# UNIFORM RATES OF APPROXIMATION BY SHORT ASYMPTOTIC EXPANSIONS IN THE CLT FOR QUADRATIC FORMS OF SUMS OF I.I.D. RANDOM VECTORS

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

Abstract. Let $X, X_1, X_2, \dots$ be i.i.d. $\mathbb{R}^d$-valued real random vectors. Assume that $\mathbf{E}\, X = 0$ 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 investigate the distributions of non-degenerate quadratic forms $\mathbb{Q}[S_N]$ of the normalized sums $S_N = N^{-1/2}\,(X_1 + \dots + X_N)$ and show that, without any additional conditions, for any $a \in \mathbb{R}^d$,

$$\Delta_N^{(a)} \overset{\text{def}}{=} \sup_x \Big| \mathbf{P}\big\{\mathbb{Q}[S_N - a] \le x\big\} - \mathbf{P}\big\{\mathbb{Q}[G - a] \le x\big\} - E_a(x) \Big| = \mathcal{O}\big(N^{-1}\big),$$

provided that $d \ge 5$ and the fourth moment of $X$ exists. Here $E_a(x)$ is the Edgeworth type correction of order $\mathcal{O}\big(N^{-1/2}\big)$. Furthermore, we provide explicit bounds of order $\mathcal{O}\big(N^{-1}\big)$ for $\Delta_N^{(a)}$ and for the concentration function of the random variable $\mathbb{Q}[S_N + a]$, $a \in \mathbb{R}^d$. Our results extend the corresponding results of Bentkus and Götze (1997a) ($d \ge 9$) to the case $d \ge 5$.

## 1. Introduction and results

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

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 $\mathbf{E}\,\|X\|^2 < \infty$. Let $G$ be a mean zero Gaussian random 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

---

*Date*: August 16, 2009.

1991 *Mathematics Subject Classification.* Primary 60F05; secondary 62E20.

*Key words and phrases.* Central Limit Theorem, concentration inequalities, convergence rates, multidimensional spaces, quadratic forms, ellipsoids, hyperboloids, lattice point problem, theta-series.

[1]Research supported by the SFB 701 in Bielefeld and by grant RFBR-DFG 09-01-91331.

[2]Research supported by grant RFBR 09-01-12180, by a program by the Alexander von Humboldt Foundation and by a program of fundamental researches of Russian Academy of Sciences "Modern problems of fundamental mathematics".

distributions $\mathcal{L}(S_N)$ of sums

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

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 shall say that $\mathbb{Q}$ is non-degenerate if $\ker \mathbb{Q} = \big\{0\big\}$.

Denote, for $q > 0$,

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

Introduce the distribution functions

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

Write

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

We shall provide explicit bounds for $\Delta_N$.

**Theorem 1.1.** *Let $\mathbf{E}\,X = 0$. Assume that $\mathbb{Q}$ and $\mathbb{C}$ are non-degenerate 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 1.1 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 $d \ge 5$.

The distribution function of $\|S_N\|^2$ (for bounded $X$ with values in $\mathbb{R}^d$) may have jumps of order $\mathcal{O}\big(N^{-1}\big)$, for all $1 \le d \le \infty$. See, e.g., BG (1996, p. 468). Therefore, the bound of Theorem 1.1 is optimal.

Theorem 1.1 and the method of its 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 $x \ne 0$. Let $\operatorname{vol} E_r$ be the volume of the ellipsoid

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

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 of Götze (2004) is related to Theorem 1.1 (see also BG (1995a, 1997b)).

**Theorem 1.2.** (Götze (2004)) *For all dimensions $d \ge 5$,*

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

*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 1.2 solves the lattice point problem for $d \geq 5$, and 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 Esseen (1945) for the CLT for ellipsoids with axes parallel to coordinate axes. A related result for indefinite forms may be found in Götze and Margulis (2009).

For Hilbert spaces the order of error under the conditions of Theorem 1.1 had been investigated intensively. See Zalesskiĭ, Sazonov and Ulyanov (1988) and Nagaev (1989) for the optimal (with respect to eigenvalues of $\mathbb{C}$) bound of order $\mathcal{O}(N^{-1/2})$ under the assumption of finiteness of the third moment. For a more detailed discussion see Bentkus, Götze, Paulauskas and Račkauskas (1990), BG (1995b, 1996, 1997a) and Senatov (1989).

Under more restrictive moment and dimension conditions the estimate $\mathcal{O}(N^{-1+\varepsilon})$, with $\varepsilon \downarrow 0$ as $d \uparrow \infty$, was obtained by Götze (1979), by an application of a result for bivariate $U$-statistics. The symmetrization inequality for characteristic functions introduced in Götze (1979) and its extensions play the crucial role in the proofs of bounds in the CLT on ellipsoids and hyperboloids in finite and infinite dimensional cases. This inequality is related to Weyl's (1915/16) inequality for trigonometric sums. Under some special smoothness assumptions, error bounds $\mathcal{O}(N^{-1})$ (and even 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) ($d \geq 9$, a more precise dependence of constants on the eigenvalues of $\mathbb{C}$) and Bogatyrev, Götze and Ulyanov (2006) (non-uniform 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.

Additional restrictions like the diagonality of $\mathbb{Q}$, $\mathbb{C}$ and the independence of first five coordinates allowed already to reduce the dimension requirement for the bound $\mathcal{O}(N^{-1})$ to $d \geq 5$, see Bentkus and Götze (1996). The independence assumption in BG (1996) allowed to apply an adaption of the Hardy–Littlewood circle method. For the general case described in Theorem 1.1, we have to develop a new tool. 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

$\mathcal{O}(N^{-3/4}\,\log^\delta N)$ for $\Delta_N$. Hardy (1916) conjectured that up to logarithmic factors this is the optimal order.

To formulate the results we need more notation repeating most part of the notation used in BG (1997a). Write $\sigma^2 \overset{\text{def}}{=} \beta_2$. Let $\sigma_1^2 \geq \sigma_2^2 \geq \dots$ be the eigenvalues of $\mathbb{C}$, counting their multiplicities. We have $\sigma^2 = \sigma_1^2 + \sigma_2^2 + \cdots$.

Throughout $\mathcal{S} = \{e_1, \dots, e_s\} \subset \mathbb{R}^d$ denotes a finite set of cardinality $s$. We shall write $\mathcal{S}_o$ instead of $\mathcal{S}$ if the system $\{e_1, \dots, e_s\}$ is orthonormal.

Let $p > 0$ and $\delta \geq 0$. Following BG (1997a), we introduce the non-degeneracy condition for the distribution of a $d$-dimensional vector $Y$:

$$\mathcal{N}(p, \delta, \mathcal{S}, Y): \qquad \mathbf{P}\big\{\, \|Y - e\| \leq \delta \big\} \geq p, \qquad \text{for all } e \in \mathcal{S} \cup \mathbb{Q}\mathcal{S}. \tag{1.4}$$

We shall refer to condition (1.4) as condition $\mathcal{N}(p, \delta, \mathcal{S}, Y) = \mathcal{N}(p, \delta, \mathcal{S}, Y; \mathbb{Q})$.

A particular case where explicit lower bounds for $p$ in (1.4) can be given in terms of eigenvalues of $\mathbb{C}$ and $\mathbb{Q}$, is described in BG (1997a). Assume that there exists an orthonormal system $\mathcal{S}_o = \{e_1, \dots, e_s\}$ of eigenvectors of $\mathbb{C}$ such that $\mathbb{Q}\mathcal{S}_o$ is again a system of eigenvectors of $\mathbb{C}$. Then we shall say that condition $\mathcal{B}(\mathcal{S}_o, \mathbb{C}) = \mathcal{B}(\mathcal{S}_o, \mathbb{C}; \mathbb{Q})$ is fulfilled. In this case we shall write

$$\lambda_s^2 = \min_{e \in \mathcal{S}_o \cup \mathbb{Q}\mathcal{S}_o} \sigma_e^2, \tag{1.5}$$

where $\sigma_e^2$ is the eigenvalue of $\mathbb{C}$ which corresponds to the eigenvector $e$. In particular, such a system $\mathcal{S}_o$ exists if we assume that $\mathbb{Q}$ and $\mathbb{C}$ are diagonal in a common orthonormal basis. If, in addition, $\mathbb{Q}$ is isometric, then we can choose $\mathcal{S}_o$ such that $\lambda_s^2 = \sigma_s^2$.

Introduce truncated random vectors

$$X^\diamond = X\,\mathbf{I}\big\{\, \|X\| \leq \sigma\sqrt{N}\,\big\}, \qquad X_\diamond = X\,\mathbf{I}\big\{\, \|X\| > \sigma\sqrt{N}\,\big\}, \tag{1.6}$$

$$X^\bullet = X\,\mathbf{I}\big\{\, \|\mathbb{C}^{-1/2}\,X\| \leq \sqrt{N}\,\big\}, \qquad X_\bullet = X\,\mathbf{I}\big\{\, \|\mathbb{C}^{-1/2}\,X\| > \sqrt{N}\,\big\}, \tag{1.7}$$

and their moments (for $q > 0$)

$$\Lambda_4^\diamond = \frac{1}{\sigma^4\,N}\,\mathbf{E}\,\|X^\diamond\|^4, \qquad \Pi_q^\diamond = \frac{N}{(\sigma\sqrt{N})^q}\,\mathbf{E}\,\|X_\diamond\|^q, \tag{1.8}$$

$$\Lambda_4^\bullet = \frac{1}{\sigma^4\,N}\,\mathbf{E}\,\|X^\bullet\|^4, \qquad \Pi_q^\bullet = \frac{N}{(\sigma\sqrt{N})^q}\,\mathbf{E}\,\|X_\bullet\|^q. \tag{1.9}$$

Here and below $\mathbf{I}\big\{A\big\}$ denotes the indicator of an event $A$. Clearly, we have

$$X^\diamond + X_\diamond = X^\bullet + X_\bullet = X, \quad \|X^\diamond\|\,\|X_\diamond\| = \|X^\bullet\|\,\|X_\bullet\| = 0. \tag{1.10}$$

Generally speaking, $X^\bullet$ 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 (1.7) was there applied to the vector $X^\diamond$. The use of $X^\bullet$ is more natural for the estimation of constants in the case $d < \infty$. It is easy to see that

$$\mathbf{P}\big\{\, \|X^\bullet\| \leq \sigma_1\sqrt{N}\,\big\} = \mathbf{P}\big\{\, \|X_\bullet\| > \sigma_d\sqrt{N}\,\big\} = 1. \tag{1.11}$$

In Sections 3 and 4 we shall denote

$$X' = X^{\bullet} - \mathbf{E}\, X^{\bullet} + W, \tag{1.12}$$

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 2.3.

Introduce the notation which will be used in the sequel. By $c, c_1, c_2, \dots$ we shall denote absolute positive constants. If a constant depends on, say, $s$, then we shall 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., Theorems 1.3, 1.4, 1.5 and 2.1) we write $c_0$ for an *arbitrary* positive absolute constant, for example one may choose $c_0 = 1$. We shall 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 shall also write $A \asymp_s B$ if $A \ll_s B \ll_s A$. For $a \in \mathbb{R}^d$ we shall always denote $b = \sqrt{N}\, a$. By $\lceil \alpha \rceil$ we shall denote the integer part of a number $\alpha$.

Throughout we assume that all random vectors and variables are independent in aggregate, if the contrary is not clear from the context. By $\overline{X}$ and $X_1, X_2, \dots$ we shall denote independent copies of a random vector $X$. Similarly, $\overline{G}, G_1, G_2, \dots$ are independent copies of $G$ and so on. By $\mathcal{L}(X)$ we shall 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 $Z_N = X_1 + \dots + X_N$. Then $S_N = N^{-1/2} Z_N$. Similarly, by $Z_N^{\diamond}$ (resp. $Z_N^{\bullet}$ and $Z_N'$) we shall denote sums of $N$ independent copies of $X^{\diamond}$ (resp. $X^{\bullet}$ and $X'$). For example, $Z_N' = X_1' + \dots + X_N'$.

We shall 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.

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

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

given all random vectors but $Y$.

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

$$\widehat{F}(t) = \int_{-\infty}^{\infty} \mathrm{e}\{tx\}\, dF(x) \tag{1.13}$$

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

$$F_a(x) \overset{\mathrm{def}}{=} \mathbf{P}\big\{\mathbb{Q}[S_N - a] \le x\big\}, \qquad H_a(x) \overset{\mathrm{def}}{=} \mathbf{P}\big\{\mathbb{Q}[G - a] \le x\big\}, \qquad a \in \mathbb{R}^d. \tag{1.14}$$

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

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

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

$$\widehat{E}_a(t)=\frac{2\,(it)^2}{3\sqrt{N}}\,\mathbf{E}\;\mathrm{e}\big\{t\mathbb{Q}[Y]\big\}\big(3\,\langle\mathbb{Q}X,Y\rangle\,\langle\mathbb{Q}X,X\rangle+2it\,\langle\mathbb{Q}X,Y\rangle^3\big),\quad Y=G-a. \tag{1.15}$$

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

$$E_a(x)\overset{\mathrm{def}}{=}\Theta_b(Nx)\overset{\mathrm{def}}{=}\frac{1}{6\sqrt{N}}\,\chi(A_x),\qquad A_x=\big\{u\in\mathbb{R}^d:\ \mathbb{Q}[u-a]\le x\big\}, \tag{1.16}$$

with the signed measure

$$\chi(A)\overset{\mathrm{def}}{=}\int_A\mathbf{E}\,p'''(x)X^3\,dx,\qquad\qquad \text{for the Borel sets } A\subset\mathbb{R}^d, \tag{1.17}$$

and where

$$p'''(x)u^3=p(x)\big(3\langle\mathbb{C}^{-1}u,u\rangle\langle\mathbb{C}^{-1}x,u\rangle-\langle\mathbb{C}^{-1}x,u\rangle^3\big) \tag{1.18}$$

denotes the third Frechet derivative of $p$ in the 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.

We can write similar representations for $E_a^{\bullet}(x)=\Theta_b^{\bullet}(Nx)$, $E_a^{\diamond}(x)=\Theta_b^{\diamond}(Nx)$ and $E_a'(x)=\Theta_b'(Nx)$ just replacing $X$ by $X^{\bullet}$, $X^{\diamond}$ and $X'$ in (1.15) or (1.17).

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

$$\Psi_b(x)\overset{\mathrm{def}}{=}\mathbf{P}\big\{\mathbb{Q}[Z_N-b]\le x\big\}, \tag{1.19}$$

and

$$\Phi_b(x)\overset{\mathrm{def}}{=}\mathbf{P}\big\{\mathbb{Q}[\sqrt{N}\,G-b]\le x\big\}. \tag{1.20}$$

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

$$\Delta_N^{(a)}\overset{\mathrm{def}}{=}\sup_{x\in\mathbb{R}}\big|F_a(x)-H_a(x)-E_a(x)\big|=\sup_{x\in\mathbb{R}}\big|\Psi_b(x)-\Phi_b(x)-\Theta_b(x)\big|, \tag{1.21}$$

where $E_a(x)$ is the Edgeworth correction (see (1.14), (1.16), (1.19) and (1.20) to justify the last equality in (1.21)). We write $\Delta_{N,\bullet}^{(a)}$ and $\Delta_{N,\diamond}^{(a)}$ replacing $E_a$ by $E_a^{\bullet}$ and $E_a^{\diamond}$ in (1.21).

The aim of this paper is to derive for $\Delta_N^{(a)}$ bounds of order $\mathcal{O}(N^{-1})$ without any additional smoothness type assumptions. Theorem 1.3 (which was proved in BG (1997a)) solved this problem in the case $13\le d\le\infty$.

In Theorems 1.3, 1.4 and 1.5 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)).

**Theorem 1.3.** (BG (1997a, Theorem 1.3)) *Let* $\mathbf{E}\, X = 0$, $\delta = 1/300$, $\mathbb{Q}^2 = \mathbb{I}_d$, $s = 13$ *and* $13 \le d \le \infty$. *Then we have:*

($i$) *Assume that condition* $\mathcal{N}(p, \delta, \mathcal{S}_o, c_0 G/\sigma)$ *holds. Then*

$$\Delta_N^{(a)} \le C\left(\Pi_3^\diamond + \Lambda_4^\diamond\right)\left(1 + \|a/\sigma\|^6\right) \tag{1.22}$$

*and*

$$\Delta_{N,\diamond}^{(a)} \le C\left(\Pi_2^\diamond + \Lambda_4^\diamond\right)\left(1 + \|a/\sigma\|^6\right) \tag{1.23}$$

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

($ii$) *Assume that condition* $\mathcal{B}(\mathcal{S}_o, \mathbb{C})$ *is fulfilled. Then the constant in* (1.22) *and* (1.23) *satisfies* $C \le \exp\bigl\{c\sigma^2\lambda_{13}^{-2}\bigr\}$.

Unfortunately, we cannot apply Theorem 1.3 for $d = 5, 6, \dots, 12$. The main result of the paper is Theorem 1.4. It is valid for $5 \le d < \infty$ in finite-dimensional spaces $\mathbb{R}^d$ only. However, the bounds of Theorem 1.4 depend on the smallest $\sigma_d$. This makes them unstable if one of coordinates of $X$ degenerates.

**Theorem 1.4.** *Let* $\mathbf{E}\, X = 0$, $\delta = 1/300$, $\mathbb{Q}^2 = \mathbb{I}_d$, $s = 5$ *and* $5 \le d < \infty$. *Then we have:*

($i$) *Assume that condition* $\mathcal{N}(p, \delta, \mathcal{S}_o, c_0\, G/\sigma)$ *holds. Then*

$$\Delta_N^{(a)} \le C\left(\sigma_d\sigma^{-1}\,\Pi_3^\bullet + \Lambda_4^\bullet\right)\left(1 + \|a/\sigma\|^3\right), \tag{1.24}$$

*and*

$$\Delta_{N,\bullet}^{(a)} \le C\left(\sigma_d^2\sigma^{-2}\,\Pi_2^\bullet + \Lambda_4^\bullet\right)\left(1 + \|a/\sigma\|^3\right), \tag{1.25}$$

*with* $C = c_d\, p^{-3}\,(\sigma/\sigma_d)^4$.

