The Annals of Probability
2014, Vol. 42, No. 1, 354–397
DOI: 10.1214/13-AOP839

EXPLICIT RATES OF APPROXIMATION IN THE CLT FOR QUADRATIC FORMS

BY FRIEDRICH GÖTZE[1] AND ANDREI YU. ZAITSEV[1,2]

Universität Bielefeld and St. Petersburg Department of Steklov Mathematical Institute

Let $X, X_1, X_2, \dots$ be i.i.d. $\mathbb{R}^d$-valued real random vectors. Assume that $\mathbf{E}X = 0$, $\operatorname{cov} X = \mathbb{C}$, $\mathbf{E}\|X\|^2 = \sigma^2$ and that X is not concentrated in a proper subspace of $\mathbb{R}^d$. Let G be a mean zero Gaussian random vector with the same covariance operator as that of X. We study the distributions of nondegenerate quadratic forms $\mathbb{Q}[S_N]$ of the normalized sums $S_N = N^{-1/2}(X_1 + \cdots + X_N)$ and show that, without any additional conditions,

$$\Delta_N \stackrel{\text{def}}{=} \sup_x \bigl|\mathbf{P}\{\mathbb{Q}[S_N] \le x\} - \mathbf{P}\{\mathbb{Q}[G] \le x\}\bigr| = \mathcal{O}(N^{-1}),$$

provided that $d \ge 5$ and the fourth moment of X exists. Furthermore, we provide explicit bounds of order $\mathcal{O}(N^{-1})$ for Δ_N for the rate of approximation by short asymptotic expansions and for the concentration functions of the random variables $\mathbb{Q}[S_N + a]$, $a \in \mathbb{R}^d$. The order of the bound is optimal. It extends previous results of Bentkus and Götze [*Probab. Theory Related Fields* **109** (1997a) 367–416] (for $d \ge 9$) to the case $d \ge 5$, which is the smallest possible dimension for such a bound. Moreover, we show that, in the finite dimensional case and for isometric $\mathbb{Q}$, the implied constant in $\mathcal{O}(N^{-1})$ has the form $c_d \sigma^d (\det \mathbb{C})^{-1/2} \mathbf{E}\|\mathbb{C}^{-1/2} X\|^4$ with some c_d depending on d only. This answers a long standing question about optimal rates in the central limit theorem for quadratic forms starting with a seminal paper by Esséen [*Acta Math.* **77** (1945) 1–125].

1. Introduction. Let $\mathbb{R}^d$ be the d-dimensional space of real vectors $x = (x_1, \dots, x_d)$ with scalar product $\langle x, y\rangle = x_1 y_1 + \cdots + x_d y_d$ and norm $\|x\| = \langle x, x\rangle^{1/2}$. We also denote by $\mathbb{R}^\infty$ a separable Hilbert space consisting of all real sequences $x = (x_1, x_2, \dots)$ such that $\|x\|^2 = x_1^2 + x_2^2 + \cdots < \infty$.

Let $X, X_1, X_2, \dots$ be a sequence of i.i.d. $\mathbb{R}^d$-valued random vectors. Assume that $\mathbf{E}X = 0$ and $\sigma^2 \stackrel{\text{def}}{=} \mathbf{E}\|X\|^2 < \infty$. Let G be a mean zero Gaussian random

Received December 2011; revised February 2013.
[1]Supported by the SFB 701 in Bielefeld and by Grant RFBR-DFG 09-01-91331.
[2]Supported by Grants NSh-1216.2012.01, RFBR 09-01-12180, 10-01-00242 and 11-01-12104, by the Alexander von Humboldt Foundation and by a program of fundamental researches of Russian Academy of Sciences “Modern problems of fundamental mathematics.”
MSC2010 subject classifications. Primary 60F05; secondary 62E20.
Key words and phrases. Central Limit theorem, concentration functions, convergence rates, multidimensional spaces, quadratic forms, ellipsoids, hyperboloids, lattice point problem, theta-series.

vector such that its covariance operator $\mathbb{C} = \operatorname{cov} G : \mathbb{R}^d \to \mathbb{R}^d$ is equal to $\operatorname{cov} X$. It is well known that the distributions $\mathcal{L}(S_N)$ of sums

$$S_N \overset{\text{def}}{=} N^{-1/2}(X_1 + \cdots + X_N) \tag{1.1}$$

converge weakly to $\mathcal{L}(G)$.

Let $\mathbb{Q} : \mathbb{R}^d \to \mathbb{R}^d$ be a linear symmetric bounded operator, and let $\mathbb{Q}[x] = \langle \mathbb{Q}x, x\rangle$ be the corresponding quadratic form. We say that $\mathbb{Q}$ is nondegenerate if $\ker \mathbb{Q} = \{0\}$.

Denote, for $q > 0$,

$$\beta_q \overset{\text{def}}{=} \mathbf{E}\|X\|^q, \qquad \beta \overset{\text{def}}{=} \beta_4.$$

Introduce the distribution functions

$$F(x) \overset{\text{def}}{=} \mathbf{P}\{\mathbb{Q}[S_N] \le x\}, \qquad H(x) \overset{\text{def}}{=} \mathbf{P}\{\mathbb{Q}[G] \le x\}. \tag{1.2}$$

Write

$$\Delta_N \overset{\text{def}}{=} \sup_{x \in \mathbb{R}} |F(x) - H(x)|. \tag{1.3}$$

THEOREM 1.1. *Assume that $\mathbb{Q}$ and $\mathbb{C}$ are nondegenerate and that $d \ge 5$ or $d = \infty$. Then*

$$\Delta_N \le c(\mathbb{Q}, \mathbb{C})\beta/N.$$

The constant $c(\mathbb{Q}, \mathbb{C})$ in this bound depends on $\mathbb{Q}$ and $\mathbb{C}$ only.

THEOREM 1.2. *Let the conditions of Theorem 1.1 be satisfied, and let $5 \le d < \infty$. Assume that the operator $\mathbb{Q}$ is isometric. Then*

$$\Delta_N \le c_d \sigma^d (\det \mathbb{C})^{-1/2} \mathbf{E}\|\mathbb{C}^{-1/2} X\|^4/N.$$

The constant c_d in this bound depends on d only.

Theorems 1.1 and 1.2 are simple consequences of the main result of this paper, Theorem 2.2; see also Theorem 2.1. Theorem 1.1 was proved in Götze and Zaitsev (2008). It confirms a conjecture of Bentkus and Götze (1997a) [below BG (1997a)]. It generalizes to the case $d \ge 5$ the corresponding result of BG (1997a). In their Theorem 1.1, it was assumed that $d \ge 9$, while our Theorem 1.1 is proved for $d \ge 5$. Theorem 1.2 yields an explicit bound in terms of the distribution $\mathcal{L}(X)$.

The distribution function of $\|S_N\|^2$ (for bounded X with values in $\mathbb{R}^d$) may have jumps of order $\mathcal{O}(N^{-1})$, for all $1 \le d \le \infty$; see, for example, BG [(1997a), page 468]. Therefore, the bounds of Theorems 1.1 and 1.2 are optimal with respect to the order in N.

Theorems 1.1, 1.2 and the method of their proof are closely related to the lattice point problem in number theory. Suppose that $d < \infty$ and that $\langle \mathbb{Q}x, x\rangle > 0$, for $x \neq 0$. Let $\operatorname{vol} E_r$ be the volume of the ellipsoid

$$E_r = \{x \in \mathbb{R}^d : \mathbb{Q}[x] \le r^2\} \qquad \text{for } r \ge 0.$$

Write $\operatorname{vol}_{\mathbb{Z}} E_r$ for the number of points in $E_r \cap \mathbb{Z}^d$, where $\mathbb{Z}^d \subset \mathbb{R}^d$ is the standard lattice of points with integer coordinates.

The following result due to Götze (2004) is related to Theorems 1.1 and 1.2; see also BG (1995a, 1997b).

THEOREM 1.3. *For all dimensions $d \ge 5$,*

$$\sup_{a\in\mathbb{R}^d} \left| \frac{\operatorname{vol}_{\mathbb{Z}}(E_r + a) - \operatorname{vol} E_r}{\operatorname{vol} E_r} \right| = \mathcal{O}(r^{-2}) \qquad \textit{for } r \ge 1,$$

where the constant in $\mathcal{O}(r^{-2})$ depends on the dimension d and on the lengths of axes of the ellipsoid E_1 only.

Theorem 1.3 solves the lattice point problem for $d \ge 5$. It improves the classical estimate $\mathcal{O}(r^{-2d/(d+1)})$ due to Landau (1915), just as Theorem 1.1 improves the bound $\mathcal{O}(N^{-d/(d+1)})$ by Esséen (1945) in the CLT for ellipsoids with axes parallel to coordinate axes. A related result for indefinite forms may be found in Götze and Margulis (2010).

Work on the estimation of the rate of approximation under the conditions of Theorem 1.1 for Hilbert spaces started in the second half of the last century. See Zalesskiĭ, Sazonov and Ulyanov (1988) and Nagaev (1989) for optimal bounds of order $\mathcal{O}(N^{-1/2})$ (with respect to eigenvalues of $\mathbb{C}$) assuming finiteness of the third moment. For a more detailed discussion see Yurinskii (1982), Zalesskiĭ, Sazonov and Ulyanov (1991), Bentkus, Götze, Paulauskas and Račkauskas (1991), BG (1995b, 1996, 1997a) and Senatov (1997, 1998).

Under some more restrictive moment and dimension conditions the estimate of order $\mathcal{O}(N^{-1+\varepsilon})$, with $\varepsilon \downarrow 0$ as $d \uparrow \infty$, was obtained by Götze (1979). The proof in Götze (1979) was based on a new symmetrization inequality for characteristic functions of quadratic forms. This inequality is related to Weyl's (1916) inequality for trigonometric sums. This inequality and its extensions (see Lemma 6.1) play a crucial role in the proofs of bounds in the CLT for ellipsoids and hyperboloids in finite and infinite dimensional cases. Under some additional smoothness assumptions, error bounds $\mathcal{O}(N^{-1})$ (and, moreover, Edgeworth type expansions) were obtained in Götze (1979), Bentkus (1984), Bentkus, Götze and Zitikis (1993). BG (1995b, 1996, 1997a) established the bound of order $\mathcal{O}(N^{-1})$ without smoothness-type conditions. Similar bounds for the rate of infinitely divisible approximations were obtained by Bentkus, Götze and Zaitsev (1997). Among recent publications, we should mention the papers of Nagaev and Chebotarev (1999, 2005) ($d \ge 13$, providing a more precise dependence of constants on the eigenvalues of $\mathbb{C}$) and Bogatyrev, Götze and Ulyanov (2006) (nonuniform bounds for $d \geq 12$); see also Götze and Ulyanov (2000). The proofs of bounds of order $\mathcal{O}(N^{-1})$ are based on discretization (i.e., a reduction to lattice valued random vectors) and the symmetrization techniques mentioned above.

Assuming the matrices $\mathbb{Q}$ and $\mathbb{C}$ to be diagonal, and the independence of the first five coordinates of X, BG (1996) have already reduced the dimension requirement for the bound $\mathcal{O}(N^{-1})$ to $d \geq 5$. The independence assumption in BG (1996) allowed to apply an adaption of the Hardy–Littlewood circle method. For the general case considered in Theorem 1.1, one needs to develop new techniques. Some yet unpublished results of Götze (1994) provide the rate $\mathcal{O}(N^{-1})$ for sums of two independent *arbitrary* quadratic forms (each of rank $d \geq 3$). Götze and Ulyanov (2003) obtained bounds of order $\mathcal{O}(N^{-1})$ for some ellipsoids in $\mathbb{R}^d$ with $d \geq 5$ in the case of lattice distributions of X.

The optimal possible dimension condition for this rate is just $d \geq 5$, due to the lower bounds of order $\mathcal{O}(N^{-1} \log N)$ for dimension $d = 4$ in the corresponding lattice point problem. The question about precise convergence rates in dimensions $2 \leq d \leq 4$ still remains completely open (even in the simplest case where $\mathbb{Q}$ is the identity operator $\mathbb{I}_d$, and for random vectors with independent Rademacher coordinates). It should be mentioned that, in the case $d = 2$, a precise convergence rate would imply a solution of the famous circle problem. Known lower bounds in the circle problem correspond to the bound of order $\mathcal{O}(N^{-3/4} \log^\delta N)$, $\delta > 0$, for Δ_N. Hardy (1916) conjectured that up to logarithmic factors this is the optimal order.

Now we describe the most important elements of the proof. We have to mention that a big part of the proof repeats the arguments of BG (1997a); see BG (1997a) for the description and application of symmetrization inequality and discretization procedure. In our proof we do not use the multiplicative inequalities of BG (1997a). Here we replace those techniques by arguments from the geometry of numbers, developed in Götze (2004), combined with effective equidistribution results by Götze and Margulis (2010) for suitable actions of unipotent subgroups of $\mathrm{SL}(2, \mathbb{R})$; see Lemma 8.2. These new techniques (compared to previous results) are mainly concentrated in Sections 5–8.

Using the Fourier inversion formula [see (4.2) and (4.3)], we have to estimate some integrals of the absolute values of differences of characteristic functions of quadratic forms. In Section 6, we reduce the estimation of characteristic functions to the estimation of a theta-series; see Lemma 6.5 and inequality (6.28). To this end, we write the expectation with respect to Rademacher random variables as a sum with binomial weights $p(m)$ and $p(\overline{m})$. Then we estimate $p(m)$ and $p(\overline{m})$ from above by discrete Gaussian exponential weights $c_s q(m)$ and $c_s q(\overline{m})$; see (6.16), (6.19), (6.21) and (6.22). Together with the nonnegativity of some characteristic functions [see (6.20) and (6.24)], this allows us to apply then the Poisson summation formula from Lemma 6.4. This formula reduces the problem to an estimation of integrals of theta-series. Section 7 is devoted to some facts from number theory. We consider the lattices, their α-characteristics [which are defined in (7.5) and (7.6)] and Minkowski's successive minima. In Section 8, we reduce the estimation of integrals of theta-series to some integrals of α-characteristics. An application of the crucial Lemma 8.2, mentioned above, ends the proof.

2. Results. To formulate the results we need more notation repeating most part of the notation used in BG (1997a). Let $\sigma_1^2 \ge \sigma_2^2 \ge \cdots$ be the eigenvalues of $\mathbb{C}$, counting their multiplicities. We have $\sigma^2 = \sigma_1^2 + \sigma_2^2 + \cdots$.

We identify the linear operators and corresponding matrices. By $\mathbb{I}_d : \mathbb{R}^d \to \mathbb{R}^d$ we denote the identity operator and, simultaneously, the diagonal matrix with entries 1 on the diagonal. By $\mathbb{O}_d$ we denote the $(d \times d)$ matrix with zero entries.

Throughout $\mathcal{S} = \{e_1, \dots, e_s\} \subset \mathbb{R}^d$ denotes a finite set of cardinality s. We write $\mathcal{S}_o$ instead of $\mathcal{S}$ if the system $\{e_1, \dots, e_s\}$ is orthonormal. Let $p > 0$ and $\delta \ge 0$. Denote

$$P(\delta, \mathcal{S}, Y) = \min_{e \in \mathcal{S}} \mathbf{P}\{\|Y - e\| \le \delta\}. \tag{2.1}$$

Similarly to BG (1997a), we use the following nondegeneracy condition for the distribution of a d-dimensional vector Y:

$$P_{\mathbb{Q}}(\delta, \mathcal{S}, Y) \stackrel{\text{def}}{=} \min\{P(\delta, \mathcal{S}, Y), P(\delta, \mathbb{Q}\mathcal{S}, Y)\} \ge p, \tag{2.2}$$

where $p > 0$ is a parameter involved in the condition. Note that

$$P(\delta, \mathcal{S}, Y) = P_{\mathbb{I}_d}(\delta, \mathcal{S}, Y). \tag{2.3}$$

Introduce truncated random vectors

$$X^\diamond = X\mathbf{I}\{\|X\| \le \sigma\sqrt{N}\}, \qquad X_\diamond = X\mathbf{I}\{\|X\| > \sigma\sqrt{N}\}, \tag{2.4}$$

$$X^\square = X\mathbf{I}\{\|\mathbb{C}^{-1/2}X\| \le \sqrt{dN}\}, \qquad X_\square = X\mathbf{I}\{\|\mathbb{C}^{-1/2}X\| > \sqrt{dN}\}, \tag{2.5}$$

and their moments (for $q > 0$)

$$\Lambda_4^\diamond = \frac{1}{\sigma^4 N}\mathbf{E}\|X^\diamond\|^4, \qquad \Pi_q^\diamond = \frac{N}{(\sigma\sqrt{N})^q}\mathbf{E}\|X_\diamond\|^q, \tag{2.6}$$

$$\Lambda_4^\square = \frac{1}{d^2 N}\mathbf{E}\|\mathbb{C}^{-1/2}X^\square\|^4, \qquad \Pi_q^\square = \frac{N}{(\sqrt{dN})^q}\mathbf{E}\|\mathbb{C}^{-1/2}X_\square\|^q. \tag{2.7}$$

Here and below $\mathbf{I}\{A\}$ denotes the indicator of an event A. Of course, definitions (2.5) and (2.7) have sense if $d < \infty$ and the covariance operator $\mathbb{C}$ is nondegenerate.

Clearly, we have

$$X^\diamond + X_\diamond = X^\square + X_\square = X, \qquad \|X^\diamond\|\|X_\diamond\| = \|X^\square\|\|X_\square\| = 0. \tag{2.8}$$

Generally speaking, $X^{\square}$ and $X^{\diamond}$ are different truncated vectors. In BG (1997a) the i.i.d. copies of the vectors $X^{\diamond}$ and $X_{\diamond}$ only were involved. Truncation (2.5) was there applied to the vector $X^{\diamond}$. The use of $X^{\square}$ is more natural for the estimation of constants in the case $d < \infty$. It is easy to see that

$$(\mathbb{C}^{-1/2}X)^{\diamond} = (\mathbb{C}^{-1/2}X)^{\square} = \mathbb{C}^{-1/2}X^{\square} \tag{2.9}$$

and

$$(\mathbb{C}^{-1/2}X)_{\diamond} = (\mathbb{C}^{-1/2}X)_{\square} = \mathbb{C}^{-1/2}X_{\square}. \tag{2.10}$$

Equalities (2.9) and (2.10) provide a possibility to apply auxiliary results obtained in BG (1997a) for truncated vectors $X^{\diamond}$ and $X_{\diamond}$ to truncated vectors $\mathbb{C}^{-1/2}X^{\square}$ and $\mathbb{C}^{-1/2}X_{\square}$. However, one should take into account that σ^2, $\Lambda_4^{\diamond}$, $\Pi_q^{\diamond}$, $G, \ldots$ have to be replaced by corresponding objects related to the vector $\mathbb{C}^{-1/2}X$ (i.e., by d, $\Lambda_4^{\square}$, $\Pi_q^{\square}$, $\mathbb{C}^{-1/2}G, \ldots$).

By $c, c_1, c_2, \ldots$ we denote absolute positive constants. If a constant depends on, say, s, then we point out the dependence writing c_s or $c(s)$. We denote by c universal constants which might be different in different places of the text. Furthermore, in the conditions of theorems and lemmas (see, e.g., Theorem 2.1 and the proofs of Theorems 2.2, 2.4 and 2.5) we write c_0 for an *arbitrary* positive absolute constant; for example, one may choose $c_0 = 1$. We write $A \ll B$ if there exists an absolute constant c such that $A \le cB$. Similarly, $A \ll_s B$ if $A \le c(s)B$. We also write $A \asymp_s B$ if $A \ll_s B \ll_s A$. By $\lfloor \alpha \rfloor$ we denote the largest integer not greater than α.

Throughout we assume that all random vectors and variables are independent in aggregate if the contrary is not clear from the context. By $X_1, X_2, \ldots$ we shall denote independent copies of a random vector X. Similarly, $G_1, G_2, \ldots$ are independent copies of G and so on. By $\mathcal{L}(X)$ we denote the distribution of X. Define the symmetrization $\widetilde{X}$ of a random vector X as a random vector with distribution $\mathcal{L}(\widetilde{X}) = \mathcal{L}(X_1 - X_2)$.

Instead of normalized sums S_N, it is sometimes more convenient to consider the sums $Z_N = X_1 + \cdots + X_N$. Then $S_N = N^{-1/2}Z_N$. Similarly, by $Z_N^{(\diamond)}$ (resp., $Z_N^{(\square)}$) we shall denote sums of N independent copies of $X^{\diamond}$ (resp., $X^{\square}$). For example, $Z_N^{(\square)} = X_1^{\square} + \cdots + X_N^{\square}$.

The expectation $\mathbf{E}_Y$ with respect to a random vector Y we define as the conditional expectation

$$\mathbf{E}_Y f(X, Y, Z, \ldots) = \mathbf{E}\big(f(X, Y, Z \ldots) | X, Z, \ldots\big)$$

given all random vectors but Y.

