y|^2)^{3/2}} \\ &\ge \frac{4}{n}\, \mathbb{E}\, \frac{|X|^2\, |Y|^2}{(|X|^2 + |Y|^2)^{3/2}}. \end{aligned}$$

This gives

$$\mathbb{E}\,\frac{(|X|^2+|Y|^2)^{1/2}}{8\sqrt{n}}\,\Big(\xi^2-\frac{1}{n}\Big) \geq \frac{1}{8n^{3/2}}\,\mathbb{E}\Big[\frac{4\,|X|^2\,|Y|^2}{(|X|^2+|Y|^2)^{3/2}}-(|X|^2+|Y|^2)^{1/2}\Big]$$

$$= -\frac{1}{8n^{3/2}}\,\mathbb{E}\,\frac{(|X|^2-|Y|^2)^2}{(|X|^2+|Y|^2)^{3/2}} \;=\; -\mathbb{E}R_0.$$

Thus, the summand $\mathbb{E}R_0$ in (6.3) neutralizes the second expectation, and we are left with the term containing ξ^4. □

Proof of Theorem 1.2 We apply Proposition 6.1. By the assumption, $\mathbb{E}\rho^2=1$ and $\mathrm{Var}(\rho^2)=\frac{1}{2n}\,\sigma_4^2$, where $\sigma_4^2=\frac{1}{n}\,\mathrm{Var}(|X|^2)$. Using

$$2\,\langle X,Y\rangle = |X|^2+|Y|^2-|X-Y|^2, \qquad \xi = 1-\frac{|X-Y|^2}{|X|^2+|Y|^2},$$

we have

$$\xi^4 \geq (1-\alpha)^4\,1_{\{|X-Y|^2\,\leq\,\alpha\,(|X|^2+|Y|^2)\}}$$

$$\geq (1-\alpha)^4\,1_{\{|X-Y|^2\,\leq\,\alpha\lambda n,\ |X|^2+|Y|^2\,\geq\,\lambda n\}}, \quad 0<\alpha,\lambda<1.$$

On the set $|X|^2+|Y|^2\geq \lambda n$, we necessarily have $\rho^2\geq\frac{\lambda}{2}$, so

$$\mathbb{E}\,\rho\,\xi^4 \geq \frac{(1-\alpha)^4}{\sqrt{2}}\,\sqrt{\lambda}\,\mathbb{P}\Big\{|X-Y|^2\leq\alpha\lambda n,\ |X|^2+|Y|^2\geq\lambda n\Big\}$$

$$\geq \frac{(1-\alpha)^4}{\sqrt{2}}\,\sqrt{\lambda}\,\Big(\mathbb{P}\{|X-Y|^2\leq\alpha\lambda n\}-\mathbb{P}\{|X|^2+|Y|^2\leq\lambda n\}\Big).$$

But, by Chebyshev's inequality

$$\mathbb{P}\big\{|X|^2\leq\lambda n\big\} \;=\; \mathbb{P}\big\{n-|X|^2\geq(1-\lambda)\,n\big\} \;\leq\; \frac{\mathrm{Var}(|X|^2)}{(1-\lambda)^2\,n^2} \;=\; \frac{\sigma_4^2}{(1-\lambda)^2\,n},$$

implying

$$\mathbb{P}\big\{|X|^2+|Y|^2\leq\lambda n\big\} \;\leq\; \Big(\mathbb{P}\big\{|X|^2\leq\lambda n\big\}\Big)^2 \;\leq\; \frac{1}{(1-\lambda)^4}\,\frac{\sigma_4^4}{n^2}.$$

Hence

$$\mathbb{E}\,\rho\,\xi^4 \;\geq\; \frac{(1-\alpha)^4}{\sqrt{2}}\,\sqrt{\lambda}\,\Big(\mathbb{P}\{|X-Y|^2\leq\alpha\lambda n\}-\frac{1}{(1-\lambda)^4}\,\frac{\sigma_4^4}{n^2}\Big).$$

Choosing, for example, $\alpha = \lambda = \frac{1}{2}$, we get

$$\mathbb{E}\,\rho\,\xi^4 \ \geq \ \frac{1}{32}\,\mathbb{P}\Big\{|X-Y|^2 \leq \frac{1}{4}\,n\Big\} - \frac{\sigma_4^4}{2n^2}.$$

It remains to apply (6.1) with $b = 1$ and replace F with Φ on the basis of (2.3). □

7 Lipschitz Systems

While upper bounds of order $n^{-1/2}$ for the L^2-distance $\omega(F_\theta, F)$ on average are provided in (1.2) and in the more general inequality (3.5) of Proposition 3.1, in this section we focus on the conditions that provide similar lower bounds, as a consequence of Theorem 1.2.

Let L be a fixed measurable function on the underlying probability space $(\Omega, \mathfrak{F}, \mathbb{P})$. We will say that the system $X_1, \dots, X_n$ of random variables on $(\Omega, \mathfrak{F}, \mathbb{P})$, or the random vector $X = (X_1, \dots, X_n)$ in $\mathbb{R}^n$ satisfies a Lipschitz condition with a parameter function L, if

$$\max_{1 \leq k \leq n} |X_k(t) - X_k(s)| \leq n\,|L(t) - L(s)|, \qquad t, s \in \Omega. \tag{7.1}$$

When Ω is an interval of the real line (finite or not), and $L(t) = Lt$, $L > 0$, this condition means that every function X_k in the system has a Lipschitz semi-norm at most Ln.

As before, we use the variance functional $\sigma_4^2 = \frac{1}{n}\,\mathrm{Var}(|X|^2)$.

Proposition 7.1 *Suppose that* $\mathbb{E}\,|X|^2 = n$. *If the random vector* X *satisfies the Lipschitz condition with a parameter function* L, *then*

$$\mathbb{E}_\theta\,\omega^2(F_\theta, F) \ \geq \ \frac{c_L}{n} - \frac{c_0\,(1+\sigma_4^4)}{n^2} \tag{7.2}$$

with some absolute constant $c_0 > 0$ *and with a constant* c_L *depending on the distribution of* L *only. Moreover, if* L *has finite second moment, then with some absolute constant* $c_1 > 0$

$$\mathbb{E}_\theta\,\omega^2(F_\theta, F) \ \geq \ \frac{c_1}{n\sqrt{\mathrm{Var}(L)}} - \frac{c_0\,(1+\sigma_4^4)}{n^2}. \tag{7.3}$$

Note that, if $X_1, \dots, X_n$ form an orthonormal system in $L^2(\Omega, \mathfrak{F}, \mathbb{P})$, i.e., the random vector X is isotropic, and if L has finite second moment $\|L\|_2^2 = \mathbb{E}L^2$, then this moment has to be bounded from below by a multiple of $1/n^2$. Indeed, the projection of the function $\eta(t) = 1$ in $L^2(\Omega, \mathfrak{F}, \mathbb{P})$ to the linear hull H of

$X_1, \dots, X_n$ has the form $\mathrm{Proj}_H(\eta) = \sum_{k=1}^n \langle \eta, X_k \rangle X_k$, and we have Bessel's inequality

$$1 = \|\eta\|_2^2 \geq \|\mathrm{Proj}_H(\eta)\|_2^2 = \sum_{k=1}^n \langle \eta, X_k \rangle^2 = \sum_{k=1}^n (\mathbb{E}X_k)^2$$

(where we used the canonical innde product $\langle \cdot, \cdot \rangle$ in $L^2(\Omega, \mathfrak{F}, \mathbb{P})$). By the Lipschitz assumption, $|X_k(t) - X_k(s)|^2 \leq n^2 |L(t) - L(s)|^2$. Integrating this inequality over the product measure $\mathbb{P}(dt) \otimes \mathbb{P}(ds)$, we obtain a lower bound

$$n^2 \, \mathrm{Var}(L) \geq \mathrm{Var}(X_k) = 1 - (\mathbb{E}X_k)^2.$$

One may now perform summation over $k = 1, \dots, n$, which together with Bessel's inequality leads to

$$\mathrm{Var}(L) \geq \frac{n-1}{n^3} \geq \frac{1}{2n^2} \quad (n \geq 2).$$

The Lipschitz condition (7.1) guarantees the validity of the following property, which can be combined with Theorem 1.2 to obtain (7.2) and (7.3).

Lemma 7.2 *Suppose that the random vector $X = (X_1, \dots, X_n)$ satisfies the Lipschitz condition with the parameter function L. If Y is an independent copy of X, then*

$$\mathbb{P}\big\{|X - Y|^2 \leq \lambda n\big\} \geq \frac{c\sqrt{\lambda}}{n}, \qquad 0 \leq \lambda \leq 1,$$

where the constant $c > 0$ depends on the distribution of L only. Moreover, if L has finite second moment, then

$$\mathbb{P}\big\{|X - Y|^2 \leq \lambda n\big\} \geq \frac{\sqrt{\lambda}}{6n\sqrt{\mathrm{Var}(L)}}, \qquad 0 \leq \lambda \leq n^2 \, \mathrm{Var}(L).$$

In turn, this lemma is based on the following general observation.

Lemma 7.3 *If η is an independent copy of a random variable ξ, then for any $\varepsilon_0 > 0$,*

$$\mathbb{P}\{|\xi - \eta| \leq \varepsilon\} \geq c\varepsilon, \qquad 0 \leq \varepsilon \leq \varepsilon_0,$$

with some constant $c > 0$ independent of ε. Moreover, if the standard deviation $\sigma = \sqrt{\mathrm{Var}(\xi)}$ is finite, then

$$\mathbb{P}\{|\xi - \eta| \leq \varepsilon\} \geq \frac{1}{6\sigma}\,\varepsilon, \qquad 0 \leq \varepsilon \leq \sigma.$$

Proof The difference $\xi - \eta$ has a non-negative characteristic function $h(t) = |\psi(t)|^2$, where ψ is the characteristic function of ξ. Denoting by H the distribution function of $\xi - \eta$, we start with a general identity

$$\int_{-\infty}^{\infty} \hat{p}(x)\, dH(x) = \int_{-\infty}^{\infty} p(t)h(t)\, dt, \tag{7.4}$$

which is valid for any integrable function $p(t)$ on the real line with Fourier transform $\hat{p}(x) = \int_{-\infty}^{\infty} e^{itx}\, p(t)\, dt$, $x \in \mathbb{R}$. Given $\varepsilon > 0$, here we take a standard pair

$$p(t) = \frac{1}{2\pi}\Big(\frac{\sin\frac{\varepsilon t}{2}}{\frac{\varepsilon t}{2}}\Big)^2, \qquad \hat{p}(x) = \frac{1}{\varepsilon}\Big(1 - \frac{|x|}{\varepsilon}\Big)^+,$$

where we use the notation $a^+ = \max\{a, 0\}$. In this case,

$$\int_{-\infty}^{\infty} \hat{p}(x)\, dH(x) \ \le\ \frac{1}{\varepsilon}\int_{[-\varepsilon,\varepsilon]} dH(x) \ =\ \frac{1}{\varepsilon}\,\mathbb{P}\{|\xi - \eta| \le \varepsilon\}.$$

On the other hand, since the function $\frac{\sin u}{u}$ is decreasing in $0 < u < \frac{\pi}{2}$, we have

$$\int_{-\infty}^{\infty} p(t)h(t)\, dt \ \ge\ \frac{1}{2\pi}\,\big(2\sin(1/2)\big)^2 \int_{-1/\varepsilon}^{1/\varepsilon} h(t)\, dt \ \ge\ \frac{1}{7}\int_{-1/\varepsilon}^{1/\varepsilon} h(t)\, dt.$$

Hence, whenever $0 < \varepsilon \le \varepsilon_0$, by (7.4),

$$\mathbb{P}\{|\xi - \eta| \le \varepsilon\} \ \ge\ \frac{\varepsilon}{7}\int_{-1/\varepsilon}^{1/\varepsilon} h(t)\, dt \ \ge\ \frac{\varepsilon}{7}\int_{-1/\varepsilon_0}^{1/\varepsilon_0} h(t)\, dt.$$

Since $h(t)$ is bounded away from zero near the origin, the first assertion follows.

One may quantify this statement in terms of the variance $\sigma^2 = \mathrm{Var}(\xi)$ by using Taylor's expansion for $h(t)$ about zero. Indeed, it gives $1 - h(t) \le \sigma^2 t^2$, and thus for $\varepsilon \le \varepsilon_0 = \sigma$,

$$\int_{-1/\varepsilon}^{1/\varepsilon} h(t)\, dt \ \ge\ \int_{-1/\sigma}^{1/\sigma} (1 - \sigma^2 t^2)\, dt \ =\ \frac{4}{3\sigma}.$$

Since $\frac{\varepsilon}{7} \cdot \frac{4}{3\sigma} \ge \frac{1}{6\sigma}\,\varepsilon$, the lemma is proved. □

Proof of Lemma 7.2 Let us equip the product space $\Omega^2 = \Omega \times \Omega$ with the product measure $\mathbb{P}^2 = \mathbb{P} \otimes \mathbb{P}$ and redefine X on this new probability space as $X(t, s) =$

$X(t)$, $(t,s) \in \Omega^2$. Then one can introduce an independent copy of X in the form $Y(t,s) = X(s)$. By the Lipschitz condition,

$$|X(t,s) - Y(t,s)|^2 \, = \, \sum_{k=1}^n |X_k(t) - X_k(s)|^2 \, \le \, n^3 \, |L(t) - L(s)|^2.$$

Hence, if η is an independent copy of the random variable $\xi = L$, then

$$\mathbb{P}\big\{|X - Y|^2 \le \lambda n\big\} \, \ge \, \mathbb{P}\big\{n^3 \, |\xi - \eta|^2 \le \lambda n\big\} \, = \, \mathbb{P}\bigg\{|\xi - \eta| \le \frac{\sqrt{\lambda}}{n}\bigg\}.$$

But, by Lemma 7.3 with $\varepsilon_0 = 1$, the latter probability is at least $c\,\frac{\sqrt{\lambda}}{n}$, where the constant c depends on L only (via its distribution). An application of the second inequality of Lemma 7.3 yields the second assertion. □

To include more examples, let us now give a bit more general form of Lemma 7.2, assuming that $(\Omega, \mathbb{P}) = (\Omega_1 \times \Omega_2, \mathbb{P}_1 \otimes \mathbb{P}_2)$ is a product probability space.