($ii$) *Assume that condition* $\mathcal{B}(\mathcal{S}_o, \mathbb{C})$ *is fulfilled. Then the constant in* (1.24) *and* (1.25) *may be estimated as* $C \le c_d\sigma^4\sigma_d^{-4}\exp\bigl\{c\sigma^2/\lambda_5^2\bigr\}$.

Theorems 1.3 and 1.4 yield Theorem 1.1, using that $E_0(x) \equiv 0$,

$$\sigma_d\sigma^{-1}\,\Pi_3^\bullet + \Lambda_4^\bullet \le \Pi_3^\diamond + \Lambda_4^\diamond \le \beta_4/(\sigma^4 N), \tag{1.26}$$

and

$$\sigma_d^2\sigma^{-2}\,\Pi_2^\bullet + \Lambda_4^\bullet \le \Pi_2^\diamond + \Lambda_4^\diamond \le \beta_4/(\sigma^4 N). \tag{1.27}$$

Theorem 1.4 extends to the case $d \ge 5$ Theorem 1.5 of BG (1997a) which contains the corresponding bounds for $d \ge 9$. Moreover, the first inequalities in (1.26) and (1.27) show that Theorem 1.4 is a little bit sharper than Theorem 1.5 of BG (1997a) even in the case $9 \le d < \infty$. In BG (1997a), $\sigma_d\sigma^{-1}\,\Pi_3^\bullet + \Lambda_4^\bullet$ and $\sigma_d^2\sigma^{-2}\,\Pi_2^\bullet + \Lambda_4^\bullet$ were replaced in (1.24) and (1.25) by $\Pi_3^\diamond + \Lambda_4^\diamond$ and $\Pi_2^\diamond + \Lambda_4^\diamond$ respectively.

If, in the conditions of Theorem 1.4, the distribution of $X$ is symmetric or $a = 0$, then the Edgeworth corrections $E_a(x)$ and $E_a^\bullet(x)$ vanish and

$$\Delta_N^{(a)} = \Delta_{N,\bullet}^{(a)} \le C\left(\sigma_d^2\sigma^{-2}\,\Pi_2^\bullet + \Lambda_4^\bullet\right)\left(1 + \|a/\sigma\|^3\right), \qquad C = c_d\, p^{-3}\,(\sigma/\sigma_d)^4. \tag{1.28}$$

The corresponding inequality from Theorem 1.4 of BG (1997a) provides in the case $s = 9$ and $9 \le d \le \infty$ the bound

$$\Delta_N^{(a)} \le C\left(\Pi_2^\diamond + \Lambda_4^\diamond\right)\left(1 + \|a/\sigma\|^4\right), \qquad C = cp^{-4}. \tag{1.29}$$

It is clear that sometimes the bound (1.28) may be sharper than (1.29), but unfortunately, it depends on the smallest eigenvalue $\sigma_d$.

In Götze and Zaitsev (2008) we proved Theorem 1.4 in the case $a = 0$ and hence, Theorem 1.1. First direct attempts to prove similar bounds for $a \ne 0$, assuming $d \ge 5$ instead of $d \ge 9$ in Theorems 1.4 and 1.5 of BG (1997a) failed. The main problem was that Lemma 3.2 of BG (1997a) allowed us to integrate the remainder terms of expansions for $\alpha < s/2$ only. Our Lemma 3.2 provides the bounds for $\alpha \ge s/2$ too. Note, however, that this difficulty was already successively avoided in the proof of the main result of BG (1996) (without estimation of constants).

The bounds for constants in Theorems 1.3 and 1.4 are not optimal. See the papers of Nagaev and Chebotarev (1999), (2005), Götze and Ulyanov (2000), and Bogatyrev, Götze and Ulyanov (2006) for more precise estimates of constants in the case $d \ge 9$. One can find lower bounds for $\Delta_N^{(a)}$ under different conditions on $a$ and $\mathcal{L}(X)$ in Götze and Ulyanov (2000).

Note that, in the proof of Theorem 1.3 in BG (1997a), inequalities (1.22) and (1.23) were derived for the Edgeworth correction $E_a(x)$ defined by (1.15). However, from Theorems 1.3 and Theorem 1.4 it follows that, at least for $13 \le d < \infty$, definitions (1.15) and (1.16) determine the same function $E_a(x)$. Indeed, both functions may be represented as $N^{-1/2}\,K(x)$, where $K(x)$ are some functions of bounded variation which are independent of $N$. Furthermore, inequalities (1.22) and (1.24) provide both bounds of order $\mathcal{O}(N^{-1})$. This is possible if the Edgeworth corrections $E_a(x)$ are the same in these inequalities.

On the other hand, it is proved that definition (1.15) determine a function of bounded variation for $d \ge 9$ only (see BG (1997a, Lemma 5.7)) while definition (1.16) has no sense for $d = \infty$.

Introduce the concentration function

$$Q(X;\,\lambda) = Q(X;\,\lambda;\,\mathbb{Q}) = \sup_{a,\,x\in\mathbb{R}^d} \mathbf{P}\big\{x \le \mathbb{Q}[X-a] \le x+\lambda\big\}, \quad \text{for } \lambda \ge 0. \tag{1.30}$$

It should be mentioned that the supremum in (1.30) is taken not only over all $x$, but over all $x$ and $a \in \mathbb{R}^d$. Usually, one defines the concentration function of the random variable $\mathbb{Q}[X-a]$ taking the supremum over all $x \in \mathbb{R}^d$ only.

**Theorem 1.5.** *Let $\mathbb{Q}^2 = \mathbb{I}_d$, $5 \le s \le d \le \infty$ and $0 \le \delta \le 1/(5s)$. Then we have:*

*(i) If condition $\mathcal{N}(p,\delta,\mathcal{S}_o,\widetilde{X})$ is fulfilled with some $p > 0$, then*

$$Q(Z_N;\,\lambda) \le c_s\,(pN)^{-1}\max\big\{1;\,\lambda\big\}, \qquad \lambda \ge 0. \tag{1.31}$$

*(ii) If, for some $m$, condition $\mathcal{N}(p,\delta,\mathcal{S}_o,m^{-1/2}\widetilde{Z}_m)$ is fulfilled, then*

$$Q(Z_N;\,\lambda)\le c_s\,(pN)^{-1}\max\{m;\,\lambda\},\qquad \lambda\ge 0. \tag{1.32}$$

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 $\mathbb{R}^d$, $1\le d\le\infty$, then, for any $\delta>0$ and $\mathcal{S}$ there exists a natural number $m$ such that the condition $\mathcal{N}(p,\delta,\mathcal{S},m^{-1/2}\widetilde{Z}_m)$ holds with some $p>0$.

Theorem 1.5 and more explicit Theorem 2.1 extend to the case $d\ge 5$ Theorems 1.6 and 2.1 of BG (1997a) which were proved for $d\ge 9$.

We conclude the Introduction by a brief description of the basic 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, discretization procedure and double large sieve. We do not use here multiplicative inequalities of BG (1997a). We replace here their application by some arguments coming from the number theory. The original part of our proof is concentrated at Sections 4–7. Similarly to BG (1997a), in Section 2, we prove bounds for concentration functions. The proofs is technically simpler as those of Theorem 1.4, but it already contains all the principal ideas. These proofs repeats almost literally the corresponding proofs of BG (1997a). The only difference consists in the use of new Lemma 7.3 which allows us to estimate characteristic functions for large values of argument $t$. In Sections 3 and 4 Theorem 1.4 is proved. We shall replace Lemma 9.4 of BG (1997a) by its improvement, Lemmas 4.1. Another difference is in another choice of $k\asymp_d\sigma_d^{-3}N^{1/4}\,\overline{\beta}^{3/4}$ in (4.30) and (4.31) instead of $k\asymp_d\sigma_d^{-2}\sqrt{N\overline{\beta}}$ in BG (1997a).

In Sections 5–7 we prove estimates for characteristic functions. Section 5 is started with results from BG (1997a) (Lemmas 5.1–5.3). Their proofs in BG (1997a) are based on conditioning, discretization, as well as on the double large sieve.

Let $\varepsilon_1,\varepsilon_2,\dots$ denote i.i.d. symmetric Rademacher random variables. Let $\delta>0$ and $\mathcal{S}=\{e_1,\dots,e_s\}\subset\mathbb{R}^d$. We shall write $\mathcal{L}(Y)\in\mathbf{\Gamma}(\delta;\,\mathcal{S})$ if a discrete random vector $Y$ is distributed as $\varepsilon_1z_1+\dots+\varepsilon_sz_s$, with some (non-random) $z_j\in\mathbb{R}^d$ such that $\|z_j-e_j\|\le\delta$, for all $1\le j\le s$.

Define the function

$$\mathcal{M}(t;N)=1/\sqrt{|\,t\,|\,N},\ \text{ for } |\,t\,|\le N^{-1/2},\qquad \mathcal{M}(t;N)=\sqrt{|\,t\,|},\ \text{ for } |\,t\,|\ge N^{-1/2}. \tag{1.33}$$

It is easy to see that, for $s>0$,

$$2^{-1}\big(|\,tN|^{-s/2}+|\,t\,|^{s/2}\big)\le\mathcal{M}^s(t;N)\le|\,tN|^{-s/2}+|\,t\,|^{s/2}. \tag{1.34}$$

Assuming the condition $\mathcal{N}(p,\delta,\mathcal{S}_o,\widetilde{X})$ with $0\le\delta\le 1/(5s)$ and an orthonormal system $\mathcal{S}_o$, we can use Lemma 5.2 which implies that, for any $a\in\mathbb{R}^d$ and $t\in\mathbb{R}$,

$$\big|\,\widehat{\Psi}_a(t)\,\big|=\big|\mathbf{E}\ \mathrm{e}\big\{t\,\mathbb{Q}[Z_N-a]\big\}\,\big|\ll_s\mathcal{M}^s(t;k),\quad k=pN. \tag{1.35}$$

Moreover, by Lemma 5.4, for any $0 < A \le B$, $a \in \mathbb{R}^d$ and $\gamma > 0$,

$$\int_A^B \big|\,\widehat{\Psi}_a(t)\,\big|\,\frac{dt}{|\,t\,|}\;\le c_\gamma(s)\,(pN)^{-\gamma}\,\log\frac{B}{A}+\sup_{\Gamma}\int_A^B\sqrt{\mathbf{E}\,\mathrm{e}\big\{\,t\,\langle\widetilde{W},\widetilde{W}'\rangle/2\,\big\}}\;\frac{dt}{|\,t\,|}\,, \tag{1.36}$$

where $W = V_1 + \dots + V_n$ and $W' = V_1' + \dots + V_n'$ are independent sums of independent copies of random vectors $V$ and $V'$ respectively, and the supremum $\sup_\Gamma$ is taken over all $\mathcal{L}(V), \mathcal{L}(V') \in \mathbf{\Gamma}(\delta;\, \mathcal{S}_o)$.

Comparing (1.36) and (5.4) (used by BG (1997a)), we see that inequality (5.4) is related to sums of *non-i.i.d.* vectors $\{V_j\}$ and $\{V_j'\}$ while inequality (1.36) deals with i.i.d. vectors. Nevertheless, we shall derive (1.36) from (5.4).

Inequalities of type (1.35) allow to prove in Theorem 1.1 error bounds $\mathcal{O}(N^{-\alpha})$ only, for some $\alpha < 1$. This is due to possible oscillations of $\big|\,\widehat{\Psi}_a(t)\,\big|$ between 0 and 1, as $|\,t\,| \sim N^{-\varepsilon}$ with small $\varepsilon \ge 0$. In Section 5, we reduce the estimation of $\mathbf{E}\,\mathrm{e}\big\{\,t\,\langle\widetilde{W},\widetilde{W}'\rangle/2\,\big\}$ to the estimation of a theta-series (see Lemma 5.6 and inequalities (5.28) and (5.29)). 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 (5.13), (5.16), (5.18) and (5.19). Together with the non-negativity of some characteristic functions (see (5.17) and (5.21)), this allows us to apply then the Poisson summation formula from Lemma 5.5. This formula reduces the problem to an estimation of integrals of theta-series. Section 6 is devoted to some facts from Number Theory. We consider the lattices, their $\alpha$-characteristics and Minkowski's successive minima. In Section 7 we reduce the estimation of integrals of theta-series to some integrals of $\alpha$-characteristics. An application of a new Lemma 7.2 proved by Götze and Margulis (2009) ends the proof.

**Acknowledgment** We would like to thank V.V. Ulyanov for helpful discussions.

## 2. Proofs of bounds for concentration functions; truncation

We start Section 2 with Theorem 2.1 which (under additional restrictions) provides more explicit bounds for the concentration than those of Theorem 1.5. In the next theorem, $c_0$ is an arbitrary positive absolute constant. Recall as well that we write $\beta = \beta_4 = \mathbf{E}\,\|X\|^4$.

**Theorem 2.1.** *Assume that* $5 \le d \le \infty$ *and that the operator* $\mathbb{Q}$ *is isometric. Then, for any random vector* $X$ *such that* $\mathbf{E}\,X = 0$ *and* $\sigma^2 < \infty$, *we have*:

(*i*) *Assume condition* $\mathcal{N}(p, \delta, \mathcal{S}_o, c_0 G/\sigma)$ *to be satisfied with* $s = 5$ *and* $\delta = 1/200$. *Then*

$$Q(Z_N;\, \lambda) \ll p^{-2} \max\big\{\,\Pi_2^{\diamond} + \Lambda_4^{\diamond};\; \lambda\sigma^{-2}N^{-1}\,\big\}, \qquad \lambda \ge 0. \tag{2.1}$$

*In particular,* $Q(Z_N;\, \lambda) \ll p^{-2}N^{-1}\max\big\{\,\beta/\sigma^4;\; \lambda/\sigma^2\,\big\}$.

*(ii) Let condition* $\mathcal{B}(\mathcal{S}_o, \mathbb{C})$ *(see (1.5)) be fulfilled with* $s = 5$*. Then*

$$Q(Z_N;\,\lambda) \le \exp\bigl\{ c\sigma^2/\lambda_5^2 \bigr\} \max\bigl\{ \Pi_2^\diamond + \Lambda_4^\diamond;\, \lambda\sigma^{-2} N^{-1} \bigr\}, \qquad \lambda \ge 0, \tag{2.2}$$

*where c denotes a sufficiently large absolute constant.*

*Proof of Theorems* 1.5 *and* 2.1. Below we shall prove the assertions (1.31); (1.31) $\Longrightarrow$ (1.32); (1.32) $\Longrightarrow$ (2.1) and (2.1) $\Longrightarrow$ (2.2). 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 7.3 in the proof of Lemma 2.2. □

For $T \ge t_0, t_1 \ge 0$ and $a \in \mathbb{R}^d$, define the integrals

$$I_0 = \int_{-t_1}^{t_1} \bigl| \widehat{\Psi}_a(t) \bigr| \, dt, \qquad\qquad I_1 = \int_{t_0 \le |t| \le T} \bigl| \widehat{\Psi}_a(t) \bigr| \, \frac{dt}{|t|}\,,$$

where

$$\widehat{\Psi}_a(t) = \mathbf{E}\ \mathrm{e}\bigl\{ t\,\mathbb{Q}[Z_N - a] \bigr\} \tag{2.3}$$

denotes the Fourier–Stieltjes transform of the distribution function $\Psi_a$ of $\mathbb{Q}[Z_N - a]$ (see (1.13) and (1.19)). Note that $\bigl|\widehat{\Psi}_a(-t)\bigr| = \bigl|\widehat{\Psi}_a(t)\bigr|$.

**Lemma 2.2.** *Assume condition* $\mathcal{N}(p, \delta, \mathcal{S}_o, \widetilde{X})$ *with some* $0 \le \delta \le 1/(5s)$ *and* $s \ge 5$*. Let*

$$t_0 = c_1(s)\,(pN)^{-1+2/s}, \qquad t_1 = c_2(s)\,(pN)^{-1/2}, \qquad c_3(s) \le T \le c_4(s) \tag{2.4}$$

*with some positive constants* $c_j(s)$*,* $1 \le j \le 4$*. Then*

$$I_0 \ll_s (pN)^{-1}, \qquad\qquad I_1 \ll_s (pN)^{-1}. \tag{2.5}$$

*Proof.* Denote $k = pN$. Without loss of generality we assume that $k \ge c_s$, for a sufficiently large constant $c_s$. Indeed, if $k \le c_s$, then one can prove (2.5) using $|\widehat{\Psi}_a| \le 1$. Choosing $c_s$ to be large enough, we ensure that $k \ge c_s$ implies $1/k \le t_0 \le t_1 \le T$.

Let us prove inequality (2.5) for $I_0$. By Lemma 5.2, we have

$$\bigl|\widehat{\Psi}_a(t)\bigr| \ll_s \mathcal{M}^s(t;k), \qquad\qquad k = pN. \tag{2.6}$$

Taking into account that $|\widehat{\Psi}_a| \le 1$, we see that $\bigl|\widehat{\Psi}_a(t)\bigr| \ll_s \min\bigl\{1;\, \mathcal{M}^s(t;k)\bigr\}$. Furthermore, denoting $t_2 = k^{-1/2}\max\{1; c_2(s)\}$ and using definition (1.33) of the function $\mathcal{M}$, we obtain

$$I_0 \ll_s \int_0^{1/k} dt + \int_{1/k}^{\infty} \frac{dt}{(t\,k)^{s/2}} + \int_0^{t_2} t^{s/2}\, dt = \frac{1}{k} + \frac{c_s}{k} + \frac{c_s}{k^{(s+2)/4}} \ll_s \frac{1}{k}\,,$$

proving (2.5) for $I_0$.