Throughout we write $\mathrm{e}\{x\} \stackrel{\text{def}}{=} \exp\{ix\}$. By

$$\widehat{F}(t) = \int_{-\infty}^{\infty} \mathrm{e}\{tx\}\, dF(x), \tag{2.11}$$

we denote the Fourier–Stieltjes transform of a function F of bounded variation or, in other words, the Fourier transform of the measure which has the distribution function F.

Introduce the distribution functions

$$F_a(x) \stackrel{\text{def}}{=} \mathbf{P}\{\mathbb{Q}[S_N - a] \le x\}, \qquad H_a(x) \stackrel{\text{def}}{=} \mathbf{P}\{\mathbb{Q}[G - a] \le x\}, \tag{2.12}$$
$$a \in \mathbb{R}^d,\ x \in \mathbb{R}.$$

Furthermore, define, for $d = \infty$ and $a \in \mathbb{R}^d$, the Edgeworth correction

$$E_a(x) = E_a(x; \mathbb{Q}, X)$$

as a function of bounded variation such that $E_a(-\infty) = 0$ and its Fourier–Stieltjes transform is given by

$$\widehat{E}_a(t) = \frac{2(it)^2}{3\sqrt{N}} \mathbf{E}\, \mathrm{e}\{t\mathbb{Q}[Y]\}\big(3\langle \mathbb{Q}X, Y\rangle\langle \mathbb{Q}X, X\rangle + 2it\langle \mathbb{Q}X, Y\rangle^3\big), \tag{2.13}$$
$$Y = G - a.$$

In finite dimensional spaces (for $1 \le d < \infty$) we define the Edgeworth correction as follows; see Bhattacharya and Rao (1986). Let ϕ denote the standard normal density in $\mathbb{R}^d$. Then $p(y) = \phi(\mathbb{C}^{-1/2} y)/\sqrt{\det \mathbb{C}}$, $y \in \mathbb{R}^d$, is the density of G, and, for $a \in \mathbb{R}^d$, $b = \sqrt{N} a$, we have

$$E_a(x) \stackrel{\text{def}}{=} \Theta_b(Nx) \stackrel{\text{def}}{=} \frac{1}{6\sqrt{N}} \chi(A_x), \qquad A_x = \{u \in \mathbb{R}^d : \mathbb{Q}[u - a] \le x\}, \tag{2.14}$$

with the signed measure

$$\chi(A) \stackrel{\text{def}}{=} \int_A \mathbf{E} p'''(y) X^3 \, dy \qquad \text{for the Borel sets } A \subset \mathbb{R}^d, \tag{2.15}$$

and where

$$p'''(y)u^3 = p(y)\big(3\langle \mathbb{C}^{-1}u, u\rangle\langle \mathbb{C}^{-1}y, u\rangle - \langle \mathbb{C}^{-1}y, u\rangle^3\big) \tag{2.16}$$

denotes the third Frechet derivative of p in direction u.

Notice that $E_a = 0$ if $a = 0$ or if $\mathbf{E}\langle X, y\rangle^3 = 0$, for all $y \in \mathbb{R}^d$. In particular, $E_a = 0$ if X is symmetric [i.e., $\mathcal{L}(X) = \mathcal{L}(-X)$].

We can write similar representations for $E_a^{\square}(x) = \Theta_b^{\square}(Nx)$ and $E_a^{\diamond}(x) = \Theta_b^{\diamond}(Nx)$ just replacing X by $X^{\square}$ and $X^{\diamond}$ in (2.13) or (2.15) with $Y = G - a$.

For $b \in \mathbb{R}^d$, introduce the distribution functions

$$\Psi_b(x) \stackrel{\text{def}}{=} \mathbf{P}\{\mathbb{Q}[Z_N - b] \le x\} = F_a(x/N) \tag{2.17}$$

and

$$\Phi_b(x) \stackrel{\text{def}}{=} \mathbf{P}\{\mathbb{Q}[\sqrt{N} G - b] \le x\} = H_a(x/N). \tag{2.18}$$

Define, for $a \in \mathbb{R}^d$, $b = \sqrt{N}a$,

$$\Delta_N^{(a)} \stackrel{\text{def}}{=} \sup_{x\in\mathbb{R}} \left|F_a(x) - H_a(x) - E_a(x)\right| = \sup_{x\in\mathbb{R}} \left|\Psi_b(x) - \Phi_b(x) - \Theta_b(x)\right|; \tag{2.19}$$

see (2.12), (2.14), (2.17) and (2.18) to justify the last equality in (2.19). We write $\Delta_{N,\square}^{(a)}$ and $\Delta_{N,\diamond}^{(a)}$ replacing E_a by $E_a^{\square}$ and $E_a^{\diamond}$ in (2.19).

The aim of this paper is to derive for $\Delta_N^{(a)}$ explicit bounds of order $\mathcal{O}(N^{-1})$ without any additional smoothness type assumptions. Theorem 2.1 [which was proved in BG (1997a)] solved this problem in the case $13 \le d \le \infty$.

In Theorems 2.1–2.5 we assume that the symmetric operator $\mathbb{Q}$ is isometric, that is, that $\mathbb{Q}^2$ is the identity operator $\mathbb{I}_d$. This does not restrict generality; see Remark 1.7 in BG (1997a). Indeed, any symmetric operator $\mathbb{Q}$ may be decomposed as $\mathbb{Q} = \mathbb{Q}_1\mathbb{Q}_0\mathbb{Q}_1$, where $\mathbb{Q}_0$ is symmetric and isometric and $\mathbb{Q}_1$ is symmetric bounded and nonnegative, that is, $\langle \mathbb{Q}_1 x, x\rangle \ge 0$, for all $x \in \mathbb{R}^d$. Thus, for any symmetric $\mathbb{Q}$, we can apply all our bounds replacing the random vector X by $\mathbb{Q}_1 X$, the Gaussian random vector G by $\mathbb{Q}_1 G$, the shift a by $\mathbb{Q}_1 a$, etc. In the case of concentration functions (see Theorems 2.4 and 2.5), we have $Q(X;\lambda;\mathbb{Q}) = Q(\mathbb{Q}_1 X;\lambda;\mathbb{Q}_0)$, and we may apply the results provided $\mathbb{Q}_1 X$ (instead of X) satisfies the conditions.

THEOREM 2.1 [BG (1997a), Theorem 1.3]. *Assume that* $\delta = 1/300$, $\mathbb{Q}^2 = \mathbb{I}_d$, $s = 13$ *and* $13 \le d \le \infty$. *Let* $P_{\mathbb{Q}}(\delta, \mathcal{S}_o, c_0 G/\sigma) \ge p > 0$, *where* c_0 *is an arbitrary positive absolute constant. Then*

$$\Delta_N^{(a)} \le C\left(\Pi_3^{\diamond} + \Lambda_4^{\diamond}\right)\left(1 + \|a/\sigma\|^6\right) \tag{2.20}$$

and

$$\Delta_{N,\diamond}^{(a)} \le C\left(\Pi_2^{\diamond} + \Lambda_4^{\diamond}\right)\left(1 + \|a/\sigma\|^6\right) \tag{2.21}$$

with $C = cp^{-6} + c(\sigma/\theta_8)^8$, *where* $\theta_1^4 \ge \theta_2^4 \ge \cdots$ *are the eigenvalues of* $(\mathbb{C}\mathbb{Q})^2$.

Unfortunately, we cannot apply Theorem 2.1 for $d = 5, 6, \ldots, 12$. Moreover, the quantity C depends on p which is exponentially small with respect to eigenvalues of $\mathbb{C}$.

The main result of the paper is Theorem 2.2. It is valid for $5 \le d < \infty$ in finite-dimensional spaces $\mathbb{R}^d$ only. However, the bounds of Theorem 2.2 depend on the smallest σ_j's. This makes them unstable if one or more of coordinates of X degenerates. In our finite dimensional results, Theorems 2.2, 2.4, 2.5 and Corollary 2.3, we always assume that the covariance operator $\mathbb{C}$ is nondegenerate.

THEOREM 2.2. *Let* $\mathbb{Q}^2 = \mathbb{I}_d$, $5 \le d < \infty$. *Then*

$$\Delta_N^{(a)} \le C\left(\Pi_3^{\square} + \Lambda_4^{\square}\right)\left(1 + \|a/\sigma\|^3\right) \tag{2.22}$$

and

$$\Delta_{N,\square}^{(a)} \leq C\big(\Pi_2^{\square} + \Lambda_4^{\square}\big)\big(1 + \|a/\sigma\|^3\big), \tag{2.23}$$

with $C = c_d \sigma^d (\det \mathbb{C})^{-1/2}$.

In Götze and Zaitsev (2010) [see also a preprint of Götze and Zaitsev (2009) which is available in Internet], an analogue of Theorem 2.2 was proved in the case $s = 5$ and $5 \leq d < \infty$ with bounds for constants which are not optimal. It extends to the case $d \geq 5$ Theorem 1.5 of BG (1997a) which contains the corresponding bounds for $d \geq 9$. Unfortunately, in both papers, the quantity C depends on p which is exponentially small with respect to σ_9/σ^2 [in BG (1997a)] and to σ_5/σ^2 [in Götze and Zaitsev (2010)]. Under some additional conditions, C may be estimated from above by $c_d \exp(c\sigma^2 \sigma_9^{-2})$ and by $c_d \exp(c\sigma^2\sigma_5^{-2})$, respectively. The case $a = 0$ was considered earlier in Götze and Zaitsev (2008). As a consequence, we have proved Theorem 1.1.

It is easy to see that, according to (2.5) and (2.7),

$$\Pi_3^{\square} + \Lambda_4^{\square} \leq \mathbf{E}\big\|\mathbb{C}^{-1/2}X\big\|^{3+\delta} / \big(d^{(3+\delta)/2} N^{(1+\delta)/2}\big) \qquad \text{for } 0 \leq \delta \leq 1 \tag{2.24}$$

and

$$\Pi_2^{\square} + \Lambda_4^{\square} \leq \mathbf{E}\big\|\mathbb{C}^{-1/2}X\big\|^{2+\delta} / \big(d^{(2+\delta)/2} N^{\delta/2}\big) \qquad \text{for } 0 \leq \delta \leq 2. \tag{2.25}$$

Therefore, Theorem 2.2 implies the following Corollary 2.3.

COROLLARY 2.3. *Let* $\mathbb{Q}^2 = \mathbb{I}_d$, $5 \leq d < \infty$. *Then*

$$\Delta_N^{(a)} \ll_d C\big(1 + \|a/\sigma\|^3\big)\mathbf{E}\big\|\mathbb{C}^{-1/2}X\big\|^{3+\delta} / N^{(1+\delta)/2} \qquad \textit{for } 0 \leq \delta \leq 1 \tag{2.26}$$

and

$$\Delta_{N,\square}^{(a)} \ll_d C\big(1 + \|a/\sigma\|^3\big)\mathbf{E}\big\|\mathbb{C}^{-1/2}X\big\|^{2+\delta} / N^{\delta/2} \qquad \textit{for } 0 \leq \delta \leq 2, \tag{2.27}$$

with $C = \sigma^d (\det \mathbb{C})^{-1/2}$. *In particular,*

$$\max\big\{\Delta_N^{(a)}, \Delta_{N,\square}^{(a)}\big\} \ll_d C\big(1 + \|a/\sigma\|^3\big)\mathbf{E}\big\|\mathbb{C}^{-1/2}X\big\|^4 / N. \tag{2.28}$$

Theorem 2.1 and Corollary 2.3 yield Theorems 1.1 and 1.2, using that $E_0(x) \equiv 0$, $\mathbf{E}\|\mathbb{C}^{-1/2}X\|^4 \leq \beta/\sigma_d^4$, and $\Pi_2^{\diamond} + \Lambda_4^{\diamond} \leq \Pi_3^{\diamond} + \Lambda_4^{\diamond} \leq \beta/(\sigma^4 N)$.

Comparing Theorem 2.2 and Corollary 2.3 with the main results of BG (1997a) and Götze and Zaitsev (2010), we see that the constants in Theorem 2.2 and Corollary 2.3 are written explicitly in terms of moment characteristics of $\mathcal{L}(X)$. In the case of nonpositive definite quadratic forms $\mathbb{Q}$ such kind of estimates were unknown.

If, in the conditions of Theorem 2.2, the distribution of X is symmetric or $a = 0$, then the Edgeworth corrections $E_a(x)$ and $E_a^{\square}(x)$ vanish and

$$\Delta_N^{(a)} = \Delta_{N,\square}^{(a)} \leq C\big(\Pi_2^{\square} + \Lambda_4^{\square}\big)\big(1 + \|a/\sigma\|^3\big), \qquad C = c_d \sigma^d (\det \mathbb{C})^{-1/2}. \tag{2.29}$$

The corresponding inequality from Theorem 1.4 of BG (1997a) yields in the case $s=9$ and $9 \le d \le \infty$ under the condition $P_{\mathbb{Q}}(\delta, \mathcal{S}_o, c_0 G/\sigma) \ge p > 0$ with $\delta = 1/300$ the bound

$$\Delta_N^{(a)} \le C\big(\Pi_2^{\diamond} + \Lambda_4^{\diamond}\big)\big(1 + \|a/\sigma\|^4\big), \qquad C = cp^{-4}. \tag{2.30}$$

It is clear that sometimes the bound (2.30) may be sharper than (2.29) but, unfortunately, it depends on p which is usually exponentially small with respect to σ_9/σ^2.

Several authors have obtained more precise estimates of constants in the case of d-dimensional balls with $d \ge 12$, including the case $d = \infty$. For balls, $\mathbb{Q} = \mathbb{I}_d$. In the papers mentioned above, the authors have used the aproach of BG (1997a) and obtained bounds with constants depending on $s \le d$ largest eigenvalues $\sigma_1^2 \ge \sigma_2^2 \ge \cdots \ge \sigma_s^2$ of the covariance operator $\mathbb{C}$; see Nagaev and Chebotarev (1999, 2005), with $d \ge s = 13$, and Götze and Ulyanov (2000), and Bogatyrev, Götze and Ulyanov (2006), with $d \ge s = 12$. It should be mentioned, that, in a particular case, where $\mathbb{Q} = \mathbb{I}_d$ and $d \ge 12$, these results may be sharper than (2.22), for some covariance operators $\mathbb{C}$. The lower bounds for $\Delta_N^{(a)}$ under different conditions on a and $\mathcal{L}(X)$ are given in Götze and Ulyanov (2000). See the upper bounds for $\Delta_N^{(a)}$ with $s = 12$ and $d = \infty$ in Ulyanov and Götze (2011), where the dependence on the eigenvalues of $\mathbb{C}$ is given in the upper bound in an explicit form which coincides with that in the lower bound. See also the review of recent results for "almost" quadratic forms in Prokhorov and Ulyanov (2013).

Thus we see that the statement of Theorem 2.2 is especially interesting for $d = 5, \ldots, 11$. It is new even in the case of d-dimensional balls. It is plausible that the bounds for constants in Theorem 2.2 could be also improved for balls with $d \ge 5$, especially in the case where d is large. It seems, however, that this is impossible in the case of general $\mathbb{Q}$ even if $\mathbb{Q}^2 = \mathbb{I}_d$. For example, consider the operator $\mathbb{Q}$ such that $\mathbb{Q}e_j = e_{d-j+1}$, where $\mathbb{C}e_j = \sigma_j^2 e_j$, $j = 1, 2, \ldots, d$, are eigenvectors of $\mathbb{C}$. Following the proof of Theorem 2.2, we see that the bounds for the modulus of the characteristic function $|\widehat{\Psi}_b(t)| = |\mathbf{E}\,\mathrm{e}\{t\mathbb{Q}[Z_N - b]\}|$ behave as the bounds for the modulus of the characteristic function $|\mathbf{E}\,\mathrm{e}\{t\mathbb{I}_d[Z_N - b]\}|$, but with eigenvalues of the covariance operator $\sigma_1\sigma_d, \sigma_2\sigma_{d-1}, \sigma_3\sigma_{d-2}, \ldots$ which may be essentially smaller than $\sigma_1^2 \ge \sigma_2^2 \ge \sigma_3^2 \ge \cdots$. Therefore, it is natural that the bounds for constants in Theorem 2.2 depends on the smallest eigenvalues of the covariance operator $\mathbb{C}$.

Note that, in the proof of Theorem 2.1 in BG (1997a), inequalities (2.20) and (2.21) were derived for the Edgeworth correction $E_a(x)$ defined by (2.13). However, from Theorems 2.1 and 2.2 it follows that, at least for $13 \le d < \infty$, definitions (2.13) and (2.14) determine the same function $E_a(x)$. Indeed, both functions may be represented as $N^{-1/2}K_j(x)$, where $K_j(x)$ are some functions of bounded variation which are independent of N. Furthermore, inequalities (2.20) and (2.22) provide both bounds of order $\mathcal{O}(N^{-1})$. This is possible only if the Edgeworth corrections $E_a(x)$ are the same in these inequalities.

On the other hand, it is proved (for $d \geq 9$) that definition (2.13) determines a function of bounded variation [see BG (1997a, Lemma 5.7)], while definition (2.14) has no sense for $d = \infty$.

Introduce the concentration function

$$Q(X;\lambda) = Q(X;\lambda;\mathbb{Q}) \\ = \sup_{a\in\mathbb{R}^d, x\in\mathbb{R}} \mathbf{P}\{x \leq \mathbb{Q}[X-a] \leq x+\lambda\} \qquad \text{for } \lambda \geq 0. \tag{2.31}$$

Note that, evidently, $Q(X+Y;\lambda) \leq Q(X;\lambda)$, for any Y which is independent of X.

We say that a random vector Y is concentrated in $\mathbb{L} \subset \mathbb{R}^d$ if $\mathbf{P}\{Y \in \mathbb{L}\} = 1$. In BG [(1997a), item (iii) of Theorem 1.6] it was shown that if $\widetilde{X}$ is not concentrated in a proper closed linear subspace of $\mathbb{R}^d$, $1 \leq d \leq \infty$, then for any $\delta > 0$ and $\mathcal{S}$, there exists a natural number m such that the condition $P_{\mathbb{Q}}(\delta, \mathcal{S}, m^{-1/2}\widetilde{Z}_m) \geq p$ holds with some $p > 0$.

In this paper, we shall prove the following Theorems 2.4 and 2.5.

THEOREM 2.4. *Let* $\mathbb{Q}^2 = \mathbb{I}_d$, $5 \leq s = d < \infty$ *and* $0 \leq \delta \leq 1/(5s)$. *Then*:

(i)

$$Q(Z_N;\lambda) \ll_d (pN)^{-1} \max\{1; \lambda\sigma^{-2}\}\sigma^d(\det\mathbb{C})^{-1/2} \qquad \textit{for all } \lambda \geq 0, \tag{2.32}$$

if $P(\delta, \mathcal{S}_o, \mathbb{C}^{-1/2}\widetilde{X}) \geq p$ *for some* $\mathcal{S}_o$ *and* $p > 0$.

(ii)

$$Q(Z_N;\lambda) \ll_d (pN)^{-1} \max\{m; \lambda\sigma^{-2}\}\sigma^d(\det\mathbb{C})^{-1/2} \qquad \textit{for all } \lambda \geq 0, \tag{2.33}$$

if, for some $\mathcal{S}_o$ *and positive integer* m, $P(\delta, \mathcal{S}_o, m^{-1/2}\mathbb{C}^{-1/2}\widetilde{Z}_m) \geq p > 0$.

THEOREM 2.5. *Assume that* $5 \leq d < \infty$ *and that* $\mathbb{Q}^2 = \mathbb{I}_d$. *Then*

$$Q(Z_N;\lambda) \ll_d \max\{\Pi_2^{\square} + \Lambda_4^{\square}; \lambda\sigma^{-2}N^{-1}\}\sigma^d(\det\mathbb{C})^{-1/2} \\ \textit{for all } \lambda \geq 0. \tag{2.34}$$

In particular, $Q(Z_N;\lambda) \ll_d N^{-1}\max\{\mathbf{E}\|\mathbb{C}^{-1/2}X\|^4; \lambda\sigma^{-2}\}\sigma^d(\det\mathbb{C})^{-1/2}$.

Theorems 2.4 and 2.5 yield more explicit versions of Theorems 1.5 and 2.1 from Götze and Zaitsev (2010) [which extend to the case $5 \leq d \leq \infty$ Theorems 1.6 and 2.1 of BG (1997a) which were proved for $9 \leq d \leq \infty$]. We should mention that the results of Götze and Zaitsev (2010) do not follow from Theorems 2.2, 2.4 and 2.5. For example, they may be sharper than Theorems 2.2, 2.4 and 2.5, in a particular case, where $\mathbb{Q} = \mathbb{I}_d$ and $\sigma_5 \asymp_d \sigma$. Under some additional conditions, $\sigma^d(\det\mathbb{C})^{-1/2}$ is replaced by $\exp(c\sigma^2\sigma_5^{-2}) \asymp_d 1$. On the other hand,

$\sigma^d(\det \mathbb{C})^{-1/2}$ provides a power-type dependence on eigenvalues of $\mathbb{C}$ and the results are valid for $\mathbb{Q}$ which might be not positive definite.