Lemma 7.4 *Let $X = (X_1, \dots, X_n) : \Omega \rightarrow \mathbb{R}^n$ be a random vector such that, for some measurable functions L_1 and L_2 defined on Ω_1 and Ω_2 respectively,*

$$\max_{1 \le k \le n} |X_k(t_1,t_2) - X_k(s_1,s_2)| \, \le \, n\,|L_1(t_1) - L_1(s_1)| + |L_2(t_2) - L_2(s_2)| \qquad (7.5)$$

for all $(t_1,t_2), (s_1,s_2) \in \Omega$. If Y is an independent copy of X, then

$$\mathbb{P}\big\{|X - Y|^2 \le \lambda n\big\} \ge \frac{c\lambda}{n}, \qquad 0 \le \lambda \le 1, \qquad (7.6)$$

where the constant $c > 0$ depends on the distributions of L_1 and L_2 only.

Proof Again, let us equip the product space $\Omega^2 = \Omega \times \Omega$ with the product measure $\mathbb{P}^2 = \mathbb{P} \otimes \mathbb{P}$ and put $X(t,s) = X(t)$, $Y(t,s) = X(s)$ for $t = (t_1,t_2) \in \Omega$ and $s = (s_1,s_2) \in \Omega$, so that Y is an independent copy of X. By the Lipschitz condition (7.5), for any $k \le n$,

$$|X_k(t) - X_k(s)|^2 \, \le \, 2n^2 \, |L_1(t_1) - L_1(s_1)| + 2\,|L_2(t_2) - L_2(s_2)|^2,$$

so

$$\begin{aligned} |X(t) - Y(s)|^2 &= \sum_{k=1}^n |X_k(t) - X_k(s)|^2 \\ &\le 2n^3 \, |L_1(t_1) - L_1(s_1)|^2 + 2n \, |L_2(t_2) - L_2(s_2)|^2. \end{aligned}$$

Putting $L_1(t_1, t_2) = L_1(t_1)$ and $L_2(t_1, t_2) = L_2(t_2)$, one may treat L_1 and L_2 as independent random variables. If L_1' is an independent copy of L_1 and L_2' is an independent copy of L_2, we obtain that

$$\begin{aligned}\mathbb{P}\{|X - Y|^2 \le \lambda n\} &\ge \mathbb{P}\Big\{n^2 \,|L_1 - L_1'|^2 + |L_2 - L_2'|^2 \le \frac{\lambda}{2}\Big\} \\ &\ge \mathbb{P}\Big\{n^2 \,|L_1 - L_1'|^2 \le \frac{\lambda}{4}\Big\}\, \mathbb{P}\Big\{|L_2 - L_2'|^2 \le \frac{\lambda}{4}\Big\} \\ &= \mathbb{P}\Big\{|L_1 - L_1'| \le \frac{1}{2n}\sqrt{\lambda}\Big\}\, \mathbb{P}\Big\{|L_2 - L_2'| \le \frac{1}{2}\sqrt{\lambda}\Big\}.\end{aligned}$$

It remains to apply Lemma 7.3. □

Let us now combine the inequality (1.8) of Theorem 1.2 with the inequality (7.6) applied with $\lambda = \frac{1}{4}$. Then we obtain the following generalization of Proposition 7.1.

Proposition 7.5 *Under the Lipschitz condition* (7.5)*, we have*

$$\mathbb{E}_\theta\, \omega^2(F_\theta, F) \;\ge\; \frac{c}{n} - \frac{c_0\,(1 + \sigma_4^4)}{n^2},$$

where $c_0 > 0$ is an absolute constant, while $c > 0$ depends on the distributions of L_1 and L_2. A similar estimate also holds when F is replaced with the normal distribution function Φ.

The last assertion follows from the inequality (2.3), cf. Corollary 2.2.

8 Berry-Esseen-Type Bounds

We now turn to the study of the Kolmogorov distance

$$\rho(F_\theta, F) = \sup_x |F_\theta(x) - F(x)|, \quad \theta \in \mathbb{S}^{n-1},$$

between the distribution functions F_θ of the weighted sums $S_\theta = \langle X, \theta\rangle$ and the typical distribution function $F = \mathbb{E}_\theta F_\theta$. We are mostly interested in bounding the second moment $\mathbb{E}_\theta\, \rho^2(F_\theta, F)$. As in the case of the L^2-distance, our basic tool will be a Fourier analytic approach relying upon a general Berry-Esseen-type bound

$$c\,\rho(U, V) \;\le\; \int_0^T \frac{|\hat U(t) - \hat V(t)|}{t}\,dt + \frac{1}{T}\int_0^T |\hat V(t)|\,dt, \qquad T > 0, \tag{8.1}$$

where U and V may be arbitrary distribution functions on the line with characteristic functions $\hat U$ and $\hat V$ respectively (cf. e.g. [3, 23, 24]).

As before, we denote by f_θ and f the characteristic functions associated to F_θ and F. Recall that σ_{2p}-functionals were defined in (2.2).

Lemma 8.1 *If $T \geq T_0 \geq 1$, then for all $p \geq 1$,*

$$c_p\,\mathbb{E}_\theta\,\rho^2(F_\theta, F) \leq \int_0^1 \frac{\mathbb{E}_\theta\,|f_\theta(t) - f(t)|^2}{t^2}\,dt + \log T\,\int_0^{T_0} \frac{\mathbb{E}_\theta\,|f_\theta(t) - f(t)|^2}{t}\,dt$$
$$+\ \log T\,\int_{T_0}^T \frac{\mathbb{E}_\theta\,|f_\theta(t)|^2}{t}\,dt + \frac{1}{T^2} + \frac{1+\sigma_{2p}^{2p}}{n^p}, \tag{8.2}$$

where the constants $c_p > 0$ depend on p only.

Proof By (8.1), for any $\theta \in \mathbb{S}^{n-1}$,

$$c\,\rho(F_\theta, F)\ \leq\ \int_0^T \frac{|f_\theta(t) - f(t)|}{t}\,dt + \frac{1}{T}\int_0^T |f(t)|\,dt,$$

and squaring it, we get

$$c\,\rho^2(F_\theta, F)\ \leq\ \Big(\int_0^T \frac{|f_\theta(t) - f(t)|}{t}\,dt\Big)^2 + \frac{1}{T^2}\Big(\int_0^T |f(t)|\,dt\Big)^2.$$

Let us split integration in the first integral into the intervals $[0, 1]$ and $[1, T]$. By Cauchy's inequality,

$$\Big(\int_0^1 \frac{|f_\theta(t) - f(t)|}{t}\,dt\Big)^2\ \leq\ \int_0^1 \frac{|f_\theta(t) - f(t)|^2}{t^2}\,dt,$$

while

$$\Big(\int_1^T \frac{|f_\theta(t) - f(t)|}{t}\,dt\Big)^2\ \leq\ \log T\,\int_1^T \frac{|f_\theta(t) - f(t)|^2}{t}\,dt.$$

Hence

$$c\,\rho^2(F_\theta, F) \leq \int_0^1 \frac{|f_\theta(t) - f(t)|^2}{t^2}\,dt$$
$$+\ \log T\,\int_1^T \frac{|f_\theta(t) - f(t)|^2}{t}\,dt + \frac{1}{T^2}\Big(\int_0^T |f(t)|\,dt\Big)^2.$$

Without an essential loss one may extend integration in the second integral to the larger interval $[0, T]$. Moreover, taking the expectation over θ, we then get

$$c\,\mathbb{E}_\theta\,\rho^2(F_\theta, F) \le \int_0^1 \frac{\mathbb{E}_\theta\,|f_\theta(t) - f(t)|^2}{t^2}\,dt$$
$$+\ \log T \int_0^T \frac{\mathbb{E}_\theta\,|f_\theta(t) - f(t)|^2}{t}\,dt + \frac{1}{T^2}\Big(\int_0^T |f(t)|\,dt\Big)^2.$$

Again, one may split integration in the second last integral to the two intervals $[0, T_0]$ and $[T_0, T]$, so that to consider separately sufficiently large values of t for which $|f_\theta(t)|$ is small enough (with high probability). More precisely, since $f(t) = \mathbb{E}_\theta\, f_\theta(t)$ and

$$|f_\theta(t) - f(t)|^2 \le 2\,|f_\theta(t)|^2 + 2\,|f(t)|^2,$$

we have $|f(t)|^2 \le \mathbb{E}_\theta\,|f_\theta(t)|^2$ and therefore

$$\mathbb{E}_\theta\,|f_\theta(t) - f(t)|^2 \le 4\,\mathbb{E}_\theta\,|f_\theta(t)|^2.$$

It remains to apply Lemma 2.4. □

In order to control the last integral in (8.2), one may apply the upper bound (2.8) on J_n in the representation (3.3) to get that, for all $t \in \mathbb{R}$,

$$\mathbb{E}_\theta\,|f_\theta(t)|^2 \ \le\ 5\,\mathbb{E}\,e^{-t^2|X-Y|^2/2n} + 4\,e^{-n/12},$$

where Y is an independent copy of the random vector X. Splitting the last expectation to the event $A = \{|X - Y|^2 \le \frac{1}{4}\,n\}$ and its complement leads to

$$\mathbb{E}_\theta\,|f_\theta(t)|^2 \ \le\ 5\,e^{-t^2/8} + 4\,e^{-n/12} + 5\,\mathbb{P}(A). \tag{8.3}$$

The latter probability may further be estimated by using the moment functionals such as m_p.

To recall the argument (cf. also [8], Proposition 2.5), first note that, by (2.9) with $\lambda = \frac{3}{4}$,

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

On the other hand, by Markov's inequality, assuming that $p \ge 1$ is integer, we have

$$\mathbb{P}\Big\{|\,\langle X, Y\rangle\,| \ge \frac{1}{4}\,n\Big\} \ \le\ \frac{4^{2p}\,\mathbb{E}\,\langle X, Y\rangle^{2p}}{n^{2p}} \ =\ \frac{4^{2p}\,m_{2p}^{2p}}{n^p}.$$

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

$$\Big\{|X-Y|^2 \le \frac{1}{4}\Big\} \subset \Big\{|X| + |Y|^2 \le \frac{1}{4}\,n\Big\} \cup \Big\{\,\langle X, Y\rangle > \frac{1}{4}\,n\Big\},$$

and it follows that

$$\mathbb{P}(A) \ \le\ \mathbb{P}\Big\{|X|^2 + |Y|^2 \le \frac{3}{4}\,n\Big\} + \mathbb{P}\Big\{\,\langle X, Y\rangle > \frac{1}{4}\,n\Big\} \ \le\ \frac{4^{2p}}{n^p}\,(m_{2p}^{2p} + \sigma_{2p}^{2p}).$$

Returning to (8.3) and noting that necessarily $m_{2p} \ \ge\ m_2 \ \ge\ 1$ under the assumption that $\mathbb{E}\,|X|^2 = n$, we thus obtain that

$$c_p\,\mathbb{E}_\theta\,|f_\theta(t)|^2 \ \le\ \frac{m_{2p}^{2p} + \sigma_{2p}^{2p}}{n^p} + e^{-t^2/8}.$$

Using this bound, the inequality (8.2) is simplified:

Lemma 8.2 *If the random vector X in $\mathbb{R}^n$ satisfies $\mathbb{E}\,|X|^2 = n$, then for all $T \ge T_0 \ge 1$ and any integer $p \ge 1$,*

$$c_p\,\mathbb{E}_\theta\,\rho^2(F_\theta, F) \le \int_0^1 \frac{\mathbb{E}_\theta\,|f_\theta(t) - f(t)|^2}{t^2}\,dt + \log T\,\int_0^{T_0} \frac{\mathbb{E}_\theta\,|f_\theta(t) - f(t)|^2}{t}\,dt$$
$$+ \frac{m_{2p}^{2p} + \sigma_{2p}^{2p}}{n^p}\,(1 + \log T)^2 + \frac{1}{T^2} + e^{-T_0^2/8}\,\log T \qquad (8.4)$$

with constants c_p depending on p only.

9 Quantitative Forms of Sudakov's Theorem for the Kolmogorov Distance

Let us specialize Lemma 8.2 to the value $p = 1$, assuming that the random vector X is isotropic in $\mathbb{R}^n$ (so that $m_2 = 1$). If σ_2 is bounded, then choosing

$$T = 4n, \quad T_0 = 4\sqrt{\log n},$$

the last three terms in (8.4) produce a quantity of order at most $(\log n)^2/n$. In order to bound the integrals in (8.4), one may apply the classical Poincaré inequality on the unit sphere $\mathbb{S}^{n-1}$

$$\mathbb{E}_\theta|u(\theta)|^2 \ \le\ \frac{1}{n-1}\,\mathbb{E}_\theta\,|\nabla u(\theta)|^2 \qquad (9.1)$$

to the mean zero functions $u_t(\theta) = f_\theta(t) - f(t)$. They are well defined and smooth on $\mathbb{R}^n$ for any fixed value $t \in \mathbb{R}$ and have gradients (by differentiating in (3.1)) given by

$$\langle \nabla u_t(\theta), w\rangle \;=\; it\,\mathbb{E}\,\langle X, w\rangle\, e^{it\langle X,\theta\rangle}, \quad w \in \mathbb{C}^n,$$

where we use the canonical inner product in the product complex space. By the isotropy assumption,

$$|\,\langle \nabla u_t(\theta), w\rangle\,| \le |t|\,\mathbb{E}\,|\,\langle X, w\rangle\,| \le |t|\,|w|$$

for all w. Hence $|\nabla u_t(\theta)|^2 \le t^2$ for any $\theta \in \mathbb{R}^n$, so that by (9.1),

$$\mathbb{E}_\theta\,|f_\theta(t) - f(t)|^2 \le \frac{t^2}{n-1}. \tag{9.2}$$

Applying this inequality in (8.4) together with the first bound in (2.3) in order to replace F with Φ, we obtain:

Proposition 9.1 *Given an isotropic random vector X in $\mathbb{R}^n$,*

$$\mathbb{E}_\theta\,\rho^2(F_\theta, \Phi) \;\le\; c\,(1+\sigma_2^2)\,\frac{(\log n)^2}{n}.$$

Since $\sigma_2 \le \sigma_4$, we thus have

$$\big(\mathbb{E}_\theta\,\rho^2(F_\theta, \Phi)\big)^{1/2} \;\le\; c\,(1+\sigma_4)\,\frac{\log n}{\sqrt{n}} \tag{9.3}$$

which sharpens (1.1). The latter bound will be an essential step in the proof of Theorem 1.3, while (1.1) is not strong enough.