It remains to estimate $I_1$. Using (1.34), (2.4) and (2.6), it is easy to verify that

$$\int_{t_0}^{k^{-2/s}} \bigl|\widehat{\Psi}_a(t)\bigr| \,\frac{dt}{t} \ll_s \int_{t_0}^{\infty} \frac{dt}{t\,(t\,k)^{s/2}} + \int_0^{k^{-2/s}} t^{s/2-1}\, dt \ll_s \frac{1}{k}\,. \tag{2.7}$$

Furthermore, $T \asymp_s 1$ (see (2.4)). Lemma 7.3 implies now that

$$\int_{c_5(s)k^{-1}}^{T} \big|\,\widehat{\Psi}_a(t)\big|\,\frac{dt}{t} \ll_s \frac{1}{k}\,, \tag{2.8}$$

with some $c_5(s)$, and $c_5(s)\,k^{-1} < k^{-2/s}$, if $k \ge c_s$ with sufficiently large $c_s$. The second inequality in (2.5) follows from (2.7) and (2.8). $\square$

*Proof of* (1.31). Using a well-known inequality for concentration functions (see, for example, Petrov (1975, Lemma 3 of Ch. 3)), we have

$$Q(Z_N;\,\lambda) \le 2 \sup_{a\in\mathbb{R}^d} \max\Big\{\lambda;\ \frac{1}{T}\Big\} \int_{-T}^{T} \big|\,\widehat{\Psi}_a(t)\big|\,dt, \tag{2.9}$$

for any $T > 0$. To estimate the integral in (2.9) we shall apply Lemma 2.2. Let us choose $T = 1$. Without loss of generality we assume that $pN \ge 1$. Then, using $1 \le 1/|\,t\,|$, for $|\,t\,| \le 1$, we have

$$\int_{-T}^{T} \big|\,\widehat{\Psi}_a(t)\big|\,dt \le \int_{|t|\le(pN)^{-1/2}} \big|\,\widehat{\Psi}_a(t)\big|\,dt + \int_{(pN)^{-1/2}\le|t|\le 1} \big|\,\widehat{\Psi}_a(t)\big|\,\frac{dt}{|\,t\,|} \overset{\text{def}}{=} J_0 + J_1.$$

Lemma 2.2 implies $J_0 \ll_s 1/(pN)$, $J_1 \ll_s 1/(pN)$ and, hence, (1.31). $\square$

*Proof of* (1.31) $\Longrightarrow$ (1.32). 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 $W_k = Y_1 + \dots + Y_k$. Then $\mathcal{L}(Z_N) = \mathcal{L}(\sqrt{m}\,W_k + y)$ with $k = \lceil N/m \rceil$ and with some $y$ independent of $W_k$. Therefore, we have $Q(Z_N;\lambda) \le Q(W_k;\lambda/m)$. In order to estimate $Q(W_k;\lambda/m)$ we can apply (1.31) replacing $Z_N$ by $W_k$. We obtain

$$Q(W_k;\lambda/m) \ll_s (pk)^{-1}\max\{1;\,\lambda/m\} \ll_s (pN)^{-1}\max\{m;\,\lambda\}. \qquad \square$$

Recall that truncated random vectors and their moments were defined by (1.6)–(1.9) and that $\mathbb{C} = \operatorname{cov} X = \operatorname{cov} G$.

**Lemma 2.3.** *The random vectors $X^\bullet$, $X_\bullet$ satisfy*

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

*There exist independent centered Gaussian vectors $G_*$ and $W$ such that $\mathcal{L}(G) = \mathcal{L}(G_* + W)$ and*

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

*Furthermore,* $\mathbf{E}\,\|G\|^2 = \mathbf{E}\,\|G_*\|^2 + \mathbf{E}\,\|W\|^2$ *and* $\mathbf{E}\,\|W\|^2 \le 2\sigma^2\,\Pi_2^\bullet$.

We omit the simple proof of this lemma (see BG (1997a, Lemma 2.4) for the same statement with $\diamond$ instead of $\bullet$). Lemma 2.3 allows us to define the vector $X'$ by (1.12).

**Lemma 2.4.** (BG (1997a, Lemma 5.3)) *Let $\delta > 0$. Assume that $X$ is symmetric. Then there exists an absolute positive constant $c$ such that the condition $\mathcal{N}(2p, \delta, \mathcal{S}, G)$ implies the condition $\mathcal{N}(p, 2\delta, \mathcal{S}, S_m)$, for $m \geq c\beta/(p\delta^4)$.*

**Lemma 2.5.** (BG (1997a, Lemma 5.4)) *Assume that $0 < 4\varepsilon \leq \delta \leq 1$. Let $e \in \mathbb{R}^d$, $\|e\| = 1$, be an eigenvector of the covariance operator $\mathbb{C} = \operatorname{cov} G$, so that $\mathbb{C}e = \sigma_e e$ with some $\sigma_e > 0$. Then the probability $p_e = \mathbf{P}\big\{ \|\varepsilon\sigma^{-1}G - e\| \leq \delta \big\}$ satisfies the inequality $p_e \geq \exp\big\{ -c\sigma^2\varepsilon^{-2}\sigma_e^{-2} \big\}$ with some positive absolute constant $c$.*

Recall that $Z_N^\bullet$ and $Z_N^\diamond$ denote sums of $N$ independent copies of $X^\bullet$ and $X^\diamond$ respectively.

**Lemma 2.6.** *Let $\varepsilon > 0$. There exist absolute positive constants $c$ and $c_1$ such that*

*$(i)$ the condition $\Pi_2^\bullet \leq c_1 p\delta^2/(\varepsilon^2\sigma^2)$ implies that*

$$\mathcal{N}(p, \delta, \mathcal{S}, \varepsilon G) \Longrightarrow \mathcal{N}(p/4, 4\delta, \mathcal{S}, \varepsilon(2m)^{-1/2}\widetilde{Z_m^\bullet}),$$

*for $m \geq c\varepsilon^4\sigma^4 N\, \Lambda_4^\bullet/(p\delta^4)$.*

*$(ii)$ the statement of item $(i)$ remains true with replacing $\bullet$ by $\diamond$.*

*Proof.* Statement $(ii)$ is proved in BG (1997a, Lemma 2.5)). We repeat this proof concerning item $(i)$.

Statement $(i)$ follows from relations (2.10)–(2.11) which are written out below, since $p$, $\delta$, $\varepsilon$ in these relations are arbitrary and $\mathbf{E}\,\|\widetilde{X^\bullet}\|^4 \leq 16\,\mathbf{E}\,\|X^\bullet\|^4 = 16\,N\sigma^4\Lambda_4^\bullet$.

For the Gaussian random vector $G_*$ defined in Lemma 2.3, we have

$$\mathcal{N}(2p, \delta, \mathcal{S}, \varepsilon G) \Longrightarrow \mathcal{N}(p, 2\delta, \mathcal{S}, \varepsilon G_*), \qquad \text{provided that } \Pi_2^\bullet \leq p\delta^2/(2\varepsilon^2\sigma^2). \tag{2.10}$$

If $m \geq c\varepsilon^4\mathbf{E}\,\|\widetilde{X^\bullet}\|^4/(p\delta^4)$ with a sufficiently large absolute constant $c$, then

$$\mathcal{N}(2p, \delta, \mathcal{S}, \varepsilon G_*) \Longrightarrow \mathcal{N}(p, 2\delta, \mathcal{S}, \varepsilon\,(2m)^{-1/2}\,\widetilde{Z_m^\bullet}). \tag{2.11}$$

Let us prove (2.10). For $e \in \mathbb{R}^d$ define $p_e$ by $2p_e = \mathbf{P}\big\{ \|\varepsilon G - e\| < \delta \big\}$. Assuming that $\Pi_2^\bullet \leq p_e\delta^2/(2\varepsilon^2\sigma^2)$, it suffices to prove that $\mathbf{P}\big\{ \|\varepsilon G_* - e\| < 2\delta \big\} \geq p_e$. Replacing $\delta$ by $\delta/\varepsilon$ and $e$ by $e/\varepsilon$, we see that we can assume that $\varepsilon = 1$. Applying Lemma 2.3, we obtain

$$\begin{aligned} \mathbf{P}\big\{ \|G_* - e\| < 2\delta \big\} &\geq \mathbf{P}\big\{ \|W\| + \|G - e\| < 2\delta \big\} \\ &\geq \mathbf{P}\big\{ \|W\| < \delta,\ \|W\| + \|G - e\| < 2\delta \big\} \\ &\geq \mathbf{P}\big\{ \|W\| < \delta \text{ and } \|G - e\| < \delta \big\} \\ &\geq 2p_e - \mathbf{P}\big\{ \|W\| \geq \delta \big\} \\ &\geq 2p_e - \delta^{-2}\mathbf{E}\,\|W\|^2 \geq 2p_e - 2\delta^{-2}\sigma^2\Pi_2^\bullet \geq p_e. \end{aligned} \tag{2.12}$$

This implies (2.10).

Let us prove (2.11). Notice that $\operatorname{cov}\big(\varepsilon\widetilde{X^\bullet}/\sqrt{2}\big) = \operatorname{cov}(\varepsilon G_*)$. Therefore, to prove (2.11) it suffices to apply Lemma 5.3 of BG (1997a), replacing in that lemma $X$ by $\varepsilon\widetilde{X^\bullet}/\sqrt{2}$. $\square$

*Proof* (1.32) $\Longrightarrow$ (2.1). By a well known truncation argument, we have

$$\big|\mathbf{P}\big\{ Z_N \in A\big\} - \mathbf{P}\big\{ Z_N^{\diamond} \in A\big\}\big| \le N\,\mathbf{P}\big\{ \|X\| > \sigma\sqrt{N}\,\big\} \le \Pi_2^{\diamond}, \tag{2.13}$$

for any measurable set $A$, and

$$Q(Z_N, \lambda) \le \Pi_2^{\diamond} + Q(Z_N^{\diamond}, \lambda). \tag{2.14}$$

Recall that we are proving (2.1) assuming that $s = 5$ and $\delta = 1/200$. Hence, $4\delta = 1/50 < 1/(5s)$. Write $K = \varepsilon/\sqrt{2}$ with $\varepsilon = c_0/\sigma$. Then, by Lemma 2.6, we have

$$\mathcal{N}(p, \delta, \mathcal{S}_o, \varepsilon G) \Longrightarrow \mathcal{N}(p/4, 4\delta, \mathcal{S}_o, m^{-1/2}\, K\,\widetilde{Z_m^{\diamond}}), \tag{2.15}$$

provided that

$$\Pi_2^{\diamond} \le c_1 p, \qquad m \ge cN\,\Lambda_4^{\diamond}/p. \tag{2.16}$$

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

The non-degeneracy condition (2.15) for $K\,\widetilde{Z_m^{\diamond}}$ allows to apply (1.32) of Theorem 1.5, and we obtain

$$Q(Z_N^{\diamond}, \lambda) = Q(K\,Z_N^{\diamond}, K^2\lambda) \ll (pN)^{-1} \max\{m;\ K^2\lambda\}, \tag{2.17}$$

for any $m$ such that (2.16) is fulfilled. Choosing the minimal $m$ in (2.16), we obtain

$$Q(Z_N^{\diamond}, \lambda) \ll p^{-2} \max\{\Lambda_4^{\diamond};\ \lambda/(\sigma^2 N)\}. \tag{2.18}$$

Combining the estimates (2.14) and (2.18), we conclude the proof. $\square$

*Proof* (2.1) $\Longrightarrow$ (2.2). Note that the bound (2.1) holds with a probability $p$ of condition $\mathcal{N}(p, \delta, \mathcal{S}_o, c_0 G/\sigma)$. Let us choose $4c_0 = \delta = 1/200$. Then, using Lemma 2.5 and the assumption $\mathcal{B}(\mathcal{S}_o, \mathbb{C})$, the effective lower bound $p \ge \exp\big\{c\sigma^2/\lambda_5^2\big\}$ follows. $\square$

## 3. Auxiliary lemmas

In Sections 3 and 4 we shall prove Theorem 1.4. Therefore, we shall assume that its conditions are satisfied. We consider the case $d < \infty$ assuming that the following conditions are fulfilled

$$\mathbb{Q}^2 = \mathbb{I}_d, \quad \sigma^2 = 1, \quad s = 5, \quad \delta = 1/300, \quad b = \sqrt{N}\,a, \quad \mathcal{N}(p, \delta, \mathcal{S}_o, c_0 G). \tag{3.1}$$

Notice that the assumption $\sigma^2 = 1$ does not restrict generality since from Theorem 1.4 with $\sigma^2 = 1$ we can derive the general result replacing $X$, $G$ by $X/\sigma$, $G/\sigma$, etc. Other assumptions in (3.1) are included as conditions in Theorem 1.4. Section 3 is devoted to some auxiliary lemmas which are similar to corresponding lemmas of BG (1997a).

In several places, the proof of Theorem 1.4 repeats almost literally the proof of Theorem 1.5 in BG (1997a). Note, however, that we shall use truncated vectors $X_j^{\bullet}$, while in BG (1997a) the vectors $X_j^{\diamond}$ were involved. We start with an application of the Fourier transform to the functions $\Psi_b$ and $\Phi_b$, where $b = \sqrt{N}\,a$. We shall estimate integrals over the Fourier transforms using results of Sections 2, 5–7 and some technical lemmas of

BG (1997a). We shall apply as well some methods of estimation of the rate of approximation in the CLT in multidimensional spaces (cf., e.g., Bhattacharya and Rao (1986)).

We shall use the following formulas for the Fourier inversion (see BG (1997a)). A smoothing inequality of Prawitz (1972) implies (see BG (1996, Section 4)) that

$$F(x) = \frac{1}{2} + \frac{i}{2\pi}\ \mathrm{V.P.}\int_{|t|\le K} \mathrm{e}\big\{-xt\big\}\widehat{F}(t)\ \frac{dt}{t} + R, \tag{3.2}$$

for any $K > 0$ and any distribution function $F$ with characteristic function $\widehat{F}$ (see (1.13)), where

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

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.

For any function $F:\mathbb{R}\to\mathbb{R}$ of bounded variation such that $F(-\infty)=0$ and $2F(x) = F(x+)+F(x-)$, for all $x\in\mathbb{R}$, the following Fourier–Stieltjes inversion formula is valid (see, e.g., Chung (1974))

$$F(x) = \frac{1}{2}\,F(\infty) + \frac{i}{2\pi}\lim_{M\to\infty}\mathrm{V.P.}\int_{|t|\le M} \mathrm{e}\{-xt\}\,\widehat{F}(t)\,\frac{dt}{t}\,. \tag{3.3}$$

This formula is well-known for distribution functions. For functions of bounded variation, it may be justified by linearity arguments.

Recall that the random vectors $X^\bullet$, $X'$ are defined in (1.7) and (1.12) and $Z_N^\bullet$, $Z_N'$ are sums of $N$ their independent copies. Write $\Psi_b^\bullet$ and $\Psi_b'$ for the distribution function of $\mathbb{Q}[Z_N^\bullet - b]$ and $\mathbb{Q}[Z_N' - b]$ respectively. For $0\le k\le N$ introduce the the distribution function

$$\Psi_b^{(k)}(x) = \mathbf{P}\big\{\mathbb{Q}\big[G_1+\dots+G_k+X_{k+1}'+\dots+X_N'-b\big]\le x\big\}. \tag{3.4}$$

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 $X_j^\bullet$ instead of $X_j^\diamond$.

**Lemma 3.1.** *Let $\Pi_2^\bullet \le c_1 p$ and let an integer $1\le m\le N$ satisfy $m\ge c_2 N\Lambda_4^\bullet/p$, with some sufficiently small (resp. large) positive absolute constant $c_1$ (resp. $c_2$). Let $c_3$ be an absolute constant. Write*

$$K = c_0^2/(2m), \qquad\qquad t_1 = c_3\,(pN/m)^{-1/2}.$$

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

$$F(x) = \frac{1}{2} + \frac{i}{2\pi}\ \mathrm{V.P.}\int_{|t|\le t_1} \mathrm{e}\{-xtK\}\,\widehat{F}(tK)\ \frac{dt}{t} + R_1, \tag{3.5}$$

*with $|R_1| \ll (pN)^{-1}m$.*

*Proof.* Let us prove (3.5). We shall combine (3.2) and Lemma 2.2. Changing the variable $t=\tau K$ in the approximate Fourier inversion formula (3.2), we obtain

$$F(x)=\frac{1}{2}+\frac{i}{2\pi}\,\mathrm{V.P.}\int_{|t|\le 1}\mathrm{e}\{-xtK\}\,\widehat{F}(tK)\,\frac{dt}{t}+R, \tag{3.6}$$

where

$$|R|\le\int_{|t|\le 1}\big|\widehat{F}(tK)\big|\,dt. \tag{3.7}$$

Notice that $\Psi^\bullet$, $\Psi'$, $\Psi^{(k)}$, $\Phi_b$ are distribution functions of random variables which may be written in the form:

$$\mathbb{Q}[V+T],\qquad V\overset{\mathrm{def}}{=}G_1+\dots+G_k+X^\bullet_{k+1}+\dots+X^\bullet_N,$$

with some $k$, $0\le k\le N$, and some random vector $T$ which is independent of $X^\bullet_j$ 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 $K^{1/2}X^\bullet$. Let $Y_1,Y_2,\dots$ be independent copies of $Y$. Then we can write

$$K^{1/2}V\overset{\mathcal{D}}{=}Y_1+\dots+Y_l+T_1 \tag{3.8}$$

with $l=\lceil N/(2m)\rceil$ and some random $T_1$ independent of $Y_1,\dots,Y_l$. By Lemma 2.6, we have

$$\mathcal{N}(p,\delta,\mathcal{S},c_0G)\Longrightarrow\mathcal{N}(p/4,4\delta,\mathcal{S},\widetilde{Y}) \tag{3.9}$$

provided that

$$\Pi^\bullet_2\le c_1p\qquad\text{and}\qquad m\ge c_2N\Lambda^\bullet_4/p. \tag{3.10}$$

The inequalities in (3.10) are just conditions of Lemma 3.1. Due to (3.1), (3.8) and (3.9), we can apply Lemma 2.2 in order to estimate the integrals in (3.6) and (3.7). Replacing in Lemma 2.2 $X$ by $Y$ and $N$ by $l$, we obtain (3.5) in the case $k<N/2$.