In Theorems 2.2 and 2.5, we do not assume conditions $P(\cdot) \geq p > 0$ or $P_{\mathbb{Q}}(\cdot) \geq p > 0$. In the proofs, we use, however, that, for any fixed absolute positive constant c_0 and any positive quantity c_d depending on d only, condition $P(\delta, \mathcal{S}_o, c_0\mathbb{C}^{-1/2}G) \geq p$ is fulfilled with $s = d$, $\delta = c_d$ and $p \asymp_d 1$, for any orthonormal system $\mathcal{S}_o$.

Similarly to BG (1997a), in Section 3, we prove bounds for concentration functions. The proof is technically simpler as that of Theorem 2.2, but it shows how to apply the principal ideas. This proof repeats almost literally the corresponding proof of BG (1997a). The only difference consists in the use of new Lemma 8.3 which allows us to estimate characteristic functions of quadratic forms for relatively large values of argument t. In Sections 4 and 5, Theorem 2.2 is proved. We replace Lemma 9.4 of BG (1997a) by its improvement, Lemma 5.1. Another difference is in another choice of k in (5.31) and (5.32) in comparison with that in BG (1997a). In Sections 6–8 we prove estimates for characteristic functions which were discussed in Section 1.

3. Proofs of bounds for concentration functions.

Proof of Theorems 2.4 *and* 2.5. Below we prove assertions (2.32); (2.32) $\Longrightarrow$ (2.33) and (2.33) $\Longrightarrow$ (2.34). The proof repeats almost literally the corresponding proof of BG (1997a). It is given here for the sake of completeness. The only essential difference is in the use of Lemma 8.3 in the proof of Lemma 3.1. We have also to replace everywhere 9 by 5 and $\diamond$ by $\square$.

For $0 \leq t_0 \leq T$ and $b \in \mathbb{R}^d$, define the integrals

$$I_0 = \int_{-T}^{T} |\widehat{\Psi}_b(t)|\, dt, \qquad I_1 = \int_{t_0 \leq |t| \leq T} |\widehat{\Psi}_b(t)| \frac{dt}{|t|},$$

where

$$\widehat{\Psi}_b(t) = \mathbf{E}\, \mathrm{e}\{t\mathbb{Q}[Z_N - b]\} \tag{3.1}$$

denotes the Fourier–Stieltjes transform of the distribution function Ψ_b of $\mathbb{Q}[Z_N - b]$. Note that $|\widehat{\Psi}_b(-t)| = |\widehat{\Psi}_b(t)|$.

LEMMA 3.1. *Assume that* $P(\delta, \mathcal{S}_o, \mathbb{C}^{-1/2}\widetilde{X}) \geq p > 0$ *with some* $0 \leq \delta \leq 1/(5s)$ *and* $5 \leq s = d < \infty$. *Let* $\sigma^2 = 1$ *and*

$$t_0 = c_1(s)\sigma_1^{-2}(pN)^{-1+2/s}, \qquad c_2(s)\sigma_1^{-2} \leq T \leq c_3(s)\sigma_1^{-2} \tag{3.2}$$

with some positive constants $c_j(s)$, $1 \le j \le 3$. *Then*

$$I_0 \ll_s (\det \mathbb{C})^{-1/2}(pN)^{-1}, \qquad I_1 \ll_s (\det \mathbb{C})^{-1/2}(pN)^{-1}. \tag{3.3}$$

PROOF. Note that the condition $\sigma^2 = 1$ implies that

$$T \asymp_s \sigma_1^2 \asymp_s \sigma^2 = 1 \quad \text{and} \quad \det \mathbb{C} \le 1. \tag{3.4}$$

Denote $k = pN$. Without loss of generality we assume that $k \ge c_s$, for a sufficiently large quantity c_s depending on s only. Indeed, if $k \le c_s$, then one can prove (3.3) using (3.4) and $|\widehat{\Psi}_b| \le 1$. Choosing c_s to be large enough, we ensure that $k \ge c_s$ implies $1/k \le t_0 \le T$.

Lemma 8.3 and (3.4) imply now that

$$\int_{c_4(s)k^{-1+2/s}}^{T} |\widehat{\Psi}_b(t)| \frac{dt}{t} \ll_s \frac{(\det \mathbb{C})^{-1/2}}{k} \tag{3.5}$$

for any $c_4(s)$ depending on s only. Inequalities (3.4) and (3.5) imply (3.3) for I_1.

Let us prove inequality (3.2) for I_0. By (3.4) and by Lemma 8.1, for any $\gamma > 0$ and any fixed $t \in \mathbb{R}$ satisfying $k^{1/2}|t| \le c_5(s)$, where $c_5(s)$ is an arbitrary quantity depending on s only, we have (taking into account that $|\widehat{\Psi}_b| \le 1$)

$$|\widehat{\Psi}_b(t)| \ll_{\gamma,s} \min\{1; k^{-\gamma} + k^{-s/2}|t|^{-s/2}(\det \mathbb{C})^{-1/2}\}, \qquad k = pN. \tag{3.6}$$

Furthermore, choosing an appropriate γ and using (3.4)–(3.6), we obtain

$$(\det \mathbb{C})^{1/2} I_0 \ll_s \int_0^{1/k} dt + \frac{1}{k} + \int_{1/k}^{\infty} \frac{dt}{(tk)^{s/2}} \ll_s \frac{1}{k}, \tag{3.7}$$

proving (3.2) for I_0. □

PROOF OF (2.32). Let $\sigma^2 = 1$. Using a well-known inequality for concentration functions [see, e.g., Petrov (1975), Lemma 3 of Chapter 3], we have

$$Q(Z_N; \lambda) \le 4 \sup_{b \in \mathbb{R}^d} \max\{\lambda; 1\} \int_0^1 |\widehat{\Psi}_b(t)|\, dt. \tag{3.8}$$

To estimate the integral in (3.8) we apply Lemma 3.1 which implies that

$$Q(Z_N; \lambda) \ll_d \max\{\lambda; 1\}(pN)^{-1}(\det \mathbb{C})^{-1/2}, \tag{3.9}$$

proving (2.32) in the case $\sigma^2 = 1$. If $\sigma^2 \ne 1$, we obtain (2.32) applying (3.9) to Z_N/σ. □

PROOF OF (2.32) $\Longrightarrow$ (2.33). Without loss of generality we can assume that $N/m \ge 2$. Let $Y_1, Y_2, \dots$ be independent copies of $m^{-1/2} Z_m$. Denote $W_k = Y_1 + \cdots + Y_k$. Then $\mathcal{L}(Z_N) = \mathcal{L}(\sqrt{m} W_k + y)$, where $k = \lfloor N/m \rfloor$ is the largest integer not greater than N/m and y is independent of W_k. Therefore, $Q(Z_N; \lambda) \le$

$Q(W_k;\lambda/m)$. In order to estimate $Q(W_k;\lambda/m)$ we apply (2.32) replacing Z_N by W_k. We have

$$\begin{aligned} Q(W_k;\lambda/m) &\ll_s (pk)^{-1}\max\{1;\lambda\sigma^{-2}/m\}\sigma^d(\det\mathbb{C})^{-1/2} \\ &\ll (pN)^{-1}\max\{m;\lambda\sigma^{-2}\}\sigma^d(\det\mathbb{C})^{-1/2}. \end{aligned} \tag{3.10}$$ □

Recall that truncated random vectors and their moments are defined by (2.4)–(2.7) and that $\mathbb{C} = \operatorname{cov} X = \operatorname{cov} G$.

LEMMA 3.2. *The random vectors* $X^{\square}$, $X_{\square}$ *satisfy*

$$\langle \mathbb{C}x, x\rangle = \langle \operatorname{cov} X^{\square}x, x\rangle + \mathbf{E}\langle X_{\square}, x\rangle^2 + \langle \mathbf{E}X^{\square}, x\rangle^2.$$

There exist independent centered Gaussian vectors G_* *and* W *such that*

$$\mathcal{L}(G) = \mathcal{L}(G_* + W)$$

and

$$2\operatorname{cov} G_* = 2\operatorname{cov} X^{\square} = \operatorname{cov}\widetilde{X}^{\square}, \qquad \langle \operatorname{cov} Wx, x\rangle = \mathbf{E}\langle X_{\square}, x\rangle^2 + \langle \mathbf{E}X^{\square}, x\rangle^2.$$

Furthermore,

$$\mathbf{E}\|\mathbb{C}^{-1/2}G\|^2 = d = \mathbf{E}\|\mathbb{C}^{-1/2}G_*\|^2 + \mathbf{E}\|\mathbb{C}^{-1/2}W\|^2$$

and $\mathbf{E}\|\mathbb{C}^{-1/2}W\|^2 \le 2d\Pi_2^{\square}$.

We omit the simple proof of this lemma; see BG [(1997a), Lemma 2.4] for the same statement with $\diamond$ instead of $\square$.

Recall that $Z_N^{(\square)}$ and $Z_N^{(\diamond)}$ denote sums of N independent copies of $X^{\square}$ and $X^{\diamond}$, respectively.

LEMMA 3.3. *Let* $\varepsilon > 0$. *There exist absolute positive constants* c *and* c_1 *such that the condition* $\Pi_2^{\square} \le c_1 p\delta^2/(d\varepsilon^2)$ *implies that*

$$P(\delta, \mathcal{S}, \varepsilon\mathbb{C}^{-1/2}G) \ge p \quad \Longrightarrow \quad P(4\delta, \mathcal{S}, \varepsilon(2m)^{-1/2}\mathbb{C}^{-1/2}\widetilde{Z}_m^{(\square)}) \ge p/4$$

for $m \ge c\varepsilon^4 d^2 N\Lambda_4^{\square}/(p\delta^4)$.

Lemmas 3.2 and 3.3 are in fact the statements of Lemmas 2.4 and 2.5 from BG (1997a) applied to the vectors $\mathbb{C}^{-1/2}X$ instead of the vectors X. We use in this connection equalities (2.3), (2.9) and (2.10) replacing in the formulation σ^2, $\Lambda_4^{\diamond}$, $\Pi_q^{\diamond}$, G, $Z_m^{(\diamond)}, \dots$ by d, $\Lambda_4^{\square}$, $\Pi_q^{\square}$, $\mathbb{C}^{-1/2}G$, $Z_m^{(\square)}, \dots$, respectively.

PROOF OF (2.33) $\Longrightarrow$ (2.34). By a standard truncation argument, we have

$$\left|\mathbf{P}\{Z_N \in A\} - \mathbf{P}\{Z_N^{(\square)} \in A\}\right| \le N\mathbf{P}\{\|\mathbb{C}^{-1/2}X\| > \sqrt{dN}\} \le \Pi_2^{\square} \tag{3.11}$$

for any Borel set A, and

$$Q(Z_N,\lambda) \le \Pi_2^{\square} + Q\big(Z_N^{(\square)},\lambda\big). \tag{3.12}$$

Recall that we are proving (2.34) assuming that $5 \le d < \infty$. It is easy to see that, for any absolute positive constant c_0 and for any orthonormal system $\mathcal{S}_o = \{e_1,\dots,e_s\} \subset \mathbb{R}^d$, condition

$$P\big(\delta,\mathcal{S}_o,c_0\mathbb{C}^{-1/2}G\big) \ge p \quad \text{with } p \asymp_d 1, 5 \le s = d < \infty, \delta = 1/(20s) \tag{3.13}$$

is in fact fulfilled automatically since the vector $\mathbb{C}^{-1/2}G$ has standard Gaussian distribution in $\mathbb{R}^d$ and, therefore,

$$\mathbf{P}\big\{\|c_0\mathbb{C}^{-1/2}G - e\| \le \delta\big\} = \mathbf{P}\big\{\|\mathbb{C}^{-1/2}G - c_0^{-1}e\| \le c_0^{-1}\delta\big\} = c(d,c_0)$$

for any vector $e \in \mathbb{R}^d$ with $\|e\| = 1$. For any fixed c_0, the $c(d,c_0)$ may be considered as a quantity depending on d only. Clearly, $4\delta = 1/(5s)$. Write $K = \varepsilon/\sqrt{2}$ with $\varepsilon = c_0$. Then, by (3.13) and Lemma 3.3, we have

$$P\big(\delta,\mathcal{S}_o,\varepsilon\mathbb{C}^{-1/2}G\big) \ge p \quad \Longrightarrow \quad P\big(4\delta,\mathcal{S}_o,m^{-1/2}K\mathbb{C}^{-1/2}\widetilde{Z}_m^{(\square)}\big) \ge p/4, \tag{3.14}$$

provided that

$$\Pi_2^{\square} \le c_1(d), \qquad m \ge c_2(d)N\Lambda_4^{\square}. \tag{3.15}$$

Without loss of generality we may assume that $\Pi_2^{\square} \le c_1(d)$, since otherwise the result follows easily from the trivial inequality $Q(Z_N;\lambda) \le 1$.

The nondegeneracy condition (3.14) for $K\widetilde{Z}_m^{(\square)}$ allows us to apply inequality (2.33) of Theorem 2.4, and, using (3.13), we obtain

$$\begin{aligned} Q\big(Z_N^{(\square)},\lambda\big) &= Q\big(KZ_N^{(\square)},K^2\lambda\big) \\ &\ll_d N^{-1}\max\{m;\, K^2\lambda/K^2\sigma^2\}\sigma^d(\det\mathbb{C})^{-1/2} \end{aligned} \tag{3.16}$$

for any m such that (3.15) is fulfilled. Choosing the minimal m in (3.15), we obtain

$$Q\big(Z_N^{(\square)},\lambda\big) \ll_d \max\big\{\Lambda_4^{\square};\,\lambda/(\sigma^2N)\big\}\sigma^d(\det\mathbb{C})^{-1/2}. \tag{3.17}$$

Combining the estimates (3.12) and (3.17), we complete the proof. $\square$

4. Auxiliary lemmas. In Sections 4 and 5 we prove Theorem 2.2. Therefore, we assume that its conditions are satisfied. We consider the case $d < \infty$ assuming that the following conditions are satisfied:

$$\mathbb{Q}^2 = \mathbb{I}_d, \qquad \sigma^2 = 1, \qquad d \ge 5, \qquad b = \sqrt{N}a. \tag{4.1}$$

Notice that the assumption $\sigma^2 = 1$ does not restrict generality since from Theorem 2.2 with $\sigma^2 = 1$, we can derive the general result replacing X, G by X/σ, G/σ, etc. Other assumptions in (4.1) are included as conditions in Theorem 2.2.

Section 4 is devoted to some auxiliary lemmas which are similar to corresponding lemmas of BG (1997a).

In several places, the proof of Theorem 2.2 repeats almost literally the proof of Theorem 1.5 in BG (1997a). Note, however, that we use truncated vectors $X_j^{\square}$, while in BG (1997a) the vectors $X_j^{\diamond}$ were involved. We start with an application of the Fourier transform to the functions Ψ_b and Φ_b, where $b = \sqrt{N}a$. We estimate integrals over the Fourier transforms using results of Sections 3, 6–8 and some technical lemmas of BG (1997a). We also apply some methods of estimation of the rate of approximation in the CLT in multidimensional spaces; cf., for example, Bhattacharya and Rao (1986).

Below we use the following formula for the Fourier inversion; see, for example, BG (1997a). A smoothing inequality of Prawitz (1972) implies [see BG (1996), Section 4] that

$$F(x) = \frac{1}{2} + \frac{i}{2\pi}\,\mathrm{V.P.}\int_{|t|\le K} \mathrm{e}\{-xt\}\widehat{F}(t)\frac{dt}{t} + R \tag{4.2}$$

for any $K > 0$ and any distribution function F with characteristic function $\widehat{F}$ [see (2.11)], where

$$|R| \le \frac{1}{K}\int_{|t|\le K} \bigl|\widehat{F}(t)\bigr|\,dt. \tag{4.3}$$

Here $\mathrm{V.P.}\int f(t)\,dt = \lim_{\varepsilon\to 0}\int_{|t|>\varepsilon} f(t)\,dt$ denotes the principal value of the integral.

In Sections 4 and 5, we denote

$$X' = X^{\square} - \mathbf{E}X^{\square} + W, \tag{4.4}$$

where W is a centered Gaussian random vector which is independent of all other random vectors and variables and is chosen so that $\operatorname{cov} X' = \operatorname{cov} G$. Such a vector W exists by Lemma 3.2. We define $E_a'(x) = \Theta_b'(Nx)$ replacing X by X' in (2.13) or (2.15) with $Y = G - a$.

Recall that the random vector $X^{\square}$ is defined in (2.5) and $Z_N^{(\square)}$ is a sum of its N independent copies. Similarly, $Z_N' = X_1' + \cdots + X_N'$. Write $\Psi_b^{\square}$ and Ψ_b' for the distribution function of $\mathbb{Q}[Z_N^{(\square)} - b]$ and $\mathbb{Q}[Z_N' - b]$, respectively. For $0 \le k \le N$ introduce the distribution function

$$\Psi_b^{(k)}(x) = \mathbf{P}\bigl\{\mathbb{Q}\bigl[G_1 + \cdots + G_k + X_{k+1}' + \cdots + X_N' - b\bigr] \le x\bigr\}. \tag{4.5}$$

Notice that $\Psi_b^{(0)} = \Psi_b'$, $\Psi_b^{(N)} = \Phi_b$.

The proof of the following lemma repeats the proof of Lemma 3.1 of BG (1997a). The difference is that here we use the truncated vectors $X_j^{\square}$ instead of $X_j^{\diamond}$.

LEMMA 4.1. *Let c_d be a quantity depending on d only. There exist positive quantities $c_1(d)$ and $c_2(d)$ depending on d only such that the following statement* *is valid. Let* $\Pi_2^\square \le c_1(d)p$ *and let an integer* $1 \le m \le N$ *satisfy* $m \ge c_2(d)N\Lambda_4^\square/p$. *Write*

$$K = c_0^2/(2m), \qquad t_1 = c_d(pN/m)^{-1+2/d}.$$

Let F *denote any of the functions* $\Psi_b^\square$, Ψ_b', $\Psi_b^{(k)}$ *or* Φ_b. *Then we have*

$$F(x) = \frac{1}{2} + \frac{i}{2\pi}\,\mathrm{V.P.}\int_{|t|\le t_1} \mathrm{e}\{-xtK\}\widehat{F}(tK)\frac{dt}{t} + R_1, \tag{4.6}$$

with $|R_1| \ll_d (pN)^{-1}m(\det\mathbb{C})^{-1/2}$.

PROOF. We assume that $(pN)^{-1}m \le c_3(d)$ with sufficiently small $c_3(d)$ since otherwise the statement of Lemma 4.1 is trivial; see (3.4), (4.2) and (4.3). Let us prove (4.6). We combine (4.2) and Lemma 3.1. Changing the variable $t = \tau K$ in formula (4.2), we obtain

$$F(x) = \frac{1}{2} + \frac{i}{2\pi}\,\mathrm{V.P.}\int_{|t|\le 1} \mathrm{e}\{-xtK\}\widehat{F}(tK)\frac{dt}{t} + R, \tag{4.7}$$

where

$$|R| \le \int_{|t|\le 1} \bigl|\widehat{F}(tK)\bigr|\,dt. \tag{4.8}$$

Notice that $\Psi_b^\square$, Ψ_b', $\Psi_b^{(k)}$ and Φ_b are distribution functions of random variables which may be written in the following form:

$$\mathbb{Q}[V+T], \qquad V \stackrel{\text{def}}{=} G_1 + \cdots + G_k + X_{k+1}^\square + \cdots + X_N^\square,$$

with some k, $0 \le k \le N$, and some random vector T which is independent of $X_j^\square$ and G_j, for all j. Let us consider separately two possible cases, $k \ge N/2$ and $k < N/2$.

The case $k < N/2$. Let Y denote a sum of m independent copies of $K^{1/2}X^\square$. Let $Y_1, Y_2, \dots$ be independent copies of Y. Then we have

$$\mathcal{L}\bigl(K^{1/2}V\bigr) = \mathcal{L}(Y_1 + \cdots + Y_l + T_1) \tag{4.9}$$

with $l = \lfloor N/(2m)\rfloor$ and some random T_1 independent of $Y_1, \dots, Y_l$. By (3.13) and by Lemma 3.3, we have

$$P\bigl(\delta, \mathcal{S}, c_0\mathbb{C}^{-1/2}G\bigr) \ge p \quad\Longrightarrow\quad P\bigl(4\delta, \mathcal{S}, \mathbb{C}^{-1/2}\widetilde{Y}\bigr) \ge p/4 \tag{4.10}$$

provided that

$$\Pi_2^\square \ll p/d^3 \quad\text{and}\quad m \gg d^6N\Lambda_4^\square/p. \tag{4.11}$$

The inequalities in (4.11) follow from conditions of Lemma 4.1 if we choose some sufficiently small (resp., large) $c_1(d)$ [resp. $c_2(d)$]. Due to (3.13), (4.1), (4.9) and (4.10), we can apply Lemma 3.1 in order to estimate the integrals in (4.7) and (4.8). Replacing in Lemma 3.1 X by Y and N by l, we obtain (4.6) in the case $k < N/2$.

The case $k \geq N/2$. We can argue as in the previous case defining now Y as a sum of m independent copies of $K^{1/2}G$. Condition $P(4\delta, \mathcal{S}, \mathbb{C}^{-1/2}\widetilde{Y}) \geq p/4$ is satisfied by (3.13), since now $\mathcal{L}(\widetilde{Y}) = \mathcal{L}(c_0 G)$.