Let us now consider another scenario in Lemma 8.2, where the distribution of X is supported on the sphere $\sqrt{n}\,\mathbb{S}^{n-1}$. In this case,

$$\begin{aligned}\mathbb{E}_\theta\,|f_\theta(t) - f(t)|^2 &= \mathbb{E}_\theta\,|f_\theta(t)|^2 - |f(t)|^2 \\ &= \mathbb{E}J_n(t|X-Y|) - J_n(t\sqrt{n})^2\end{aligned}$$

according to (3.3), while $\sigma_4 = 0$. Hence, in (8.4) with $p = 2$ we arrive at the following preliminary bound which is needed for the proof of Theorem 1.1 in its second part. Here we use again that $m_4 \ge m_2 \ge 1$.

Corollary 9.2 *Suppose that $|X| = \sqrt{n}$ a.s., and Y is an independent copy of X. Then*

$$c\,\mathbb{E}_\theta\,\rho^2(F_\theta, F) \;\le\; \int_0^1 \frac{\Delta_n(t)}{t^2}\,dt + \log n \int_0^{4\sqrt{\log n}} \frac{\Delta_n(t)}{t}\,dt + \frac{(\log n)^2}{n^2}\,m_4^4, \tag{9.4}$$

where

$$\Delta_n(t) = \mathbb{E} J_n(t|X-Y|) - J_n(t\sqrt{n})^2. \tag{9.5}$$

10 Proof of Theorem 1.1 for the Kolmogorov Distance

To study the integrals in (9.4), assume additionally that the random vector X in $\mathbb{R}^n$ is isotropic with mean zero and put

$$\xi = \frac{\langle X, Y\rangle}{n},$$

where Y is an independent copy of X. Note that $\frac{1}{n^2}\, m_4^4 = \mathbb{E}\xi^4$ which is present in the last term on the right-hand side of (9.4).

Focusing on the first integral, we need to develop an asymptotic bound on $\Delta_n(t)$ for $t \in [0,1]$. Since $|X-Y|^2 = 2n(1-\xi)$, (9.5) becomes

$$\Delta_n(t) = \mathbb{E} J_n\big(t\sqrt{2n(1-\xi)}\,\big) - \big(J_n(t\sqrt{n})\big)^2.$$

We use the asymptotic formula (2.7),

$$J_n\big(t\sqrt{n}\big) = \Big(1 - \frac{t^4}{4n}\Big)\, e^{-t^2/2} + \varepsilon_n(t), \quad t \in \mathbb{R}, \tag{10.1}$$

where $\varepsilon_n(t)$ denotes a quantity of the form $O\big(n^{-2}\min(1,t^4)\big)$ with a universal constant in O. It implies a similar representation

$$\big(J_n\big(t\sqrt{n}\big)\big)^2 = \Big(1 - \frac{t^4}{2n}\Big)\, e^{-t^2} + \varepsilon_n(t). \tag{10.2}$$

Since $|\xi| \le 1$ a.s., we also have

$$J_n\big(t\sqrt{2n(1-\xi)}\,\big) = \Big(1 - \frac{t^4}{n}\,(1-\xi)^2\Big)\, e^{-t^2(1-\xi)} + \varepsilon_n(t).$$

Hence, subtracting from $e^{t^2\xi}$ the linear term $1 + t^2\xi$ and adding, one may write

$$\begin{aligned}
\Delta_n(t) &= e^{-t^2}\, \mathbb{E}\,\Big(\Big(1 - \frac{t^4}{n}\,(1-\xi)^2\Big)\, e^{t^2\xi} - \Big(1 - \frac{t^4}{2n}\Big)\Big) + \varepsilon_n(t) \\
&= e^{-t^2}\, \mathbb{E}\,(U + V) + \varepsilon_n(t)
\end{aligned}$$

with

$$U = \frac{t^4}{n}\Big(\frac{1}{2} - (1-\xi)^2\Big) + \Big(1 - \frac{t^4}{n}(1-\xi)^2\Big)\cdot t^2\xi,$$

$$V = \Big(1 - \frac{t^4}{n}(1-\xi)^2\Big)(e^{t^2\xi} - 1 - t^2\xi).$$

Using $\mathbb{E}\xi = 0$, $\mathbb{E}\xi^2 = \frac{1}{n}$ and hence $\mathbb{E}\,|\xi|^3 \le \mathbb{E}\xi^2 \le \frac{1}{n}$, we find that in the interval $0 \le t \le 1$,

$$\mathbb{E}\,U = -\frac{t^4}{2n} - \frac{t^4}{n^2} + \frac{2t^6}{n^2} - \frac{t^6}{n}\,\mathbb{E}\xi^3 = -\frac{t^4}{2n} + \varepsilon_n(t).$$

Next write

$$V = W - \frac{t^4}{n}(1-\xi)^2\,W, \qquad W = e^{t^2\xi} - 1 - t^2\xi.$$

Using $|e^x - 1 - x| \le 2x^2$ for $|x| \le 1$, we have $|W| \le 2t^4\xi^2$. Hence, the expected value of the second term in the representation for V does not exceed $8t^8/n^2$. Moreover, by Taylor's expansion,

$$W = \frac{1}{2}\,t^4\xi^2 + \frac{1}{6}\,t^6\xi^3 + Rt^8\xi^4, \qquad R = \sum_{k=4}^{\infty}\frac{t^{2k-8}}{k!}\,\xi^{k-4},$$

implying that

$$\mathbb{E}\,W = \frac{t^4}{2n} + \frac{t^6}{6}\,\mathbb{E}\xi^3 + Ct^8\,\mathbb{E}\xi^4,$$

where C is bounded by an absolute constant. Summing the two expansions, we arrive at

$$\mathbb{E}\,(U+V) = \frac{t^6}{6}\,\mathbb{E}\xi^3 + Ct^8\,\mathbb{E}\xi^4 + \varepsilon_n(t)$$

and therefore

$$\int_0^1 \frac{\Delta_n(t)}{t^2}\,dt \le \mathbb{E}\xi^3 + c\,\mathbb{E}\xi^4 + O(n^{-2}).$$

Here $\mathbb{E}\xi^4 \ge (\mathbb{E}\xi^2)^2 = n^{-2}$, so the term $O(n^{-2})$ may be absorbed by the 4-th moment of ξ. Since $\mathbb{E}\xi^3 \ge 0$, the bound (9.4) may be simplified to

$$c\,\mathbb{E}_\theta\,\rho^2(F_\theta, F) \le \mathbb{E}\xi^3 + \mathbb{E}\xi^4 + \log n\int_0^{4\sqrt{\log n}} \frac{\Delta_n(t)}{t}\,dt + \frac{(\log n)^2}{n^2}\,m_4^4,$$

that is,

$$c\,\mathbb{E}_\theta\, \rho^2(F_\theta, F) \,\le\, \log n \int_0^{4\sqrt{\log n}} \frac{\Delta_n(t)}{t}\,dt + \mathbb{E}\xi^3 + (\log n)^2\,\mathbb{E}\xi^4. \tag{10.3}$$

Turning to the remaining integral (which is most important), let us express it in terms of the functions $g_n(t) = J_n(t\sqrt{2n})$ and

$$\psi(\alpha) = \int_0^T \frac{g_n(\alpha t) - g_n(t)}{t}\,dt, \qquad 0 \le \alpha \le \sqrt{2},\ T > 1,$$

which will be needed with $T = 4\sqrt{\log n}$ and $\alpha = \sqrt{1-\xi}$. Namely, we have

$$\int_0^T \frac{\Delta_n(t)}{t}\,dt \,=\, \mathbb{E}\,\psi\big(\sqrt{1-\xi}\big) + \int_0^T \frac{J_n(t\sqrt{2n}) - (J_n(t\sqrt{n}))^2}{t}\,dt. \tag{10.4}$$

To proceed, we need to develop a Taylor expansion for $\xi \to \psi\big(\sqrt{1-\xi}\big)$ around zero in powers of ξ. Recall that $g_n(t)$ represents the characteristic function of the random variable $\sqrt{2n}\,\theta_1$ on the probability space $(\mathbb{S}^{n-1}, \mathfrak{s}_{n-1})$. This already ensures that $|g_n(t)| \le 1$ and

$$|g_n'(t)| \,\le\, \sqrt{2n}\,\mathbb{E}\,|\theta_1| \,\le\, \sqrt{2n}\,(\mathbb{E}\,\theta_1^2)^{1/2} \,=\, \sqrt{2}$$

for all $t \in \mathbb{R}$. Hence

$$|g_n(\alpha t) - g_n(t)| \le \sqrt{2}\,|\alpha - 1|\,|t| \le 2\,|t|,$$

so that

$$\begin{aligned} |\psi(\alpha)| &\le \int_0^1 \frac{|g_n(\alpha t) - g_n(t)|}{t}\,dt + \int_1^T \frac{|g_n(\alpha t) - g_n(t)|}{t}\,dt \\ &\le 2 + 2\log T \,<\, 4\log T \end{aligned} \tag{10.5}$$

(since $T > e$). In addition, $\psi(1) = 0$ and

$$\psi'(\alpha) = \int_0^T g_n'(\alpha t)\,dt = \frac{1}{\alpha}\,(g_n(\alpha T) - 1).$$

Therefore, we arrive at another expression

$$\psi(\alpha) = \int_1^\alpha \frac{g_n(Tx) - 1}{x}\,dx = \int_1^\alpha \frac{g_n(Tx)}{x}\,dx - \log\alpha.$$

For $|\varepsilon| \leq 1$, let

$$v(\varepsilon) = \int_1^{(1-\varepsilon)^{1/2}} \frac{g_n(Tx)}{x}\, dx,$$

$$u(\varepsilon) = \psi\big((1-\varepsilon)^{1/2}\big) \, = \, v(\varepsilon) - \frac{1}{2}\, \log(1-\varepsilon),$$

so that $\mathbb{E}\, \psi\big(\sqrt{1-\xi}\big) \, = \, \mathbb{E}\, u(\xi)$. Applying the non-uniform bound $|g_n(t)| \leq 5\,(e^{-t^2} + e^{-n/12})$, cf. (2.8), we have that, for $-1 \leq \varepsilon \leq \frac{1}{2}$,

$$|v(\varepsilon)| \leq \sup_{\frac{1}{\sqrt{2}} \leq x \leq \sqrt{2}} |g_n(Tx)| \int_{\frac{1}{\sqrt{2}}}^{\sqrt{2}} \frac{1}{x}\, dx$$

$$\leq \sup_{z \geq T/\sqrt{2}} |g_n(z)|\, \log 2 \, \leq \, 5 \log 2\, (e^{-T^2/2} + e^{-n/12}) \, \leq \, \frac{c}{n^8},$$

where the last inequality is specialized to the choice $T = 4\sqrt{\log n}$. Using the Taylor expansion on the same interval for the log-function, we also have $-\log(1-\varepsilon) \leq \varepsilon + \frac{1}{2}\, \varepsilon^2 + \frac{1}{3}\, \varepsilon^3 + \frac{2}{3}\, \varepsilon^4$. Combining the two inequalities, we get

$$u(\varepsilon) \leq \frac{1}{2}\, \varepsilon + \frac{1}{4}\, \varepsilon^2 + \frac{1}{6}\, \varepsilon^3 + \frac{1}{3}\, \varepsilon^4 + \frac{c}{n^8}, \quad -1 \leq \varepsilon \leq \frac{1}{2}. \tag{10.6}$$

In order to involve the remaining interval $\frac{1}{2} \leq \varepsilon \leq 1$ in the inequality of a similar type, recall that, by (10.5), $|u(\varepsilon)| \leq 4 \log T$ for all $|\varepsilon| \leq 1$. Hence, the inequality (10.6) will hold automatically for this interval, if we increase the coefficient in front of ε^4 to a suitable multiple of $\log T$. As a result, we obtain the desired inequality on the whole segment, that is,

$$u(\varepsilon) \leq \frac{1}{2}\, \varepsilon + \frac{1}{4}\, \varepsilon^2 + \frac{1}{6}\, \varepsilon^3 + (c \log T)\, \varepsilon^4 + \frac{c}{n^8}. \quad -1 \leq \varepsilon \leq 1.$$

In particular,

$$\psi\big(\sqrt{1-\xi}\big) \, \leq \, \frac{1}{2}\, \xi + \frac{1}{4}\, \xi^2 + \frac{1}{6}\, \xi^3 + (c \log T)\, \xi^4 + \frac{c}{n^8},$$

and taking the expectation, we get

$$\mathbb{E}\, \psi\big(\sqrt{1-\xi}\big) \leq \frac{1}{4n} + \frac{1}{6}\, \mathbb{E}\xi^3 + (c \log T)\, \mathbb{E}\xi^4, \tag{10.7}$$

where the term cn^{-8} was absorbed by the 4-th moment of ξ.

Now, let us turn to the integral

$$I_n = \int_0^T \frac{J_n(t\sqrt{2n}) - (J_n(t\sqrt{n}))^2}{t}\,dt,$$

appearing in (10.4), and recall the asymptotic formulas (10.1) and (10.2). After integration, the remainder term $\varepsilon_n(t) = O\big(n^{-2}\min(1,t^4)\big)$ will create an error of order at most $n^{-2}\log T$, up to which I_n is equal to

$$-\int_0^T \frac{t^4}{2n}\,e^{-t^2}\,\frac{dt}{t} = -\frac{1}{4n}\Big(1-(T^2+1)\,e^{-T^2}\Big) = -\frac{1}{4n} + o(n^{-15}).$$

Thus,

$$I_n = -\frac{1}{4n} + O(n^{-2}\log T).$$

Applying this expansion together with (10.7) in (10.4), we therefore obtain that

$$\int_0^T \frac{\Delta_n(t)}{t}\,dt \ \le\ \frac{1}{6}\,\mathbb{E}\xi^3 + c\log T\,\mathbb{E}\xi^4.$$

One can now apply this estimate in (10.3), and then we eventually arrive at

$$\mathbb{E}_\theta\,\rho^2(F_\theta,F) \ \le\ c_1\,(\log n)\,\mathbb{E}\xi^3 + c_2\,(\log n)^2\,\mathbb{E}\xi^4.$$

By (2.3) with $p=\infty$, a similar inequality remains to hold for the standard normal distribution function Φ in place of F. This proves the inequality (1.4). □

11 Relations Between L^1, L^2 and Kolmogorov Distances

Given a random vector X in $\mathbb{R}^n$, let us now compare the L^2 and L^∞ distances on average, between the distributions F_θ of the weighted sums $\langle X,\theta\rangle$ and the typical distribution $F = \mathbb{E}_\theta F_\theta$. Such information will be needed to derive appropriate lower bounds on $\mathbb{E}_\theta\,\rho(F_\theta,F)$.