*The case* $k\ge N/2$. We can argue as in the previous case defining now $Y$ as a sum of $m$ independent copies of $K^{1/2}G$. Condition $\mathcal{N}(p/4,4\delta,\mathcal{S}_o,\widetilde{Y})$ is satisfied by (3.1), since now $\mathcal{L}(\widetilde{Y})=\mathcal{L}(c_0G)$. $\square$

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

$$\varkappa(t;N,\mathcal{L}(X))=\sup_{x\in\mathbb{R}^d}\big|\mathbf{E}\,\mathrm{e}\big\{t\mathbb{Q}[Z_j]+\langle x,Z_j\rangle\big\}\big|,\qquad Z_j=X_1+\dots+X_j, \tag{3.11}$$

with $j=\big\lceil(N-2)/14\big\rceil$. In the sequel, we shall use that

$$\varkappa(t;N,\mathcal{L}(X'),\mathcal{L}(G))\le\varkappa(t;N,\mathcal{L}(X^\bullet),\mathcal{L}(G)). \tag{3.12}$$

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

**Lemma 3.2.** *Assume the conditions of Lemma* 3.1. *Then*

$$\int_{|t|\le t_1} \big(|t|K\big)^\alpha \varkappa\big(tK; N, \mathcal{L}(X^\bullet), \mathcal{L}(G)\big)\ \frac{dt}{|t|} \ll_\alpha \begin{cases} (Np)^{-\alpha}, & \textit{for } 0\le\alpha<s/2,\\ (Np)^{-\alpha}\big(1+|\log(Np/m)|\big), & \textit{for } \alpha=s/2,\\ (Np)^{-\alpha}\big(1+(Np/m)^{(2\alpha-s)/4}\big), & \textit{for } \alpha>s/2. \end{cases} \tag{3.13}$$

Lemma 3.2 is a generalization of Lemma 3.2 from BG (1997a) which contains the same bound for $0\le\alpha<s/2$. In this paper, we have to estimate the left hand side of (3.13) in the case $s/2\le\alpha$ too.

*Proof.* By (3.1) and (3.9), the condition $\mathcal{N}(p/4, 4\delta, \mathcal{S}_o, K^{1/2}\widetilde{Z_m^\bullet})$ is fulfilled. Therefore, collecting independent copies of $K^{1/2}X^\bullet$ in groups as in (3.8), we can apply Theorem 5.2. By this theorem,

$$\varkappa(tK; N, \mathcal{L}(X^\bullet)) \ll \mathcal{M}^s(t; pN/m).$$

A similar upper bound is valid for $\varkappa(tK; N, \mathcal{L}(G))$ (cf. the proof of (3.5) in the case $k>N/2$). Using definition (1.33) of the function $\mathcal{M}(\cdot,\cdot)$ and (1.34), we get

$$\varkappa(tK; N, \mathcal{L}(X^\bullet), \mathcal{L}(G)) \ll_s \min\Big\{1;\ \big(m/(tpN)\big)^{s/2}\Big\}, \qquad \text{for} \quad |t|\le t_1.$$

Integrating this bound (cf. the estimation of $I_1$ in Lemma 2.2), we obtain (3.13). □

## 4. Proof of Theorem 1.4

To simplify notation, in Section 4 we write $\Pi=\Pi_2^\bullet$ and $\Lambda=\Lambda_4^\bullet$. The assumption $\sigma^2=1$ and relations (1.9) and (1.11) imply

$$\Pi+\Lambda\,N\gg 1, \qquad \Pi+\Lambda\le 1, \qquad \sigma_j^2\le 1, \qquad \lambda_j^2\le 1. \tag{4.1}$$

Recall that $\Delta_N^{(a)}$ and functions $\Psi_b$, $\Phi_b$ and $\Theta_b$ are defined by (1.16) and (1.19)–(1.21). Note now that $\Theta_b^\bullet(x)=E_a^\bullet(x/N)$ and, according to (1.21),

$$\Delta_N^{(a)} \le \Delta_{N,\bullet}^{(a)} + \sup_{x\in\mathbb{R}} \big|\Theta_b(x)-\Theta_b^\bullet(x)\big|, \tag{4.2}$$

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

$$\Delta_{N,\bullet}^{(a)} = \sup_{x\in\mathbb{R}} \big|\Psi_b(x)-\Phi_b(x)-\Theta_b^\bullet(x)\big|. \tag{4.3}$$

Let us verify that

$$\sup_{x\in\mathbb{R}} \big|\Theta_b(x)-\Theta_b^\bullet(x)\big| \ll_d \sigma_d^{-3}\,\Pi_3^\bullet. \tag{4.4}$$

To this end we shall apply representation (1.16)–(1.17) of the Edgeworth correction as a signed measure and estimate the variation of that measure. Indeed, using (1.16), we have

$$\sup_{x\in\mathbb{R}} \big|\Theta_b(x)-\Theta_b^\bullet(x)\big| \ll N^{-1/2} I, \qquad I \overset{\text{def}}{=} \int_{\mathbb{R}^d} \big|\mathbf{E}\,p'''(x)X^3 - \mathbf{E}\,p'''(x)X^{\bullet 3}\big|\,dx. \tag{4.5}$$

By the explicit formula (1.18), the function $u \mapsto p'''(x)u^3$ is a 3-linear form in the variable $u$. Therefore, using $\|\mathbb{C}^{-1/2}u\| \le \sigma_d^{-1}\|u\|$, $X = X^\bullet + X_\bullet$ and $\|X^\bullet\|\,\|X_\bullet\| = 0$, we have $p'''(x)X^3 - p'''(x)X^{\bullet 3} = p'''(x)X_\bullet^3$, and

$$N^{-1/2} I \le 3\Pi_3^\bullet \sigma_d^{-3} \int_{\mathbb{R}^d} \big(\|\mathbb{C}^{-1/2}x\| + \|\mathbb{C}^{-1/2}x\|^3\big)\, p(x)\, dx = c_d \Pi_3^\bullet \sigma_d^{-3}. \tag{4.6}$$

Inequalities (4.5) and (4.6) imply now (4.4).

To prove the statement $(i)$ of Theorem 1.4, we have to derive that

$$\Delta_{N,\bullet}^{(a)} \ll_d p^{-3}\, \sigma_d^{-2}\, (\Pi + \Lambda\, \sigma_d^{-2})(1 + \|a\|)^3. \tag{4.7}$$

While proving (4.7) we assume that

$$\Pi \le c_d p\, \sigma_d^2, \qquad\qquad \Lambda \le c_d p\, \sigma_d^4, \tag{4.8}$$

with a sufficiently small positive constant $c_d$ depending on $d$ only. These assumptions do not restrict generality. Indeed, we have $\big|\Psi_b(x) - \Phi_b(x)\big| \le 1$. If conditions (4.8) do not hold, then the estimate

$$\sup_{x\in\mathbb{R}} \big|\Theta_b^\bullet(x)\big| \ll_d N^{-1/2} \mathbf{E}\, \|\mathbb{C}^{-1/2} X^\bullet\|^3 \ll_d \sigma_d^{-2} \Lambda^{1/2} \tag{4.9}$$

immediately implies (4.7). In order to prove (4.9) we can use representation (1.16)–(1.17) of the Edgeworth correction. Estimating the variation of that measure and using

$$\beta_3^2 \le \sigma^2 \beta, \qquad \|\mathbb{C}^{-1/2}u\| \le \sigma_d^{-1}\|u\|, \qquad \mathbf{E}\, \|\mathbb{C}^{-1/2} X^\bullet\|^2 \le \mathbf{E}\, \|\mathbb{C}^{-1/2} X\|^2 = d, \tag{4.10}$$

we obtain (4.9).

It is clear that

$$\Delta_{N,\bullet}^{(a)} \le \sup_{x\in\mathbb{R}} \Big( \big|\Psi_b(x) - \Psi_b'(x)\big| + \big|\Theta_b^\bullet(x) - \Theta_b'(x)\big| + \big|\Psi_b'(x) - \Phi_b(x) - \Theta_b'(x)\big| \Big). \tag{4.11}$$

Similarly to (4.5), we have

$$\sup_{x\in\mathbb{R}} \big|\Theta_b^\bullet(x) - \Theta_b'(x)\big| \ll N^{-1/2} J, \qquad J \stackrel{\text{def}}{=} \int_{\mathbb{R}^d} \big|\mathbf{E}\, p'''(x) X^{\bullet 3} - \mathbf{E}\, p'''(x) X'^3\big|\, dx. \tag{4.12}$$

Recall that vector $X'$ is defined in (1.12). By Lemma 2.3, we have $\mathbf{E}\,\|W\|^2 \le 2\Pi$ (hence, $\mathbf{E}\,\|W\|^q \ll \Pi^{q/2}$, for $0 \le q \le 2$). Moreover, representing $W$ as a sum of a large number of i.i.d. Gaussian summands and using the Rosenthal inequality (see BG (1997a, inequality (1.24)), it is easy to see that

$$\mathbf{E}\,\|W\|^q \ll_q \big(\mathbf{E}\,\|W\|^2\big)^{q/2} \ll_q \Pi^{q/2}, \quad q \ge 0. \tag{4.13}$$

Furthermore, according to (1.9), (1.11) and (4.8),

$$\mathbf{E}\,\|X_\bullet\| \le \Pi\, \sigma_d^{-1} N^{-1/2} \ll \Pi^{1/2} N^{-1/2}. \tag{4.14}$$

Hence, by (1.9), (1.12), (4.13) and (4.14),

$$\overline{\beta} \stackrel{\text{def}}{=} \mathbf{E}\,\|X'\|^4 \ll N\Lambda + \Pi^2. \tag{4.15}$$

Using (1.18), (4.8), (4.10) and (4.12)–(4.14), we get

$$\begin{aligned} N^{-1/2} J &\ll \Pi^{1/2}(N^{-1/2}\Pi+\Lambda^{1/2})\sigma_d^{-3}\int_{\mathbb{R}^d}\big(\|\mathbb{C}^{-1/2}x\|+\|\mathbb{C}^{-1/2}x\|^3\big)\,p(x)\,dx \\ &\ll_d (\Pi+\Lambda\sigma_d^{-2})\sigma_d^{-2}. \end{aligned} \tag{4.16}$$

Thus, according to (4.12) and (4.16),

$$\sup_{x\in\mathbb{R}}\big|\Theta_b^\bullet(x)-\Theta_b'(x)\big| \ll_d (\Pi+\Lambda\sigma_d^{-2})\sigma_d^{-2}. \tag{4.17}$$

The same approach is applicable for the estimation of $\big|\Theta_b'\big|$. Using (1.12), (1.16)–(1.18), (4.10) and (4.13), we get

$$\begin{aligned} \sup_{x\in\mathbb{R}}\big|\Theta_b'(x)\big| &\ll N^{-1/2}\int_{\mathbb{R}^d}\big|\mathbf{E}\,p'''(x)X'^3\big|\,dx \\ &\ll_d \Lambda^{1/2}\sigma_d^{-2}+N^{-1/2}\Pi^{3/2}\sigma_d^{-3}. \end{aligned} \tag{4.18}$$

Let us prove that

$$\sup_{x\in\mathbb{R}}\big|\Psi_b(x)-\Psi_b'(x)\big| \ll p^{-2}(\Pi\sigma_d^{-2}+\Lambda)(1+\|a\|^2). \tag{4.19}$$

Using truncation (see (1.11) and (2.13)), we have $|\Psi_b-\Psi_b^\bullet|\le\Pi\sigma_d^{-2}$, and

$$\sup_{x\in\mathbb{R}}\big|\Psi_b(x)-\Psi_b'(x)\big| \le \Pi\sigma_d^{-2}+\sup_{x\in\mathbb{R}}\big|\Psi_b^\bullet(x)-\Psi_b'(x)\big|. \tag{4.20}$$

In order to estimate $|\Psi_b^\bullet-\Psi_b'|$, we shall apply Lemmas 3.1 and 3.2. The number $m$ in these Lemmas exists and $N\Lambda/p\gg 1$, as it follows from (4.1) and (4.8). Let us choose the minimal $m$, that is, $m\asymp N\Lambda/p$. Then $(pN)^{-1}m\ll\Lambda/p^2$ and $m/N\ll\Lambda/p$. Therefore, using Lemma 3.1, we have

$$\sup_x\big|\Psi_b^\bullet(x)-\Psi_b'(x)\big| \ll p^{-2}\Lambda+\int_{|t|\le t_1}\big|\widehat{\Psi}_b^\bullet(\tau)-\widehat{\Psi}_b'(\tau)\big|\,\frac{dt}{|t|}\,,\qquad \tau=tK. \tag{4.21}$$

We shall prove that

$$\big|\widehat{\Psi}_b^\bullet(\tau)-\widehat{\Psi}_b'(\tau)\big| \ll \varkappa\Pi\sigma_d^{-1}\,|\tau|\,N\,(1+|\tau|\,N)(1+\|a\|^2), \tag{4.22}$$

with $\varkappa=\varkappa(\tau;N,\mathcal{L}(X^\bullet))$. Combining (4.20)–(4.22), using $\tau=tK$ and integrating inequality (4.22) with the help of Lemma 3.2, we derive (4.19).

Let us prove (4.22). Recall that $X'=X^\bullet-\mathbf{E}\,X^\bullet+W$, where $W$ denotes a centered Gaussian random vector which is independent of all other random vectors and such that $\operatorname{cov}X'=\operatorname{cov}G$ (see Lemma 2.3). Writing $D=Z_N^\bullet-\mathbf{E}\,Z_N^\bullet-b$, we have

$$Z_N^\bullet-b=D+\mathbf{E}\,Z_N^\bullet,\qquad\qquad Z_N'-b\overset{\mathcal{D}}{=}D+\sqrt{N}\,W,$$

and

$$\big|\widehat{\Psi}_b^\bullet(\tau)-\widehat{\Psi}_b'(\tau)\big| \le \big|f_1(\tau)\big|+\big|f_2(\tau)\big| \tag{4.23}$$

with

$$f_1(\tau) = \mathbf{E}\, \mathrm{e}\big\{\tau \mathbb{Q}[D + \sqrt{N} W]\big\} - \mathbf{E}\, \mathrm{e}\big\{\tau \mathbb{Q}[D]\big\},$$
$$f_2(t) = \mathbf{E}\, \mathrm{e}\big\{\tau \mathbb{Q}[D + \mathbf{E}\, Z_N^\bullet]\big\} - \mathbf{E}\, \mathrm{e}\big\{\tau \mathbb{Q}[D]\big\}. \tag{4.24}$$

Now we have to prove that both $\big|f_1(\tau)\big|$ and $\big|f_2(\tau)\big|$ may be estimated by the right hand side of (4.22).

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 $B = N\mathbb{Q}[W]\big\}$. Taylor's expansions of the exponent in (4.24) in powers of $i\tau B$ and $i\tau A$ with remainders $\mathcal{O}(\tau B)$ and $\mathcal{O}(\tau^2 A^2)$ respectively imply (recall that $\mathbf{E}\, W = 0$ and $\mathbb{Q}^2 = \mathbb{I}_d$)

$$\big|f_1(\tau)\big| \ll \varkappa |\tau| N \mathbf{E}\, \|W\|^2 + \varkappa \tau^2 N \mathbf{E}\, \|W\|^2\, \mathbf{E}\, \|D\|^2, \tag{4.25}$$

where $\varkappa = \varkappa(\tau; N, \mathcal{L}(X^\bullet))$. 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 $\sigma^2 = 1$, $\mathbf{E}\, \|W\|^2 \ll \Pi$ and $\mathbf{E}\, \|D\|^2 \ll N(1 + \|a\|^2)$, we derive from (4.25) that

$$\big|f_1(\tau)\big| \ll \varkappa \Pi |\tau| N \big(1 + |\tau| N\big)\big)(1 + \|a\|^2). \tag{4.26}$$

Note that $\mathbf{E}\, Z_N^\bullet = N\, \mathbf{E}\, X^\bullet = -N\, \mathbf{E}\, X_\bullet$. Expanding the exponent $\mathrm{e}\big\{\tau \mathbb{Q}[D + \mathbf{E}\, Z_N^\bullet]\big\}$, using (4.14) and proceeding similarly to the proof of (4.26), we obtain

$$\big|f_2(\tau)\big| \ll \varkappa \Pi \sigma_d^{-1} |\tau| N (1 + \|a\|). \tag{4.27}$$

Inequalities (4.23), (4.26) and (4.27) imply now (4.22).

It remains to estimate $\big|\Psi_b' - \Phi_b - \Theta_b'\big|$. Recall that the distribution functions $\Psi_b^{(l)}(x)$, for $0 \le l \le N$, are defined in (3.4).