Following BG (1997a), introduce the upper bound $\varkappa(t; N, X)$ for the characteristic function of quadratic forms; cf. Bentkus (1984) and Bentkus, Götze and Zitikis (1993). We define $\varkappa(t; N, X) = \varkappa^*(t; N, X) + \varkappa^*(t; N, G)$, where

$$\varkappa^*(t; N, X) = \sup_{x \in \mathbb{R}^d} \left| \mathbf{E}\, e\{t\mathbb{Q}[Z_j] + \langle x, Z_j\rangle\}\right|, \qquad Z_j = X_1 + \cdots + X_j, \tag{4.12}$$

with $j = \lfloor (N-2)/14 \rfloor$. Note that $|\mathbf{E}\, e\{t\mathbb{Q}[Z_j] + \langle x, Z_j\rangle\}| = |\mathbf{E}\, e\{t\mathbb{Q}[Z_j - y]\}|$ with $y = -\mathbb{Q}x/(2t)$. In the sequel, we use that

$$\varkappa(t; N, X') \leq \varkappa(t; N, X^{\square}). \tag{4.13}$$

For the proof, it suffices to note that $X' = X^{\square} - \mathbf{E}X^{\square} + W$ and W is independent of $X^{\square}$.

LEMMA 4.2. *Let the conditions of Lemma 4.1 be satisfied. Then*

$$\begin{aligned} &\int_{|t| \leq t_1} (|t|K)^{\alpha} \varkappa(tK; N, X^{\square}) \frac{dt}{|t|} \\ &\qquad \ll_{\alpha,d} (\det \mathbb{C})^{-1/2} \begin{cases} (Np)^{-\alpha}, & \textit{for } 0 \leq \alpha < d/2, \\ (Np)^{-\alpha}\big(1 + |\log(Np/m)|\big), & \\ & \textit{for } \alpha = d/2, \\ (Np)^{-\alpha}\big(1 + (Np/m)^{(2\alpha - d)/d}\big), & \\ & \textit{for } \alpha > d/2. \end{cases} \end{aligned} \tag{4.14}$$

Lemma 4.2 is a generalization of Lemma 3.2 from BG (1997a) which contains the same bound for $0 \leq \alpha < d/2$. In this paper, we have to estimate the left-hand side of (4.14) in the case $d/2 \leq \alpha$ too.

PROOF. We assume again that $(pN)^{-1}m \leq c_3(d)$ with sufficiently small $c_3(d)$ since otherwise (4.14) is an easy consequence of $|\varkappa| \leq 1$.

By (3.13) and (4.10), the condition $P(4\delta, \mathcal{S}_o, K^{1/2}\mathbb{C}^{-1/2}\widetilde{Z}_m^{(\square)}) \geq p/4$ is fulfilled. Therefore, collecting independent copies of $K^{1/2}X^{\square}$ in groups as in (4.9), we can apply inequality (8.34) of Lemma 8.1. By (3.4), (3.13) and (8.34), for any $\gamma > 0$ and $|t| \leq t_1$,

$$\varkappa^*(tK; N, X^{\square}) \ll_{\gamma,d} (pN/m)^{-\gamma} + \min\{1; (Np/m)^{-d/2}|t|^{-d/2}(\det \mathbb{C})^{-1/2}\}.$$

We have used that $\sigma^2 = 1$ implies $\sigma_1^2 \asymp_d 1$. A similar upper bound is valid for the quantity $\varkappa^*(tK; N, G)$; cf. the proof of (4.6) for $k > N/2$. Thus we get for any $\gamma > 0$ and $|t| \leq t_1$,

$$\varkappa(tK; N, X^{\square}) \ll_{\gamma,d} (pN/m)^{-\gamma} + \min\{1; (\det \mathbb{C})^{-1/2}(m/(|t|pN))^{d/2}\}.$$

Integrating this bound (cf. the estimation of I_1 in Lemma 3.1), we obtain (4.14). $\square$

5. Proof of Theorem 2.2. To simplify notation, in Section 5 we write $\Pi = \Pi_2^\square$ and $\Lambda = \Lambda_4^\square$. The assumption $\sigma^2 = 1$ and equalities $\mathbf{E}\|\mathbb{C}^{-1/2}X\|^2 = d$, (2.5) and (2.7) imply

$$\Pi + \Lambda N \gg 1, \qquad \Pi + \Lambda \le 1, \qquad \sigma_j^2 \le 1, \qquad \det \mathbb{C} \le 1. \tag{5.1}$$

Recall that $\Delta_N^{(a)}$ and functions Ψ_b, Φ_b and Θ_b are defined in (2.14) and (2.17)–(2.19). Note now that $\Theta_b^\square(x) = E_a^\square(x/N)$ and, according to (2.19),

$$\Delta_N^{(a)} \le \Delta_{N,\square}^{(a)} + \sup_{x\in\mathbb{R}} \bigl|\Theta_b(x) - \Theta_b^\square(x)\bigr|, \tag{5.2}$$

where $b = \sqrt{N}a$ and

$$\Delta_{N,\square}^{(a)} = \sup_{x\in\mathbb{R}} \bigl|\Psi_b(x) - \Phi_b(x) - \Theta_b^\square(x)\bigr|. \tag{5.3}$$

Let us verify that

$$\sup_{x\in\mathbb{R}} \bigl|\Theta_b(x) - \Theta_b^\square(x)\bigr| \ll_d \Pi_3^\square. \tag{5.4}$$

To this end we apply representation (2.14)–(2.15) of the Edgeworth correction as a signed measure and estimate the variation of that measure. Indeed, using (2.14)–(2.15), we have

$$\begin{gathered} \sup_{x\in\mathbb{R}} \bigl|\Theta_b(x) - \Theta_b^\square(x)\bigr| \ll N^{-1/2} I, \\ I \stackrel{\text{def}}{=} \int_{\mathbb{R}^d} \bigl|\mathbf{E}p'''(x)X^3 - \mathbf{E}p'''(x)X^{\square 3}\bigr|\, dx. \end{gathered} \tag{5.5}$$

By the explicit formula (2.16), the function $u \mapsto p'''(x)u^3$ is a 3-linear form in the variable u. Therefore, using $X = X^\square + X_\square$ and $\|X^\square\|\|X_\square\| = 0$, we have $p'''(x)X^3 - p'''(x)X^{\square 3} = p'''(x)X_\square^3$, and

$$N^{-1/2} I \le 3d^{3/2}\Pi_3^\square \int_{\mathbb{R}^d} \bigl(\|\mathbb{C}^{-1/2}x\| + \|\mathbb{C}^{-1/2}x\|^3\bigr) p(x)\, dx = c_d \Pi_3^\square. \tag{5.6}$$

Inequalities (5.5) and (5.6) imply now (5.4).

To prove the statement of Theorem 2.2, we have to derive that

$$\Delta_{N,\square}^{(a)} \ll_d (\Pi + \Lambda)\bigl(1 + \|a\|\bigr)^3 (\det \mathbb{C})^{-1/2}. \tag{5.7}$$

While proving (5.7) we assume that

$$\Pi \le c_d \quad \text{and} \quad \Lambda \le c_d, \tag{5.8}$$

with a sufficiently small positive constant c_d depending on d only. These assumptions do not restrict generality. Indeed, we have $|\Psi_b(x) - \Phi_b(x)| \le 1$. If conditions (5.8) do not hold, then the estimate

$$\sup_{x\in\mathbb{R}}\bigl|\Theta_b^{\square}(x)\bigr| \ll_d N^{-1/2}\mathbf{E}\bigl\|\mathbb{C}^{-1/2}X^{\square}\bigr\|^3 \ll_d \Lambda^{1/2} \tag{5.9}$$

immediately implies (5.7). In order to prove (5.9) we can use (2.7) and representation (2.14)–(2.15) of the Edgeworth correction. Estimating the variation of that measure and using

$$\mathbf{E}\bigl\|\mathbb{C}^{-1/2}X^{\square}\bigr\|^2 \le \mathbf{E}\bigl\|\mathbb{C}^{-1/2}X\bigr\|^2 = d, \tag{5.10}$$

$$\bigl(\mathbf{E}\bigl\|\mathbb{C}^{-1/2}X^{\square}\bigr\|^3\bigr)^2 \le \mathbf{E}\bigl\|\mathbb{C}^{-1/2}X^{\square}\bigr\|^2\mathbf{E}\bigl\|\mathbb{C}^{-1/2}X^{\square}\bigr\|^4, \tag{5.11}$$

we obtain (5.9).

It is clear that

$$\begin{aligned}\Delta_{N,\square}^{(a)} \le \sup_{x\in\mathbb{R}}\bigl(&\bigl|\Psi_b(x) - \Psi_b'(x)\bigr| + \bigl|\Theta_b^{\square}(x) - \Theta_b'(x)\bigr| \\ &+ \bigl|\Psi_b'(x) - \Phi_b(x) - \Theta_b'(x)\bigr|\bigr).\end{aligned} \tag{5.12}$$

Similarly to (5.5), we have

$$\begin{gathered}\sup_{x\in\mathbb{R}}\bigl|\Theta_b^{\square}(x) - \Theta_b'(x)\bigr| \ll N^{-1/2}J, \\ J \overset{\text{def}}{=} \int_{\mathbb{R}^d}\bigl|\mathbf{E}p'''(x)X^{\square 3} - \mathbf{E}p'''(x)X'^3\bigr|\,dx.\end{gathered} \tag{5.13}$$

Recall that vector X' is defined in (4.4). By Lemma 3.2, we have $\mathbf{E}\|\mathbb{C}^{-1/2}W\|^2 \le 2d\Pi$ (hence, $\mathbf{E}\|\mathbb{C}^{-1/2}W\|^q \ll_d \Pi^{q/2}$, for $0 \le q \le 2$). Using the well-known equivalence of moments of Gaussian random vectors, we conclude that

$$\mathbf{E}\bigl\|\mathbb{C}^{-1/2}W\bigr\|^q \ll_q \bigl(\mathbf{E}\bigl\|\mathbb{C}^{-1/2}W\bigr\|^2\bigr)^{q/2} \ll_{q,d} \Pi^{q/2}, \qquad q \ge 0. \tag{5.14}$$

Furthermore, according to (2.5), (2.7) and (5.8),

$$\mathbf{E}\bigl\|\mathbb{C}^{-1/2}X_{\square}\bigr\| \ll_d \Pi N^{-1/2} \ll_d \Pi^{1/2}N^{-1/2}. \tag{5.15}$$

Hence, by (2.7), (4.4), (5.1), (5.14) and (5.15),

$$\mathbf{E}\bigl\|X'\bigr\|^4 \ll \overline{\beta} \overset{\text{def}}{=} \mathbf{E}\bigl\|\mathbb{C}^{-1/2}X'\bigr\|^4 \ll_d N\Lambda + \Pi^2. \tag{5.16}$$

Using (2.16), (5.1), (5.8), (5.10) and (5.13)–(5.15), we get

$$\begin{aligned}N^{-1/2}J &\ll_d \Pi^{1/2}\bigl(N^{-1/2}\Pi + \Lambda^{1/2}\bigr)\int_{\mathbb{R}^d}\bigl(\bigl\|\mathbb{C}^{-1/2}x\bigr\| + \bigl\|\mathbb{C}^{-1/2}x\bigr\|^3\bigr)p(x)\,dx \\ &\ll_d \Pi + \Lambda.\end{aligned} \tag{5.17}$$

Thus, according to (5.13) and (5.17),

$$\sup_{x\in\mathbb{R}}\bigl|\Theta_b^{\square}(x)-\Theta_b'(x)\bigr|\ll_d \Pi+\Lambda. \tag{5.18}$$

The same approach is applicable for the estimation of $|\Theta_b'|$. Using (2.14)–(2.16), (4.4), (5.1), (5.10), (5.11), (5.14) and (5.15), we get

$$\begin{aligned}\sup_{x\in\mathbb{R}}\bigl|\Theta_b'(x)\bigr| &\ll N^{-1/2}\int_{\mathbb{R}^d}\bigl|\mathbf{E}p'''(x)X'^3\bigr|\,dx\\ &\ll_d \Lambda^{1/2}+N^{-1/2}\Pi^{3/2}.\end{aligned} \tag{5.19}$$

Let us prove that

$$\sup_{x\in\mathbb{R}}\bigl|\Psi_b(x)-\Psi_b'(x)\bigr|\ll(\det\mathbb{C})^{-1/2}p^{-2}(\Pi+\Lambda)\bigl(1+\|a\|^2\bigr). \tag{5.20}$$

Using truncation [see (3.11)], we have $|\Psi_b-\Psi_b^{\square}|\le\Pi$, and

$$\sup_{x\in\mathbb{R}}\bigl|\Psi_b(x)-\Psi_b'(x)\bigr|\le\Pi+\sup_{x\in\mathbb{R}}\bigl|\Psi_b^{\square}(x)-\Psi_b'(x)\bigr|. \tag{5.21}$$

In order to estimate $|\Psi_b^{\square}-\Psi_b'|$, we apply Lemmas 4.1 and 4.2. The number m in these Lemmas exists and $N\Lambda/p\gg_d 1$, as it follows from (5.1) and (5.8). Let us choose the minimal m, that is, $m\asymp_d N\Lambda/p$. Then $(pN)^{-1}m\ll_d\Lambda/p^2$ and $m/N\ll_d\Lambda/p$. Therefore, using Lemma 4.1, we have

$$\begin{aligned}&\sup_x\bigl|\Psi_b^{\square}(x)-\Psi_b'(x)\bigr|\\ &\quad\ll_d p^{-2}\Lambda(\det\mathbb{C})^{-1/2}+\int_{|t|\le t_1}\bigl|\widehat{\Psi}_b^{\square}(\tau)-\widehat{\Psi}_b'(\tau)\bigr|\frac{dt}{|t|},\qquad \tau=tK.\end{aligned} \tag{5.22}$$

We shall prove that

$$\bigl|\widehat{\Psi}_b^{\square}(\tau)-\widehat{\Psi}_b'(\tau)\bigr|\ll_d \varkappa\Pi|\tau|N\bigl(1+|\tau|N\bigr)\bigl(1+\|a\|^2\bigr) \tag{5.23}$$

with $\varkappa=\varkappa(\tau;N,X^{\square})$. Combining (5.21)–(5.23), using $\tau=tK$ and integrating inequality (5.23) with the help of Lemma 4.2, we derive (5.20).

Let us prove (5.23). Writing $D=Z_N^{(\square)}-\mathbf{E}Z_N^{(\square)}-b$, we have

$$Z_N^{(\square)}-b=D+\mathbf{E}Z_N^{(\square)},\qquad \mathcal{L}\bigl(Z_N'-b\bigr)=\mathcal{L}(D+\sqrt{N}W)$$

and

$$\bigl|\widehat{\Psi}_b^{\square}(\tau)-\widehat{\Psi}_b'(\tau)\bigr|\le\bigl|f_1(\tau)\bigr|+\bigl|f_2(\tau)\bigr| \tag{5.24}$$

with

$$\begin{aligned}f_1(\tau)&=\mathbf{E}\,\mathrm{e}\bigl\{\tau\mathbb{Q}[D+\sqrt{N}W]\bigr\}-\mathbf{E}\,\mathrm{e}\bigl\{\tau\mathbb{Q}[D]\bigr\},\\ f_2(\tau)&=\mathbf{E}\,\mathrm{e}\bigl\{\tau\mathbb{Q}\bigl[D+\mathbf{E}Z_N^{(\square)}\bigr]\bigr\}-\mathbf{E}\,\mathrm{e}\bigl\{\tau\mathbb{Q}[D]\bigr\}.\end{aligned} \tag{5.25}$$

Now we have to prove that both $|f_1(\tau)|$ and $|f_2(\tau)|$ may be estimated by the right-hand side of (5.23).

Let us consider f_1. We can write $\mathbb{Q}[D+\sqrt{N}W]=\mathbb{Q}[D]+A+B$ with $A=2\sqrt{N}\langle\mathbb{Q}D,W\rangle$ and $B=N\mathbb{Q}[W]$. Taylor's expansions of the exponent in (5.25) in powers of $i\tau B$ and $i\tau A$ with remainders $\mathcal{O}(\tau B)$ and $\mathcal{O}(\tau^2A^2)$, respectively, imply (recall that $\mathbf{E}W=0$ and $\mathbb{Q}^2=\mathbb{I}_d$)

$$|f_1(\tau)|\ll\varkappa|\tau|N\mathbf{E}\|W\|^2+\varkappa\tau^2N\mathbf{E}\|W\|^2\mathbf{E}\|D\|^2, \tag{5.26}$$

where $\varkappa=\varkappa(\tau;N,X^\square)$. The estimation of the remainders of these expansions is based on the splitting and conditioning techniques described in Section 9 of BG (1997a); see also Bentkus, Götze and Zaitsev (1997). Using the relations $\mathbf{E}\|W\|^2\ll\mathbf{E}\|\mathbb{C}^{-1/2}W\|^2\ll_d\Pi$, $\sigma^2=1$ and $\mathbf{E}\|D\|^2\ll N(1+\|a\|^2)$, we derive from (5.26) that

$$|f_1(\tau)|\ll_d\varkappa\Pi|\tau|N\big(1+|\tau|N\big)\big(1+\|a\|^2\big). \tag{5.27}$$

Note that $\mathbf{E}Z_N^{(\square)}=N\mathbf{E}X^\square=-N\mathbf{E}X_\square$. Expanding the exponent $\mathrm{e}\{\tau\mathbb{Q}[D+\mathbf{E}Z_N^{(\square)}]\}$, using (5.15) and proceeding similarly to the proof of (5.27), we obtain

$$|f_2(\tau)|\ll_d\varkappa\Pi|\tau|N\big(1+\|a\|\big). \tag{5.28}$$

Inequalities (5.24), (5.27) and (5.28) imply now (5.23).

It remains to estimate $|\Psi_b'-\Phi_b-\Theta_b'|$. Recall that the distribution functions $\Psi_b^{(l)}(x)$, for $0\le l\le N$, are defined in (4.5).

Fix an integer k, $1\le k\le N$. Clearly, we have

$$\sup_{x\in\mathbb{R}}\big|\Psi_b'(x)-\Phi_b(x)-\Theta_b'(x)\big|\le I_1+I_2+I_3, \tag{5.29}$$

where

$$I_1=\sup_{x\in\mathbb{R}}\big|\Psi_b^{(k)}(x)-\Phi_b(x)-(N-k)\Theta_b'(x)/N\big|, \tag{5.30}$$

$$I_2=\sup_{x\in\mathbb{R}}\big|\Psi_b'(x)-\Psi_b^{(k)}(x)\big| \tag{5.31}$$

and

$$I_3=\sup_{x\in\mathbb{R}}kN^{-1}\big|\Theta_b'(x)\big|. \tag{5.32}$$

Let estimate I_1. Define the distributions

$$\begin{aligned}\mu(A)&=\mathbf{P}\Big\{U_k+\sum_{j=k+1}^{N}X_j'\in\sqrt{N}A\Big\},\\ \mu_0(A)&=\mathbf{P}\{U_N\in\sqrt{N}A\}=\mathbf{P}\{G\in A\},\end{aligned} \tag{5.33}$$

where $U_l = G_1 + \cdots + G_l$. Introduce the measure χ' replacing X by X' in (2.15). For the Borel sets $A \subset \mathbb{R}^d$ define the Edgeworth correction (to the distribution μ) as

$$\mu_1^{(k)}(A) = (N-k)N^{-3/2}\chi'(A)/6. \tag{5.34}$$

Introduce the signed measure

$$\nu = \mu - \mu_0 - \mu_1^{(k)}. \tag{5.35}$$

It is easy to see that a re-normalization of random vectors implies [see relations (2.14), (2.17)–(2.19), (4.5) and (5.33)–(5.35)]

$$\begin{aligned} \big|\Psi_b^{(k)}(x) - \Phi_b(x) - (N-k)\Theta_b'(x)/N\big| &= \nu\big(\{u \in \mathbb{R}^d : \mathbb{Q}[u-a] \le x/N\}\big) \\ &\le \delta_N \stackrel{\text{def}}{=} \sup_{A \subset \mathbb{R}^d} \big|\nu(A)\big|. \end{aligned} \tag{5.36}$$

LEMMA 5.1. *Assume that $d < \infty$ and $1 \le k \le N$. Then there exists a $c(d)$ depending on d only and such that δ_N defined in* (5.36) *satisfies the inequality*

$$\delta_N \ll_d \frac{\overline{\beta}}{N} + \frac{N^{d/2}}{k^{d/2}} \exp\{-c(d)k/\overline{\beta}\} \tag{5.37}$$

with $\overline{\beta} = \mathbf{E}\|\mathbb{C}^{-1/2}X'\|^4$.

An outline of the proof. We repeat and slightly improve the proof of Lemma 9.4 in BG (1997a); cf. the proof of Lemma 2.5 in BG (1997a). We shall prove (5.37) assuming that $\operatorname{cov} X = \operatorname{cov} X' = \operatorname{cov} G = \mathbb{I}_d$. Applying it to $\mathbb{C}^{-1/2}X'$ and $\mathbb{C}^{-1/2}G$, we obtain (5.37) in general case.