Proposition 11.1 *If $|X| \le b\sqrt{n}$ a.s., then, for any $\alpha \in [1,2]$,*

$$b^{-\alpha/2}\,\mathbb{E}_\theta\,\omega^\alpha(F_\theta,F) \ \le\ 14\,(\log n)^{\alpha/4}\,\mathbb{E}_\theta\,\rho^\alpha(F_\theta,F) + \frac{8}{n^4}. \tag{11.1}$$

As will be clear from the proof, at the expense of a larger coefficient in front of $\log n$, the last term n^{-4} can be replaced by $n^{-\beta}$ for any prescribed value of β.

A relation similar to (11.1) is also true for the Kantorovich or L^1-distance

$$W(F_\theta, F) = \int_{-\infty}^{\infty} |F_\theta(x) - F(x)|\, dx$$

in place of L^2. We state it for the case $\alpha = 1$.

Proposition 11.2 *If $|X| \leq b\sqrt{n}$ a.s., then*

$$\mathbb{E}_\theta\, W(F_\theta, F) \ \leq\ 14\, b\sqrt{\log n}\, \mathbb{E}_\theta\, \rho(F_\theta, F) + \frac{8b}{n^4}. \tag{11.2}$$

Proof Put $R_\theta(x) = F_\theta(-x) + (1 - F_\theta(x))$ for $x > 0$ and define similarly R on the basis of F. Using

$$(F_\theta(-x) - F(-x))^2 \leq F_\theta(-x)^2 + F(-x)^2,$$
$$(F_\theta(x) - F(x))^2 \leq (1 - F_\theta(x))^2 + (1 - F(x))^2,$$

we have

$$(F_\theta(-x) - F(-x))^2 + (F_\theta(x) - F(x))^2 \leq R_\theta(x)^2 + R(x)^2.$$

Hence, given $T > 0$ (to be specified later on), we have

$$\begin{aligned} \omega^2(F_\theta, F) &= \int_{-T}^{T} (F_\theta(x) - F(x))^2\, dx + \int_{|x| \geq T} (F_\theta(x) - F(x))^2\, dx \\ &\leq 2T\rho^2(F_\theta, F) + \int_T^\infty R_\theta(x)^2\, dx + \int_T^\infty R(x)^2\, dx. \end{aligned}$$

It follows that, for any $\alpha \in [1, 2]$,

$$\omega^\alpha(F_\theta, F) \leq (2T)^{\frac{\alpha}{2}}\, \rho^\alpha(F_\theta, F) + \Big(\int_T^\infty R_\theta(x)^2\, dx\Big)^{\frac{\alpha}{2}} + \Big(\int_T^\infty R(x)^2\, dx\Big)^{\frac{\alpha}{2}}$$

and therefore, by Jensen's inequality,

$$\begin{aligned} \mathbb{E}_\theta\, \omega^\alpha(F_\theta, F) &\leq (2T)^{\frac{\alpha}{2}}\, \mathbb{E}_\theta\, \rho^\alpha(F_\theta, F) \\ &\quad + \Big(\int_T^\infty \mathbb{E}_\theta\, R_\theta(x)^2\, dx\Big)^{\frac{\alpha}{2}} + \Big(\int_T^\infty R(x)^2\, dx\Big)^{\frac{\alpha}{2}}. \end{aligned}$$

Next, by Markov's inequality, for any $x > 0$ and $p \geq 1$,

$$R_\theta(x)^2 \leq \Big(\frac{\mathbb{E}\, |\, \langle X, \theta\rangle\, |^p}{x^p}\Big)^2 \leq \frac{\mathbb{E}\, |\, \langle X, \theta\rangle\, |^{2p}}{x^{2p}}$$

and

$$\mathbb{E}_\theta R_\theta(x)^2 \le \Big(\frac{\mathbb{E}\,|\langle X,\theta\rangle|^p}{x^p}\Big)^2 \le \frac{\mathbb{E}_\theta\,\mathbb{E}\,|\langle X,\theta\rangle|^{2p}}{x^{2p}}.$$

Since $R = \mathbb{E}_\theta R_\theta$, a similar inequality holds true for R as well (by Cauchy's inequality). Hence

$$\mathbb{E}_\theta\,\omega^\alpha(F_\theta,F) \ \le\ (2T)^{\frac{\alpha}{2}}\,\mathbb{E}_\theta\,\rho^\alpha(F_\theta,F) + 2\Big(\mathbb{E}_\theta\,\mathbb{E}\,|\langle X,\theta\rangle|^{2p}\int_T^\infty \frac{1}{x^{2p}}\,dx\Big)^{\frac{\alpha}{2}}.$$

When $\theta = (\theta_1,\dots,\theta_n)$ is treated as a random vector with distribution $\mathfrak{s}_{n-1}$, which is independent of X, the inner product $\langle X,\theta\rangle$ has the same distribution as the random variable $|X|\,\theta_1$. Therefore, recalling Lemma 2.5 and using the assumption $|X| \le b\sqrt{n}$ a.e., we have

$$\mathbb{E}_\theta\,\mathbb{E}\,|\langle X,\theta\rangle|^{2p} \ =\ \mathbb{E}\,|X|^{2p}\,\mathbb{E}_\theta\,|\theta_1|^{2p} \ \le\ 2\,(2b^2p)^p,$$

so that

$$2\Big(\mathbb{E}_\theta\int_T^\infty \frac{\mathbb{E}\,|\langle X,\theta\rangle|^{2p}}{x^{2p}}\,dx\Big)^{\frac{\alpha}{2}} \ \le\ \frac{2^{\frac{\alpha}{2}+1}}{(2p-1)^{\frac{\alpha}{2}}}\,\frac{(2b^2p)^{\frac{\alpha p}{2}}}{T^{\frac{\alpha(2p-1)}{2}}}.$$

Thus,

$$\mathbb{E}_\theta\,\omega^\alpha(F_\theta,F) \ \le\ (2T)^{\frac{\alpha}{2}}\,\mathbb{E}_\theta\,\rho^\alpha(F_\theta,F) + \frac{2^{\frac{\alpha}{2}+1}}{(2p-1)^{\frac{\alpha}{2}}}\,T^{\frac{\alpha}{2}}\Big(\frac{2b^2p}{T^2}\Big)^{\frac{\alpha p}{2}}.$$

Let us choose $T = 2b\sqrt{p}$ in which case the above inequality becomes

$$\mathbb{E}_\theta\,\omega^\alpha(F_\theta,F) \ \le\ (4b\sqrt{p})^{\frac{\alpha}{2}}\,\mathbb{E}_\theta\,\rho^\alpha(F_\theta,F) + \frac{2^{\alpha+1}}{(2p-1)^{\frac{\alpha}{2}}}\,(b\sqrt{p})^{\frac{\alpha}{2}}\,2^{-\frac{\alpha p}{2}}.$$

To simplify, one can use $\sqrt{p} \le 2p-1$ for $p \ge 1$ together with $2^{\alpha+1} \le 8$ and $2^{-\frac{\alpha p}{2}} \le 2^{-\frac{p}{2}}$ (since $1 \le \alpha \le 2$), which leads to

$$\mathbb{E}_\theta\,\omega^\alpha(F_\theta,F) \ \le\ (4b\sqrt{p})^{\frac{\alpha}{2}}\,\mathbb{E}_\theta\,\rho^\alpha(F_\theta,F) + 8\,b^{\frac{\alpha}{2}}\,2^{-p/2}.$$

Finally, choosing $p = p_n = (8\,\log n)/\log 2$, we arrive at (11.1).

Now, turning to (11.2), we use the same functions R_θ and R as before and write

$$W(F_\theta, F) = \int_{-T}^{T} |F_\theta(x) - F(x)|\,dx + \int_{|x|\ge T} |F_\theta(x) - F(x)|\,dx$$
$$\le 2T\rho(F_\theta, F) + \int_T^\infty R_\theta(x)\,dx + \int_T^\infty R(x)\,dx,$$

which gives

$$\mathbb{E}_\theta\, W(F_\theta, F) \ \le\ 2T\,\mathbb{E}_\theta\,\rho(F_\theta, F) + 2\int_T^\infty R(x)\,dx.$$

By Markov's inequality, for any $x > 0$ and $p > 1$,

$$R_\theta(x) \le \frac{\mathbb{E}\,|\,\langle X, \theta\rangle\,|^p}{x^p}, \quad R(x) = \mathbb{E}_\theta R_\theta(x) \le \frac{\mathbb{E}_\theta\,\mathbb{E}\,|\,\langle X, \theta\rangle\,|^p}{x^p}.$$

Hence

$$\mathbb{E}_\theta\, W(F_\theta, F) \ \le\ 2T\,\mathbb{E}_\theta\,\rho(F_\theta, F) + 2\,\mathbb{E}_\theta\,\mathbb{E}\,|\,\langle X, \theta\rangle\,|^p \int_T^\infty \frac{1}{x^p}\,dx.$$

Here, one may use once more the bound (2.10), which yields

$$\mathbb{E}_\theta\,\mathbb{E}\,|\,\langle X, \theta\rangle\,|^p \ =\ \mathbb{E}\,|X|^p\,\mathbb{E}_\theta\,|\theta_1|^p \ \le\ 2\,\big(b^2 p\big)^{p/2}$$

and

$$\mathbb{E}_\theta\, W(F_\theta, F) \ \le\ 2T\,\mathbb{E}_\theta\,\rho(F_\theta, F) + \frac{4}{p-1}\,\frac{(b^2 p)^{p/2}}{T^{p-1}}.$$

Let us take $T = 2b\sqrt{p}$ in which case the above inequality becomes

$$\mathbb{E}_\theta\, W(F_\theta, F) \ \le\ 4b\sqrt{p}\,\mathbb{E}_\theta\,\rho(F_\theta, F) + 8b\,\frac{\sqrt{p}}{p-1}\,2^{-p}.$$

Here we arrive at (11.2), by choosing again $p = p_n$ and using $\sqrt{p_n} < p_n - 1$. □

12 Lower Bounds: Proof of Theorem 1.3

A lower bound on $\mathbb{E}_\theta\,\rho^2(F_\theta, \Phi)$ which would be close to the upper bound (1.4) may be given with the help of the lower bound on $\mathbb{E}_\theta\,\omega^2(F_\theta, \Phi)$. More precisely, this can be done in the case where the quantity $\frac{1}{n^{3/2}}\,m_3^3 + \frac{1}{n^2}\,m_4^4$ asymptotically dominates

n^{-2} (in particular, when m_4 is essentially larger than 1). Combining the asymptotic expansion (1.3) of Theorem 1.1 with the bound (11.1) of Proposition 11.1 for $\alpha = 2$ and $b = 1$, and recalling the second relation in (2.3) on the normal approximation for the typical distribution F, we therefore obtain:

Proposition 12.1 *If X is an isotropic random vector in $\mathbb{R}^n$ with mean zero and such that $|X| = \sqrt{n}$ a.s., then*

$$\sqrt{\log n}\, \mathbb{E}_\theta\, \rho^2(F_\theta, \Phi) \ \geq\ \frac{c_1}{n^{3/2}}\, m_3^3 + \frac{c_2}{n^2}\, m_4^4 - \frac{c_3}{n^2}. \tag{12.1}$$

The relation (11.2) for the Kantorovich distance W may be used to answer the following question: Is it possible to sharpen the lower bound (12.1) by replacing $\mathbb{E}_\theta\, \rho^2(F_\theta, \Phi)$ with $\mathbb{E}_\theta\, \rho(F_\theta, \Phi)$? To this aim, we will need an additional information about moments of $\omega(F_\theta, F)$ of order higher than 2.

Lemma 12.2 *If X is isotropic and satisfies $|X| \leq b\sqrt{n}$, then*

$$c\, \Big(\mathbb{E}_\theta\, \omega^3(F_\theta, F)\Big)^{1/3} \ \leq\ (1 + \sigma_4)\sqrt{b}\ \frac{(\log n)^{5/4}}{\sqrt{n}}. \tag{12.2}$$

Proof For any distribution function G with finite first absolute moment, the function on the unit sphere $\mathbb{S}^{n-1}$ of the form $g(\theta) = W(F_\theta, G)$ has a Lipschitz semi-norm $\|g\|_{\mathrm{Lip}} \leq 1$. Therefore, it admits a subgaussian large deviation bound

$$\mathfrak{s}_{n-1}\big\{W(F_\theta, G) \geq m + r\big\} \ \leq\ e^{-(n-1)r^2/2}, \qquad r \geq 0, \tag{12.3}$$

where $m = \mathbb{E}_\theta\, W(F_\theta, G)$. Indeed, consider the elementary representation

$$\begin{aligned} W(F_\theta, G) &\equiv \int_{-\infty}^{\infty} |F_\theta(x) - G(x)|\, dx \\ &= \sup_u \bigg[\int_{-\infty}^{\infty} u\, dF_\theta - \int_{-\infty}^{\infty} u\, dG\bigg], \end{aligned}$$

where the supremum is running over all functions u on $\mathbb{R}$ with $\|u\|_{\mathrm{Lip}} \leq 1$. For any such u,

$$H_u(\theta) = \int_{-\infty}^{\infty} u\, dF_\theta \ =\ \mathbb{E}\, u(\langle X, \theta\rangle)$$

is Lipschitz on $\mathbb{R}^n$ and therefore on $\mathbb{S}^{n-1}$. Moreover, $\|g\|_{\mathrm{Lip}} \leq \sup_u \|H_u\|_{\mathrm{Lip}} \leq 1$.

Hence, (12.3) is fulfilled as a consequence of fact that the logarithmic Sobolev constant for the uniform distribution on the unit sphere is equal to $n - 1$ (cf. [21]). In particular, for any $r \geq 0$,

$$\mathfrak{s}_{n-1}\big\{W(F_\theta, F) \geq m + r\big\} \ \leq\ e^{-(n-1)r^2/2}$$

with $m = \mathbb{E}_\theta\, W(F_\theta, F)$. In turn, the latter ensures that, for any $p \geq 2$,

$$\Big(\mathbb{E}_\theta\, W(F_\theta, F)^p\Big)^{1/p} \leq m + \frac{\sqrt{p}}{\sqrt{n-1}}. \tag{12.4}$$

For the proof, put $\xi = (W(F_\theta, F) - m)^+$. Using $\Gamma(x+1) \leq x^x$ with $x = p/2 \geq 1$, we have

$$\begin{aligned}\mathbb{E}_\theta\, \xi^p &= \int_0^\infty \mathfrak{s}_{n-1}\{\xi \geq r\}\, dr^p \;\leq\; \int_0^\infty e^{-(n-1)\, r^2/2}\, dr^p \\ &= \Big(\frac{\sqrt{2}}{\sqrt{n-1}}\Big)^p\, \Gamma\Big(\frac{p}{2}+1\Big) \;\leq\; \Big(\frac{\sqrt{p}}{\sqrt{n-1}}\Big)^p \;\equiv\; A^p \quad (A \geq 0).\end{aligned}$$

Thus, $\|\xi\|_p = (\mathbb{E}_\theta\, \xi^p)^{1/p} \leq A$. Since $W(F_\theta, F) \leq \xi + m$, we conclude, by the triangle inequality, that

$$\|W(F_\theta, F)\|_p \leq \|\xi\|_p + m \leq A + m,$$

that is, (12.4) holds.