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

$$\sup_{x \in \mathbb{R}} \big|\Psi_b'(x) - \Phi_b(x) - \Theta_b'(x)\big| \le I_1 + I_2 + I_3, \tag{4.28}$$

where

$$I_1 = \sup_{x \in \mathbb{R}} \big|\Psi_b^{(k)}(x) - \Phi_b(x) - (N - k)\, \Theta_b'(x)/N\big|, \tag{4.29}$$

$$I_2 = \sup_{x \in \mathbb{R}} \big|\Psi_b'(x) - \Psi_b^{(k)}(x)\big|, \tag{4.30}$$

and

$$I_3 = \sup_{x \in \mathbb{R}} \ k N^{-1} \big|\Theta_b'(x)\big|. \tag{4.31}$$

Let estimate $I_1$. Define the distributions

$$\mu(A) = \mathbf{P}\Big\{U_k + \sum_{j=k+1}^{N} X_j' \in \sqrt{N} A\Big\}, \qquad \mu_0(A) = \mathbf{P}\big\{U_N \in \sqrt{N} A\big\} = \mathbf{P}\big\{G \in A\big\}, \tag{4.32}$$

where $U_l = G_1 + \cdots + G_l$. Introduce the measure $\chi'$ replacing $X$ by $X'$ in (1.17). For the Borel sets $A \subset \mathbb{R}^d$ define the Edgeworth correction (to the distribution $\mu$) as

$$\mu_1^{(k)}(A) = (N-k)\, N^{-3/2} \chi'(A)/6. \tag{4.33}$$

Introduce the signed measure

$$\nu = \mu - \mu_0 - \mu_1^{(k)}. \tag{4.34}$$

It is easy to see that a re-normalization of random vectors implies (see relations (1.16), (1.19)–(1.21), (3.4) and (4.32)–(4.34))

$$\begin{aligned} \bigl|\Psi_b^{(k)}(x) - \Phi_b(x) - (N-k)\,\Theta_b'(x)/N\bigr| &= \nu\bigl(\bigl\{u \in \mathbb{R}^d : \mathbb{Q}[u-a] \le x/N\bigr\}\bigr) \\ &\le \delta_N \overset{\text{def}}{=} \sup_{A \subset \mathbb{R}^d} \bigl|\nu(A)\bigr|. \end{aligned} \tag{4.35}$$

**Lemma 4.1.** *Assume that $d < \infty$ and $1 \le k \le N$. Then there exists $c_d$ such that $\delta_N$ defined in* (4.35) *satisfies the inequality*

$$\delta_N \ll_d \frac{\overline{\beta}}{\sigma_d^4 N} + \frac{N^{d/2}}{k^{d/2}} \exp\{-c_d\, k\, \sigma_d^4/\overline{\beta}\} \tag{4.36}$$

*with $\overline{\beta} = \mathbf{E}\,\|X'\|^4$.*

*An outline of the proof.* We repeat and slightly improve the proof of Lemma 9.3 in BG (1997a) (cf. the proof of Lemma 2.5 in BG (1996)). Assuming that $\operatorname{cov} X = \operatorname{cov} X' = \operatorname{cov} G = \mathbb{I}_d$, we shall prove that

$$\delta_N \ll_d \frac{\overline{\beta}}{N} + \frac{N^{d/2}}{k^{d/2}} \exp\{-c_d\, k/\overline{\beta}\}. \tag{4.37}$$

Applying (4.37) to $\mathbb{C}^{-1/2} X'$ and $\mathbb{C}^{-1/2} G$ and estimating $\|\mathbb{C}^{-1/2}\| \le 1/\sigma_d$, we obtain (4.36).

While proving (4.37) we can assume that $\overline{\beta}/N \le c_d$ and $N \ge 1/c_d$ with a sufficiently small positive constant $c_d$. Otherwise (4.37) follows from the trivial 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

$$g(t) = \mathbf{E}\ \mathrm{e}\bigl\{\langle N^{-1/2} t, G\rangle\bigr\}, \quad h(t) = \mathbf{E}\ \mathrm{e}\bigl\{\langle N^{-1/2} t, X'\rangle\bigr\}, \tag{4.38}$$

$$f(t) = \mathbf{E}\ \mathrm{e}\bigl\{\langle N^{-1/2} t, Z_n'\rangle\bigr\} = h^n(t), \qquad f_0(t) = \mathbf{E}\ \mathrm{e}\bigl\{\langle N^{-1/2} t, U_n'\rangle\bigr\} = g^n(t), \tag{4.39}$$

$$f_1(t) = n\, m(t)\, f_0(t), \quad \text{where} \quad m(t) = \frac{1}{6\, N^{3/2}}\, \mathbf{E}\,\langle it, X'\rangle^3, \tag{4.40}$$

$$\widehat{\nu}(t) = (f(t) - f_0(t) - f_1(t))\, g(\rho t), \quad \rho^2 = k. \tag{4.41}$$

It is easy to see that

$$\widehat{\nu}(t) = \int_{\mathbb{R}^d} \mathrm{e}\{\langle t, x\rangle\}\, \nu(dx). \tag{4.42}$$

As a consequence of the truncation we have

$$\mathbf{E}\,\big\|\,Z_l'/\sqrt{N}\,\big\|^\gamma \ll_{\gamma,d} 1, \qquad \gamma>0, \quad 1\le l\le N. \tag{4.43}$$

By a slight extension of the proof of Lemma 11.6 in Bhattacharya and Rao (1986), see as well the proof of Lemma 2.5 in BG (1996), we obtain

$$\delta_N \ll_d \max_{|\alpha|\le 2d} \int_{t\in\mathbb{R}^d} \big|\,\partial^\alpha \widehat{\nu}(t)\,\big|\,dt. \tag{4.44}$$

Here $|\alpha| = |\alpha_1| + \cdots + |\alpha_d|$, $\alpha = (\alpha_1,\dots,\alpha_d)$, $\alpha_j \in \mathbb{Z}$, $\alpha_j \ge 0$. In order to derive (4.37) from (4.44), it suffices to prove that, for $|\alpha| \le 2d$,

$$\big|\,\partial^\alpha \widehat{\nu}(t)\,\big| \ll_d g(c_1\rho t), \tag{4.45}$$

$$\big|\,\partial^\alpha \widehat{\nu}(t)\,\big| \ll_d \overline{\beta}\,N^{-1}\,(1+\|t\|^6)\,\exp\{-c_2\,\|t\|^2\}, \quad \text{for } \|t\|^2 \le c_3(d)\,N/\overline{\beta}. \tag{4.46}$$

Indeed, using (4.45) and denoting $T = \sqrt{c_3(d)\,N/\overline{\beta}}$, we obtain

$$\int\limits_{\|t\|\ge T} \big|\,\partial^\alpha \widehat{\nu}(t)\,\big|\,dt \ll_d \int\limits_{\|t\|\ge T} g(c_1\rho t)\,dt \ll_d \frac{N^{d/2}}{\rho^d}\ \exp\Big\{-\frac{c_1^2\rho^2T^2}{8N}\Big\}\int\limits_{\mathbb{R}^d}\exp\{-c_1^2\,\|t\|^2/8\}\,dt, \tag{4.47}$$

and it is easy to see that the right hand side of (4.47) is bounded from above by the second summand in the right hand side of (4.37). Similarly, using (4.46), we can integrate $\big|\,\partial^\alpha \widehat{\nu}(t)\,\big|$ over $\|t\| \le T$, and the integral is bounded from above by $c_d\,\overline{\beta}/N$.

In the proof of (4.45)–(4.47) we applied standard methods of estimation which are provided in Bhattacharya and Rao (1986). In particular, we used a Bergström type identity

$$f - f_0 - f_1 = \sum_{j=0}^{n-1}(h-g-m)\,h^j\,g^{n-j-1} + \sum_{j=0}^{n-1} m \sum_{l=0}^{j-1}(h-g)\,h^l\,g^{n-l-1}, \tag{4.48}$$

relations (4.38)–(4.43), $1\le k\le N$, $\big|\,\partial^\alpha \exp\{-c_4\,\|t\|^2\}\,\big| \ll_\alpha \exp\{-c_5\,\|t\|^2\}$, $\sqrt{N}/\overline{\beta}^{1/2} \gg_d 1$ and $y^{c_d}\exp\{-y\} \ll_d 1$, for $y>0$. □

Applying (4.29), (4.35) and Lemma 4.1, we get

$$I_1 \ll_d \frac{\overline{\beta}}{\sigma_d^4 N} + \frac{N^{d/2}}{k^{d/2}}\ \exp\{-c_d\,k\,\sigma_d^4/\overline{\beta}\}. \tag{4.49}$$

For the estimation of $I_2$ we shall use Lemma 4.2 which is an easy consequence of BG (1997a, Lemma 9.3) and (3.12).

**Lemma 4.2.** *We have*

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

*where* $\varkappa = \varkappa(t; N, \mathcal{L}(X^\bullet), \mathcal{L}(G))$ *(cf.* (3.11)*).*

As in the proof of (4.21), applying Lemma 3.1 (choosing $m \asymp N(\Lambda+\Pi)/p$), we obtain

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

Applying Lemma 4.2 and replacing in that Lemma $t$ by $\tau$, we have

$$\big|\,\widehat{\Psi}_b'(\tau) - \widehat{\Psi}_b^{(k)}(\tau)\,\big| \ll \varkappa\tau^2 k\,\big(\overline{\beta} + |\tau|\,N\,\overline{\beta} + |\tau|\,N\sqrt{N\,\overline{\beta}}\big)(1+\|a\|^3). \tag{4.50}$$

Integrating with the help of Lemma 3.2, we obtain

$$I_2 \ll p^{-2}(\Pi+\Lambda) + p^{-3}kN^{-2}\big(\overline{\beta} + \sqrt{N\,\overline{\beta}}\big)\big(1 + p^{1/2}/(\Pi+\Lambda)^{1/4}\big)(1+\|a\|^3). \tag{4.51}$$

Let us choose $k \asymp_d \sigma_d^{-3} N^{1/4}\,\overline{\beta}^{3/4}$. Such $k \le N$ exists by $\overline{\beta} \gg \sigma^4 = 1$, by (4.15) and by assumption (4.8). Then (4.49) and (4.51) turn into

$$I_1 \ll_d \frac{\overline{\beta}}{\sigma_d^4 N} + \Big(\frac{\sigma_d^4 N}{\overline{\beta}}\Big)^{3d/8} \exp\Big\{-c_d\,\Big(\frac{\sigma_d^4 N}{\overline{\beta}}\Big)^{1/4}\Big\} \ll_d \frac{\overline{\beta}}{\sigma_d^4 N}\,, \tag{4.52}$$

and

$$I_2 \ll_d p^{-2}(\Pi+\Lambda) + \frac{1}{\sigma_d^3 p^3}\,\Big(\Big(\frac{\overline{\beta}}{N}\Big)^{5/4} + \Big(\frac{\overline{\beta}}{N}\Big)^{7/4}\Big)\big(1+p^{1/2}/(\Pi+\Lambda)^{1/4}\big)(1+\|a\|^3). \tag{4.53}$$

Using (4.8), (4.15) and (4.53), we get

$$I_2 \ll_d p^{-2}(\Pi+\Lambda) + \frac{\overline{\beta}}{\sigma_d^3 p^3 N}\,(1+\|a\|^3). \tag{4.54}$$

Finally, by (4.8), (4.15), (4.18) and (4.31),

$$I_3 \ll_d \frac{k}{N}\,\big(\Lambda^{1/2}\sigma_d^{-2} + N^{-1/2}\Pi^{3/2}\sigma_d^{-3}\big) \ll \Lambda\sigma_d^{-4} + \Pi\sigma_d^{-2}. \tag{4.55}$$

Inequalities (4.8), (4.11), (4.15), (4.17), (4.19), (4.28), (4.52), (4.54) and (4.55) imply now (4.7) (and, hence, (1.25)) by an application of $\sigma_d \le 1$, and $\Pi + \Lambda \le 1$. Note that, by (1.9) and (1.11), we have $\sigma_d \Pi \le \Pi_3^{\bullet}$. Together with (4.2) and (4.4), inequality (4.7) yields (1.24). The statement $(i)$ of Theorem 1.4 is proved.

Let us prove $(i) \Longrightarrow (ii)$. To obtain $p \ge \exp\{c\lambda_5^{-2}\}$ we can use $(i)$ in the case when the condition $\mathcal{N}(p,\delta,\mathcal{S}_o,c_0G)$ is fulfilled with $c_0 = \delta/4 = 1/1200$. Indeed, the condition $\mathcal{B}(\mathcal{S}_o,\mathbb{C})$ guarantees that $e \in \mathcal{S}_o \cup \mathbb{Q}\mathcal{S}_o$ are eigenvectors of the covariance operator $\mathbb{C}$, and we can get the lower bound for $p$ by an application of Lemma 2.5 using $c_0 = \delta/4$. $\square$

## 5. From Probability to Number Theory

In Section 5 we shall reduce the estimation of the integrals of the modulus of characteristic functions $\widehat{\Psi}_a(t)$ to the estimation the integrals of some theta-series. We shall use the following lemmas.

**Lemma 5.1.** (BG (1997a, Lemma 5.1)) *Let $L, C \in \mathbb{R}^d$. Let $Z, U, V$ and $W$ denote independent random vectors taking values in $\mathbb{R}^d$. Denote by*

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

*a real-valued polynomial of second order. Then*

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

**Lemma 5.2.** (BG (1997a, Theorem 7.1)) *Assume that $\mathbb{Q}^2 = \mathbb{I}_d$ and that the condition $\mathcal{N}(p, \delta, \mathcal{S}_o, \widetilde{X})$ holds with some $0 < p \le 1$ and $0 \le \delta \le 1/(5s)$. Then, for any $a \in \mathbb{R}^d$ and $t \in \mathbb{R}$,*

$$\big|\,\widehat{\Psi}_a(t)\big| \ll_s \mathcal{M}^s\big(t; pN\big),$$

*where the function $\mathcal{M}$ and $\widehat{\Psi}_a(t)$ are defined by* (1.33) *and* (2.3) *respectively.*

Let $\varepsilon_1, \varepsilon_2, \dots$ denote i.i.d. symmetric Rademacher random variables. Let $\delta > 0$ and $\mathcal{S} = \{e_1, \dots, e_s\} \subset \mathbb{R}^d$. We shall write $\mathcal{L}(Y) \in \mathbf{\Gamma}(\delta; \mathcal{S})$ if a discrete random vector $Y$ is distributed as $\varepsilon_1 z_1 + \cdots + \varepsilon_s z_s$, with some (non-random) $z_j \in \mathbb{R}^d$ such that $\|z_j - e_j\| \le \delta$, for all $1 \le j \le s$. Recall that $\mathcal{S}_o = \{e_1, \dots, e_s\} \subset \mathbb{R}^d$ denotes an orthonormal system.

**Lemma 5.3.** (BG (1997a, Corollary 6.3)) *Assume that $\mathbb{Q}^2 = \mathbb{I}_d$ and that the condition $\mathcal{N}(p, \delta, \mathcal{S}, \widetilde{X})$ holds with some $0 < p \le 1$ and $\delta > 0$. Write $n = \lceil pN/(5s)\rceil$. Then, for any $0 < A \le B$, $a \in \mathbb{R}^d$ and $\gamma > 0$, we have*

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

*with*

$$I = \sup_{\Gamma}\sup_{b\in\mathbb{R}^d}\int_A^B \sqrt{\varphi(t/4)}\,\frac{dt}{|t|}, \qquad \varphi(t) \overset{\mathrm{def}}{=} \Big|\mathbf{E}\,\mathrm{e}\big\{\,t\,\mathbb{Q}[Y + \mathbb{Q}Y' + b]\big\}\Big|^2, \tag{5.1}$$

*where $Y = U_1 + \cdots + U_n$ and $Y' = U_1' + \cdots + U_n'$ denote sums of independent* (*non-i.i.d.*) *vectors, and $\sup\limits_{\Gamma}$ is taken over all $\big\{\mathcal{L}(U_j), \mathcal{L}(U_j') :\ 1 \le j \le n\big\} \subset \mathbf{\Gamma}(\delta; \mathcal{S})$.*

**Lemma 5.4.** *Assume that $\mathbb{Q}^2 = \mathbb{I}_d$ and that the condition $\mathcal{N}(p, \delta, \mathcal{S}, \widetilde{X})$ holds with some $0 < p \le 1$ and $\delta > 0$. Let*

$$n \overset{\mathrm{def}}{=} \lceil pN/(11s)\rceil \ge 1. \tag{5.2}$$

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

$$\int_A^B \big|\,\widehat{\Psi}_a(t)\,\big|\,\frac{dt}{|t|} \le c_\gamma(s)\,(pN)^{-\gamma}\,\log\frac{B}{A} + \sup_\Gamma \int_A^B \sqrt{\mathbf{E}\,\mathrm{e}\big\{t\,\langle\widetilde{W},\widetilde{W'}\rangle/2\big\}}\,\frac{dt}{|t|}\,, \tag{5.3}$$

*where* $W = V_1 + \cdots + V_n$ *and* $W' = V_1' + \cdots + V_n'$ *are independent sums of independent copies of random vectors* $V$ *and* $V'$ *respectively, and the supremum* $\sup_\Gamma$ *is taken over all* $\mathcal{L}(V), \mathcal{L}(V') \in \mathbf{\Gamma}(\delta;\mathcal{S})$.

Note that this lemma will be proved for general $\mathcal{S}$, but in this paper we need $\mathcal{S} = \mathcal{S}_o$ only.

*Proof of Lemma* 5.4. Let us show that

$$\int_A^B \big|\,\widehat{\Psi}_a(t)\,\big|\,\frac{dt}{|t|} \le c_\gamma(s)\,(pN)^{-\gamma}\,\log\frac{B}{A} + \sup_\Gamma \int_A^B \sqrt{\mathbf{E}\,\mathrm{e}\big\{t\,\langle\widetilde{W},\widetilde{W'}\rangle/2\big\}}\,\frac{dt}{|t|}\,, \tag{5.4}$$

where $W = V_1 + \cdots + V_n$ and $W' = V_1' + \cdots + V_n'$ are independent sums of of independent (*non-i.i.d.*) vectors, and $\sup_\Gamma$ is taken over all $\big\{\mathcal{L}(V_j), \mathcal{L}(V_j') :\ 1 \le j \le n\big\} \subset \mathbf{\Gamma}(\delta;\mathcal{S})$. Comparing (5.3) and (5.4), we see that inequality (5.4) is related to sums of *non-i.i.d.* vectors $\{V_j\}$ and $\{V_j'\}$ while inequality (5.3) deals with i.i.d. vectors. Nevertheless, we shall derive (5.3) from (5.4).

While proving (5.4) we can assume that $pN \ge c_s$ with a sufficiently large constant $c_s$, since otherwise (5.4) is trivially fulfilled.

Let $\varphi(t)$ be defined in (5.1), where $Y = U_1 + \cdots + U_n$ and $Y' = U_1' + \cdots + U_n'$ denote sums of independent (non-i.i.d.) vectors with $\big\{\mathcal{L}(U_j), \mathcal{L}(U_j') :\ 1 \le j \le n\big\} \subset \mathbf{\Gamma}(\delta;\mathcal{S})$.

We shall apply the symmetrization Lemma 5.1. Split $Y = T + T_1$ and $Y' + \mathbb{Q}b = R + R_1 + R_2$ into sums of independent sums of independent summands so that each of the sums $T$, $R$ and $R_1$ contains $n = \lceil pN/(11\,s)\rceil$ independent summands $U_j$ and $U_j'$ respectively. Such an $n$ exists since $pN \ge c_s$ with a sufficiently large $c_s$. Lemma 5.1 and symmetry of $\mathbb{Q}$ imply

$$2|\varphi(t)|^2 \le \mathbf{E}\,\mathrm{e}\big\{2\,t\,\langle\widetilde{T},\mathbb{Q}^2\widetilde{R}\rangle\big\} + \mathbf{E}\,\mathrm{e}\big\{2\,t\,\langle\widetilde{T},\mathbb{Q}^2\widetilde{R}_1\rangle\big\}. \tag{5.5}$$

Recall that $\mathbb{Q}^2 = \mathbb{I}_d$. Inequality (5.4) follows now from (5.5) and Lemma 5.3.