While proving (5.37) we assume that $\overline{\beta}/N \le c_d$ and $N \ge 1/c_d$ with a sufficiently small positive constant c_d. Otherwise (5.37) follows from the obvious bounds $\overline{\beta} \ge \sigma^4 = d^2$ and

$$\delta_N \ll_d 1 + (\overline{\beta}/N)^{1/2} \int_{\mathbb{R}^d} \|x\|^3 p(x)\,dx \ll_d 1 + (\overline{\beta}/N)^{1/2}.$$

Set $n = N - k$. Denoting by Z_j' and U_j' sums of j independent copies of X' and G', respectively, introduce the multidimensional characteristic functions

$$g(t) = \mathbf{E}\,\mathrm{e}\{\langle N^{-1/2}t, G\rangle\}, \qquad h(t) = \mathbf{E}\,\mathrm{e}\{\langle N^{-1/2}t, X'\rangle\}, \tag{5.38}$$

$$\begin{aligned} f(t) &= \mathbf{E}\,\mathrm{e}\{\langle N^{-1/2}t, Z_n'\rangle\} = h^n(t), \\ f_0(t) &= \mathbf{E}\,\mathrm{e}\{\langle N^{-1/2}t, U_n'\rangle\} = g^n(t), \end{aligned} \tag{5.39}$$

$$f_1(t) = n\,m(t) f_0(t) \qquad \text{where } m(t) = \frac{1}{6N^{3/2}} \mathbf{E}\langle it, X'\rangle^3, \tag{5.40}$$

$$\widehat{\nu}(t) = \big(f(t) - f_0(t) - f_1(t)\big) g(\rho t), \qquad \rho^2 = k. \tag{5.41}$$

It is easy to see that

$$\widehat{\nu}(t) = \int_{\mathbb{R}^d} \mathrm{e}\{\langle t, x\rangle\}\nu(dx). \tag{5.42}$$

Using a truncation, we obtain

$$\mathbf{E}\|Z_l'/\sqrt{N}\|^{\gamma} \ll_{\gamma,d} 1, \qquad \gamma > 0, 1 \le l \le N. \tag{5.43}$$

By an extension of the proof of Lemma 11.6 in Bhattacharya and Rao (1986) [see also the proof of Lemma 2.5 in BG (1996)], we obtain

$$\delta_N \ll_d \max_{|\alpha|\le 2d} \int_{\mathbb{R}^d} \bigl|\partial^{\alpha}\widehat{\nu}(t)\bigr|\, dt. \tag{5.44}$$

Here $|\alpha| = |\alpha_1| + \cdots + |\alpha_d|$, $\alpha = (\alpha_1, \ldots, \alpha_d)$, $\alpha_j \in \mathbb{Z}$, $\alpha_j \ge 0$. In order to derive (5.37) from (5.44), it suffices to prove that, for $|\alpha| \le 2d$,

$$\bigl|\partial^{\alpha}\widehat{\nu}(t)\bigr| \ll_d g(c_1\rho t), \tag{5.45}$$

$$\bigl|\partial^{\alpha}\widehat{\nu}(t)\bigr| \ll_d \overline{\beta} N^{-1}(1 + \|t\|^6)\exp\{-c_2\|t\|^2\} \qquad \text{for } \|t\|^2 \le c_3(d)N/\overline{\beta}. \tag{5.46}$$

Indeed, using (5.45) and denoting $T = \sqrt{c_3(d)N/\overline{\beta}}$, we obtain

$$\begin{aligned}\int_{\|t\|\ge T} \bigl|\partial^{\alpha}\widehat{\nu}(t)\bigr|\, dt &\ll_d \int_{\|t\|\ge T} g(c_1\rho t)\, dt \\ &\ll_d \frac{N^{d/2}}{\rho^d} \exp\biggl\{-\frac{c_1^2\rho^2T^2}{8N}\biggr\} \int_{\mathbb{R}^d} \exp\{-c_1^2\|t\|^2/8\}\, dt,\end{aligned} \tag{5.47}$$

and it is easy to see that the right-hand side of (5.47) is bounded from above by the second summand on the right-hand side of (5.37). Similarly, using (5.46), we can integrate $|\partial^{\alpha}\widehat{\nu}(t)|$ over $\|t\| \le T$, and the integral is bounded from above by $c_d\overline{\beta}/N$.

In the proof of (5.45)–(5.47) we applied standard methods of estimation which are provided in Bhattacharya and Rao (1986). In particular, we used a Bergström type identity

$$f - f_0 - f_1 = \sum_{j=0}^{n-1}(h - g - m)h^j g^{n-j-1} + \sum_{j=0}^{n-1} m \sum_{l=0}^{j-1}(h - g)h^l g^{n-l-1}, \tag{5.48}$$

relations (5.38)–(5.43), $1 \le k \le N$, $|\partial^{\alpha}\exp\{-c_4\|t\|^2\}| \ll_{\alpha} \exp\{-c_5\|t\|^2\}$, $\sqrt{N}/\overline{\beta}^{1/2} \gg_d 1$ and $y^{c_d}\exp\{-y\} \ll_d 1$, for $y > 0$.

Applying (5.30), (5.36) and Lemma 5.1, we get

$$I_1 \ll_d \frac{\overline{\beta}}{N} + \frac{N^{d/2}}{k^{d/2}}\exp\{-c(d)k/\overline{\beta}\}. \tag{5.49}$$

For the estimation of I_2 we shall use Lemma 5.2 which is an easy consequence of BG [(1997a), Lemma 9.3], (4.13) and (5.16).

LEMMA 5.2. *We have*

$$\big|\widehat{\Psi}_b'(t)-\widehat{\Psi}_b^{(l)}(t)\big| \ll \varkappa t^2 l\big(\overline{\beta}+|t|N\overline{\beta}+|t|N\sqrt{N\overline{\beta}}\big)\big(1+\|a\|^3\big) \qquad \text{for } 0\le l\le N,$$

where $\varkappa=\varkappa(t;N,X^{\square})$; *cf.* (4.12).

As in the proof of (5.22), applying Lemma 4.1 [choosing $m \asymp_d N(\Lambda+\Pi)/p$] and using (3.13), we obtain

$$I_2 \ll_d (\Lambda+\Pi)(\det\mathbb{C})^{-1/2}+\int_{|t|\le t_1}\big|\widehat{\Psi}_b'(\tau)-\widehat{\Psi}_b^{(k)}(\tau)\big|\,dt/|t|, \qquad \tau=tK.$$

The existence of such an m is ensured by (3.13), (5.1) and (5.8). Applying Lemma 5.2 and replacing in that lemma t by τ, we have

$$\big|\widehat{\Psi}_b'(\tau)-\widehat{\Psi}_b^{(k)}(\tau)\big| \ll \varkappa\tau^2 k\big(\overline{\beta}+|\tau|N\overline{\beta}+|\tau|N\sqrt{N\overline{\beta}}\big)\big(1+\|a\|^3\big). \tag{5.50}$$

Integrating with the help of Lemma 4.2 and using (3.13), we obtain

$$\begin{aligned} I_2 \ll_d\ &(\det\mathbb{C})^{-1/2}\big(\Pi+\Lambda+kN^{-2}(\overline{\beta}+\sqrt{N\overline{\beta}})\\ &\times\big(1+(\Pi+\Lambda)^{-1/d}\big)\big(1+\|a\|^3\big)\big). \end{aligned} \tag{5.51}$$

Let us choose $k\asymp_d N^{1/4}\overline{\beta}^{3/4}$. Such $k\le N$ exists by $\overline{\beta}\gg_d \sigma^4=1$, by (5.16) and by assumption (5.8). Then (5.49) and (5.51) turn into

$$I_1 \ll_d \frac{\overline{\beta}}{N}+\left(\frac{N}{\overline{\beta}}\right)^{3d/8}\exp\left\{-c_d\left(\frac{N}{\overline{\beta}}\right)^{1/4}\right\} \ll_d \frac{\overline{\beta}}{N} \tag{5.52}$$

and

$$\begin{aligned} I_2 \ll_d\ &(\det\mathbb{C})^{-1/2}\left(\Pi+\Lambda+\left(\left(\frac{\overline{\beta}}{N}\right)^{5/4}+\left(\frac{\overline{\beta}}{N}\right)^{7/4}\right)\right.\\ &\left.\times\big(1+(\Pi+\Lambda)^{-1/d}\big)\big(1+\|a\|^3\big)\right). \end{aligned} \tag{5.53}$$

Using (3.13), (5.8), (5.16) and (5.53), we get

$$I_2 \ll_d (\det\mathbb{C})^{-1/2}\left(\Pi+\Lambda+\frac{\overline{\beta}}{N}\big(1+\|a\|^3\big)\right). \tag{5.54}$$

Finally, by (5.8), (5.16), (5.20) and (5.32),

$$I_3 \ll_d \frac{k}{N}\big(\Lambda^{1/2}+N^{-1/2}\Pi^{3/2}\big) \ll \Lambda+\Pi. \tag{5.55}$$

Inequalities (5.8), (5.12), (5.16), (5.18), (5.20), (5.29), (5.52), (5.54) and (5.55) imply now (5.7) [and, hence, (2.23)] by an application of $\Pi+\Lambda\le 1$. Note that, by (2.7), we have $\Pi\le\Pi_3^{\square}$. Together with (5.2) and (5.4), inequality (5.7) yields (2.22). The statement of Theorem 2.2 is proved. □

6. From probability to number theory. In Section 6 we reduce the estimation of the integrals of the modulus of characteristic functions $\widehat{\Psi}_b(t)$ to the estimation the integrals of some theta-series. We shall use the following lemmas.

LEMMA 6.1 [BG (1997a), Lemma 5.1]. *Let $L, C \in \mathbb{R}^d$ and let $\mathbb{Q}: \mathbb{R}^d \to \mathbb{R}^d$ be a symmetric linear operator. Let Z, U, V and W denote independent random vectors taking values in $\mathbb{R}^d$. Denote by*

$$P(x) = \langle \mathbb{Q}x, x\rangle + \langle L, x\rangle + C, \qquad x \in \mathbb{R}^d,$$

a real-valued polynomial of second order. Then

$$2\big|\mathbf{E}\,\mathrm{e}\{t P(Z+U+V+W)\}\big|^2 \le \mathbf{E}\,\mathrm{e}\{2t\langle \mathbb{Q}\widetilde{Z}, \widetilde{U}\rangle\} + \mathbf{E}\,\mathrm{e}\{2t\langle \mathbb{Q}\widetilde{Z}, \widetilde{V}\rangle\}.$$

Let $\delta > 0$, $\mathcal{S} = \{e_1, \dots, e_s\} \subset \mathbb{R}^d$ and let $\mathbb{D}: \mathbb{R}^d \to \mathbb{R}^d$ be a linear operator. Usually, we take $\mathbb{D} = \mathbb{C}^{-1/2}$. Denote

$$\mathbf{\Gamma}(\delta; \mathbb{D}, \mathcal{S}) = \big\{(z_1, \dots, z_s) : z_j \in \mathbb{R}^d, \|\mathbb{D}z_j - e_j\| \le \delta, \text{ for all } 1 \le j \le s\big\}. \tag{6.1}$$

Recall that $\mathcal{S}_o = \{e_1, \dots, e_s\} \subset \mathbb{R}^d$ denotes an orthonormal system.

Let $\{\varepsilon_{jk}, j = 1, 2 \dots, s; k = 1, 2 \dots\} \cup \{\varepsilon'_{jk}, j = 1, 2 \dots, s; k = 1, 2 \dots\}$ be i.i.d. symmetric Rademacher random variables.

LEMMA 6.2. *Assume that $\mathbb{Q}^2 = \mathbb{I}_d$ and that the condition $P(\delta, \mathcal{S}, \mathbb{D}\widetilde{X}) \ge p$ holds with some $p > 0$ and $\delta > 0$. Write $m = \lfloor pN/(5s)\rfloor$. Then, for any $0 < A \le B$, $b \in \mathbb{R}^d$ and $\gamma > 0$, we have*

$$\int_A^B \big|\widehat{\Psi}_b(t)\big| \frac{dt}{|t|} \le I + c_\gamma(s)(pN)^{-\gamma} \log \frac{B}{A}, \tag{6.2}$$

with

$$I = \sup_{\Gamma} \sup_{b \in \mathbb{R}^d} \int_A^B \sqrt{\varphi_b(t/4)} \frac{dt}{|t|}, \qquad \varphi_b(t) \overset{\text{def}}{=} \big|\mathbf{E}\,\mathrm{e}\{t\mathbb{Q}[Y+b]\}\big|^2, \tag{6.3}$$

where $Y = \sum_{k=1}^m U_k$ denote a sum of independent (non i.i.d.) vectors $U_k = \sum_{j=1}^s \varepsilon_{jk} z_{jk}$, and $\sup_\Gamma$ is taken over all $\{(z_{1k}, \dots, z_{sk}) \in \mathbf{\Gamma}(\delta; \mathbb{D}, \mathcal{S}), k = 1, \dots, m\}$.

Lemma 6.2 is an analogue of Corollary 6.3 from BG (1997a). Its proof is even simpler than that in BG (1997a). Therefore it is omitted.

LEMMA 6.3. *Assume that $\mathbb{Q}^2 = \mathbb{I}_d$ and that the condition $P(\delta, \mathcal{S}, \mathbb{D}\widetilde{X}) \ge p$ holds with some $p > 0$ and $\delta > 0$. Let*

$$n \overset{\text{def}}{=} \lfloor pN/(16s)\rfloor \ge 1. \tag{6.4}$$

Then, for any $0 < A \le B$, $b \in \mathbb{R}^d$ *and* $\gamma > 0$,

$$(6.5)\qquad \int_A^B |\widehat{\Psi}_b(t)| \frac{dt}{|t|} \le c_\gamma(s)(pN)^{-\gamma} \log\frac{B}{A} + \sup_\Gamma \int_A^B \sqrt{\mathbf{E}\,\mathrm{e}\{t\langle \mathbb{Q}\widetilde{W}, \widetilde{W}'\rangle/2\}} \frac{dt}{|t|},$$

and for any fixed $t \in \mathbb{R}$,

$$(6.6)\qquad |\widehat{\Psi}_b(t)| \le c_\gamma(s)(pN)^{-\gamma} + \sup_\Gamma \sqrt{\mathbf{E}\,\mathrm{e}\{t\langle \mathbb{Q}\widetilde{W}, \widetilde{W}'\rangle/2\}},$$

where $W = V_1 + \cdots + V_n$ *and* $W' = V_1' + \cdots + V_n'$ *are independent sums of independent copies of random vectors* $V = \sum_{j=1}^s \varepsilon_{j1} z_j$ *and* $V' = \sum_{j=1}^s \varepsilon_{j1}' z_j'$, *and* $\sup_\Gamma$ *is taken over all* $(z_1, \dots, z_s), (z_1', \dots, z_s') \in \mathbf{\Gamma}(\delta; \mathbb{D}, \mathcal{S})$.

Note that this lemma will be proved for general $\mathcal{S}$, but in this paper we need $\mathcal{S} = \mathcal{S}_o$ only. Moreover, a more careful estimation of binomial probabilities could allow us to replace $c_\gamma(s)(pN)^{-\gamma}$ in (6.2), (6.5) and (6.6) by $c(s)\exp\{-cpN\}$; see, for example, Nagaev and Chebotarev (2005). However, we do not need to use this improvement.

PROOF OF LEMMA 6.3. Inequality (6.6) is an analogue of the statement of Lemma 7.3 from BG (1997a). Its proof is even simpler than that in BG (1997a). Therefore it is omitted.

Let us show that

$$(6.7)\qquad \int_A^B |\widehat{\Psi}_b(t)| \frac{dt}{|t|} \le c_\gamma(s)(pN)^{-\gamma} \log\frac{B}{A} + \sup_\Gamma \int_A^B \sqrt{\mathbf{E}\,\mathrm{e}\{t\langle \mathbb{Q}\widetilde{W}, \widetilde{W}'\rangle/2\}} \frac{dt}{|t|},$$

where $W = V_1 + \cdots + V_n$ and $W' = V_1' + \cdots + V_n'$ are independent sums of independent (non i.i.d.) vectors $V_k = \sum_{j=1}^s \varepsilon_{jk} z_{jk}$, and $V_k' = \sum_{j=1}^s \varepsilon_{jk}' z_{jk}'$, respectively, while $\sup_\Gamma$ is taken over all $\{(z_{1k}, \dots, z_{sk}), (z_{1k}', \dots, z_{sk}') \in \mathbf{\Gamma}(\delta; \mathbb{D}, \mathcal{S}), k = 1, \dots, n\}$.

Comparing (6.5) and (6.7), we see that inequality (6.7) is related to sums of *non i.i.d.* vectors $\{V_j\}$ and $\{V_j'\}$ while inequality (6.5) deals with i.i.d. vectors. Nevertheless, we derive (6.5) from (6.7).

While proving (6.7) we can assume that $pN \ge c_s$ with a sufficiently large constant c_s, since otherwise (6.7) is obviously valid.

Let $\varphi_b(t)$ be defined in (6.3), where $Y = \sum_{k=1}^m U_k$ is a sum of independent (non i.i.d.) vectors $U_k = \sum_{j=1}^s \varepsilon_{jk} z_{jk}$, where $\{(z_{1k}, \dots, z_{sk}) \subset \mathbf{\Gamma}(\delta; \mathbb{D}, \mathcal{S}), k = 1, \dots, m\}$, $m = \lfloor pN/(5s) \rfloor$.

We shall apply the symmetrization Lemma 6.1. Split $Y = T + T_1 + T_2$ into sums of independent sums of independent summands so that each of the sums T, T_1 and T_2 contains $n = \lfloor pN/(16s) \rfloor$ independent summands U_j. Such an n exists since $pN \ge c_s$ with a sufficiently large c_s. Lemma 6.1 implies that

$$(6.8)\qquad 2\varphi_b(t) \le \mathbf{E}\,\mathrm{e}\{2t\langle \mathbb{Q}\widetilde{T}, \widetilde{T}_1\rangle\} + \mathbf{E}\,\mathrm{e}\{2t\langle \mathbb{Q}\widetilde{T}, \widetilde{T}_2\rangle\}.$$

Inequality (6.7) follows now from (6.8) and Lemma 6.2.

Let now $W = V_1 + \dots + V_n$ and $W' = V_1' + \dots + V_n'$ be independent sums of independent vectors $V_k = \sum_{j=1}^{s} \varepsilon_{jk} z_{jk}$, and $V_k' = \sum_{j=1}^{s} \varepsilon_{jk}' z_{jk}'$, respectively, with $\{(z_{1k}, \dots, z_{sk}), (z_{1k}', \dots, z_{sk}') \in \mathbf{\Gamma}(\delta; \mathbb{D}, \mathcal{S}), k = 1, \dots, n\}$.

Using that all random vectors $\widetilde{V}_k$ are symmetrized and have nonnegative characteristic functions and applying Hölder's inequality, we obtain, for each t,

$$\mathbf{E}\,\mathrm{e}\{t\langle \mathbb{Q}\widetilde{W}, \widetilde{W}'\rangle\} = \mathbf{E}_{\widetilde{W}'}\left(\prod_{k=1}^{n} \mathbf{E}_{\widetilde{V}_k}\,\mathrm{e}\{t\langle \mathbb{Q}\widetilde{V}_k, \widetilde{W}'\rangle\}\right) \tag{6.9}$$

$$\leq \left(\prod_{k=1}^{n} \mathbf{E}_{\widetilde{W}'}(\mathbf{E}_{\widetilde{V}_k}\,\mathrm{e}\{t\langle \mathbb{Q}\widetilde{V}_k, \widetilde{W}'\rangle\})^n\right)^{1/n} \tag{6.10}$$

$$= \left(\prod_{k=1}^{n} \mathbf{E}_{\widetilde{W}'}(\mathbf{E}_{\widetilde{T}_k}\,\mathrm{e}\{t\langle \mathbb{Q}\widetilde{T}_k, \widetilde{W}'\rangle\})\right)^{1/n} \tag{6.11}$$

$$= \left(\prod_{k=1}^{n} \mathbf{E}\,\mathrm{e}\{t\langle \mathbb{Q}\widetilde{T}_k, \widetilde{W}'\rangle\}\right)^{1/n}, \tag{6.12}$$

where $\widetilde{T}_k \overset{\text{def}}{=} \sum_{l=1}^{n} \widetilde{V}_{kl}$ denotes a sum of i.i.d. copies $\widetilde{V}_{kl}$ of $\widetilde{V}_k$ which are independent of all other random vectors and variables.