Let us proceed with one elementary general inequality, connecting the three distances,

$$\begin{aligned}\omega^2(F_\theta, F) &= \int_{-\infty}^{\infty} (F_\theta(x) - F(x))^2\, dx \\ &\leq \int_{-\infty}^{\infty} \sup_x |F_\theta(x) - F(x)|\, |F_\theta(x) - F(x)|\, dx \;=\; \rho(F_\theta, F)\, W(F_\theta, F).\end{aligned}$$

Putting $\omega = \omega(F_\theta, F)$, $W = W(F_\theta, F)$, $\rho = \rho(F_\theta, F)$, we thus have $\omega^3 \leq W^{3/2}\rho^{3/2}$ and, by Hölder's inequality with exponents $p = 4$ and $q = 4/3$,

$$\|\omega\|_3 = \big(\mathbb{E}_\theta\, \omega^3\big)^{1/3} \leq \big(\mathbb{E}_\theta\, W^6\big)^{1/12}\, \big(\mathbb{E}_\theta\, \rho^2\big)^{1/4}.$$

By (12.4) with $p = 6$, we have

$$\big(\mathbb{E}_\theta\, W^6\big)^{1/6} \leq \mathbb{E}_\theta\, W + \frac{4}{\sqrt{n}},$$

so that

$$\|\omega\|_3 \leq \Big(\mathbb{E}_\theta\, W + \frac{4}{\sqrt{n}}\Big)^{1/2}\, \big(\mathbb{E}_\theta\, \rho^2\big)^{1/4}.$$

Applying Proposition 11.2 and noting that necessarily $b \geq 1$ in the isotrpic case, we get

$$\|\omega\|_3 \leq 4\sqrt{b}\,\Big(\sqrt{\log n}\;\mathbb{E}_\theta\,\rho + \frac{1}{\sqrt{n}}\Big)^{1/2}\,\big(\mathbb{E}_\theta\,\rho^2\big)^{1/4}.$$

Here we employ the inequality (9.3) with F in place of Φ, i.e.

$$\mathbb{E}_\theta\,\rho(F_\theta, F) \leq \big(\mathbb{E}_\theta\,\rho^2(F_\theta, F)\big)^{1/2} \leq c\,(1+\sigma_4)\,\frac{\log n}{\sqrt{n}}.$$

Since the last expression dominates the term $\frac{1}{\sqrt{n}}$, it follows that

$$\|\omega\|_3 \leq c\sqrt{b}\,\Big(\sqrt{\log n}\;(1+\sigma_4)\,\frac{\log n}{\sqrt{n}}\Big)^{1/2}\,\Big((1+\sigma_4)\,\frac{\log n}{\sqrt{n}}\Big)^{1/2},$$

and we arrive at the upper bound (12.2). □

Let us now explain how this bound can be used to refine the lower bound (12.1). The argument is based on the following general elementary observation. Given a random variable ξ, introduce the L^p-norms $\|\xi\|_p = (\mathbb{E}\,|\xi|^p)^{1/p}$.

Lemma 12.3 *If* $\xi \geq 0$ *with* $0 < \|\xi\|_3 < \infty$*, then*

$$\mathbb{E}\,\xi \geq \frac{1}{\mathbb{E}\,\xi^3}\,(\mathbb{E}\,\xi^2)^2. \tag{12.5}$$

Moreover,

$$\mathbb{P}\Big\{\xi \geq \frac{1}{\sqrt{2}}\,\|\xi\|_2\Big\} \;\geq\; \frac{1}{8}\,\Big(\frac{\|\xi\|_2}{\|\xi\|_3}\Big)^6. \tag{12.6}$$

Thus, in the case where $\|\xi\|_2$ and $\|\xi\|_3$ are equivalent within not too large factors, $\|\xi\|_1$ will be of a similar order. Moreover, ξ cannot be much smaller than its mean $\mathbb{E}\xi$ on a large part of the probability space (where it was defined).

Proof Let ξ be defined on the probability space $(\Omega, \mathfrak{F}, \mathbb{P})$. By homogeneity with respect to ξ, we may assume that $\mathbb{E}\xi = 1$, so that $dQ = \xi d\mathbb{P}$ is a probability measure. Then, (12.5) follows from the Cauchy inequality $(\mathbb{E}_Q\xi)^2 \leq \mathbb{E}_Q\xi^2$ on the space $(\Omega, \mathfrak{F}, Q)$.

To prove (12.6), given $r > 0$, let $p = \mathbb{P}\{\xi \geq r\}$. By Hölder's inequality with exponents $3/2$ and 3,

$$\mathbb{E}\,\xi^2\,1_{\{\xi \geq r\}} \leq \big(\mathbb{E}\,\xi^3\big)^{2/3}\,p^{1/3}.$$

Hence, choosing $r = \frac{1}{\sqrt{2}} \|\xi\|_2$, we get

$$\mathbb{E}\,\xi^2 = \mathbb{E}\,\xi^2\, 1_{\{\xi \geq r\}} + \mathbb{E}\,\xi^2\, 1_{\{\xi < r\}}$$
$$\leq \big(\mathbb{E}\,\xi^3\big)^{2/3} p^{1/3} + r^2 \;=\; \big(\mathbb{E}\,\xi^3\big)^{2/3} p^{1/3} + \frac{1}{2}\,\mathbb{E}\,\xi^2.$$

Hence $p^{1/3} \geq \frac{1}{2\,(\mathbb{E}\,\xi^3)^{2/3}}\,\mathbb{E}\,\xi^2$ which is the desired bound (12.6). □

We now combine Lemma 12.2 with Lemma 12.3 which is applied on the unit sphere to $\xi(\theta) = \omega(F_\theta, F)$ viewed as a random variable on the probability space $(\mathbb{S}^{n-1}, \mathfrak{s}_{n-1})$. Recall that $b \geq 1$ in the isotropic case.

Proposition 12.4 *Let X be an isotropic random vector in $\mathbb{R}^n$ such that $|X| \leq b\sqrt{n}$ a.s. Assume that*

$$\mathbb{E}_\theta\, \omega^2(F_\theta, F) \;\geq\; \frac{D}{n}$$

with some $D > 0$. Then

$$\mathbb{E}_\theta\, \omega(F_\theta, F) \;\geq\; \frac{c}{(1+\sigma_4)^3\, b^{\frac{3}{2}}}\, \frac{D^2}{(\log n)^{\frac{15}{4}} \sqrt{n}}. \tag{12.7}$$

Moreover,

$$\mathfrak{s}_{n-1}\Big\{\omega(F_\theta, F) \geq \frac{1}{\sqrt{2n}}\sqrt{D}\Big\} \;\geq\; \frac{c}{(1+\sigma_4)^6\, b^3}\, \frac{D^3}{(\log n)^{\frac{15}{2}}}.$$

Proof of Theorem 1.3 The lower bound (12.7) implies a similar assertion about the Kolmogorov distance. Indeed, by Proposition 11.1 with $\alpha = 1$, we have

$$\frac{1}{\sqrt{b}}\,\mathbb{E}_\theta\, \omega(F_\theta, F) \;\leq\; 14\,(\log n)^{1/4}\, \mathbb{E}_\theta\, \rho(F_\theta, F) + \frac{8}{n^4}.$$

Using $\frac{8}{n^4} < \frac{1}{n^3} \cdot 14\,(\log n)^{1/4}$, we therefore obtain that

$$\mathbb{E}_\theta\, \rho(F_\theta, F) \geq \frac{1}{14\sqrt{b}\,(\log n)^{1/4}}\, \mathbb{E}_\theta\, \omega(F_\theta, F) - \frac{1}{n^3}$$
$$\geq \frac{c}{(1+\sigma_4)^3\, b^2}\, \frac{D^2}{(\log n)^4 \sqrt{n}} - \frac{1}{n^3}.$$

To replace F with Φ, it remains to recall the bound $\rho(F, \Phi) \leq \frac{c}{n}\,(1 + \sigma_4^2)$, cf. (2.3). □

In the isotropic case with $|X|^2 = n$ a.s., the above lower bound is further simplified to

$$\mathbb{E}_\theta\, \rho(F_\theta, F) \ \geq \ \frac{cD^2}{(\log n)^4\sqrt{n}} - \frac{1}{n^3}.$$

On the other hand, let us note that the rates for the normal approximation of F_θ that are better than $1/n$ (on average) cannot be obtained under the support assumption as above. That is, if $|X| = \sqrt{n}$ a.s., then

$$\mathbb{E}_\theta\, \rho(F_\theta, \Phi) \ \geq \ \frac{c}{n}.$$

Indeed, using the convexity of the distance function $G \to \rho(G, \Phi)$ and applying Jensen's inequality, we have that $\mathbb{E}_\theta\, \rho(F_\theta, \Phi) \geq \rho(F, \Phi)$. It remains to appeal to Proposition 2.6.

13 Functional Examples

13.1. For the trigonometric system as in item (i) of the Introduction (with n even), the linear forms

$$\langle X, \theta\rangle = \sqrt{2}\sum_{k=1}^{\frac{n}{2}} \big(\theta_{2k-1}\cos(kt) + \theta_{2k}\sin(kt)\big), \qquad \theta = (\theta_1, \dots, \theta_n) \in \mathbb{S}^{n-1},$$

represent trigonometric polynomials of degree at most $\frac{n}{2}$. The normalization $\sqrt{2}$ is chosen in order to meet the requirement that the random vector X is isotropic with respect to the normalized Lebesgue measure $\mathbb{P}$ on $\Omega = (-\pi, \pi)$. Moreover, in this case $|X| = \sqrt{n}$, so that $\sigma_4 = 0$. Hence, by Theorem 1.1, we have the upper bounds (1.6). On the other hand, since for all $k \leq \frac{n}{2}$

$$|X_k(t) - X_k(s)| \leq k\sqrt{2}\,|t-s| \leq \frac{n}{\sqrt{2}}\,|t-s|, \qquad t, s \in \Omega,$$

the Lipschitz condition (7.1) is fulfilled with $L(t) = \frac{t}{\sqrt{2}}$. Hence, Proposition 7.1 is applicable and yields the lower bound

$$\mathbb{E}_\theta\, \omega^2(F_\theta, \Phi) \ \geq \ \frac{c_1}{n} - \frac{c_2}{n^2} \ \geq \ \frac{c_3}{n},$$

where in the last inequality we assume that $n \geq n_0$ for some universal integer n_0. This restriction may be dropped, since the distances $\omega^2(F_\theta, \Phi)$

are bounded away from zero for $n < n_0$ uniformly over all $\theta \in \mathbb{S}^{n-1}$, just due to the property that the distributions F_θ are supported on the bounded interval $[-\sqrt{n_0}, \sqrt{n_0}]$. Note that the above lower estimate (may also be obtained by applying Theorem 1.1. Thus, for all $n \geq 2$,

$$\frac{c_0}{n} \leq \mathbb{E}_\theta\, \omega^2(F_\theta, \Phi) \leq \frac{c_1}{n}. \tag{13.1}$$

Applying Proposition 12.4, we obtain similar bounds for the L^1-norm (modulo logarithmic factors). Namely, it gives

$$\frac{c_0}{(\log n)^{\frac{15}{4}}\sqrt{n}} \leq \mathbb{E}_\theta\, \omega(F_\theta, \Phi) \leq \frac{c_1}{\sqrt{n}}. \tag{13.2}$$

We also get an analogous pointwise lower bound on the "essential" part of the unit sphere.

A similar statement is also true for the Kolmogorov distance. Here, the upper bound is provided in Proposition 9.1, while the lower bound is obtained when combining Theorem 1.3 with the left inequality in (13.1). That is,

$$\frac{c_0}{(\log n)^4\sqrt{n}} \leq \mathbb{E}_\theta\, \rho(F_\theta, \Phi) \leq \big(\mathbb{E}_\theta\, \rho^2(F_\theta, \Phi)\big)^{1/2} \leq \frac{c_1 \log n}{\sqrt{n}}. \tag{13.3}$$

13.2. Analogous results remain true for the cosine trigonometric system $X = (X_1, \dots, X_n)$ as in item (ii). Due to the normalization $\sqrt{2}$, the distribution of X is isotropic in $\mathbb{R}^n$. The property $|X| = \sqrt{n}$ is not true anymore; however, there is a pointwise bound $|X| \leq \sqrt{2n}$. In addition, the variance functional σ_4^2 does not depend on n. Indeed, write

$$X_k^2 = 2\cos^2(kt) = 1 + \cos(2kt) = 1 + \frac{e^{2ikt} + e^{-2ikt}}{2},$$

so that

$$2\,(|X|^2 - n) = \sum_{0<|k|\leq n} e^{2ikt}, \qquad 4\,(|X|^2 - n)^2 = \sum_{0<|k|,|l|\leq n} e^{2i(k+l)t}.$$

It follows that

$$4\,\mathrm{Var}(|X|^2) = \sum_{0<|k|,|l|\leq n} \mathbb{E}\, e^{2i(k+l)t} = \sum_{0<|k|\leq n,\, l=-k} 1 = 2n.$$

Hence

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

As before, the Lipschitz condition is fulfilled with the function $L(t) = t\sqrt{2}$. Therefore, with similar arguments we obtain all the bounds (13.1)–(13.3).

Let us also note that the sums $\sum_{k=1}^n \cos(kt)$ remain bounded for growing n (for any fixed $0 < t < \pi$). Hence the normalized sums

$$S_n = \frac{1}{\sqrt{n}} \sum_{k=1}^n X_k = \frac{\sqrt{2}}{\sqrt{n}} \sum_{k=1}^n \cos(kt),$$

which correspond to $\langle X, \theta \rangle$ with equal coefficients, are convergent to zero pointwise on Ω as $n \to \infty$. In particular, they fail to satisfy the central limit theorem.