Let now $W = V_1 + \cdots + V_n$ and $W' = V_1' + \cdots + V_n'$ be independent sums of of independent (non-i.i.d.) vectors with $\big\{\mathcal{L}(V_j), \mathcal{L}(V_j') :\ 1 \le j \le n\big\} \subset \mathbf{\Gamma}(\delta;\mathcal{S})$. Using that all $\mathcal{L}(\widetilde{V}_j)$ are symmetrized and have non-negative characteristic functions and applying

Hölder's inequality, we obtain, for each $t$,

$$\mathbf{E}\,\mathrm{e}\big\{t\,\langle\widetilde{W},\widetilde{W}'\rangle\big\} = \mathbf{E}_{\widetilde{W}'}\Big(\prod_{j=1}^{n}\mathbf{E}_{\widetilde{V}_j}\,\mathrm{e}\big\{t\,\langle\widetilde{V}_j,\widetilde{W}'\rangle\big\}\Big) \tag{5.6}$$

$$\leq \Big(\prod_{j=1}^{n}\mathbf{E}_{\widetilde{W}'}\big(\mathbf{E}_{\widetilde{V}_j}\,\mathrm{e}\big\{t\,\langle\widetilde{V}_j,\widetilde{W}'\rangle\big\}\big)^{n}\Big)^{1/n} \tag{5.7}$$

$$= \Big(\prod_{j=1}^{n}\mathbf{E}_{\widetilde{W}'}\big(\mathbf{E}_{\widetilde{T}_j}\,\mathrm{e}\big\{t\,\langle\widetilde{T}_j,\widetilde{W}'\rangle\big\}\big)\Big)^{1/n} \tag{5.8}$$

$$= \Big(\prod_{j=1}^{n}\mathbf{E}\,\mathrm{e}\big\{t\,\langle\widetilde{T}_j,\widetilde{W}'\rangle\big\}\Big)^{1/n}, \tag{5.9}$$

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

Repeating the steps (5.6)–(5.9) for each factor $\mathbf{E}\,\mathrm{e}\big\{t\,\langle\widetilde{T}_j,\widetilde{W}'\rangle\big\}$ instead of the expectation $\mathbf{E}\,\mathrm{e}\big\{t\,\langle\widetilde{W},\widetilde{W}'\rangle\big\}$ on the right hand side separately, we get (with $\widetilde{T}'_k \overset{\mathrm{def}}{=} \sum_{l=1}^{n}\widetilde{V}'_{kl}$, where $\widetilde{V}'_{kl}$ are i.i.d. copies of $\widetilde{V}'_k$ independent of all other random vectors)

$$\mathbf{E}\,\mathrm{e}\big\{t\,\langle\widetilde{W},\widetilde{W}'\rangle\big\} \leq \Big(\prod_{j=1}^{n}\prod_{k=1}^{n}\mathbf{E}\,\mathrm{e}\big\{t\,\langle\widetilde{T}_j,\widetilde{T}'_k\rangle\big\}\Big)^{1/n^2}. \tag{5.10}$$

Thus, using (5.10) and the arithmetic-geometric mean inequality, we have

$$\begin{aligned}\int_A^B \sqrt{\mathbf{E}\,\mathrm{e}\big\{t\,\langle\widetilde{W},\widetilde{W}'\rangle/2\big\}}\,\frac{dt}{|t|} &\leq \int_A^B \Big(\prod_{j=1}^{n}\prod_{k=1}^{n}\mathbf{E}\,\mathrm{e}\big\{t\,\langle\widetilde{T}_j,\widetilde{T}'_k\rangle/2\big\}\Big)^{1/2n^2}\,\frac{dt}{|t|}\\ &\leq \frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n}\int_A^B\Big(\mathbf{E}\,\mathrm{e}\big\{t\,\langle\widetilde{T}_j,\widetilde{T}'_k\rangle/2\big\}\Big)^{1/2}\,\frac{dt}{|t|}\\ &\leq \sup_{\Gamma}\int_A^B\sqrt{\mathbf{E}\,\mathrm{e}\big\{t\,\langle\widetilde{T},\widetilde{T}'\rangle/2\big\}}\,\frac{dt}{|t|}\,,\end{aligned} \tag{5.11}$$

where $T = U_1+\cdots+U_n$ and $T' = U'_1+\cdots+U'_n$ are independent sums of independent copies of random vectors $U$ and $U'$ respectively, and the supremum $\sup_{\Gamma}$ is taken over all $\mathcal{L}(U),\mathcal{L}(U')\in\mathbf{\Gamma}(\delta;\mathcal{S})$. Inequalities (5.4) and (5.11) imply now the statement of the lemma. □

The following Lemma 5.5 provides a Poisson summation formula.

**Lemma 5.5.** *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 non-degenerate linear operator. Then*

$$\sum_{m\in\mathbb{Z}^s} \exp\bigl\{-z\,\mathbb{S}[m+a] + 2\pi i\,\langle m, b\rangle\bigr\}$$
$$= \bigl(\det(\mathbb{S}/\pi)\bigr)^{-1/2} z^{-s/2} \exp\bigl\{-2\pi i\,\langle a, b\rangle\bigr\} \sum_{l\in\mathbb{Z}^s} \exp\Bigl\{-\frac{\pi^2}{z}\,\mathbb{S}^{-1}[l+b] - 2\pi i\,\langle a, l\rangle\Bigr\},$$

*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), p. 116, or Mumford (1983), p. 189, formula (5.1); and p. 197, formula (5.9). □

Let the conditions of Lemma 5.4 be satisfied. Introduce one-dimensional lattice probability distributions $H_n = \mathcal{L}(\xi_n)$ with integer valued $\xi_n$ setting

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

It is easy to see that $1 \ll A_n \ll 1$. Moreover, by Lemma 5.5,

$$\widehat{H}_n(t) \geq 0, \qquad \text{for all } t \in \mathbb{R}. \tag{5.12}$$

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

$$q(m) \overset{\text{def}}{=} \mathbf{P}\bigl\{\,\zeta_n = m\bigr\} = A_n^s\, n^{-s/2} \exp\bigl\{-\|m\|^2/2n\bigr\}\,. \tag{5.13}$$

**Lemma 5.6.** *Let* $W = V_1 + \dots + V_n$ *and* $W' = V_1' + \dots + V_n'$ *denote independent sums of independent copies of random vectors* $V$ *and* $V'$ *such that*

$$V = \varepsilon_1 z_1 + \dots + \varepsilon_s z_s, \qquad V' = \varepsilon_{s+1} z_1' + \dots + \varepsilon_{2s} z_s',$$

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

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

*where* $\zeta_n'$ *are independent copies of* $\zeta_n$ *and* $c$ *is an absolute constant.*

*Proof.* Without loss of generality, we shall assume that $n \geq c$, where $c$ is an absolute constant which is so large as it will be needed for the validity of arguments below. 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

$$\mathbf{E}\ \mathrm{e}\bigl\{\,t\,\langle \widetilde{W}, \widetilde{W}'\rangle/4\bigr\} = \mathbf{E}\ \mathrm{e}\bigl\{\,\langle \mathbb{B}_t R, T\rangle\,\bigr\}, \qquad \text{for all } t \in \mathbb{R}, \tag{5.14}$$

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

$$\begin{aligned}\mathbf{E}\,\mathrm{e}\big\{\langle\mathbb{B}_t R, T\rangle\big\} &= \mathbf{E}\,\mathbf{E}_R\,\mathrm{e}\big\{\langle\mathbb{B}_t R, T\rangle\big\}\\ &= \sum_{\overline{m}\in\mathbb{Z}^s} p(\overline{m}) \sum_{m\in\mathbb{Z}^s} p(m)\,\mathrm{e}\big\{\langle\mathbb{B}_t m, \overline{m}\rangle\big\},\end{aligned}\tag{5.15}$$

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

$$p(m)=\mathbf{P}\big\{R=m\big\}=\prod_{j=1}^{s}\mathbf{P}\big\{R_j=m_j\big\}=\prod_{j=1}^{s}2^{-2n}\binom{2n}{m_j+n},\tag{5.16}$$

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

$$\mathbf{E}_R\,\mathrm{e}\big\{\langle\mathbb{B}_t R, T\rangle\big\}=\sum_{m\in\mathbb{Z}^s} p(m)\,\mathrm{e}\big\{\langle\mathbb{B}_t m,\overline{m}\rangle\big\}\ge 0\tag{5.17}$$

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 absolute constants $c_1$ and $c_2$ such that

$$\mathbf{P}\big\{R_j=m_j\big\}\ll n^{-1/2}\exp\big\{-m_j^2/2n\big\},\qquad \text{for } |m_j|\le c_1 n,\tag{5.18}$$

and

$$\mathbf{P}\big\{|R_j|\ge c_1 n\big\}\ll\exp\{-c_2 n\}.\tag{5.19}$$

Using (5.15)–(5.19), we obtain

$$\begin{aligned}\mathbf{E}\,\mathrm{e}\big\{\langle\mathbb{B}_t R, T\rangle\big\} &\ll_s \sum_{\overline{m}\in\mathbb{Z}^s} q(\overline{m}) \sum_{m\in\mathbb{Z}^s} p(m)\,\mathrm{e}\big\{\langle\mathbb{B}_t m,\overline{m}\rangle\big\}+\exp\{-c_2 n\}\\ &= \sum_{m\in\mathbb{Z}^s} p(m) \sum_{\overline{m}\in\mathbb{Z}^s} q(\overline{m})\,\mathrm{e}\big\{\langle\mathbb{B}_t m,\overline{m}\rangle\big\}+\exp\{-c_2 n\}\\ &= \mathbf{E}\,\mathbf{E}_{\zeta_n}\,\mathrm{e}\big\{\langle\mathbb{B}_t R,\zeta_n\rangle\big\}+\exp\{-c_2 n\}\\ &= \mathbf{E}\,\mathrm{e}\big\{\langle\mathbb{B}_t R,\zeta_n\rangle\big\}+\exp\{-c_2 n\}.\end{aligned}\tag{5.20}$$

Now we repeat our previous arguments, noting that

$$\mathbf{E}_{\zeta_n}\,\mathrm{e}\big\{\langle\mathbb{B}_t R,\zeta_n\rangle\big\}=\sum_{\overline{m}\in\mathbb{Z}^s} q(\overline{m})\;\mathrm{e}\big\{\langle\mathbb{B}_t R,\overline{m}\rangle\big\}\ge 0\tag{5.21}$$

is a value of the non-negative characteristic function of the random vector $\zeta_n$ (see (5.12)). Using again (5.18) and (5.19), we obtain

$$\mathbf{E}\,\mathrm{e}\big\{\langle\mathbb{B}_t R,\zeta_n\rangle\big\}\ll_s \mathbf{E}\,\mathrm{e}\big\{\langle\mathbb{B}_t\zeta_n,\zeta_n'\rangle\big\}+\exp\{-c_2 n\}.\tag{5.22}$$

Relations (5.14), (5.20) and (5.22) imply the statement of the lemma. □

Let us estimate the expectation $\mathbf{E}\,\mathrm{e}\big\{\langle\mathbb{B}_t\zeta_n,\zeta_n'\rangle\big\}$ under the conditions of Lemmas 5.4 and 5.6, assuming that $\delta\le 1/(5s)$, $n\ge c$, where $c$ is a sufficiently large absolute constant, and

$$\|z_j-e_j\|\le\delta,\qquad \|z_j'-e_j\|\le\delta,\qquad \text{for }\ 1\le j\le s, \tag{5.23}$$

with an orthonormal system $\mathcal{S}=\mathcal{S}_o=\big\{e_1,e_2,\dots,e_s\big\}$ involved in the conditions of Lemma 5.4. We can rewrite $\mathbf{E}\,\mathrm{e}\big\{\langle\mathbb{B}_t\zeta_n,\zeta_n'\rangle\big\}$ as

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

Thus, by (5.13),

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

Applying Lemma 5.5 with $\mathbb{S}=\mathbb{I}_s$, $z=1/2n$, $a=0$, $b=(2\pi)^{-1}\,\mathbb{B}_t\,\overline{m}$, we obtain

$$\mathbf{E}\,\mathrm{e}\big\{\langle\mathbb{B}_t\zeta_n,\zeta_n'\rangle\big\}\ll_s n^{-s/2}\sum_{l,m\in\mathbb{Z}^s}\exp\big\{-2\pi^2n\,\|\,l+(2\pi)^{-1}\mathbb{B}_t\,m\|^2-\|m\|^2/2n\big\}. \tag{5.24}$$

Note that the vectors $\zeta_n$ and $\zeta_n'$ are i.i.d. Hence,

$$\mathbf{E}\,\mathrm{e}\big\{\langle\mathbb{B}_t\zeta_n,\zeta_n'\rangle\big\}=\mathbf{E}\,\mathrm{e}\big\{\langle\mathbb{B}_t^*\zeta_n,\zeta_n'\rangle\big\}, \tag{5.25}$$

where $\mathbb{B}_t^*$ denotes the adjoint operator for $\mathbb{B}_t$. Similarly to (5.24), we could derive the inequality

$$\mathbf{E}\,\mathrm{e}\big\{\langle\mathbb{B}_t^*\zeta_n,\zeta_n'\rangle\big\}\ll_s n^{-s/2}\sum_{l,m\in\mathbb{Z}^s}\exp\big\{-2\pi^2n\,\|\,l+(2\pi)^{-1}\mathbb{B}_t^*m\|^2-\|m\|^2/2n\big\}. \tag{5.26}$$

Denote

$$r=\sqrt{2\pi^2n}. \tag{5.27}$$

By (5.24)–(5.27), we have

$$\begin{aligned}\mathbf{E}\,\mathrm{e}\big\{\langle\mathbb{B}_t\zeta_n,\zeta_n'\rangle\big\}&=\Big(\mathbf{E}\,\mathrm{e}\big\{\langle\mathbb{B}_t\zeta_n,\zeta_n'\rangle\big\}\,\mathbf{E}\,\mathrm{e}\big\{\langle\mathbb{B}_t^*\zeta_n,\zeta_n'\rangle\big\}\Big)^{1/2}\\
&\ll_s n^{-s/2}\Big(\sum_{l,m\in\mathbb{Z}^s}\exp\big\{-2\pi^2n\,\|\,l+(2\pi)^{-1}\mathbb{B}_t\,m\|^2-\|m\|^2/2n\big\}\\
&\qquad\times\sum_{\bar{l},\overline{m}\in\mathbb{Z}^s}\exp\big\{-2\pi^2n\,\|\,\bar{l}+(2\pi)^{-1}\mathbb{B}_t^*\overline{m}\|^2-\|\overline{m}\|^2/2n\big\}\Big)^{1/2}\\
&\ll_s r^{-s}\Big(\sum_{l,m,\bar{l},\overline{m}\in\mathbb{Z}^s}\exp\big\{-r^2\,\|\,l+(2\pi)^{-1}\mathbb{B}_t\,m\|^2-\|m\|^2/r^2\\
&\qquad-r^2\,\|\,\bar{l}+(2\pi)^{-1}\mathbb{B}_t^*\overline{m}\|^2-\|\overline{m}\|^2/r^2\big\}\Big)^{1/2}.\end{aligned} \tag{5.28}$$

Now we rewrite the last sum as

$$\sum_{l,m,\bar{l},\overline{m}\in\mathbb{Z}^s} \exp\bigl\{-r^2\,\|l+(2\pi)^{-1}\mathbb{B}_t m\|^2-\|m\|^2/r^2-r^2\,\|\bar{l}+(2\pi)^{-1}\mathbb{B}_t^*\overline{m}\|^2-\|\overline{m}\|^2/r^2\bigr\}$$

$$=\sum_{m,\overline{m}\in\mathbb{Z}^{2s}} \exp\bigl\{-r^2\,\|m-t\,\mathbb{V}\overline{m}\|^2-\|\overline{m}\|^2/r^2\bigr\}, \tag{5.29}$$

where $\mathbb{V}:\mathbb{R}^{2s}\to\mathbb{R}^{2s}$ is the operator with matrix

$$\mathbb{V}=\begin{pmatrix} \mathbb{O}_s & (2\pi)^{-1}\mathbb{B}_1 \\ (2\pi)^{-1}\mathbb{B}_1^* & \mathbb{O}_s \end{pmatrix}, \tag{5.30}$$

where $\mathbb{O}_s$ denotes the $(s\times s)$ matrix with zero entries. It is important that the matrix $\mathbb{V}$ is symmetric. This will allow us to apply below Lemma 6.2 (which is related to Number Theory) to derive inequality (7.14). Note that the right-hand side of (5.29) may be considered as a theta-series.

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

$$\|\mathbb{B}_1\|\le 3/2 \quad\text{and}\quad \|\mathbb{B}_1^{-1}\|\le 2. \tag{5.31}$$

Indeed, the entries of the matrix $\mathbb{B}_1$ are $b_{ij}(1)=\langle z_i,z_j'\rangle$ with some $z_j,z_j'\in\mathbb{R}^d$ satisfying (5.23). Since $\mathcal{S}_o=\bigl\{e_1,e_2,\dots,e_s\bigr\}$ is an orthonormal system, inequalities (5.23) imply that $\mathbb{B}_1=\mathbb{I}_s+\mathbb{A}$ with some matrix $\mathbb{A}=\{a_{ij}\}$ such that $|a_{ij}|\le 2\delta+\delta^2$. Thus, we have $\|\mathbb{A}\|\le\|\mathbb{A}\|_2\le 2s\delta+s\delta^2$, 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 (5.31).

By (5.30) and (5.31), for any $x\in\mathbb{R}^{2s}$, we have

$$\|x\|\ll\|\mathbb{V}x\|+\|\mathbb{V}^{-1}x\|\ll\|x\|. \tag{5.32}$$

## 6. Some facts from Number Theory

In Section 6, 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 (5.29).