Repeating the steps (6.9)–(6.12) for each factor $\mathbf{E}\,\mathrm{e}\{t\langle \mathbb{Q}\widetilde{T}_k, \widetilde{W}'\rangle\}$ instead of the expectation $\mathbf{E}\,\mathrm{e}\{t\langle \mathbb{Q}\widetilde{W}, \widetilde{W}'\rangle\}$ on the right-hand side separately, we get (with $\widetilde{T}_i' \overset{\text{def}}{=} \sum_{l=1}^{n} \widetilde{V}_{il}'$, where $\widetilde{V}_{il}'$ are i.i.d. copies of $\widetilde{V}_i'$ independent of all other random vectors)

$$\mathbf{E}\,\mathrm{e}\{t\langle \mathbb{Q}\widetilde{W}, \widetilde{W}'\rangle\} \leq \left(\prod_{k=1}^{n}\prod_{i=1}^{n} \mathbf{E}\,\mathrm{e}\{t\langle \mathbb{Q}\widetilde{T}_k, \widetilde{T}_i'\rangle\}\right)^{1/n^2}. \tag{6.13}$$

Thus, using (6.13) and the arithmetic-geometric mean inequality, we have

$$\begin{aligned}
&\int_A^B \sqrt{\mathbf{E}\,\mathrm{e}\{t\langle \mathbb{Q}\widetilde{W}, \widetilde{W}'\rangle/2\}}\,\frac{dt}{|t|} \\
&\quad\leq \int_A^B \left(\prod_{k=1}^{n}\prod_{i=1}^{n} \mathbf{E}\,\mathrm{e}\{t\langle \mathbb{Q}\widetilde{T}_k, \widetilde{T}_i'\rangle/2\}\right)^{1/2n^2} \frac{dt}{|t|} \\
&\quad\leq \frac{1}{n^2}\sum_{k=1}^{n}\sum_{i=1}^{n}\int_A^B (\mathbf{E}\,\mathrm{e}\{t\langle \mathbb{Q}\widetilde{T}_k, \widetilde{T}_i'\rangle/2\})^{1/2}\,\frac{dt}{|t|} \\
&\quad\leq \sup_{\Gamma}\int_A^B \sqrt{\mathbf{E}\,\mathrm{e}\{t\langle \mathbb{Q}\widetilde{T}, \widetilde{T}'\rangle/2\}}\,\frac{dt}{|t|},
\end{aligned} \tag{6.14}$$

where $T = U_1 + \dots + U_n$ and $T' = U_1' + \dots + U_n'$ are independent sums of independent copies of random vectors $U = \sum_{j=1}^{s} \varepsilon_{j1} z_1$ and $U' = \sum_{j=1}^{s} \varepsilon_{j1}' z_1'$, and $\sup_{\Gamma}$ is taken over all $(z_1, \dots, z_s), (z'_1, \dots, z'_s) \in \mathbf{\Gamma}(\delta; \mathbb{D}, \mathcal{S})$. Inequalities (6.7) and (6.14) imply now the statement of the lemma. $\square$

The following Lemma 6.4 provides a Poisson summation formula.

LEMMA 6.4. *Let* $\operatorname{Re} z > 0, a, b \in \mathbb{R}^s$ *and* $\mathbb{S} : \mathbb{R}^s \to \mathbb{R}^s$ *be a positive definite symmetric nondegenerate linear operator. Then*

$$\begin{aligned}\sum_{m \in \mathbb{Z}^s} &\exp\{-z\mathbb{S}[m+a] + 2\pi i \langle m, b\rangle\} \\ &= \big(\det(\mathbb{S}/\pi)\big)^{-1/2} z^{-s/2} \exp\{-2\pi i \langle a, b\rangle\} \\ &\qquad \times \sum_{l \in \mathbb{Z}^s} \exp\Big\{-\frac{\pi^2}{z}\mathbb{S}^{-1}[l+b] - 2\pi i \langle a, l\rangle\Big\},\end{aligned}$$

where $\mathbb{S}^{-1} : \mathbb{R}^s \to \mathbb{R}^s$ *denotes the inverse positive definite operator for* $\mathbb{S}$.

PROOF. See, for example, Fricker (1982), page 116, or Mumford (1983), page 189, formula (5.1); and page 197, formula (5.9). $\square$

Let the conditions of Lemma 6.3 be satisfied. Introduce one-dimensional lattice probability distributions $H_n = \mathcal{L}(\xi_n)$ with integer valued ξ_n setting

$$\mathbf{P}\{\xi_n = k\} = A_n n^{-1/2} \exp\{-k^2/2n\} \qquad \text{for } k \in \mathbb{Z}.$$

It is easy to see that $A_n \asymp 1$. Moreover, by Lemma 6.4,

$$\widehat{H}_n(t) \ge 0 \qquad \text{for all } t \in \mathbb{R}. \tag{6.15}$$

Introduce the s-dimensional random vector ζ_n having as coordinates independent copies of ξ_n. Then, for $m = (m_1, \dots, m_s) \in \mathbb{Z}^s$, we have

$$q(m) \overset{\text{def}}{=} \mathbf{P}\{\zeta_n = m\} = A_n^s n^{-s/2} \exp\{-\|m\|^2/2n\}. \tag{6.16}$$

LEMMA 6.5. *Let* $W = V_1 + \cdots + V_n$ *and* $W' = V'_1 + \cdots + V'_n$ *denote independent sums of independent copies of random vectors* V *and* V' *such that*

$$V = \varepsilon_{11} z_1 + \cdots + \varepsilon_{s1} z_s, \qquad V' = \varepsilon'_{11} z'_1 + \cdots + \varepsilon'_{s1} z'_s,$$

with some $z_j, z'_j \in \mathbb{R}^d$. *Introduce the matrix* $\mathbb{B}_t = \{b_{ij}(t) : 1 \le i, j \le s\}$ *with* $b_{ij}(t) = t\langle \mathbb{Q} z_i, z'_j\rangle$. *Then*

$$\mathbf{E}\, e\{t\langle \mathbb{Q}\widetilde{W}, \widetilde{W}'\rangle/4\} \ll_s \mathbf{E}\, e\{\langle \mathbb{B}_t \zeta_n, \zeta'_n\rangle\} + \exp\{-cn\} \qquad \textit{for all } t \in \mathbb{R},$$

where ζ'_n *are independent copies of* ζ_n *and* c *is a positive absolute constant.*

PROOF. Without loss of generality, we assume that $n \ge c_1$, with a sufficiently large absolute constant c_1. Consider the random vector $Y = (\widetilde{\varepsilon}_1, \dots, \widetilde{\varepsilon}_s) \in \mathbb{R}^s$ with coordinates which are symmetrizations of i.i.d. Rademacher random variables. Let $R = (R_1, \dots, R_s)$ and T denote independent sums of n independent copies of $Y/2$. Then we can write

$$\mathbf{E}\,\mathrm{e}\{t\langle\mathbb{Q}\widetilde{W}, \widetilde{W}'\rangle/4\} = \mathbf{E}\,\mathrm{e}\{\langle\mathbb{B}_t R, T\rangle\} \qquad \text{for all } t \in \mathbb{R}. \tag{6.17}$$

Note that the scalar product $\langle\cdot,\cdot\rangle$ in $\mathbf{E}\,\mathrm{e}\{\langle\mathbb{B}_t R, T\rangle\}$ means the scalar product of vectors in $\mathbb{R}^s$. In order to estimate this expectation, we write it in the form

$$\begin{aligned}\mathbf{E}\,\mathrm{e}\{\langle\mathbb{B}_t R, T\rangle\} &= \mathbf{E}\mathbf{E}_R\,\mathrm{e}\{\langle\mathbb{B}_t R, T\rangle\}\\ &= \sum_{\overline{m}\in\mathbb{Z}^s} p(\overline{m}) \sum_{m\in\mathbb{Z}^s} p(m)\,\mathrm{e}\{\langle\mathbb{B}_t m, \overline{m}\rangle\},\end{aligned} \tag{6.18}$$

with summing over $m = (m_1, \dots, m_s) \in \mathbb{Z}^s$, $\overline{m} = (\overline{m}_1, \dots, \overline{m}_s) \in \mathbb{Z}^s$ and

$$p(m) = \mathbf{P}\{R = m\} = \prod_{j=1}^{s} \mathbf{P}\{R_j = m_j\} = \prod_{j=1}^{s} 2^{-2n}\binom{2n}{m_j + n}, \tag{6.19}$$

if $\max_{1\le j\le s} |m_j| \le n$ and $p(m) = 0$ otherwise. Clearly, for fixed $T = \overline{m}$,

$$\mathbf{E}_R\,\mathrm{e}\{\langle\mathbb{B}_t R, T\rangle\} = \sum_{m\in\mathbb{Z}^s} p(m)\,\mathrm{e}\{\langle\mathbb{B}_t m, \overline{m}\rangle\} \ge 0 \tag{6.20}$$

is a value of the characteristic function of symmetrized random vector $\mathbb{B}_t R$. Using Stirling's formula, it is easy to show that there exist positive absolute constants c_2 and c_3 such that

$$\mathbf{P}\{R_j = m_j\} \ll n^{-1/2}\exp\{-m_j^2/2n\} \qquad \text{for } |m_j| \le c_2 n \tag{6.21}$$

and

$$\mathbf{P}\{|R_j| \ge c_2 n\} \ll \exp\{-c_3 n\}. \tag{6.22}$$

Using (6.18)–(6.22), we obtain

$$\begin{aligned}\mathbf{E}\,\mathrm{e}\{\langle\mathbb{B}_t R, T\rangle\} &\ll_s \sum_{\overline{m}\in\mathbb{Z}^s} q(\overline{m}) \sum_{m\in\mathbb{Z}^s} p(m)\,\mathrm{e}\{\langle\mathbb{B}_t m, \overline{m}\rangle\} + \exp\{-c_3 n\}\\ &= \sum_{m\in\mathbb{Z}^s} p(m) \sum_{\overline{m}\in\mathbb{Z}^s} q(\overline{m})\,\mathrm{e}\{\langle\mathbb{B}_t m, \overline{m}\rangle\} + \exp\{-c_3 n\}\\ &= \mathbf{E}\mathbf{E}_{\zeta_n}\,\mathrm{e}\{\langle\mathbb{B}_t R, \zeta_n\rangle\} + \exp\{-c_3 n\}\\ &= \mathbf{E}\,\mathrm{e}\{\langle\mathbb{B}_t R, \zeta_n\rangle\} + \exp\{-c_3 n\}.\end{aligned} \tag{6.23}$$

Now we repeat our previous arguments, noting that

$$\mathbf{E}_{\zeta_n}\,\mathrm{e}\{\langle\mathbb{B}_t R, \zeta_n\rangle\} = \sum_{\overline{m}\in\mathbb{Z}^s} q(\overline{m})\,\mathrm{e}\{\langle\mathbb{B}_t R, \overline{m}\rangle\} \ge 0 \tag{6.24}$$

is a value of the nonnegative characteristic function of the random vector ζ_n; see (6.15). Using again (6.21) and (6.22), we obtain

$$\mathbf{E}\,\mathrm{e}\{\langle \mathbb{B}_t R, \zeta_n\rangle\} \ll_s \mathbf{E}\,\mathrm{e}\{\langle \mathbb{B}_t \zeta_n, \zeta_n'\rangle\} + \exp\{-c_3 n\}. \tag{6.25}$$

Relations (6.17), (6.23) and (6.25) imply the statement of the lemma. $\square$

Let us estimate the expectation $\mathbf{E}\,\mathrm{e}\{\langle \mathbb{B}_t \zeta_n, \zeta_n'\rangle\}$ under the conditions of Lemmas 6.3 and 6.5, assuming that $s = d$, $\mathbb{D} = \mathbb{C}^{-1/2}$, $\delta \le 1/(5s)$, $n \ge c_4$, where c_4 is a sufficiently large absolute constant, and $(z_1, \dots, z_s), (z_1', \dots, z_s') \in \mathbf{\Gamma}(\delta; \mathbb{D}, \mathcal{S})$, that is,

$$\|\mathbb{C}^{-1/2} z_j - e_j\| \le \delta, \qquad \|\mathbb{C}^{-1/2} z_j' - e_j\| \le \delta \qquad \text{for } 1 \le j \le s, \tag{6.26}$$

with an orthonormal system $\mathcal{S} = \mathcal{S}_o = \{e_1, \dots, e_s\}$ involved in the conditions of Lemma 6.3. We can rewrite $\mathbf{E}\,\mathrm{e}\{\langle \mathbb{B}_t \zeta_n, \zeta_n'\rangle\}$ as

$$\mathbf{E}\,\mathrm{e}\{\langle \mathbb{B}_t \zeta_n, \zeta_n'\rangle\} = \sum_{\overline{m} \in \mathbb{Z}^s} q(\overline{m}) \sum_{m \in \mathbb{Z}^s} q(m)\, \mathrm{e}\{\langle \mathbb{B}_t \overline{m}, m\rangle\}.$$

Thus, by (6.16),

$$\mathbf{E}\,\mathrm{e}\{\langle \mathbb{B}_t \zeta_n, \zeta_n'\rangle\} = A_n^{2s} n^{-s} \sum_{\overline{m} \in \mathbb{Z}^s} \sum_{m \in \mathbb{Z}^s} \exp\{i \langle \mathbb{B}_t \overline{m}, m\rangle - \|m\|^2/2n - \|\overline{m}\|^2/2n\}.$$

Denote

$$r = \sqrt{2\pi^2 n}. \tag{6.27}$$

Applying Lemma 6.4 with $\mathbb{S} = \mathbb{I}_s$, $z = 1/2n$, $a = 0$, $b = (2\pi)^{-1}\mathbb{B}_t \overline{m}$ and using that $A_n \asymp 1$, we obtain

$$\begin{aligned} &\mathbf{E}\,\mathrm{e}\{\langle \mathbb{B}_t \zeta_n, \zeta_n'\rangle\} \\ &\quad \ll_s n^{-s/2} \sum_{l, m \in \mathbb{Z}^s} \exp\{-2\pi^2 n \|l + (2\pi)^{-1}\mathbb{B}_t m\|^2 - \|m\|^2/2n\} \\ &\quad \ll_s r^{-s} \sum_{m, \overline{m} \in \mathbb{Z}^s} \exp\{-r^2 \|m - t\mathbb{V}\overline{m}\|^2 - \|\overline{m}\|^2/r^2\}, \end{aligned} \tag{6.28}$$

where $\mathbb{V} : \mathbb{R}^s \to \mathbb{R}^s$ is the operator with matrix

$$\mathbb{V} = (2\pi)^{-1}\mathbb{B}_1. \tag{6.29}$$

Note that the right-hand side of (6.28) may be considered as a theta-series.

Denote $y_k = \mathbb{C}^{-1/2} z_k$, $1 \le k \le s$. Let $\mathbb{Y}$ be the $(s \times s)$-matrix with entries $\langle e_j, y_k\rangle$, where index j is the number of the row, while k is the number of the column. Then the matrix $\mathbb{F} \overset{\text{def}}{=} \mathbb{Y}^* \mathbb{Y}$ has entries $\langle y_j, y_k\rangle$. Here $\mathbb{Y}^*$ is the transposed matrix for $\mathbb{Y}$. According to (6.26), we have

$$\|y_j - e_j\| \le \delta \qquad \text{for } 1 \le j \le s. \tag{6.30}$$

Let us show that [cf. BG (1997a), proof of Lemma 7.4]

$$\|\mathbb{Y}\| \le 3/2 \quad \text{and} \quad \|\mathbb{Y}^{-1}\| \le 2. \tag{6.31}$$

Since $\mathcal{S}_o = \{e_1, e_2, \dots, e_s\}$ is an orthonormal system, inequalities (6.30) imply that $\mathbb{Y} = \mathbb{I}_s + \mathbb{A}$ with some matrix $\mathbb{A} = \{a_{ij}\}$ such that $|a_{ij}| \le \delta$. Thus, we have $\|\mathbb{A}\| \le \|\mathbb{A}\|_2 \le s\delta$, where $\|\mathbb{A}\|_2$ denotes the Hilbert–Schmidt norm of the matrix $\mathbb{A}$. Therefore, the condition $\delta \le 1/(5s)$ implies $\|\mathbb{A}\| \le 1/2$ and inequalities (6.31).

The matrix $\mathbb{F}$ is symmetric and positive definite. Its determinant is the product of eigenvalues which [by (6.31)] are bounded from above and from below by some absolute positive constants. Moreover,

$$(\det \mathbb{Y})^2 = (\det \mathbb{Y}^*)^2 = \det \mathbb{F} \asymp_s 1 \asymp \|\mathbb{F}\| \asymp \|\mathbb{Y}\|. \tag{6.32}$$

Define the matrices $\overline{\mathbb{Y}}$ and $\overline{\mathbb{F}}$, replacing z_j by z_j' in the definition of $\mathbb{Y}$ and $\mathbb{F}$. Similarly to (6.32), one can show that

$$(\det \overline{\mathbb{Y}})^2 = (\det \overline{\mathbb{Y}}^*)^2 = \det \overline{\mathbb{F}} \asymp_s 1 \asymp \|\overline{\mathbb{F}}\| \asymp \|\overline{\mathbb{Y}}\|. \tag{6.33}$$

Let $\mathbb{G}$ and $\overline{\mathbb{G}}$ be the $(s \times s)$-matrices with entries $\langle e_j, \mathbb{Q} z_k\rangle$ and $\langle e_j, z_k'\rangle$, respectively. Then, clearly, $\mathbb{G} = \mathbb{Q}\mathbb{C}^{1/2}\mathbb{Y}$ and $\overline{\mathbb{G}} = \mathbb{C}^{1/2}\overline{\mathbb{Y}}$. Therefore,

$$\mathbb{B}_1 = \mathbb{G}^*\overline{\mathbb{G}} = \mathbb{Y}^*\mathbb{C}^{1/2}\mathbb{Q}\mathbb{C}^{1/2}\overline{\mathbb{Y}}. \tag{6.34}$$

Moreover, $\mathbb{Q}^2 = \mathbb{I}_d$ implies that $|\det \mathbb{Q}| = 1$ and $\|\mathbb{Q}\| = 1$. Using relations (6.29) and (6.32)–(6.34), we obtain

$$|\det \mathbb{V}| \asymp_s |\det \mathbb{B}_1| \asymp_s \det \mathbb{C} \tag{6.35}$$

and

$$\|\mathbb{V}\| \ll \|\mathbb{B}_1\| \ll \|\mathbb{C}\| \ll \sigma_1^2. \tag{6.36}$$

7. Some facts from number theory. In Section 7, we consider some facts of the geometry of numbers; see Davenport (1958) or Cassels (1959). They will help us to estimate the integrals of the right-hand side of inequality (6.28). See Götze and Margulis (2010) or Götze and Zaitsev (2010) for a more detailed version of this section.

Let $e_1, e_2, \dots, e_d$ be linearly independent vectors in $\mathbb{R}^d$. The set

$$\Lambda = \left\{ \sum_{j=1}^{d} n_j e_j : n_j \in \mathbb{Z},\ j = 1, 2, \dots, d \right\} \tag{7.1}$$

is called the lattice with basis $e_1, e_2, \dots, e_d$. The determinant $\det(\Lambda)$ of a lattice Λ is the modulus of the determinant of the matrix formed from the vectors $e_1, e_2, \dots, e_d$. If $\Lambda = \mathbb{A}\mathbb{Z}^d$, where $\mathbb{A}$ is a nondegenerate linear operator, then $\det(\Lambda) = |\det \mathbb{A}|$.

Let $F : \mathbb{R}^d \to [0, \infty)$ denote a norm on $\mathbb{R}^d$. The successive minima $M_1 \le \cdots \le M_d$ of F with respect to a lattice $\Lambda \subset \mathbb{R}^d$ are defined as follows: M_j is the infimum of $\lambda > 0$ such that the set $\{m \in \Lambda : F(m) < \lambda\}$ contains j linearly independent vectors. The following Lemma 7.1 is proved by Davenport [(1958), Lemma 1] for $\Lambda = \mathbb{Z}^d$; see also Götze and Margulis (2010).

LEMMA 7.1. *Let $M_1 \le \cdots \le M_d$ be the successive minima of a norm F with respect to a lattice $\Lambda \subset \mathbb{R}^d$. Denote $M_{d+1} = \infty$. Suppose that $1 \le j \le d$ and $M_j \le b \le M_{j+1}$, for some $b > 0$. Then*

$$\#\{m = (m_1, \dots, m_d) \in \mathbb{Z}^d : F(m) < b\} \asymp_d b^j (M_1 \cdot M_2 \cdots M_j)^{-1}. \tag{7.2}$$

Representing $\Lambda = \mathbb{A}\mathbb{Z}^d$, we see that the lattice $\Lambda = \mathbb{Z}^d$ may be replaced in Lemma 7.1 by any lattice $\Lambda \subset \mathbb{R}^d$.