13.3. An example closely related to the cosine trigonometric system is represented by the normalized Chebyshev's polynomials X_k as in item (iii), which we consider for $k = 1, 2, \dots, n$. These polynomials are orthonormal on the interval $\Omega = (-1, 1)$ with respect to the probability measure

$$\frac{d\mathbb{P}(t)}{dt} = \frac{1}{\pi\sqrt{1-t^2}}, \quad -1 < t < 1,$$

cf. e.g. [17]. Similarly to 13.2, for the random vector $X = (X_1, \dots, X_n)$ we find that

$$4\,(|X|^2 - n)^2 = \sum_{0<|k|,|l|\le n} \exp\{2i(k+l)\,\arccos t\}.$$

It follows that

$$4\,\mathrm{Var}(|X|^2) = \sum_{0<|k|,|l|\le n} \mathbb{E}\,\exp\{2i(k+l)\,\arccos t\} = \sum_{0<|k|\le n} 1 = 2n,$$

so that $\sigma_4^2 = \frac{1}{n}\,\mathrm{Var}(|X|^2) = \frac{1}{2}$. In addition, for all $k \le n$,

$$|X_k(t) - X_k(s)| \le k\sqrt{2}\,|\arccos t - \arccos s|, \qquad t, s \in \Omega,$$

which implies that the Lipschitz condition is fulfilled with the function $L(t) = \sqrt{2}\,\arccos t$. As a result, we obtain the bounds (13.1)–(13.3) as well.

13.4. Turning to item (iv), consider the functions of the form

$$X_k(t, s) = \Psi(kt + s),$$

assuming that Ψ is a 1-periodic measurable function on the real line such that

$$\int_0^1 \Psi(x)\,dx = 0 \quad \text{and} \quad \int_0^1 \Psi(x)^2\,dx = 1.$$

These conditions ensure that the random vector $X = (X_1, \dots, X_n)$ is isotropic in $\mathbb{R}^n$ with respect to the Lebesgue measure $\mathbb{P}$ on the square $\Omega = (0,1) \times (0,1)$, with $\mathbb{E}X_k = 0$. In fact, as was emphasized in [5], $\{X_k\}_{k=1}^{\infty}$ represents a strictly stationary sequence of pairwise independent random variables on Ω. The latter implies in particular that, if Ψ has finite 4-th moment on (0, 1), the variance functional

$$\sigma_4^2 = \frac{1}{n}\,\mathrm{Var}(|X|^2) = \int_0^1 \Psi(x)^4\,dx - 1$$

is finite and does not dependent on n. Hence, by Theorem 1.1, cf. (1.6), the upper bounds in (13.1)–(13.3) hold true with a constant c_1 depending on the 4-th moment of Ψ on (0, 1).

In addition, if the function Ψ has finite Lipschitz constant $\|\Psi\|_{\mathrm{Lip}}$, then for all (t_1, t_2) and (s_1, s_2) in Ω,

$$|X_k(t_1,t_2) - X_k(s_1,s_2)| \,\le\, \|\Psi\|_{\mathrm{Lip}}\,\big(k\,|t_1 - s_1| + |t_2 - s_2|\big).$$

This means that the Lipschitz condition (7.5) is fulfilled with linear functions L_1 and L_2. Hence, one may apply Proposition 7.5 giving the lower bound

$$\mathbb{E}_\theta\,\omega^2(F_\theta, F) \,\ge\, \frac{c_\Psi}{n} - \frac{c\,(1+\sigma_4^4)}{n^2}$$

in full analogy with item (i). Hence $\mathbb{E}_\theta\,\omega^2(F_\theta, \Phi) \,\ge\, \frac{c'_\Psi}{n}$ for all $n \ge n_0$, where the positive constants c_Ψ, c'_Ψ, and an integer $n_0 \ge 1$ depend on the distribution of Ψ only. Since the collection $\{F_\theta\}$ is separated from Φ in the weak sense for $n < n_0$ (by the uniform boundedness of X_k's), the latter bound holds true for all $n \ge 2$. Also, as Lipschitz functions on (0, 1) are bounded, we have $|X| \le b\sqrt{n}$ with $b = \sup_x |f(x)|$, and one may apply Theorem 1.3.

Let us summarize: *The upper bounds in* (13.1)–(13.3) *hold true, if* Ψ *has finite 4-th moment under the uniform distribution on* (0, 1). *The lower bounds hold under an additional assumption that* Ψ *has a finite Lipschitz semi-norm (with constants depending on* Ψ *only).*

Choosing, for example, $\Psi(t) = \cos t$, we obtain the system $X_k(t,s) = \cos(kt + s)$, which is closely related to the cosine trigonometric system. The main difference is however the property that X_k's are now pairwise independent. Nevertheless, the normalized sums $\frac{1}{\sqrt{n}}\sum_{k=1}^n \cos(kt+s)$ fail to satisfy the central limit theorem.

14 The Walsh System; Empirical Measures

14.1. The Walsh system on the discrete cube $\Omega = \{-1, 1\}^d$ with the uniform counting measure $\mathbb{P}$ as in item (v) in Introduction forms a complete orthonormal system in $L^2(\Omega, \mathbb{P})$. Note that each X_τ with $\tau \neq \emptyset$ is a symmetric Bernoulli random variable taking the values -1 and 1 with probability $\frac{1}{2}$. For simplicity, we exclude from this family the constant $X_\emptyset = 1$ and consider $X = \{X_\tau\}_{\tau \neq \emptyset}$ as a random vector in $\mathbb{R}^n$ of dimension $n = 2^d - 1$. As before, F_θ denotes the distribution function of the linear form

$$\langle X, \theta\rangle = \sum_{\tau \neq \emptyset} \theta_\tau X_\tau, \quad \theta = \{\theta_\tau\}_{\tau \neq \emptyset} \in \mathbb{S}^{n-1}.$$

Since $|X_\tau| = 1$ and thus $|X| = \sqrt{n}$, for the study of the asymptotic behavior of the L^2-distance $\omega(F_\theta, \Phi)$ on average, one may apply Theorem 1.1. Let Y be an independent copy of X, which we realize on the product space $\Omega^2 = \Omega \times \Omega$ with product measure $\mathbb{P}^2 = \mathbb{P} \times \mathbb{P}$ by

$$X_\tau(t, s) = \prod_{k \in \tau} t_k, \ Y_\tau(t, s) = \prod_{k \in \tau} s_k \quad t = (t_1, \dots, t_d), \ s = (s_1, \dots, s_d) \in \Omega.$$

Then the inner product

$$\langle X, Y\rangle \ = \sum_{\tau \neq \emptyset} X_\tau(t, s) Y_\tau(t, s) \ = \ -1 + \prod_{k=1}^{d} (1 + t_k s_k)$$

takes only two values, namely $2^d - 1$ in the case $t = s$, and -1 if $t \neq s$. Hence

$$\mathbb{E}\,\langle X, Y\rangle^3 \ = \ (2^d - 1)^3\, 2^{-d} + (1 - 2^{-d}) \ = \ \frac{n^3}{n+1} + \Big(1 - \frac{1}{n+1}\Big) \sim n^2$$

and

$$\mathbb{E}\,\langle X, Y\rangle^4 \ = \ (2^d - 1)^4\, 2^{-d} + (1 - 2^{-d}) \ = \ \frac{n^4}{n+1} + \Big(1 - \frac{1}{n+1}\Big) \sim n^3.$$

In other words, $m_3^3 \sim \sqrt{n}$ and $m_4^4 \sim n$ as $n \to \infty$. As a result, we may conclude that all inequalities in (13.1)–(13.3) are fulfilled for this system as well.

14.2. Here is another interesting example leading to the similar rate of normal approximation. Let $e_1, \dots, e_n$ denote the canonical basis in $\mathbb{R}^n$. Assuming that the random vector $X = (X_1, \dots, X_n)$ takes only n values, $\sqrt{n}\, e_1, \dots, \sqrt{n}\, e_n$, each with probability $1/n$, the linear form $\langle X, \theta\rangle$ also takes n values, namely, $\sqrt{n}\,\theta_1, \dots, \sqrt{n}\,\theta_n$, each with probability $1/n$, for any $\theta = (\theta_1, \dots, \theta_n) \in \mathbb{S}^{n-1}$. That is, as a measure, the distribution of $\langle X, \theta\rangle$ is described as

$$F_\theta = \frac{1}{n} \sum_{k=1}^{n} \delta_{\sqrt{n}\,\theta_k},$$

which may be viewed as an empirical measure based on the observations $Z_k = \sqrt{n}\,\theta_k$, $k = 1, \dots, n$. Each Z_k is almost standard normal, while jointly they are nearly independent (we have already considered in detail its characteristic functions $J_n(t\sqrt{n})$).

Just taking a short break, let us recall that when Z_k are indeed standard normal and independent, it is well-known that the empirical measures $G_n = \frac{1}{n}\sum_{k=1}^n \delta_{Z_k}$ approximate the standard normal law Φ with rate $1/\sqrt{n}$ with respect to the Kolmogorov distance. More precisely, $\mathbb{E}\, G_n = \Phi$ and there is a subgaussian deviation bound (cf. [22])

$$\mathbb{P}\big\{\sqrt{n}\,\rho(G_n, \Phi) \geq r\big\} \leq 2e^{-2r^2}, \qquad r \geq 0.$$

In particular, $\mathbb{E}\,\rho(G_n, \Phi) \leq \frac{c}{\sqrt{n}}$. Note that the characteristic function $g_n(t) = \frac{1}{n}\sum_{k=1}^n e^{itZ_k}$ of the measure G_n has mean $g(t) = e^{-t^2/2}$ and variance

$$\mathbb{E}\,|g_n(t) - g(t)|^2 = \frac{1}{n}\,\mathrm{Var}(e^{itZ_1}) = \frac{1}{n}\big(1 - |\mathbb{E}\, e^{itZ_1}|^2\big) = \frac{1}{n}\big(1 - e^{-t^2}\big).$$

Hence, applying Plancherel's theorem and using the identity (4.7) for the functions $\psi_r(\alpha)$ with $r = \alpha = 0$, we also have

$$\begin{aligned}\mathbb{E}\,\omega^2(G_n, \Phi) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathbb{E}\,\Big|\frac{g_n(t) - g(t)}{t}\Big|^2\,dt \\ &= \frac{1}{2\pi n}\int_{-\infty}^{\infty} \frac{1 - e^{-t^2}}{t^2}\,dt \;=\; \frac{1}{n\sqrt{\pi}}.\end{aligned}$$

Thus, on average the L^2-distance $\omega(G_n, \Phi)$ is of order $1/\sqrt{n}$ as well.

Similar properties may be expected for the random variables $Z_k = \sqrt{n}\,\theta_k$ and hence for the random vector X. Note that $|X| = \sqrt{n}$, while

$$\mathbb{E}\,\langle X, \theta\rangle^2 = \frac{1}{n}\sum_{k=1}^n (\sqrt{n}\,\theta_k)^2 = 1, \quad \theta \in \mathbb{S}^{n-1},$$

so that X is isotropic. We now involve an asymptotic formula of Corollary 5.1 which yields

$$\mathbb{E}_\theta\, \omega^2(F_\theta, \Phi) \,=\, \frac{1}{\sqrt{\pi}}\Big(1 + \frac{1}{4n}\Big)\,\mathbb{E}\,\Big(1 - (1-\xi)^{1/2}\Big) - \frac{1}{8n\sqrt{\pi}} + O\Big(\frac{1}{n^2}\Big),$$

where $\xi = \frac{\langle X, Y\rangle}{n}$ with Y being an independent copy of X. By the definition, ξ takes only two values, 1 with probability $\frac{1}{n}$ and 0 with probability $1 - \frac{1}{n}$. Hence, the last expectation is equal to $\frac{1}{n}$, and we get

$$\mathbb{E}_\theta\, \omega^2(F_\theta, \Phi) \,=\, \frac{7/8}{n\sqrt{\pi}} + O\Big(\frac{1}{n^2}\Big).$$

As for the Kolmogorov distance, one may apply again Theorem 1.3, which leads to the two-sided bound (13.3). Apparently, both logarithmic terms can be removed. Their appearance here is explained by the use of the Fourier tools (in the form of the Berry-Esseen bounds), while the proof of the Dvoretzky-Kiefer-Wolfowitz inequality on $\rho(G_n, \Phi)$ in [13] is based on the entirely different arguments.

15 Improved Rates for Lacunary Systems

An orthonormal sequence of random variables $\{X_k\}_{k=1}^\infty$ in $L^2(\Omega, \mathfrak{F}, \mathbb{P})$ is called a lacunary system of order $p > 2$, if for any sequence (a_k) in ℓ^2, the series $\sum_{k=1}^\infty a_k X_k$ converges in L^p-norm to an element of $L^p(\Omega, \mathfrak{F}, \mathbb{P})$. This property is equivalent to the validity of the Khinchine-type inequality

$$\big(\mathbb{E}\,|a_1 X_1 + \dots + a_n X_n|^p\big)^{1/p} \,\le\, M_p\,(a_1^2 + \dots + a_n^2)^{1/2} \tag{15.1}$$

for arbitrary $a_k \in \mathbb{R}$ with some constant M_p independent of n and the choice of the coefficients a_k. For basic properties of such systems we refer an interested reader to the books [16, 17].

Starting from an orthonormal lacunary system of order $p = 4$, consider the random vector $X = (X_1, \dots, X_n)$. According to Theorem 1.1, if $|X|^2 = n$ a.s. and $\mathbb{E}X = 0$, then

$$c\,\mathbb{E}_\theta\, \omega^2(F_\theta, \Phi) \,\le\, \frac{1}{n^3}\,\mathbb{E}\,\langle X, Y\rangle^3 + \frac{1}{n^4}\,\mathbb{E}\,\langle X, Y\rangle^4\,, \tag{15.2}$$

where Y is an independent copy of X. A similar bound

$$c\,\mathbb{E}_\theta\,\rho^2(F_\theta,\Phi)\ \le\ \frac{\log n}{n^3}\,\mathbb{E}\,\langle X,Y\rangle^3+\frac{(\log n)^2}{n^4}\,\mathbb{E}\,\langle X,Y\rangle^4 \tag{15.3}$$

also holds for the Kolmogorov distance. As easily follows from (15.1),

$$\mathbb{E}\,|\,\langle X,Y\rangle\,|^p\le M_p^{2p}n^{p/2}.$$

In particular,

$$\mathbb{E}\,|\,\langle X,Y\rangle\,|^3\ \le\ M_3^6\,n^{3/2},\quad \mathbb{E}\,\langle X,Y\rangle^4\ \le\ M_4^8\,n^2.$$

Hence, the bounds (15.2) and (15.3) lead to the estimates

$$c\,\mathbb{E}_\theta\,\omega^2(F_\theta,\Phi)\le\frac{1}{n^{3/2}}\,M_3^6+\frac{1}{n^2}\,M_4^8,$$
$$c\,\mathbb{E}_\theta\,\rho^2(F_\theta,\Phi)\le\frac{\log n}{n^{3/2}}\,M_3^6+\frac{(\log n)^2}{n^2}\,M_4^8.$$

Thus, if M_4 is bounded, both distances are at most of order $n^{-3/4}$ on average (modulo a logarithmic factor). Moreover, if

$$\Sigma_3(n)\ \equiv\ \mathbb{E}\,\langle X,Y\rangle^3\ =\sum_{1\le i_1,i_2,i_3\le n}\big(\mathbb{E}X_{i_1}X_{i_2}X_{i_3}\big)^2 \tag{15.4}$$

is bounded by a multiple of n, then these distances are on average at most $1/n$ (modulo a logarithmic factor in the case of ρ).