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

$$\Lambda=\Bigl\{\sum_{j=1}^d n_j e_j : n_j\in\mathbb{Z},\ j=1,2,\dots,d\Bigr\} \tag{6.1}$$

is called the lattice with basis $e_1,e_2,\dots,e_d$. The determinant $\det(\Lambda)$ of a lattice $\Lambda$ is the modulus of the determinant of the matrix formed from the vectors $e_1,e_2,\dots,e_d$:

$$\det(\Lambda)\overset{\text{def}}{=}\bigl|\det(e_1,e_2,\dots,e_d)\bigr|. \tag{6.2}$$

The determinant of a lattice does not depend on the choice of basis. Any lattice $\Lambda \subset \mathbb{R}^d$ can be represented as $\Lambda = \mathbb{A}\,\mathbb{Z}^d$, where $\mathbb{A}$ is a non-degenerate linear operator. Clearly, $\det(\Lambda) = |\det \mathbb{A}|$.

Let $m_1, \dots, m_l \in \Lambda$ be linearly independent vectors belonging to a lattice $\Lambda$. Then the set

$$\Lambda' = \Big\{ \sum_{j=1}^{l} n_j m_j : n_j \in \mathbb{Z},\ j = 1, 2, \dots, l \Big\} \tag{6.3}$$

is an $l$-dimensional sublattice of the lattice $\Lambda$. Its determinant $\det(\Lambda')$ is the determinant of the matrix formed from the coordinates of the vectors $m_1, m_2, \dots, m_l$ with respect to an orthonormal basis of the linear span of the vectors $m_1, m_2, \dots, m_l$.

Let $F : \mathbb{R}^d \to [0, \infty]$ denote a norm on $\mathbb{R}^d$, that is $F(\alpha x) = |\alpha|\, F(x)$, for $\alpha \in \mathbb{R}$, and $F(x + y) \le F(x) + F(y)$. The successive minima $M_1 \le \dots \le M_d$ of $F$ with respect to a lattice $\Lambda$ are defined as follows: Let $M_1 = \inf\big\{ F(m) : m \ne 0,\ m \in \Lambda \big\}$ and define $M_j$ as the infimum of $\lambda > 0$ such that the set $\big\{ m \in \Lambda : F(m) < \lambda \big\}$ contains $j$ linearly independent vectors. It is easy to see that these infima are attained, that is there exist linearly independent vectors $b_1, \dots, b_d \in \Lambda$ such that $F(b_j) = M_j$, $j = 1, \dots, d$. The following Lemma 6.1 is proved by Davenport (1958, Lemma 1).

**Lemma 6.1.** *Let $M_1 \le \dots \le M_d$ be the successive minima of a norm $F$ with respect to the lattice $\mathbb{Z}^d$. Suppose that $1 \le j \le d$ and $M_j \le b \le M_{j+1}$, for some $b > 0$. Then the number of $m = (m_1, \dots, m_d) \in \mathbb{Z}^d$ such that $F(m) < b$ is bounded from above by $4^j\, b^j (M_1 \cdot M_2 \cdots M_j)^{-1}$.*

Representing $\Lambda = \mathbb{A}\,\mathbb{Z}^d$, we see that the lattice $\mathbb{Z}^d$ may be replaced in Lemma 6.1 by any lattice $\Lambda \subset \mathbb{R}^d$. It suffices to apply this lemma to the norm $G(m) = F(\mathbb{A} m)$, $m \in \mathbb{Z}^d$.

Let $\|x\|_\infty = \max_{1 \le j \le d} |x_j|$, for $x = (x_1, \dots, x_d) \in \mathbb{R}^d$, and let $z(v)$ denote the distance of the number $v$ to the nearest integer.

**Lemma 6.2.** *Let $L_j(x) = \sum_{k=1}^{d} q_{jk}\, x_k$, $1 \le j \le d$, denote linear forms on $\mathbb{R}^d$ such that $q_{jk} = q_{kj}$, $j, k = 1, \dots, d$. Assume that $r \ge 1$. Let $\mu$ be the number of $m = (m_1, \dots, m_d) \in \mathbb{Z}^d$ such that*

$$z(L_j(m)) < r^{-1}, \quad |m_j| < r, \qquad \text{for all } \ 1 \le j \le d. \tag{6.4}$$

*Then*

$$\mu \ll_d (M_1 \cdot M_2 \cdots M_d)^{-1}, \tag{6.5}$$

*where $M_1 \le \dots \le M_d$ are the first $d$ of the $2d$ successive minima $M_1 \le \dots \le M_{2d}$ (with respect to $\mathbb{Z}^d$) of the norm $F : \mathbb{R}^{2d} \to [0, \infty)$ defined for vectors $y = (x, \overline{x}) \in \mathbb{R}^{2d}$, $x, \overline{x} \in \mathbb{R}^d$, $\overline{x} = (\overline{x}_1, \dots, \overline{x}_d)$, as*

$$F(y) \overset{\text{def}}{=} \max\big\{ r\,|L_1(x) - \overline{x}_1|, \dots, r\,|L_d(x) - \overline{x}_d|,\ r^{-1}\,\|x\|_\infty \big\}. \tag{6.6}$$

*Moreover,*

$$\frac{1}{2d} \le M_k M_{2d+1-k} \le (2d)^{2d-1}, \qquad 1 \le k \le 2d. \tag{6.7}$$

Here and below writing $(a,b)$, for $a \in \mathbb{R}^{d_1}$, $b \in \mathbb{R}^{d_2}$, means that $(a,b) \in \mathbb{R}^{d_1+d_2}$ and the coordinates of $(a,b)$ are the coordinates of the vectors $a$ and $b$ in the corresponding order, that is, $(a,b) = (a_1, a_2, \dots, a_{d_1}, b_1, b_2, \dots, b_{d_2})$. A similar notation will be used for more vectors $(a, b, c, \dots, z)$.

Lemma 6.2 is proved in Lemma 3 of Davenport (1958), see also Davenport (1958, formula (20), p. 113). Note that inequality (6.5) of Lemma 6.2 is, in a sense, a particular case of Lemma 6.1.

Note that, in the conditions of Lemma 6.2,

$$r^{-1} \le M_1 \le \dots \le M_d \ll_d 1, \tag{6.8}$$

where the first inequality is obvious by $F(y) \ge r^{-1} \|x\|_\infty$ (if $\|x\|_\infty = 0$ then $F(y) \ge r\,\|\overline{x}\|_\infty \ge r^{-1}\,\|\overline{x}\|_\infty$) and $M_d \ll_d 1$ follows from (6.7) for $k = d$.

**Lemma 6.3.** *Let $F_j(m)$, $j = 1, 2$, be some norms in $\mathbb{R}^d$ and $M_1 \le \dots \le M_d$ and $N_1 \le \dots \le N_d$ be the successive minima of $F_1$ with respect to a lattice $\Lambda_1$ and of $F_2$ with respect to a lattice $\Lambda_2$ respectively. Assume that $M_k \gg_d F_2(n_k)$, $k = 1, 2, \dots, d$, for some linearly independent vectors $n_1, n_2, \dots, n_d \in \Lambda_2$. Then*

$$M_k \gg_d N_k, \quad k = 1, \dots, d. \tag{6.9}$$

The proof of this lemma is elementary and therefore omitted.

**Lemma 6.4.** *Let $\Lambda$ be a lattice in $\mathbb{R}^d$ and let $c_1(d)$ and $c_2(d)$ be positive quantities depending on d only. Then*

$$\sum_{v \in \Lambda} \exp\bigl\{ -c_1(d)\,\|v\|^2 \bigr\} \asymp_d \#H, \tag{6.10}$$

*where* $H \overset{\text{def}}{=} \bigl\{ v \in \Lambda \, : \, \|v\|_\infty < c_2(d) \bigr\}$.

*Proof.* Introduce for $\mu = (\mu_1, \dots, \mu_d) \in \Xi \overset{\text{def}}{=} c_2(d)\,\mathbb{Z}^d$ the sets

$$B_\mu \overset{\text{def}}{=} \Bigl[ \mu_1 - \frac{c_2(d)}{2}\,,\, \mu_1 + \frac{c_2(d)}{2} \Bigr) \times \dots \times \Bigl[ \mu_d - \frac{c_2(d)}{2}\,,\, \mu_d + \frac{c_2(d)}{2} \Bigr)$$

such that $\mathbb{R}^d = \bigcup_\mu B_\mu$. For any fixed $w^* \in H_\mu \overset{\text{def}}{=} \bigl\{ w \in \Lambda \cap B_\mu \bigr\}$ we have

$$w - w^* \in H, \qquad \text{for any } w \in H_\mu.$$

Hence we conclude for any $\mu \in \Xi$

$$\#H_\mu \le \#H. \tag{6.11}$$

Since $x \in B_\mu$ implies $\|x\|_\infty \geq \|\mu\|_\infty/2$, we obtain by (6.11)

$$\begin{aligned} \sum_{v\in\Lambda} \exp\{-c_1(d)\,\|v\|^2\} \quad &\leq \quad \sum_{v\in\Lambda} \exp\{-c_d\,\|v\|_\infty^2\} \\ &\ll_d \quad \#H_0 + \sum_{\mu\in\Xi\setminus 0} \quad \sum_{v\in\Lambda} \mathbf{I}\{\,v \in B_\mu\}\,\exp\{-c(d)\,\|\mu\|_\infty^2\} \\ &\ll_d \quad \#H \cdot \sum_{\mu\in\Xi} \exp\{-c(d)\,\|\mu\|_\infty^2\} \\ &\ll_d \quad \#H. \end{aligned} \tag{6.12}$$

On the other hand,

$$\sum_{v\in\Lambda} \exp\{-c_1(d)\,\|v\|^2\} \geq \sum_{v\in\Lambda} \exp\{-c_d\,\|v\|_\infty^2\} \gg_d \#H. \tag{6.13}$$

This conclude the proof of Lemma 6.4. □

For a lattice $\Lambda \subset \mathbb{R}^d$, $\dim \Lambda = d$ and $1 \leq l \leq d$, we define its $\alpha_l$-characteristics by

$$\alpha_l(\Lambda) \overset{\text{def}}{=} \sup\Big\{\,\big|\det(\Lambda')\big|^{-1} : \Lambda' \subset \Lambda,\ \ l\text{-dimensional sublattice of } \Lambda\Big\}. \tag{6.14}$$

**Lemma 6.5.** *Let $F(\cdot)$ be a norm in $\mathbb{R}^d$ such that $F(\cdot) \asymp_d \|\cdot\|$. Let $M_1 \leq \cdots \leq M_d$ be the successive minima of $F$ with respect to a lattice $\Lambda \subset \mathbb{R}^d$. Then*

$$\alpha_l(\Lambda) \asymp_d (M_1 \cdot M_2 \cdots M_l)^{-1}, \quad l = 1, \ldots, d. \tag{6.15}$$

For the proof of Lemma 6.5 we shall use the following lemma formulated in Proposition (p. 517) and Remark (p. 518) in A.K. Lenstra, H.W. Lenstra and Lovász (1982).

**Lemma 6.6.** *Let $M_1 \leq \cdots \leq M_d$ be the successive minima of the standard Euclidean norm with respect to a lattice $\Lambda \subset \mathbb{R}^d$. Then there exists a basis $e_1, e_2, \ldots, e_d$ of $\Lambda$ such that*

$$M_l \asymp_d \|e_l\|, \quad l = 1, \ldots, d. \tag{6.16}$$

*Moreover,*

$$\det(\Lambda) \asymp_d \prod_{l=1}^{d} \|e_l\|. \tag{6.17}$$

*Proof of Lemma* 6.5. According to Lemma 6.3, we can replace the Euclidean norm $\|\cdot\|$ by the norm $F(\cdot)$, in the formulation of Lemma 6.6. Let $\Lambda' \subset \Lambda$ be an arbitrary $l$-dimensional sublattice of $\Lambda$ and $N_1 \leq \cdots \leq N_l$ be the successive minima of the norm $F(\cdot)$ with respect to $\Lambda'$. It is clear that $M_j \leq N_j$, $j = 1, 2, \ldots, l$. On the other hand, $M_j = F(m_j)$ for some linearly independent vectors $m_1, m_2, \ldots, m_l \in \Lambda$. In the case, where

$$\Lambda' = \Big\{\sum_{j=1}^{l} n_j m_j : n_j \in \mathbb{Z}, j = 1, 2, \ldots, l\Big\}, \tag{6.18}$$

we have $N_j = M_j$, $j = 1, 2, \dots, l$. It remains to take into account definition (6.14) and to apply Lemma 6.6. □

## 7. From Number Theory to Probability

In Section 7, we shall use these number-theoretical results of Section 6 to estimate integrals of the right-hand side of (5.29). Recall that we have assumed the conditions of Lemmas 5.4 and 5.6, $\delta \le 1/(5s)$, $n \ge c$, and (5.23), 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

$$\mathbb{D}_r \overset{\mathrm{def}}{=} \begin{pmatrix} r\,\mathbb{I}_{2s} & \mathbb{O}_{2s} \\ \mathbb{O}_{2s} & r^{-1}\,\mathbb{I}_{2s} \end{pmatrix} \in \mathrm{SL}(4s, \mathbb{R}), \quad r > 0, \tag{7.1}$$

$$\mathbb{K}_t \overset{\mathrm{def}}{=} \begin{pmatrix} \mathbb{I}_{2s} & -t\,\mathbb{I}_{2s} \\ t\,\mathbb{I}_{2s} & \mathbb{I}_{2s} \end{pmatrix}, \quad t \in \mathbb{R}, \tag{7.2}$$

$$\mathbb{U}_t \overset{\mathrm{def}}{=} \begin{pmatrix} \mathbb{I}_{2s} & -t\,\mathbb{I}_{2s} \\ \mathbb{O}_{2s} & \mathbb{I}_{2s} \end{pmatrix} \in \mathrm{SL}(4s, \mathbb{R}), \quad t \in \mathbb{R}, \tag{7.3}$$

and the lattices

$$\Lambda = \Lambda_{\mathbb{V}} \overset{\mathrm{def}}{=} \begin{pmatrix} \mathbb{I}_{2s} & \mathbb{O}_{2s} \\ \mathbb{O}_{2s} & \mathbb{V} \end{pmatrix} \mathbb{Z}^{4s} \subset \mathbb{R}^{4s}, \tag{7.4}$$

$$\Lambda_j = \mathbb{D}_j\,\mathbb{U}_{j^{-1}}\,\Lambda = \begin{pmatrix} j\,\mathbb{I}_{2s} & -\mathbb{V} \\ \mathbb{O}_{2s} & j^{-1}\,\mathbb{V} \end{pmatrix} \mathbb{Z}^{4s}, \quad j = 1, 2, \dots, \tag{7.5}$$

where the matrix $\mathbb{V}$ is defined in (5.30). Below we shall use the following simplest properties of these matrices:

$$\mathbb{D}_a\,\mathbb{D}_b = \mathbb{D}_{ab}, \quad \mathbb{U}_a\,\mathbb{U}_b = \mathbb{U}_{a+b} \quad \text{and} \quad \mathbb{D}_a\,\mathbb{U}_b = \mathbb{U}_{a^2 b}\,\mathbb{D}_a, \quad \text{for } a, b > 0. \tag{7.6}$$

In the sequel we shall apply Lemma 6.2 to linear forms

$$L_j(x) = \sum_{k=1}^{2s} t\,a_{jk}\,x_k, \qquad 1 \le j \le 2s, \tag{7.7}$$

where $t \in \mathbb{R}$ is arbitrary and $a_{jk}$ are the elements of the symmetric matrix $\mathbb{V}$. For fixed $t$, we denote the corresponding successive minima of the norm $F(\cdot)$ (defined by (6.6) and (7.7)) by $M_{j,t}$, $j = 1, \dots, 4s$. Thus, we can write

$$M_{j,t} = \bigl|L(m, \overline{m}, t)\bigr|_\infty, \tag{7.8}$$

for some $m, \overline{m} \in \mathbb{Z}^{2s}$, where

$$L(m, \overline{m}, t) = \bigl(r\,(m_1 - t\,(\mathbb{V}\overline{m})_1), \dots, r\,(m_{2s} - t\,(\mathbb{V}\overline{m})_{2s}),\ r^{-1}\overline{m}_1, \dots, r^{-1}\overline{m}_{2s}\bigr). \tag{7.9}$$

It is easy to see from the definition that $M_{j,t}$ are the successive minima of the norm $\|\cdot\|_\infty$ with respect to the lattice

$$\Xi_t \stackrel{\text{def}}{=} \begin{pmatrix} r\,\mathbb{I}_{2s} & -r\,t\,\mathbb{V} \\ \mathbb{O}_{2s} & r^{-1}\,\mathbb{I}_{2s} \end{pmatrix} \mathbb{Z}^{4s}. \tag{7.10}$$

Moreover, simultaneously, $M_{j,t}$ are the successive minima of the norm $F^*(\cdot)$ defined for $(m,\overline{m}) \in \mathbb{R}^{4s}$, $m,\overline{m} \in \mathbb{R}^{2s}$, by

$$F^*((m,\overline{m})) \stackrel{\text{def}}{=} \max\big\{\|m\|_\infty, \|\mathbb{V}^{-1}\overline{m}\|_\infty\big\} \tag{7.11}$$

with respect to the lattice

$$\Omega_t \stackrel{\text{def}}{=} \begin{pmatrix} r\,\mathbb{I}_{2s} & -r\,t\,\mathbb{V} \\ \mathbb{O}_{2s} & r^{-1}\,\mathbb{V} \end{pmatrix} \mathbb{Z}^{4s} = \mathbb{D}_r\,\mathbb{U}_t\,\Lambda. \tag{7.12}$$

Note that, according to (5.32),

$$F^*(\cdot) \asymp_s \|\cdot\|. \tag{7.13}$$

Using (5.32), (7.10)–(7.13) and Lemmas 6.2, 6.4 and 6.5, we obtain

$$\begin{aligned} \sum_{m,\overline{m}\in\mathbb{Z}^{2s}} \exp\big\{-r^2\,\|m-t\,\mathbb{V}\overline{m}\|^2 - \|\overline{m}\|^2/r^2\big\} &= \sum_{v\in\Xi_t} \exp\big\{-\|v\|^2\big\} \\ &\ll_s \#\big\{v\in\Xi_t \,:\, \|v\|_\infty < 1\big\} \\ &\ll_s (M_{1,t}\cdot M_{2,t}\cdots M_{2s,t})^{-1} \\ &\asymp_s \alpha_{2s}(\Xi_t) \asymp_s \alpha_{2s}(\Omega_t). \end{aligned} \tag{7.14}$$

Here we have used essentially that the matrix $\mathbb{V}$ is symmetric. Now, by (5.28), (5.29), (7.12) and (7.14), we have

$$\mathbf{E}\ \mathrm{e}\big\{\langle \mathbb{B}_t\zeta_n, \zeta_n'\rangle\big\} \ll_s r^{-s}\,\big(\alpha_{2s}(\Omega_t)\big)^{1/2} = r^{-s}\,\big(\alpha_{2s}(\mathbb{D}_r\,\mathbb{U}_t\,\Lambda)\big)^{1/2}. \tag{7.15}$$

Using (7.15) and Lemmas 5.4 and 5.6, we derive the following lemma.