LEMMA 7.2. *Let $F_j(m)$, $j = 1, 2$, be some norms in $\mathbb{R}^d$ and $M_1 \le \cdots \le M_d$ and $N_1 \le \cdots \le N_d$ be the successive minima of F_1 with respect to a lattice Λ_1 and of F_2 with respect to a lattice Λ_2, respectively. Let $C > 0$. Assume that $M_k \gg_d C F_2(n_k)$, $k = 1, 2, \dots, d$, for some linearly independent vectors $n_1, n_2, \dots, n_d \in \Lambda_2$. Then*

$$M_k \gg_d C N_k, \qquad k = 1, \dots, d. \tag{7.3}$$

LEMMA 7.3. *Let Λ be a lattice in $\mathbb{R}^d$ and let $c_j(d)$, $j = 1, 2, 3$, be positive quantities depending on d only. Let $F(\cdot)$ be a norm in $\mathbb{R}^d$ such that $F(\cdot) \asymp_d \|\cdot\|$. Then*

$$\begin{aligned} \sum_{v \in \Lambda} \exp\{-c_1(d)\|v\|^2\} &\asymp_d \sum_{v \in \Lambda} \exp\{-c_2(d)(F(v))^2\} \\ &\asymp_d \#\{v \in \Lambda : F(v) < c_3(d)\}. \end{aligned} \tag{7.4}$$

For a lattice $\Lambda \subset \mathbb{R}^d$ and $1 \le l \le d$, we define its α_l-characteristics by

$$\alpha_l(\Lambda) \stackrel{\text{def}}{=} \sup\{|\det(\Lambda')|^{-1} : \Lambda' \text{ is a } l\text{-dimensional sublattice of } \Lambda\}. \tag{7.5}$$

Denote

$$\alpha(\Lambda) \stackrel{\text{def}}{=} \max_{1 \le l \le d} \alpha_l(\Lambda). \tag{7.6}$$

LEMMA 7.4. *Let $F(\cdot)$ be a norm in $\mathbb{R}^d$ such that $F(\cdot) \asymp_d \|\cdot\|$. Let $c(d)$ be a positive quantity depending on d only. Let $M_1 \le \cdots \le M_d$ be the successive minima of F with respect to a lattice $\Lambda \subset \mathbb{R}^d$. Then*

$$\alpha_l(\Lambda) \asymp_d (M_1 \cdot M_2 \cdots M_l)^{-1}, \qquad l = 1, \dots, d. \tag{7.7}$$

Moreover,

$$\alpha(\Lambda) \asymp_d \#\{v \in \Lambda : \|v\| < c(d)\}, \tag{7.8}$$

provided that $M_1 \ll_d 1$.

Lemma 7.4 is an easy consequence of the following lemma formulated in proposition (page 517) and remark (page 518) in Lenstra, Lenstra and Lovász (1982).

LEMMA 7.5. *Let* $M_1 \le \cdots \le M_d$ *be the successive minima of the standard Euclidean norm with respect to a lattice* $\Lambda \subset \mathbb{R}^d$. *Then there exists a basis* $e_1, e_2, \dots, e_d$ *of* Λ *such that*

$$M_l \asymp_d \|e_l\|, \qquad l = 1, \dots, d. \tag{7.9}$$

Moreover,

$$\det(\Lambda) \asymp_d \prod_{l=1}^{d} \|e_l\|. \tag{7.10}$$

8. From number theory to probability. In Section 8, we use number-theoretical results of Section 7 to estimate integrals of the right-hand side of (6.28). Recall that we have assumed the conditions of Lemmas 6.3 and 6.5, $s = d$, $\mathbb{D} = \mathbb{C}^{-1/2}$, $\delta \le 1/(5s)$, $n \ge c_4$ and (6.15), for an orthonormal system $\mathcal{S} = \mathcal{S}_o$. The notation $\mathrm{SL}(d, \mathbb{R})$ is used below for the set of all $(d \times d)$-matrices with real entries and determinant 1.

Introduce the matrices

$$\mathbb{D}_r \stackrel{\text{def}}{=} \begin{pmatrix} r\mathbb{I}_s & \mathbb{O}_s \\ \mathbb{O}_s & r^{-1}\mathbb{I}_s \end{pmatrix} \in \mathrm{SL}(2s, \mathbb{R}), \qquad r > 0, \tag{8.1}$$

$$\mathbb{K}_t \stackrel{\text{def}}{=} \begin{pmatrix} \mathbb{I}_s & -t\mathbb{I}_s \\ t\mathbb{I}_s & \mathbb{I}_s \end{pmatrix}, \qquad t \in \mathbb{R}, \tag{8.2}$$

$$\mathbb{U}_t \stackrel{\text{def}}{=} \begin{pmatrix} \mathbb{I}_s & -t\mathbb{I}_s \\ \mathbb{O}_s & \mathbb{I}_s \end{pmatrix} \in \mathrm{SL}(2s, \mathbb{R}), \qquad t \in \mathbb{R}, \tag{8.3}$$

and the lattices

$$\Lambda \stackrel{\text{def}}{=} \begin{pmatrix} \mathbb{I}_s & \mathbb{O}_s \\ \mathbb{O}_s & \mathbb{V}_0 \end{pmatrix} \mathbb{Z}^{2s}, \tag{8.4}$$

$$\Lambda_j = \mathbb{D}_j \mathbb{U}_{j^{-1}} \Lambda = \begin{pmatrix} j\mathbb{I}_s & -\mathbb{V}_0 \\ \mathbb{O}_s & j^{-1}\mathbb{V}_0 \end{pmatrix} \mathbb{Z}^{2s}, \qquad j = 1, 2, \dots, \tag{8.5}$$

where

$$\mathbb{V}_0 = \sigma_1^{-2} \mathbb{V} \tag{8.6}$$

and the matrix $\mathbb{V}$ is defined in (6.29). Below we use the following simplest properties of these matrices:

$$\mathbb{D}_a\mathbb{D}_b = \mathbb{D}_{ab}, \qquad \mathbb{U}_a\mathbb{U}_b = \mathbb{U}_{a+b} \quad \text{and} \quad \mathbb{D}_a\mathbb{U}_b = \mathbb{U}_{a^2b}\mathbb{D}_a \quad \text{for } a, b > 0. \tag{8.7}$$

Let $\|x\|_\infty = \max_{1\le j\le d}|x_j|$, for $x = (x_1, \dots, x_d) \in \mathbb{R}^d$. Let $M_{j,t}$, $j = 1, 2, \dots, 2s$, be the successive minima of the norm $\|\cdot\|_\infty$ with respect to the lattice

$$\Xi_t \overset{\text{def}}{=} \begin{pmatrix} r\mathbb{I}_s & -rt\mathbb{V} \\ \mathbb{O}_s & r^{-1}\mathbb{I}_s \end{pmatrix} \mathbb{Z}^{2s}. \tag{8.8}$$

Moreover, simultaneously, $M_{j,t}$ are the successive minima of the norm $F^*(\cdot)$ defined for $(m, \overline{m}) \in \mathbb{R}^{2s}$, $m, \overline{m} \in \mathbb{R}^s$, by

$$F^*\bigl((m, \overline{m})\bigr) \overset{\text{def}}{=} \max\bigl\{\|m\|_\infty, \sigma_1^2\|\mathbb{V}^{-1}\overline{m}\|_\infty\bigr\} \tag{8.9}$$

with respect to the lattice

$$\Omega_t \overset{\text{def}}{=} \begin{pmatrix} r\mathbb{I}_s & -rt\mathbb{V} \\ \mathbb{O}_s & \sigma_1^{-2}r^{-1}\mathbb{V} \end{pmatrix} \mathbb{Z}^{2s} = \mathbb{D}_r\mathbb{U}_u\Lambda \qquad \text{where } u \overset{\text{def}}{=} \sigma_1^2 t. \tag{8.10}$$

Using Lemmas 7.2 and 7.5 and the equality $\det(\Xi_t) = 1$, it is easy to show that

$$M_{1,t} \ll_s 1. \tag{8.11}$$

Let $M^*_{j,t}$ be the successive minima of the Euclidean norm with respect to the lattice Ω_t. Note that, according to (6.36) and (8.9),

$$\|\cdot\| \ll_s F^*(\cdot). \tag{8.12}$$

Using (8.12) and Lemma 7.2, we obtain

$$M^*_{j,t} \ll_s M_{j,t}, \qquad j = 1, \dots, 2s. \tag{8.13}$$

According to Lemma 7.4,

$$\alpha(\Xi_t) \ll_s \alpha(\Omega_t). \tag{8.14}$$

Let us estimate $\alpha(\Omega_t)$ assuming that $r \ge 1$ and (for a moment) $t = \sigma_1^{-2}r^{-1}$. In this case

$$\Omega_t = \begin{pmatrix} r\mathbb{I}_s & -\mathbb{V}_0 \\ \mathbb{O}_s & r^{-1}\mathbb{V}_0 \end{pmatrix} \mathbb{Z}^{2s}. \tag{8.15}$$

By relation (7.8) of Lemma 7.4, we have

$$\alpha(\Omega_t) \asymp_s \#\bigl\{v \in \Omega_t : \|v\| < 1/2\bigr\} = \#K, \tag{8.16}$$

where

$$\begin{aligned} K = \bigl\{v = (m, \overline{m}) \in \mathbb{Z}^{2s} &: m, \overline{m} \in \mathbb{Z}^s, \\ &\|rm - \mathbb{V}_0\overline{m}\|^2 + \|r^{-1}\mathbb{V}_0\overline{m}\|^2 < 1/4\bigr\}. \end{aligned} \tag{8.17}$$

Let us estimate from above the right-hand side of (8.16). If $v = (m, \overline{m}) \in K$, then

$$r\|m\| \le \|rm - \mathbb{V}_0\overline{m}\| + \|\mathbb{V}_0\overline{m}\| < \frac{1}{2} + \frac{r}{2} \le r. \tag{8.18}$$

Hence $m = 0$ and $\|\mathbb{V}_0\overline{m}\| \le 1/2$. It remains to estimate the quantity

$$R \stackrel{\text{def}}{=} \#\{\overline{m} \in \mathbb{Z}^s : \|\mathbb{V}_0\overline{m}\| < 1\} \ge \#K. \tag{8.19}$$

Let $N_1 \le \cdots \le N_s$ be the successive minima of the Euclidean norm with respect to the lattice $\mathbb{V}_0\mathbb{Z}^s$. Let $e_1, e_2, \dots, e_s$ be the standard orthonormal basis of $\mathbb{Z}^s$. By (6.36) and (8.6), we have $\|\mathbb{V}_0 e_j\| \le 1$, $j = 1, 2, \dots, s$. Therefore, using Lemma 7.2, we see that $N_1 \le \cdots \le N_s \le 1$. By (6.35), (8.6), (8.19) and by Lemmas 7.1, 7.2 and 7.5,

$$R \asymp_s (N_1 \cdot N_2 \cdots N_s)^{-1} \asymp_s (\det \mathbb{V}_0)^{-1} \asymp_s \sigma_1^{2s} (\det \mathbb{C})^{-1}. \tag{8.20}$$

Hence, using (8.16), (8.19) and (8.20), we conclude that

$$\alpha(\Omega_t) \ll_s \sigma_1^{2s} (\det \mathbb{C})^{-1} \qquad \text{for } r \ge 1 \text{ and } t = \sigma_1^{-2} r^{-1}. \tag{8.21}$$

Let now $t \in \mathbb{R}$ be arbitrary. By (8.8), (8.11), (8.14) and by Lemmas 7.1, 7.3 and 7.4,

$$\begin{aligned}
&\sum_{m, \overline{m} \in \mathbb{Z}^s} \exp\{-r^2 \|m - t\mathbb{V}\overline{m}\|^2 - \|\overline{m}\|^2 / r^2\} = \sum_{v \in \Xi_t} \exp\{-\|v\|^2\} \\
&\quad \ll_s R_t \stackrel{\text{def}}{=} \#\{v \in \Xi_t : \|v\| < 1\} \\
&\quad \ll_s \alpha(\Xi_t) \ll_s \alpha(\Omega_t).
\end{aligned} \tag{8.22}$$

Now, by (6.28), (8.10) and (8.22), we have

$$\mathbf{E}\, \mathrm{e}\{\langle \mathbb{B}_t \zeta_n, \zeta_n' \rangle\} \ll_s r^{-s} \alpha(\Omega_t) = r^{-s} \alpha(\mathbb{D}_r \mathbb{U}_u \Lambda) \qquad \text{where } u = \sigma_1^2 t. \tag{8.23}$$

Let us estimate the quantity R_t, $t \in \mathbb{R}$, defined in (8.22) assuming that $r \ge 1$ and $|rt| \le c_s^* \sigma_1^{-2}$, where $c_s^* \ge 1$ is an arbitrary quantity depending on s only. By Lemma 7.3, we have

$$R_t \asymp_s \#K_0, \tag{8.24}$$

where

$$\begin{aligned}
K_0 \stackrel{\text{def}}{=} \{&v = (m, \overline{m}) \in \mathbb{Z}^{2s} : m, \overline{m} \in \mathbb{Z}^s, \\
&\|rm - rt\mathbb{V}\overline{m}\|^2 + \|r^{-1}\overline{m}\|^2 < (2c_s^*)^{-2}\}.
\end{aligned} \tag{8.25}$$

If $v = (m, \overline{m}) \in K_0$, $r \ge 1$ and $|rt| \le c_s^* \sigma_1^{-2}$, then, by (6.36) and (8.25),

$$r\|m\| \le \|rm - rt\mathbb{V}\overline{m}\| + |rt| \|\mathbb{V}\overline{m}\| < \frac{1}{2} + \frac{r}{2} \le r. \tag{8.26}$$

Hence $m = 0$ and $|rt|\|\mathbb{V}\overline{m}\| \leq (2c_s^*)^{-1} < 1$. It remains to estimate the quantity

$$S \stackrel{\text{def}}{=} \#\{\overline{m} \in \mathbb{Z}^s : |rt|\|\mathbb{V}\overline{m}\| < 1\} \geq \#K_0. \tag{8.27}$$

Let $P_1 \leq \cdots \leq P_s$ be the successive minima of the Euclidean norm with respect to the lattice $|rt|\mathbb{V}\mathbb{Z}^s$. Let $e_1, e_2, \ldots, e_s$ be the standard orthonormal basis of $\mathbb{Z}^s$. By (6.36), we have $\||rt|\mathbb{V}e_j\| \ll_s 1$, $j = 1, 2, \ldots, s$. Therefore, using Lemma 7.2, we see that $P_1 \leq \cdots \leq P_s \ll_s 1$. By (6.35), (8.27) and Lemmas 7.1 and 7.5,

$$S \asymp_s (P_1 \cdot P_2 \cdots P_s)^{-1} \asymp_s \big(\det(|rt|\mathbb{V})\big)^{-1} \asymp_s |rt|^{-s}(\det\mathbb{C})^{-1}. \tag{8.28}$$

Hence, using (8.24), (8.27) and (8.28), we conclude that

$$R_t \ll_s |rt|^{-s}(\det\mathbb{C})^{-1} \qquad \text{for } r \geq 1 \text{ and } |rt| \leq c_s^*\sigma_1^{-2}. \tag{8.29}$$

Now, by (6.28), (8.22) and (8.29), we have

$$\begin{aligned} \mathbf{E}\,\mathrm{e}\{\langle \mathbb{B}_t\zeta_n, \zeta_n'\rangle\} &\ll_s r^{-s}R_t \\ &\ll_s r^{-2s}|t|^{-s}(\det\mathbb{C})^{-1} \qquad \text{for } r \geq 1 \text{ and } |rt| \leq c_s^*\sigma_1^{-2}. \end{aligned} \tag{8.30}$$

It is easy to verify that

$$\int_{c_s\sigma_1^{-2}r^{-2+4/s}}^{\sigma_1^{-2}r^{-1}} \sqrt{r^{-2s}|t|^{-s}(\det\mathbb{C})^{-1}}\,\frac{dt}{t} \ll_s r^{-2}\sigma_1^s(\det\mathbb{C})^{-1/2} \tag{8.31}$$

for any c_s depending on s only. Note that $\sigma_1^s(\det\mathbb{C})^{-1/2} \geq 1$. Using (8.23), (8.31) and Lemmas 6.3 and 6.5, we derive the following lemma.

LEMMA 8.1. *Let the conditions of Lemma 6.3 be satisfied with $s = d$, $\mathbb{D} = \mathbb{C}^{-1/2}$, $\delta \leq 1/(5s)$ and with an orthonormal system $\mathcal{S} = \mathcal{S}_o = \{e_1, \ldots, e_s\} \subset \mathbb{R}^d$. Let c_s be an arbitrary quantity depending on s only. Then, for any $b \in \mathbb{R}^d$ and $r \geq 1$,*

$$\begin{aligned} &\int_{c_s\sigma_1^{-2}r^{-2+4/s}}^{\sigma_1^{-2}} \big|\widehat{\Psi}_b(t/2)\big|\frac{dt}{t} \\ &\qquad \ll_s (pN)^{-1}\sigma_1^s(\det\mathbb{C})^{-1/2} + r^{-s/2}\sup_{\Gamma}\int_{r^{-1}}^1 \big(\alpha(\mathbb{D}_r\mathbb{U}_u\Lambda)\big)^{1/2}\frac{du}{u}, \end{aligned} \tag{8.32}$$

where $r, \alpha(\cdot), \mathbb{D}_r$ $\mathbb{U}_t$ and the lattice Λ are defined in relations (6.4), (6.27), (6.29), (7.5), (7.6), (8.1), (8.3) and (8.4) and in Lemma 6.5. The $\sup_\Gamma$ means here the supremum over all possible values of $z_j, z_j' \in \mathbb{R}^d$ (involved in the definition of matrices $\mathbb{B}_t$ and $\mathbb{V}$) such that

$$\|\mathbb{C}^{-1/2}z_j - e_j\| \leq \delta, \qquad \|\mathbb{C}^{-1/2}z_j' - e_j\| \leq \delta \qquad \textit{for } 1 \leq j \leq s. \tag{8.33}$$

Moreover, for any $b \in \mathbb{R}^d$, $r \geq 1$ and $\gamma > 0$ and any fixed $t \in \mathbb{R}$ satisfying $|rt| \leq c_s^*\sigma_1^{-2}$, where $c_s^* \geq 1$ is an arbitrary quantity depending on s only, we have

$$\big|\widehat{\Psi}_b(t)\big| \ll_{\gamma,s} (pN)^{-\gamma} + r^{-s}|t|^{-s/2}(\det\mathbb{C})^{-1/2}. \tag{8.34}$$

Let $v=(m,\overline{m})\in\mathbb{R}^{2s}$, $m,\overline{m}\in\mathbb{R}^s$ and $t\in\mathbb{R}$. Then

$$\overline{m}+tm=\big(1+t^2\big)\overline{m}+t(m-t\overline{m}). \tag{8.35}$$

Equality (8.35) implies that

$$\|\overline{m}+tm\|\ll_s\|\overline{m}\|+\|m-t\overline{m}\|\qquad\text{for } |t|\ll_s 1. \tag{8.36}$$

Hence,

$$r\|m-t\overline{m}\|+r^{-1}\|\overline{m}+tm\|\ll_s r\|m-t\overline{m}\|+r^{-1}\|\overline{m}\| \qquad \text{for } r\gg 1,\ |t|\ll_s 1. \tag{8.37}$$

According to (8.1)–(8.3), we have

$$\begin{aligned}\mathbb{D}_r\mathbb{U}_t v&=\big(r(m-t\overline{m}),r^{-1}\overline{m}\big)\quad\text{and}\\ \mathbb{D}_r\mathbb{K}_t v&=\big(r(m-t\overline{m}),r^{-1}(\overline{m}+tm)\big).\end{aligned} \tag{8.38}$$

It is clear that the operators $\mathbb{D}_r\mathbb{U}_t$ and $\mathbb{D}_r\mathbb{K}_t$ are invertible. Therefore, using (8.37) and (8.38) and applying Lemmas 7.2 and 7.4, we derive the inequality

$$\alpha(\mathbb{D}_r\mathbb{U}_t\Omega)\ll_s\alpha(\mathbb{D}_r\mathbb{K}_t\Omega)\qquad\text{for } r\gg 1,\ |t|\ll_s 1, \tag{8.39}$$

which is valid for any lattice $\Omega\subset\mathbb{R}^{2s}$.

Let $\mathbb{T}$ be the permutation $(2s\times 2s)$-matrix which permutes the rows of a $(2s\times 2s)$-matrix $\mathbb{A}$ so that the new order (corresponding to the matrix $\mathbb{TA}$) is

$$1,s+1,2,s+2,\dots,s,2s.$$

Note that the operator $\mathbb{T}$ is isometric and $\mathbb{A}\mapsto\mathbb{A}\mathbb{T}^{-1}$ rearranges the columns of $\mathbb{A}$ in the order mentioned above. It is easy to see that

$$\alpha_j(\mathbb{T}\Omega)=\alpha_j(\Omega),\qquad j=1,\dots 2s \text{ and } \alpha(\mathbb{T}\Omega)=\alpha(\Omega) \tag{8.40}$$

for any lattice $\Omega\subset\mathbb{R}^{2s}$.