For an illustration, on the interval $\Omega=(-\pi,\pi)$ with the uniform measure $d\mathbb{P}(t)=\frac{1}{2\pi}\,dt$, consider a finite trigonometric system $X=(X_1,\dots,X_n)$ with components

$$X_{2k-1}(t)=\sqrt{2}\,\cos(m_k t),$$
$$X_{2k}(t)=\sqrt{2}\,\sin(m_k t),\qquad k=1,\dots,n/2,$$

where m_k are positive integers such that $\frac{m_{k+1}}{m_k}\ge q>1$ (assuming that n is even). Then X is an isotropic random vector satisfying $|X|^2=n$ and $\mathbb{E}X=0$, and with M_4 bounded by a function of q only. For evaluation of the moment $\Sigma_3(n)$, one may use the identities

$$\cos t=\mathbb{E}_\varepsilon\,e^{i\varepsilon t},\quad \sin t=\frac{1}{i}\,\mathbb{E}_\varepsilon\,\varepsilon\,e^{i\varepsilon t},$$

where ε is a Bernoulli random variable taking the values ± 1 with probability $\frac{1}{2}$. Let $\varepsilon_1, \varepsilon_2, \varepsilon_3$ be independent copies of ε. Using the property that $\varepsilon_1\varepsilon_3$ and $\varepsilon_2\varepsilon_3$ are independent, the first identity implies that, for all integers $1 \le n_1 \le n_2 \le n_3$,

$$\begin{aligned}
\mathbb{E}\, \cos(n_1 t)\cos(n_2 t)\cos(n_3 t) &= \mathbb{E}_\varepsilon\, \mathbb{E}\, \exp\{i(\varepsilon_1 n_1 + \varepsilon_2 n_2 + \varepsilon_3 n_3)\, t\} \\
&= \mathbb{E}_\varepsilon\, I\{\varepsilon_1 n_1 + \varepsilon_2 n_2 + \varepsilon_3 n_3 = 0\} \\
&= \mathbb{E}_\varepsilon\, I\{\varepsilon_1 n_1 + \varepsilon_2 n_2 = n_3\} \ = \ \frac{1}{4}\, I\{n_1 + n_2 = n_3\},
\end{aligned}$$

where $\mathbb{E}_\varepsilon$ means the expectation over $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$, and where $I\{A\}$ denotes the indicator of the event A. Similarly, involving also the identity for the sine function, we have

$$\begin{aligned}
\mathbb{E}\, \sin(n_1 t)\sin(n_2 t)\cos(n_3 t) &= -\mathbb{E}_\varepsilon\, \mathbb{E}\, \varepsilon_1\varepsilon_2\, \exp\{i(\varepsilon_1 n_1 + \varepsilon_2 n_2 + \varepsilon_3 n_3)\, t\} \\
&= -\mathbb{E}_\varepsilon\, \varepsilon_1\varepsilon_2\, I\{\varepsilon_1 n_1 + \varepsilon_2 n_2 + \varepsilon_3 n_3 = 0\} \\
&= -\mathbb{E}_\varepsilon\, \varepsilon_1\varepsilon_2\, I\{\varepsilon_1 n_1 + \varepsilon_2 n_2 = n_3\} \\
&= -\frac{1}{4}\, I\{n_1 + n_2 = n_3\},
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}\, \sin(n_1 t)\cos(n_2 t)\sin(n_3 t) &= -\mathbb{E}_\varepsilon\, \mathbb{E}\, \varepsilon_1\varepsilon_3\, \exp\{i(\varepsilon_1 n_1 + \varepsilon_2 n_2 + \varepsilon_3 n_3)\, t\} \\
&= -\mathbb{E}_\varepsilon\, \varepsilon_1\varepsilon_3\, I\{\varepsilon_1 n_1 + \varepsilon_2 n_2 + \varepsilon_3 n_3 = 0\} \\
&= -\mathbb{E}_\varepsilon\, \varepsilon_1\, I\{\varepsilon_1 n_1 + \varepsilon_2 n_2 = n_3\} \\
&= -\frac{1}{4}\, I\{n_1 + n_2 = n_3\},
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}\, \cos(n_1 t)\sin(n_2 t)\sin(n_3 t) &= -\mathbb{E}_\varepsilon\, \mathbb{E}\, \varepsilon_2\varepsilon_3\, \exp\{i(\varepsilon_1 n_1 + \varepsilon_2 n_2 + \varepsilon_3 n_3)\, t\} \\
&= -\mathbb{E}_\varepsilon\, \varepsilon_2\varepsilon_3\, I\{\varepsilon_1 n_1 + \varepsilon_2 n_2 + \varepsilon_3 n_3 = 0\} \\
&= -\mathbb{E}_\varepsilon\, \varepsilon_2\, I\{\varepsilon_1 n_1 + \varepsilon_2 n_2 = n_3\} \\
&= -\frac{1}{4}\, I\{n_1 + n_2 = n_3\}.
\end{aligned}$$

On the other hand, if the sine function appears in the product once or three times, such expectations will be vanishing. They are thus vanishing in all cases where $n_1 + n_2 \neq n_3$, and do not exceed $\frac{1}{4}$ in absolute value for any combination of sine and cosine terms in all cases with $n_1 + n_2 = n_3$. Therefore, the moment $\Sigma_3(n)$ in (15.4) is bounded by a multiple of

$$T_3(n) \ = \ \mathrm{card}\big\{(i_1, i_2, i_3) : 1 \le i_1 \le i_2 < i_3 \le n,\ m_{i_1} + m_{i_2} = m_{i_3}\big\}.$$

One can now involve the lacunary assumption. If $q \geq 2$, the property $i_1 \leq i_2 < i_3$ implies $m_{i_1} + m_{i_2} < m_{i_3}$, so that $T_3(n) = \Sigma_3(n) = 0$. In the case $1 < q < 2$, define A_q to be the (finite) collection of all couples (k_1, k_2) of positive integers such that

$$q^{-k_1} + q^{-k_2} \geq 1.$$

By the lacunary assumption, if $1 \leq i_1 \leq i_2 < i_3 \leq n$, we have

$$m_{i_1} + m_{i_2} \leq \big(q^{-(i_3 - i_1)} + q^{-(i_3 - i_2)}\big)\, m_{i_3} < m_{i_3},$$

as long as the couple $(i_3 - i_1, i_2 - i_1)$ is not in A_q. Hence,

$$\begin{aligned} T_3(n) &\leq \mathrm{card}\big\{(i_1, i_2, i_3) : 1 \leq i_1 \leq i_2 < i_3 \leq n,\ (i_3 - i_1, i_2 - i_1) \in A_q\big\} \\ &\leq n\, \mathrm{card}(A_q) \leq c_q n \end{aligned}$$

with constant depending on q only. Returning to (15.2) and (15.3), we then obtain:

Proposition 15.1 *For the lacunary trigonometric system X of an even length n and with parameter $q > 1$, we have*

$$\mathbb{E}_\theta\, \omega^2(F_\theta, \Phi) \leq \frac{c_q}{n^2}, \quad \mathbb{E}_\theta\, \rho^2(F_\theta, \Phi) \leq \frac{c_q\, (\log n)^2}{n^2},$$

where the constants c_q depend q only.

In this connection one should mention a classical result of Salem and Zygmund concerning distributions of the lacunary sums

$$S_n = \sum_{k=1}^{n} (a_k \cos(m_k t) + b_k \sin(m_k t))$$

with an arbitrary prescribed sequence of the coefficients $(a_k)_{k \geq 1}$ and $(b_k)_{k \geq 1}$. Assume that $\frac{m_{k+1}}{m_k} \geq q > 1$ for all k and put

$$v_n^2 = \frac{1}{2} \sum_{k=1}^{n} (a_k^2 + b_k^2) \qquad (v_n \geq 0),$$

so that the normalized sums $Z_n = S_n / v_n$ have mean zero and variance one under the measure $\mathbb{P}$. It was shown in [25] that Z_n are weakly convergent to the standard normal law, i.e., their distributions F_n under $\mathbb{P}$ satisfy $\rho(F_n, \Phi) \to 0$ as $n \to \infty$, if and only if $\frac{a_n^2 + b_n^2}{v_n^2} \to 0$ (in fact, the weak convergence was established on every subset of Ω of positive measure).

Restricting to the coefficients $\theta_{2k-1} = a_k / v_n$, $\theta_{2k} = b_k / v_n$, Salem-Zygmund's theorem may be stated as the assertion that $\rho(F_\theta, \Phi)$ is small, if and only if $\|\theta\|_\infty =$

$\max_{1\le k\le n}|\theta_k|$ is small. The latter condition naturally appears in the central limit theorem for weighted sums of independent identically distributed random variables. Thus, Proposition 15.1 complements this result in terms of the rate of convergence in the mean on the unit sphere. It would be interesting to describe explicit coefficients θ_k, for which we get a standard rate of normal approximation (perhaps, using other approaches such as the Stein method, cf. e.g. [14]).

The result of [25] was generalized in [26]; it turns out there is no need to assume that all m_k are integers, and the asymptotic normality is preserved for real m_k such that $\inf_k \frac{m_{k+1}}{m_k} > 1$. However, in this more general situation, the rate $1/n$ as in Proposition 15.1 is no longer true (although the rate $1/\sqrt{n}$ is valid). The main reason is that the means

$$\mathbb{E}X_{2k-1} = \sqrt{2}\,\mathbb{E}\cos(m_k t) = \sqrt{2}\,\frac{\sin(\pi m_k)}{\pi m_k}$$

may be non-zero. For example, choosing $m_k = 2^k + \frac{1}{2}$, we obtain an orthonormal system with $\mathbb{E}X_{2k} = 0$, while

$$\mathbb{E}X_{2k-1} = \frac{2\sqrt{2}}{\pi\,(2^{k+1}+1)}.$$

Hence

$$\mathbb{E}\,\langle X,Y\rangle = |\mathbb{E}X|^2 = \frac{8}{\pi^2}\sum_{k=1}^{n}\frac{1}{(2^{k+1}+1)^2} \to c \qquad (n\to\infty)$$

for some absolute constant $c > 0$ (where Y is an independent copy of X). In this situation, as was already mentioned in (5.3), cf. Remark 5.3, we have a lower bound

$$\mathbb{E}_\theta\,\omega^2(F_\theta,F) \ge \frac{c}{2\sqrt{\pi}\,n} + O\Big(\frac{1}{n^2}\Big).$$

Since $\mathbb{E}\,\langle X,Y\rangle^3 = O(n)$ and $\mathbb{E}\,\langle X,Y\rangle^4 = O(n^2)$, this inequality may actually be replaced with equality, according to (5.2). A similar asymptotic holds as well when F is replaced with Φ.

16 Improved Rates for Independent and Log-Concave Summands

Let $X = (X_1,\dots,X_n)$ be an isotropic random vector in $\mathbb{R}^n$ with mean zero. If the components X_k are independent, the normal approximation for the distributions F_θ of the weighted sums

$$S_\theta = \theta_1 X_1 + \dots + \theta_n X_n, \quad \theta\in\mathbb{S}^{n-1},$$

may be controlled by virtue of the Berry-Esseen theorem under the 3-rd moment assumption. Namely, this theorem provides an upper bound

$$\rho(F_\theta, \Phi) \le c \sum_{i=1}^{n} |\theta_i|^3 \, \mathbb{E}\, |X_i|^3 \tag{16.1}$$

(cf. e.g. [23, 24]). Since $\mathbb{E}\, |X_i|^3 \ge 1$, the sum in (16.1) is at least $\frac{1}{\sqrt{n}}$. On the other hand, (16.1) yields an upper estimate on average

$$\mathbb{E}_\theta\, \rho(F_\theta, \Phi) \le \frac{c\beta_3}{\sqrt{n}}, \quad \beta_3 = \max_{1 \le i \le n} \mathbb{E}\, |X_i|^3, \tag{16.2}$$

which is consistent with the standard rate.

As it turns out, the relations (16.1) and (16.2) are far from being optimal for most of θ, as the following statement due to Klartag and Sodin shows.

Theorem 16.1 ([20]) *If the random variables $X_1, \dots, X_n$ are independent, have mean zero, variance one, and finite 4-th moments, then*

$$\mathbb{E}_\theta\, \rho(F_\theta, \Phi) \le \frac{c\beta_4}{n}, \quad \beta_4 = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E} X_i^4. \tag{16.3}$$

Moreover, for any $r \ge 0$,

$$\mathfrak{s}_{n-1}\big\{n\rho(F_\theta, \Phi) \ge c\beta_4 r\big\} \le 2\, e^{-\sqrt{r}}.$$

In the i.i.d. case, $\beta_4 = \mathbb{E} X_1^4$, and we obtain an upper bound of order at most $1/n$.

In fact, in the i.i.d. case, the relation (16.3) may be further sharpened under the 5-th moment assumption, if $\mathbb{E} X_1^3 = 0$, and if $\Phi(x)$ is slightly modified to

$$G(x) = \Phi(x) - \frac{\beta_4 - 3}{8(n+2)}\, (x^3 - 3x)\, \varphi(x), \quad x \in \mathbb{R},$$

where $\varphi(x) = \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2}$ is the standard normal density.