Note now that

$$\mathbb{T}\mathbb{D}_r\mathbb{K}_t\Lambda_j=\mathbb{T}\mathbb{D}_r\mathbb{K}_t\mathbb{T}^{-1}\mathbb{T}\Lambda_j=\mathbb{W}_t\Delta_j, \tag{8.41}$$

where Δ_j is a lattice defined by

$$\Delta_j=\mathbb{T}\Lambda_j \tag{8.42}$$

and where $\mathbb{W}_t$ is $(2s\times 2s)$-matrix

$$\mathbb{W}_t=\begin{pmatrix}\mathbb{G}_{r,t} & \mathbb{O}_2 & \vdots & \mathbb{O}_2\\ \mathbb{O}_2 & \mathbb{G}_{r,t} & \vdots & \mathbb{O}_2\\ \cdots & \cdots & \cdots & \cdots\\ \mathbb{O}_2 & \mathbb{O}_2 & \vdots & \mathbb{G}_{r,t}\end{pmatrix} \tag{8.43}$$

constructed of (2×2)-matrices $\mathbb{O}_2$ (with zero entries) and

$$\mathbb{G}_{r,t} \stackrel{\text{def}}{=} \begin{pmatrix} r & -rt \\ r^{-1}t & r^{-1} \end{pmatrix}. \tag{8.44}$$

Let $|t| \le 2$ and

$$\theta = \arcsin\bigl(t(1+t^2)^{-1/2}\bigr) \qquad \text{or, equivalently } t = \tan\theta. \tag{8.45}$$

Then we have

$$\begin{aligned} |\theta| \le c^* \stackrel{\text{def}}{=} \arcsin(2/\sqrt{5}), \qquad \cos\theta = \bigl(1+t^2\bigr)^{-1/2}, \\ \sin\theta = t\bigl(1+t^2\bigr)^{-1/2}. \end{aligned} \tag{8.46}$$

It is easy to see that

$$\mathbb{G}_{r,t} = \bigl(1+t^2\bigr)^{1/2}\overline{\mathbb{D}}_r\overline{\mathbb{K}}_\theta \tag{8.47}$$

and

$$\mathbb{W}_t = \bigl(1+t^2\bigr)^{1/2}\widetilde{\mathbb{D}}_r\widetilde{\mathbb{K}}_\theta, \tag{8.48}$$

where

$$\widetilde{\mathbb{D}}_r = \begin{pmatrix} \overline{\mathbb{D}}_r & \mathbb{O}_2 & \vdots & \mathbb{O}_2 \\ \mathbb{O}_2 & \overline{\mathbb{D}}_r & \vdots & \mathbb{O}_2 \\ \cdots & \cdots & \cdots & \cdots \\ \mathbb{O}_2 & \mathbb{O}_2 & \vdots & \overline{\mathbb{D}}_r \end{pmatrix} \quad \text{and} \quad \widetilde{\mathbb{K}}_\theta = \begin{pmatrix} \overline{\mathbb{K}}_\theta & \mathbb{O}_2 & \vdots & \mathbb{O}_2 \\ \mathbb{O}_2 & \overline{\mathbb{K}}_\theta & \vdots & \mathbb{O}_2 \\ \cdots & \cdots & \cdots & \cdots \\ \mathbb{O}_2 & \mathbb{O}_2 & \vdots & \overline{\mathbb{K}}_\theta \end{pmatrix} \tag{8.49}$$

are $(2s \times 2s)$-matrices with

$$\overline{\mathbb{D}}_r \stackrel{\text{def}}{=} \begin{pmatrix} r & 0 \\ 0 & r^{-1} \end{pmatrix} \quad \text{and} \quad \overline{\mathbb{K}}_\theta \stackrel{\text{def}}{=} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \in \mathrm{SL}(2,\mathbb{R}). \tag{8.50}$$

Substituting (8.48) into equality (8.41), we obtain

$$\mathbb{T}\mathbb{D}_r\mathbb{K}_t\Lambda_j = \bigl(1+t^2\bigr)^{1/2}\widetilde{\mathbb{D}}_r\widetilde{\mathbb{K}}_\theta\Delta_j. \tag{8.51}$$

Below we also use the following crucial lemma of Götze and Margulis (2010).

LEMMA 8.2. *Let* $\widetilde{\mathbb{K}}_\theta$ *and*

$$\widetilde{\mathbb{H}} = \begin{pmatrix} \overline{\mathbb{H}} & \mathbb{O}_2 & \vdots & \mathbb{O}_2 \\ \mathbb{O}_2 & \overline{\mathbb{H}} & \vdots & \mathbb{O}_2 \\ \cdots & \cdots & \cdots & \cdots \\ \mathbb{O}_2 & \mathbb{O}_2 & \vdots & \overline{\mathbb{H}} \end{pmatrix} \tag{8.52}$$

be $(2d \times 2d)$*-matrices such that* $\overline{\mathbb{H}} \in \mathcal{G} = \mathrm{SL}(2, \mathbb{R})$ *and* $\widetilde{\mathbb{K}}_\theta$ *is defined in* (8.49) *and* (8.50). *Let* β *be a positive number such that* $\beta d > 2$. *Then, for any* $\overline{\mathbb{H}} \in \mathcal{G}$ *and any lattice* $\Delta \subset \mathbb{R}^{2d}$,

$$\int_0^{2\pi} \big(\alpha(\widetilde{\mathbb{H}}\widetilde{\mathbb{K}}_\theta \Delta)\big)^\beta \, d\theta \ll_{\beta,d} \big(\alpha(\Delta)\big)^\beta \|\overline{\mathbb{H}}\|^{\beta d - 2}. \tag{8.53}$$

Here $\|\overline{\mathbb{H}}\|$ *is the standard norm of the linear operator* $\overline{\mathbb{H}} : \mathbb{R}^2 \to \mathbb{R}^2$.

Consider, under the conditions of Lemma 8.1,

$$I_0 \stackrel{\text{def}}{=} \int_{c_s \sigma_1^{-2} r^{-2+4/s}/2}^{\sigma_1^{-2}/2} \big|\widehat{\Psi}_b(t)\big| \frac{dt}{t} = \int_{c_s \sigma_1^{-2} r^{-2+4/s}}^{\sigma_1^{-2}} \big|\widehat{\Psi}_b(t/2)\big| \frac{dt}{t}. \tag{8.54}$$

By Lemma 8.1, we have

$$I_0 \ll_s (pN)^{-1} \sigma_1^s (\det \mathbb{C})^{-1/2} + r^{-s/2} \sup_\Gamma J, \tag{8.55}$$

where

$$J = \int_{r^{-1}}^1 \big(\alpha(\mathbb{D}_r \mathbb{U}_t \Lambda)\big)^{1/2} \frac{dt}{t} \le \sum_{j=2}^{\rho} I_j, \tag{8.56}$$

with

$$I_j \stackrel{\text{def}}{=} \int_{j^{-1}}^{(j-1)^{-1}} \big(\alpha(\mathbb{D}_r \mathbb{U}_t \Lambda)\big)^{1/2} \frac{dt}{t}, \qquad j = 2, 3, \dots, \rho \stackrel{\text{def}}{=} \lfloor r \rfloor + 1. \tag{8.57}$$

Changing variable $t = vj^{-2}$ and $v = w + j$ in I_j and using the properties of matrices $\mathbb{D}_r$ and $\mathbb{U}_t$, we have

$$\begin{aligned} I_j &= \int_j^{j^2(j-1)^{-1}} \big(\alpha(\mathbb{D}_r \mathbb{U}_{vj^{-2}} \Lambda)\big)^{1/2} \frac{dv}{v} \\ &\le \int_j^{j+2} \big(\alpha(\mathbb{D}_r \mathbb{U}_{vj^{-2}} \Lambda)\big)^{1/2} \frac{dv}{v} \\ &= \int_0^2 \big(\alpha(\mathbb{D}_r \mathbb{U}_{wj^{-2}} \mathbb{U}_{j^{-1}} \Lambda)\big)^{1/2} \frac{dw}{w + j}. \end{aligned} \tag{8.58}$$

By (8.7),

$$\mathbb{D}_r \mathbb{U}_{wj^{-2}} = \mathbb{D}_{rj^{-1}} \mathbb{D}_j \mathbb{U}_{wj^{-2}} = \mathbb{D}_{rj^{-1}} \mathbb{U}_w \mathbb{D}_j. \tag{8.59}$$

According to (8.58) and (8.59),

$$I_j \ll \frac{1}{j} \int_0^2 \big(\alpha(\mathbb{D}_{rj^{-1}} \mathbb{U}_t \Lambda_j)\big)^{1/2} dt, \tag{8.60}$$

where the lattices Λ_j are defined in (8.5); see also (8.1), (8.3) and (8.4). Using (8.5), (8.15) and (8.21), we see that

$$\alpha(\Lambda_j) \ll_s \sigma_1^{2s} (\det \mathbb{C})^{-1}. \tag{8.61}$$

By (8.39), (8.40) and (8.51), we have

$$\begin{aligned} \alpha(\mathbb{D}_{rj^{-1}} \mathbb{U}_t \Lambda_j) &\ll_s \alpha(\mathbb{D}_{rj^{-1}} \mathbb{K}_t \Lambda_j) = \alpha(\mathbb{T}\mathbb{D}_{rj^{-1}} \mathbb{K}_t \Lambda_j) \\ &\ll_s \alpha(\widetilde{\mathbb{D}}_{rj^{-1}} \widetilde{\mathbb{K}}_\theta \Delta_j) \end{aligned} \tag{8.62}$$

for $|t| \ll_s 1$, $r \geq 1$, $j = 2, 3, \dots, \rho$, where Δ_j and θ are defined in (8.42) and (8.45), respectively. Using (8.45), (8.46), (8.49), (8.62) and Lemma 8.2 (with $d = s$), we obtain

$$\begin{aligned} \int_0^2 \bigl(\alpha(\mathbb{D}_{rj^{-1}} \mathbb{U}_t \Lambda_j)\bigr)^{1/2} dt &\ll_s \int_0^{c^*} \bigl(\alpha(\widetilde{\mathbb{D}}_{rj^{-1}} \widetilde{\mathbb{K}}_\theta \Delta_j)\bigr)^{1/2} \frac{d\theta}{\cos^2\theta} \\ &\ll \int_0^{2\pi} \bigl(\alpha(\widetilde{\mathbb{D}}_{rj^{-1}} \widetilde{\mathbb{K}}_\theta \Delta_j)\bigr)^{1/2} d\theta \\ &\ll_s \|\overline{\mathbb{D}}_{rj^{-1}}\|^{s/2-2} \bigl(\alpha(\Delta_j)\bigr)^{1/2}, \end{aligned} \tag{8.63}$$

if $s \geq 5$. It is clear that $\|\overline{\mathbb{D}}_{rj^{-1}}\| = rj^{-1}$. Therefore, according to (8.40), (8.42), (8.60) and (8.63),

$$I_j \ll_s \frac{1}{j} (rj^{-1})^{s/2-2} \bigl(\alpha(\Lambda_j)\bigr)^{1/2}. \tag{8.64}$$

By (8.56), (8.61) and (8.64), we obtain, for $s \geq 5$,

$$J \ll_s \sigma_1^s (\det \mathbb{C})^{-1/2} \sum_{j=2}^{\rho} \frac{1}{j} (rj^{-1})^{s/2-2} \ll_s r^{s/2-2} \sigma_1^s (\det \mathbb{C})^{-1/2}. \tag{8.65}$$

By (6.4), (6.27), (8.55) and (8.65), we have $r \asymp_s (Np)^{1/2}$ and

$$I_0 \ll_s r^{-2} \sigma_1^s (\det \mathbb{C})^{-1/2} \ll_s (Np)^{-1} \sigma_1^s (\det \mathbb{C})^{-1/2}. \tag{8.66}$$

It is clear that in a similar way we can establish that

$$\int_{\sigma_1^{-2}}^{c(s)\sigma_1^{-2}} \bigl|\widehat{\Psi}_b(t/2)\bigr| \frac{dt}{t} \ll_s r^{-2} \sigma_1^s (\det \mathbb{C})^{-1/2} \ll_s (Np)^{-1} \sigma_1^s (\det \mathbb{C})^{-1/2} \tag{8.67}$$

for any quantity $c(s)$ depending on s only. The proof will be easier due to the fact that t cannot be small in this integral.

Thus, we have proved the following lemma.

LEMMA 8.3. *Let the conditions of Lemma 6.3 be satisfied with $s = d \geq 5$, $\mathbb{D} = \mathbb{C}^{-1/2}$, $\delta \leq 1/(5s)$ and with an orthonormal system $\mathcal{S} = \mathcal{S}_o = \{e_1, \dots, e_s\} \subset$*

$\mathbb{R}^d$. Let $c_1(s)$ and $c_2(s)$ be some quantities depending on s only. Then there exists a c_s such that

$$\int_{c_1(s)\sigma_1^{-2}r^{-2+4/s}}^{c_2(s)\sigma_1^{-2}} \big|\widehat{\Psi}_b(t)\big| \frac{dt}{t} \ll_s (Np)^{-1}\sigma_1^s (\det \mathbb{C})^{-1/2}, \tag{8.68}$$

if $Np \gg_s c_s$, where r is defined in (6.4) *and* (6.27).

Acknowledgments. We would like to thank V.V. Ulyanov for helpful discussions, and two anonymous referees for useful suggestions which allowed us to improve the presentation.

REFERENCES

BENTKUS, V. Y. (1984). Asymptotic expansions for distributions of sums of independent random elements in a Hilbert space. *Lithuanian Math. J.* **24** 305–319.

BENTKUS, V. and GÖTZE, F. (1995a). On the number of lattice points in a large ellipsoid. *Russian Acad. Sc. Doklady* **343** 439–440.

BENTKUS, V. and GÖTZE, F. (1995b). Optimal rates of convergence in functional limit theorems for quadratic forms. Preprint 95-091 SFB 343, Univ. Bielefeld.

BENTKUS, V. and GÖTZE, F. (1996). Optimal rates of convergence in the CLT for quadratic forms. *Ann. Probab.* **24** 466–490. MR1387646

BENTKUS, V. and GÖTZE, F. (1997a). Uniform rates of convergence in the CLT for quadratic forms in multidimensional spaces. *Probab. Theory Related Fields* **109** 367–416. MR1481126

BENTKUS, V. and GÖTZE, F. (1997b). On the lattice point problem for ellipsoids. *Acta Arith.* **80** 101–125. MR1450919

BENTKUS, V., GÖTZE, F., PAULAUSKAS, V. and RAČKAUSKAS, A. (1991). The accuracy of Gaussian approximation in Banach spaces. In *Probability Theory*, 6 (*Russian*) 39–139. Akad. Nauk SSSR Vsesoyuz. Inst. Nauchn. i Tekhn. Inform., Moscow. MR1157207

BENTKUS, V., GÖTZE, F. and ZAITSEV, A. Y. (1997). Approximation of quadratic forms of independent random vectors by accompanying laws. *Theory Probab. Appl.* **42** 189–212.

BENTKUS, V., GÖTZE, F. and ZITIKIS, R. (1993). Asymptotic expansions in the integral and local limit theorems in Banach spaces with applications to ω-statistics. *J. Theoret. Probab.* **6** 727–780. MR1245394

BHATTACHARYA, R. N. and RANGA RAO, R. (1986). *Normal Approximation and Asymptotic Expansions*. Wiley, New York.

BOGATYREV, S. A., GÖTZE, F. and ULYANOV, V. V. (2006). Non-uniform bounds for short asymptotic expansions in the CLT for balls in a Hilbert space. *J. Multivariate Anal.* **97** 2041–2056. MR2301275

CASSELS, J. W. S. (1959). *An Introduction to the Geometry of Numbers*. Springer, Berlin. MR0157947

DAVENPORT, H. (1958). Indefinite quadratic forms in many variables. II. *Proc. Lond. Math. Soc.* (3) **8** 109–126. MR0092808

ESSEEN, C.-G. (1945). Fourier analysis of distribution functions. A mathematical study of the Laplace–Gaussian law. *Acta Math.* **77** 1–125. MR0014626

FRICKER, F. (1982). *Einführung in die Gitterpunktlehre. Lehrbücher und Monographien aus dem Gebiete der Exakten Wissenschaften* (*LMW*). *Mathematische Reihe* **73**. Birkhäuser, Basel. MR0673938

GÖTZE, F. (1979). Asymptotic expansions for bivariate von Mises functionals. *Z. Wahrsch. Verw. Gebiete* **50** 333–355. MR0554550

GÖTZE, F. (1994). Unpublished manuscript.

GÖTZE, F. (2004). Lattice point problems and values of quadratic forms. *Invent. Math.* **157** 195–226. MR2135188

GÖTZE, F. and MARGULIS, G. A. (2010). Distribution of values of quadratic forms at integral points. Preprint. Available at arXiv:1004.5123.

GÖTZE, F. and ULYANOV, V. (2000). Uniform approximations in the CLT for balls in Euclidian spaces. Preprint 00-034 SFB 343. Bielefeld Univ., Bielefeld.

GÖTZE, F. and ULYANOV, V. (2003). Asymptotic disrtribution of χ^2-type statistics. Preprint 03-033 Research group "Spectral analysis, asymptotic distributions and stochastic dynamics," Bielefeld Univ., Bielefeld.

GÖTZE, F. and ZAITSEV, A. Y. (2008). Uniform rates of convergence in the CLT for quadratic forms. Preprint 08-119 SFB 701. Bielefeld Univ., Bielefeld.

GÖTZE, F. and ZAITSEV, A. Y. (2009). Uniform rates of approximation by short asymptotic expansions in the CLT for quadratic forms of sums of i.i.d. random vectors. Preprint 09-073 SFB 701. Bielefeld Univ., Bielefeld.

GÖTZE, F. and ZAITSEV, A. Y. (2010). Uniform rates of approximation by short asymptotic expansions in the CLT for quadratic forms. *Zapiski Nauchnykh Seminarov POMI* **384** 105–153.

HARDY, G. H. (1916). The average order of the arithmetical functions $P(x)$ and $\delta(x)$. *Proc. Lond. Math. Soc.* (2) **15** 192–213. MR1576556

LANDAU, E. (1915). Zur analytischen Zahlentheorie der definiten quadratischen Formen. *Sitzber. Preuss. Akad. Wiss.* **31** 458–476.

LENSTRA, A. K., LENSTRA, H. W. JR. and LOVÁSZ, L. (1982). Factoring polynomials with rational coefficients. *Math. Ann.* **261** 515–534. MR0682664

MUMFORD, D. (1983). *Tata Lectures on Theta. I. Progress in Mathematics* **28**. Birkhäuser, Boston, MA. MR0688651

NAGAEV, S. V. (1989). On a new approach to the study of the distribution of the norm of random element in Hilbert space. In *Abstracts of the Fifth Intern. Vilnius Conf. in Probab. Theory and Math. Stat.*, **4** 77–78. Mokslas, VSP, Vilnius.

NAGAEV, S. V. and CHEBOTAREV, V. I. (1999). On the accuracy of Gaussian approximation in Hilbert space. *Acta Appl. Math.* **58** 189–215. MR1734750

NAGAEV, S. V. and CHEBOTAREV, V. I. (2005). On the accuracy of Gaussian approximation in Hilbert space. *Siberian Adv. Math.* **15** 11–73. MR2141789

PETROV, V. V. (1975). *Sums of Independent Random Variables*. Springer, New York. MR0388499

PRAWITZ, H. (1972). Limits for a distribution, if the characteristic function is given in a finite domain. *Skand. Aktuarietidskr.* **1972** 138–154. MR0375429

PROKHOROV, Y. V. and ULYANOV, V. V. (2013). Some approximation problems in statistics and probability. In *Limit Theorems in Probability, Statistics and Theory, Series: Springer Proceedings in Mathematics and Statistics*, **42**. Springer, Berlin.

SENATOV, V. V. (1997). Qualitative effects in estimates for the rate of convergence in the central limit theorem in multidimensional spaces. *Tr. Mat. Inst. Steklova* **215** 239. MR1632100

SENATOV, V. V. (1998). *Normal Approximation: New Results, Methods and Problems*. VSP, Utrecht. MR1686374

ULYANOV, V. V. and G ÖTZE, F. (2011). Short asymptotic expansions in the CLT in Euclidean spaces: A sharp estimate for its accuracy. In *Proceedings* 2011 *World Congress on Engineering and Technology. Oct.* 28–*Nov.* 2, 2011. *Shanghai, China*, **1** 260–262. IEEE Press, New York.

WEYL, H. (1916). Über die Gleichverteilung der Zahlen mod-Eins. *Math. Ann.* **77** 313–352.

YURINSKIĬ, V. V. (1982). On the accuracy of normal approximation of the probability of hitting a ball. *Theory Probab. Appl.* **27** 270–278.

ZALESSKIĬ, B. A., SAZONOV, V. V. and UL'YANOV, V. V. (1988). A sharp estimate for the accuracy of the normal approximation in a Hilbert space. *Theory Probab. Appl.* **33** 700–701.

ZALESSKIĬ, B. A., SAZONOV, V. V. and UL'YANOV, V. V. (1991). A precise estimate for the rate of convergence in the central limit theorem in a Hilbert space. *Math. USSR Sbornik* **68** 453–482.

FAKULTÄT FÜR MATHEMATIK
UNIVERSITÄT BIELEFELD
POSTFACH 100131, D-33501
BIELEFELD
GERMANY
E-MAIL: goetze@math.uni-bielefeld.de

ST. PETERSBURG DEPARTMENT
OF V. A. STEKLOV MATHEMATICAL INSTITUTE
FONTANKA 27
ST. PETERSBURG 191023
RUSSIA
E-MAIL: zaitsev@pdmi.ras.ru