Theorem 16.2 *If the random variables $X_1, \dots, X_n$ are independent, identically distributed, and have moments $\mathbb{E} X_1 = 0$, $\mathbb{E} X_1^2 = 1$, $\mathbb{E} X_1^3 = 0$, $\mathbb{E} X_1^4 = \beta_4$, $\mathbb{E}\, |X_1|^5 = \beta_5 < \infty$, then*

$$\mathbb{E}_\theta\, \rho(F_\theta, G) \le \frac{c\beta_5}{n^{3/2}}. \tag{16.4}$$

Moreover, for any $r \geq 0$,

$$\mathfrak{s}_{n-1}\Big\{n^{3/2}\rho(F_\theta, G) \geq c\beta_4 r\Big\} \leq 2\,\exp\{-r^{2/5}\}.$$

We refer an interested reader to [4] and [11]. In the i.i.d. case, both inequalities (16.3) and (16.4) are sharp in the following sense. If $\alpha_3 = \mathbb{E}X_1^3 \neq 0$ and $\beta_4 < \infty$, then, for any function G of bounded total variation, such that $G(-\infty) = 0$ and $G(\infty) = 1$, we have

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

with a constant $c > 0$ depending on α_3 and β_4. Similarly, if $\alpha_3 = 0$, $\beta_4 \neq 3$, $\beta_5 < \infty$, then

$$\mathbb{E}_\theta\, \rho(F_\theta, G) \geq \frac{c}{n^{3/2}},$$

where the constant $c > 0$ depends on β_4 and β_5 only.

In the upper bounds such as (16.3), the independence assumption may be replaced with closely related hypotheses. The random vector X is said to have a log-concave distribution, when it has a density of the form $p(x) = e^{-V(x)}$ where $V : \mathbb{R}^n \to (-\infty, \infty]$ is a convex function. Recall that the distribution of X is coordinatewise symmetric, if

$$p(\varepsilon_1 x_1, \dots, \varepsilon_n x_n) = p(x_1, \dots, x_n), \quad x_i \in \mathbb{R},$$

for any choice of signs $\varepsilon_i = \pm 1$. The following theorem sharpening (16.1) is due to Klartag.

Theorem 16.3 ([18]) *Suppose that the isotropic random vector $X = (X_1, \dots, X_n)$ in $\mathbb{R}^n$ has a coordinatewise symmetric log-concave distribution. For all $\theta = (\theta_1, \dots, \theta_n) \in \mathbb{S}^{n-1}$,*

$$\|F_\theta - \Phi\|_{\mathrm{TV}} \leq c \sum_{i=1}^n \theta_i^4. \tag{16.5}$$

Here, the total variation distance is understood in the usual sense as

$$\|F_\theta - \Phi\|_{\mathrm{TV}} = \int_{-\infty}^{\infty} |p_\theta(x) - \varphi(x)|\, dx,$$

where p_θ denotes the density of S_θ. By the assumptions, p_θ is symmetric about the origin and is log-concave for any $\theta \in \mathbb{S}^{n-1}$. Note that, by the coordinatewise symmetry, the isotropy assumption is reduced to the moment condition $\mathbb{E}X_i^2 = 1$ $(1 \leq i \leq n)$.

In particular, it follows from (16.5) that

$$\mathbb{E}_\theta\, \rho(F_\theta, \Phi) \le \mathbb{E}_\theta\, \|F_\theta - \Phi\|_{\mathrm{TV}} \le \frac{c}{n}. \tag{16.6}$$

17 Improved Rates Under Correlation-Type Conditions

Up to a logarithmically growing term, the improved rate as in the upper bound (16.3) can be achieved under more flexible correlation-type conditions (in comparison with independence). For example, one may consider an optimal value $\Lambda = \Lambda(X)$ in the relation

$$\mathrm{Var}\bigg(\sum_{i,j=1}^n a_{ij} X_i X_j\bigg) \le \Lambda \sum_{i,j=1}^n a_{ij}^2 \qquad (a_{ij} \in \mathbb{R}), \tag{17.1}$$

which we call that the random vector $X = (X_1, \dots, X_n)$ satisfies a second order correlation condition with constant Λ. This quantity is finite as long as the moment $\mathbb{E}\,|X|^4$ is finite.

To relate Λ to the moment-type characteristics which we discussed before, one may apply (17.1) with $a_{ij} = \delta_{ij}$ or (as another option) with $a_{ij} = \theta_i \theta_j$, $\theta = (\theta_1, \dots, \theta_n) \in \mathbb{S}^{n-1}$. This gives that

$$\sigma_4^2 \le \Lambda, \quad m_4^2 \le \sup_{\theta \in \mathbb{S}^{n-1}} \mathbb{E} S_\theta^4 \le 1 + \Lambda,$$

where in the last inequality we should assume that $\mathbb{E} S_\theta^2 = 1$ for all θ (i.e. X is isotropic). In the latter case, necessarily $\Lambda \ge \frac{n-1}{n}$, so that Λ is bounded away from zero.

If the distribution of X is "regular" in some sense, one may also bound Λ from above. For example, this is the case when it shares a Poincaré-type inequality

$$\lambda_1 \mathrm{Var}(u(X)) \le \mathbb{E}\,|\nabla u(X)|^2, \tag{17.2}$$

which is required to hold in the class of all bounded, smooth functions u on $\mathbb{R}^n$ with a constant $\lambda_1 > 0$ independent of u (called the spectral gap). We then have

$$\Lambda \le \frac{4}{\lambda_1^2}, \quad \Lambda \le \frac{4}{\lambda_1}, \tag{17.3}$$

where in the second inequality we assume that X is isotropic.

The following relation is established in [9].

Theorem 17.1 *If the distribution of X is isotropic and symmetric about the origin, then*

$$\mathbb{E}_\theta\, \rho(F_\theta, \Phi) \le c\Lambda\, \frac{\log n}{n}. \tag{17.4}$$

The proof is based on the second order spherical concentration phenomenon which was developed in [6] with the aim of applications to randomized central limit theorems. It indicates that the deviations of any smooth function $u(\theta)$ on $\mathbb{S}^{n-1}$ from the mean $\mathbb{E}_\theta u(\theta)$ are at most of the order $1/n$, provided that u is orthogonal in $L^2(\mathbb{R}^n, \mathfrak{s}_{n-1})$ to all linear functions and has a "bounded" Hessian (the matrix of second order partial derivatives). Being applied to the characteristic functions $u(\theta) = f_\theta(t)$, this property yields an upper bound

$$\mathbb{E}_\theta\, |f_\theta(t) - f(t)|^2 \ \le\ \frac{c\Lambda t^4}{n^2}$$

on every interval $|t| \le An^{1/5}$ with constants $c > 0$ depending on the parameter $A \ge 1$ only. This estimate can be used to bound the integrals in (8.4) to get a similar variant of (17.4).

The symmetry hypothesis in Theorem 17.1 may be dropped, if Λ is replaced by λ_1^{-1} which is a larger quantity according to (17.3). In addition, one can control large deviations of the distance $\rho(F_\theta, \Phi)$ for most of the directions θ (rather than on average). The corresponding assertions are obtained in [10].

Theorem 17.2 *Let X be an isotropic random vector in* $\mathbb{R}^n$ *with mean zero and a positive Poincaré constant* λ_1*. Then*

$$\mathbb{E}_\theta\, \rho(F_\theta, \Phi) \ \le\ c\lambda_1^{-1}\, \frac{\log n}{n}. \tag{17.5}$$

Moreover, for all $r > 0$*,*

$$\mathfrak{s}_{n-1}\Big\{\rho(F_\theta, \Phi) \ge c\lambda_1^{-1}\,\frac{\log n}{n}\, r\Big\} \ \le\ 2\, e^{-\sqrt{r}}.$$

The logarithmic term in (17.5) may be removed using the less sensitive L^2-distance:

$$\mathbb{E}_\theta\, \omega^2(F_\theta, \Phi) \le \frac{c}{\lambda_1^2 n^2}.$$

There is an extensive literature devoted to bounding the spectral gap λ_1 from below. In particular, it is positive for any log-concave probability distribution on $\mathbb{R}^n$. A well-known conjecture raised by Kannan, Lovász and Simonovits asserts that λ_1 is actually bounded away from zero, as long as the random vector X has an isotropic log-concave distribution (cf. [15]). The best known dimensional lower bound up to date is due to Klartag and Lehec [19] who showed that

$$\lambda_1 \geq \frac{c}{(\log n)^\alpha}$$

for some absolute positive constants c and α (one may take $\alpha = 10$). Applying this bound in Theorem 17.2, we therefore obtain:

Corollary 17.3 *Let X be an isotropic random vector in $\mathbb{R}^n$ with mean zero and a log-concave probability distribution. Then with some absolute positive constants c and α*

$$\mathbb{E}_\theta\, \rho(F_\theta, \Phi) \,\leq\, \frac{c(\log n)^\alpha}{n}. \tag{17.6}$$

Thus, there is a certain extension of Klartag's bound (16.6) at the expense of a logarithmic factor to the entire class of isotropic log-concave probability distributions on $\mathbb{R}^n$.

One may also argue in the opposite direction: upper bounds of the form

$$\mathbb{E}_\theta\, \rho(F_\theta, \Phi) \leq \frac{c(\log n)^\beta}{n}, \quad \beta > 0,$$

in the class of log-concave probability distributions on $\mathbb{R}^n$ imply lower bounds $\lambda_1 \geq c\,(\log n)^{-\beta'}$ with some $\beta' > 0$, cf. [9].

Acknowledgments Research was supported by SFB 1283, BSF grant 2016050, and NSF grant DMS-2154001.

References

1. M. Anttila, K. Ball, I. Perissinaki, The central limit problem for convex bodies. Trans. Am. Math. Soc. **355**(12), 4723–4735 (2003)
2. S.G. Bobkov, On concentration of distributions of random weighted sums. Ann. Probab. **31**(1), 195–215 (2003)
3. S.G. Bobkov, Closeness of probability distributions in terms of Fourier-Stieltjes transforms. Russ. Math. Surv. **71**(6), 1021–1079 (2016). Translated from: Uspekhi Matem. Nauk, vol. 71, issue 6 (432), (2016), 37–98
4. S.G. Bobkov, Edgeworth corrections in randomized central limit theorems. Geom. Aspects Funct. Anal. **2256**, 71–97 (2020)
5. S.G. Bobkov, F. Götze, Concentration inequalities and limit theorems for randomized sums. Probab. Theory Relat. Fields **137**(1–2), 49–81 (2007)
6. S.G. Bobkov, G.P. Chistyakov, F. Götze, Second-order concentration on the sphere. Commun. Contemp. Math. **19**(5), 1650058, 20pp. (2017)
7. S.G. Bobkov, G.P. Chistyakov, F. Götze, 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; reprinted in J. Math. Sci. (N.Y.) 238 (2019), no. 4, 366–376
8. S.G. Bobkov, G.P. Chistyakov, F. Götze, Berry-Esseen bounds for typical weighted sums. J. Electron. Probab. **23**(92), 1–22 (2018)
9. S.G. Bobkov, G.P. Chistyakov, F. Götze, Normal approximation for weighted sums under a second order correlation condition. Ann. Probab. **48**(3), 1202–1219 (2020)
10. S.G. Bobkov, G.P. Chistyakov, F. Götze, Poincaré-type inequalities and normal approximation for weighted sums. Electron. J. Probab. **25**, 155, 31 pp. (2020)
11. S.G. Bobkov, G.P. Chistyakov, F. Götze, *Concentration and Gaussian Approximation for Randomized Sums*. Probability Theory and Stochastic Modelling, vol. 104 (Springer Cham, 2023), 434 pp.
12. S. Brazitikos, A. Giannopoulos, P. Valettas, B.-H. Vritsiou, *Geometry of Isotropic Convex Bodies*. Mathematical Surveys and Monographs, vol. 196 (American Mathematical Society, Providence, 2014). xx+594pp.
13. A. Dvoretzky, J. Kiefer, J. Wolfowitz, Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. Ann. Math. Stat. **27**, 642–669 (1956)
14. L. Goldstein, G. Reinert, Stein's method and the zero bias transformation with application to simple random sampling. Ann. Appl. Probab. **7**(4), 935–952 (1997)
15. R. Kannan, L. Lovász, M. Simonovits, Isoperimetric problems for convex bodies and a localization lemma. Discrete Comput. Geom. **13**, 541–559 (1995)
16. S. Kaczmarz, G. Steinhaus, *Theory of Orthogonal Series* (Warszawa, Lwow, 1935); Russian ed.: Izdat. Fiz.-Mat. Lit., Moscow, 1958, 507pp.
17. B.S. Kashin, A.A Saakyan, *Orthogonal Series*. Translated from the Russian by Ralph P. Boas. Translation edited by Ben Silver. Translations of Mathematical Monographs, vol. 75 (American Mathematical Society, Providence, 1989), xii+451pp.
18. B. Klartag, A Berry-Esseen type inequality for convex bodies with an unconditional basis. Probab. Theory Relat. Fields **145**(1–2), 1–33 (2009)
19. B. Klartag, J. Lehec, Bourgain's slicing problem and KLS isoperimetry up to polylog (2022). arXiv:2203.15551v2
20. B. Klartag, S. Sodin, Variations on the Berry-Esseen theorem. Teor. Veroyatn. Primen. **56**(3), 514–533 (2011); Reprinted in: Theory Probab. Appl. 56 (2012), no. 3, 403–419
21. M. Ledoux, Concentration of measure and logarithmic Sobolev inequalities, in *Séminaire de Probabilités XXXIII*. Lecture Notes in Mathematics, vol. 1709 (Springer, Berlin, 1999), pp. 120–216
22. P. Massart, The tight constant in the Dvoretzky-Kiefer-Wolfowitz inequality. Ann. Probab. **18**(3), 1269–1283 (1990)
23. V.V. Petrov, *Sums of Independent Random Variables*. Translated from the Russian by A. A. Brown. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 82 (Springer, New York, 1975), x+346pp.
24. V.V. Petrov, *Limit Theorems for Sums of Independent Random Variables* (in Russian) (Nauka, Moscow, 1987), 318pp.
25. R. Salem, A. Zygmund, On lacunary trigonometric systems. Proc. Nat. Acad. Sci. USA **33**, 333–338 (1947)
26. R. Salem, A. Zygmund, On lacunary trigonometric series. II. Proc. Nat. Acad. Sci. USA **34**, 54–62 (1948)
27. V.N. Sudakov, Typical distributions of linear functionals in finite-dimensional spaces of high dimension (in Russian). Soviet Math. Dokl. **19**, 1578–1582 (1978); translation in: Dokl. Akad. Nauk SSSR, 243 (1978), no. 6, 1402–1405