**Lemma 7.1.** *Let the conditions of Lemma* 5.4 *be satisfied with* $\delta \le 1/(5s)$ *and with an orthonormal system* $\mathcal{S} = \mathcal{S}_o = \{e_1,\dots,e_s\} \subset \mathbb{R}^d$. *Then*

$$\int_{r^{-1}}^{1} \big|\widehat{\Psi}_a(t/2)\big|\,\frac{dt}{t} \ll_s (pN)^{-1} + r^{-s/2}\sup_{\Gamma}\int_{r^{-1}}^{1} (\alpha_{2s}(\mathbb{D}_r\,\mathbb{U}_t\,\Lambda))^{1/4}\,\frac{dt}{t}\,, \qquad \text{for all } t\in\mathbb{R}, \tag{7.16}$$

*where* $r$, $\alpha_{2s}(\cdot)$, $\mathbb{D}_r$ $\mathbb{U}_t$ *and the lattice* $\Lambda$ *are defined in relations* (5.2), (5.27), (5.30), (6.14), (7.1), (7.3) *and* (7.4) *and in Lemma* 5.6. *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*

$$\|z_j - e_j\| \le \delta, \qquad \|z_j' - e_j\| \le \delta, \qquad \text{for } 1\le j\le s. \tag{7.17}$$

Let $v=(m,\overline{m})\in\mathbb{R}^{4s}$, $m,\overline{m}\in\mathbb{R}^{2s}$. Then

$$\overline{m}+tm=(1+t^2)\,\overline{m}+t\,(m-t\overline{m}). \tag{7.18}$$

Equality (7.18) implies that

$$\|\overline{m}+tm\|\ll_s\|\overline{m}\|+\|m-t\overline{m}\|,\quad \text{for } |t|\ll_s 1. \tag{7.19}$$

Hence,

$$r\,\|m-t\overline{m}\|+r^{-1}\,\|\overline{m}+tm\|\ll_s r\,\|m-t\overline{m}\|+r^{-1}\,\|\overline{m}\|,\quad \text{for } r\gg 1,\ |t|\ll_s 1. \tag{7.20}$$

According to (7.1)–(7.3), we have

$$\mathbb{D}_r\mathbb{U}_t v=(r(m-t\overline{m}),r^{-1}\overline{m})\quad\text{and}\quad \mathbb{D}_r\mathbb{K}_t v=(r(m-t\overline{m}),r^{-1}(\overline{m}+tm)). \tag{7.21}$$

It is clear that the operators $\mathbb{D}_r\mathbb{U}_t$ and $\mathbb{D}_r\mathbb{K}_t$ are invertible. Therefore, using (7.20) and (7.21) and applying Lemmas 6.3 and 6.5, we derive the inequality

$$\alpha_{2s}(\mathbb{D}_r\mathbb{U}_t\Omega)\ll_s\alpha_{2s}(\mathbb{D}_r\mathbb{K}_t\Omega),\quad \text{for } r\gg 1,\ |t|\ll_s 1, \tag{7.22}$$

which is valid for any lattice $\Omega\subset\mathbb{R}^{4s}$.

Let $\mathbb{T}$ be the $(4s\times 4s)$ permutation matrix which permutes the rows of a $(4s\times 4s)$ matrix $\mathbb{A}$ so that the new order (corresponding to the matrix $\mathbb{T}\mathbb{A}$) is:

$$1,2s,2,2s+1,\dots,2s-1,4s.$$

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

$$\alpha_{2s}(\mathbb{T}\Omega)=\alpha_{2s}(\Omega), \tag{7.23}$$

for any lattice $\Omega\subset\mathbb{R}^{4s}$.

Note now that

$$\mathbb{T}\mathbb{D}_r\mathbb{K}_t\Lambda_j=\mathbb{T}\mathbb{D}_r\mathbb{K}_t\mathbb{T}^{-1}\mathbb{T}\Lambda_j=\mathbb{W}_t\Delta_j, \tag{7.24}$$

where $\Delta_j$ is a lattice defined by

$$\Delta_j=\mathbb{T}\Lambda_j \tag{7.25}$$

and where $\mathbb{W}_t$ is $(4s\times 4s)$-matrix

$$\mathbb{W}_t=\begin{pmatrix}\mathbb{G}_{r,t} & \mathbb{O}_2 & : & \mathbb{O}_2\\ \mathbb{O}_2 & \mathbb{G}_{r,t} & : & \mathbb{O}_2\\ \cdot\,\cdot & \cdot\,\cdot & \cdot\,\cdot & \cdot\,\cdot\\ \mathbb{O}_2 & \mathbb{O}_2 & : & \mathbb{G}_{r,t}\end{pmatrix} \tag{7.26}$$

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

$$\mathbb{G}_{r,t}\stackrel{\text{def}}{=}\begin{pmatrix} r & -rt\\ r^{-1}t & r^{-1}\end{pmatrix}. \tag{7.27}$$

Let $|t|\le 2$ and

$$\theta=\arcsin\bigl(t\,(1+t^2)^{-1/2}\bigr)\quad\text{or, equivalently,}\quad t=\tan\theta. \tag{7.28}$$

Then we have

$$|\theta| \le c^* \overset{\text{def}}{=} \arcsin(2/\sqrt{5}), \quad \cos\theta = (1+t^2)^{-1/2}, \quad \sin\theta = t\,(1+t^2)^{-1/2}. \tag{7.29}$$

It is easy to see that

$$\mathbb{G}_{r,t} = (1+t^2)^{1/2}\,\overline{\mathbb{D}}_r\,\overline{\mathbb{K}}_\theta \tag{7.30}$$

and

$$\mathbb{W}_t = (1+t^2)^{1/2}\,\widetilde{\mathbb{D}}_r\,\widetilde{\mathbb{K}}_\theta, \tag{7.31}$$

where

$$\widetilde{\mathbb{D}}_r = \begin{pmatrix} \overline{\mathbb{D}}_r & \mathbb{O}_2 & : & \mathbb{O}_2 \\ \mathbb{O}_2 & \overline{\mathbb{D}}_r & : & \mathbb{O}_2 \\ \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot \\ \mathbb{O}_2 & \mathbb{O}_2 & : & \overline{\mathbb{D}}_r \end{pmatrix} \quad \text{and} \quad \widetilde{\mathbb{K}}_\theta = \begin{pmatrix} \overline{\mathbb{K}}_\theta & \mathbb{O}_2 & : & \mathbb{O}_2 \\ \mathbb{O}_2 & \overline{\mathbb{K}}_\theta & : & \mathbb{O}_2 \\ \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot \\ \mathbb{O}_2 & \mathbb{O}_2 & : & \overline{\mathbb{K}}_\theta \end{pmatrix} \tag{7.32}$$

are $(4s \times 4s)$ matrices with

$$\overline{\mathbb{D}}_r \overset{\text{def}}{=} \begin{pmatrix} r & 0 \\ 0 & r^{-1} \end{pmatrix} \quad \text{and} \quad \overline{\mathbb{K}}_\theta \overset{\text{def}}{=} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \in \mathrm{SL}(2,\mathbb{R}). \tag{7.33}$$

Substituting (7.31) into equality (7.24), we obtain

$$\mathbb{T}\mathbb{D}_r\mathbb{K}_t\,\Lambda_j = (1+t^2)^{1/2}\,\widetilde{\mathbb{D}}_r\,\widetilde{\mathbb{K}}_\theta\,\Delta_j. \tag{7.34}$$

Below we shall also use the following crucial lemma of Götze and Margulis (2009).

**Lemma 7.2.** *Let* $\widetilde{\mathbb{K}}_\theta$ *and*

$$\widetilde{\mathbb{H}} = \begin{pmatrix} \overline{\mathbb{H}} & \mathbb{O}_2 & : & \mathbb{O}_2 \\ \mathbb{O}_2 & \overline{\mathbb{H}} & : & \mathbb{O}_2 \\ \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot \\ \mathbb{O}_2 & \mathbb{O}_2 & : & \overline{\mathbb{H}} \end{pmatrix} \tag{7.35}$$

*be* $(2d\times 2d)$*-matrices such that* $\overline{\mathbb{H}} \in G = \mathrm{SL}(2,\mathbb{R})$ *and* $\widetilde{\mathbb{K}}_\theta$ *is defined in* (7.32) *and* (7.33)*. Let* $\beta$ *is a positive number such that* $\beta d > 2$*. Then, for any* $\overline{\mathbb{H}} \in G$ *and any lattice* $\Delta \subset \mathbb{R}^{2d}$,

$$\int_0^{2\pi} (\alpha_d(\widetilde{\mathbb{H}}\,\widetilde{\mathbb{K}}_\theta\,\Delta))^\beta d\theta \ll_{\beta,d} (\alpha_d(\Delta))^\beta\,\|\overline{\mathbb{H}}\|^{\beta d-2}. \tag{7.36}$$

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

Consider, under the conditions of Lemma 7.1,

$$I_0 \overset{\text{def}}{=} \int_{r^{-1}/2}^{1/2} \big|\,\widehat{\Psi}_a(t)\,\big|\,\frac{dt}{t} = \int_{r^{-1}}^{1} \big|\,\widehat{\Psi}_a(t/2)\,\big|\,\frac{dt}{t}\,. \tag{7.37}$$

By Lemma 7.1, we have

$$I_0 \ll_s (pN)^{-1} + r^{-s/2} \sup_{\Gamma}\ J, \tag{7.38}$$

where

$$J = \int_{r^{-1}}^{1} \big(\alpha_{2s}(\mathbb{D}_r\,\mathbb{U}_t\,\Lambda)\big)^{1/4}\,\frac{dt}{t} \;\le \sum_{j=2}^{\rho} I_j, \tag{7.39}$$

with

$$I_j \overset{\mathrm{def}}{=} \int_{j^{-1}}^{(j-1)^{-1}} \big(\alpha_{2s}(\mathbb{D}_r\,\mathbb{U}_t\,\Lambda)\big)^{1/4}\,\frac{dt}{t}\,, \quad j = 2, 3, \dots, \rho \overset{\mathrm{def}}{=} \lceil r\rceil + 1. \tag{7.40}$$

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

$$\begin{aligned} I_j &= \int_{j}^{j^2(j-1)^{-1}} \big(\alpha_{2s}(\mathbb{D}_r\,\mathbb{U}_{vj^{-2}}\,\Lambda)\big)^{1/4}\,\frac{dv}{v} \\ &\le \int_{j}^{j+2} \big(\alpha_{2s}(\mathbb{D}_r\,\mathbb{U}_{vj^{-2}}\,\Lambda)\big)^{1/4}\,\frac{dv}{v} \\ &= \int_{0}^{2} \big(\alpha_{2s}(\mathbb{D}_r\,\mathbb{U}_{wj^{-2}}\,\mathbb{U}_{j^{-1}}\,\Lambda)\big)^{1/4}\,\frac{dw}{w+j}\,. \end{aligned} \tag{7.41}$$

By (7.6),

$$\mathbb{D}_r\,\mathbb{U}_{wj^{-2}} = \mathbb{D}_{rj^{-1}}\,\mathbb{D}_j\,\mathbb{U}_{wj^{-2}} = \mathbb{D}_{rj^{-1}}\,\mathbb{U}_w\,\mathbb{D}_j. \tag{7.42}$$

According to (7.41) and (7.42),

$$I_j \ll \frac{1}{j}\int_{0}^{2} \big(\alpha_{2s}(\mathbb{D}_{rj^{-1}}\mathbb{U}_t\Lambda_j)\big)^{1/4}\,dt, \tag{7.43}$$

where the lattices $\Lambda_j$ are defined in (7.5) (see also (7.1), (7.3) and (7.4)). Let $N_1^{(j)} \le \dots \le N_{4s}^{(j)}$ be the successive minima of the Euclidean norm with respect to the lattice $\Lambda_j$. Using (5.32) and (7.5), it is easy to show that

$$\det(\Lambda_j) = \big|\det\mathbb{V}\big| \asymp_s 1 \quad \text{and} \quad N_1^{(j)} \gg_s 1. \tag{7.44}$$

Therefore, by Lemmas 6.5 and 6.6, $N_k^{(j)} \asymp_s 1$, $k = 1, 2, \dots, 4s$, and

$$\alpha_{2s}(\Lambda_j) \ll_s 1, \tag{7.45}$$

By (7.22), (7.23) and (7.34), we have

$$\begin{aligned} \alpha_{2s}(\mathbb{D}_{rj^{-1}}\mathbb{U}_t\Lambda_j) &\ll_s \alpha_{2s}(\mathbb{D}_{rj^{-1}}\,\mathbb{K}_t\,\Lambda_j) = \alpha_{2s}(\mathbb{T}\,\mathbb{D}_{rj^{-1}}\,\mathbb{K}_t\,\Lambda_j) \\ &\ll_s \alpha_{2s}(\widetilde{\mathbb{D}}_{rj^{-1}}\,\widetilde{\mathbb{K}}_\theta\,\Delta_j), \end{aligned} \tag{7.46}$$

for $|t| \ll_s 1$, $r \geq 1$, $j = 2, 3, \ldots, \rho$, where $\theta$ is defined in (7.28). Using (7.28), (7.29), (7.32), (7.46) and Lemma 7.2 (with $d = 2s$), we obtain

$$\begin{aligned} \int_0^2 \big(\alpha_{2s}(\mathbb{D}_{rj^{-1}}\mathbb{U}_t\Lambda_j)\big)^{1/4}\, dt &\ll_s \int_0^{c^*} \big(\alpha_{2s}(\widetilde{\mathbb{D}}_{rj^{-1}}\,\widetilde{\mathbb{K}}_\theta\,\Delta_j)\big)^{1/4}\,\frac{d\theta}{\cos^2\theta} \\ &\ll \int_0^{2\pi} \big(\alpha_{2s}(\widetilde{\mathbb{D}}_{rj^{-1}}\,\widetilde{\mathbb{K}}_\theta\,\Delta_j)\big)^{1/4}\, d\theta \\ &\ll_s \|\overline{\mathbb{D}}_{rj^{-1}}\|^{s/2-2}\,\big(\alpha_{2s}(\Delta_j)\big)^{1/4}, \end{aligned} \tag{7.47}$$

if $s \geq 5$. It is clear that $\|\overline{\mathbb{D}}_{rj^{-1}}\| = r j^{-1}$. Therefore, according to (7.23), (7.25), (7.43) and (7.47),

$$I_j \ll_s \frac{1}{j}\,(rj^{-1})^{s/2-2}\,\big(\alpha_{2s}(\Lambda_j)\big)^{1/4}. \tag{7.48}$$

By (7.39), (7.45) and (7.48), we obtain, for $s \geq 5$,

$$J \ll_s \sum_{j=2}^{\rho} \frac{1}{j}\,(rj^{-1})^{s/2-2} \ll_s r^{s/2-2}. \tag{7.49}$$

By (5.2), (5.27), (7.38) and (7.49), we have $r \asymp_s (Np)^{1/2}$ and

$$I_0 \ll_s r^{-2} \ll_s (Np)^{-1}. \tag{7.50}$$

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

$$\int_1^{c(s)} \big|\widehat{\Psi}_a(t/2)\big|\,\frac{dt}{t} \ll_s r^{-2} \ll_s (Np)^{-1}, \tag{7.51}$$

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 7.3.** *Let the conditions of Lemma* 5.4 *be satisfied with* $s \geq 5$ *and* $c(s)$ *be a quantity depending on* $s$ *only. Then there exists a* $c_s$ *such that*

$$\int_{r^{-1}}^{c(s)} \big|\widehat{\Psi}_a(t)\big|\,\frac{dt}{t} \ll_s (Np)^{-1}, \tag{7.52}$$

*if* $Np \gg_s c_s$, *where* $r$ *is defined in* (5.2) *and* (5.27).

## References

[1] Bentkus, V., *Asymptotic expansions for distributions of sums of independent random elements in a Hilbert space*, Lithuanian Math. J., **24**, 305–319, 1984

[2] Bentkus, V. and Götze, F., *On the lattice point problem for ellipsoids*, Russian Acad. Sc. Doklady, **343**, 4, 439–440, 1995a

[3] Bentkus, V. and Götze, F., *Optimal rates of convergence in functional limit theorems for quadratic forms*, Preprint 95-091 SFB 343, Universität Bielefeld, 1995b

[4] Bentkus, V. and Götze, F., *Optimal rates of convergence in the CLT for quadratic forms*, Ann. Probab., **24**, 1, 466–490, 1996

[5] Bentkus, V. and Götze, F., *Uniform rates of convergence in the CLT for quadratic forms in multidimensional spaces*, Probab. Theory Rel. Fields, **109**, 3, 367–416, 1997a

[6] Bentkus, V. and Götze, F., *On the lattice point problem for ellipsoids*, Acta Arithmetica, **80**, 2, 101–125, 1997b

[7] Bentkus, V. and Götze, F., Paulauskas V. and Račkauskas A. *The accuracy of Gaussian approximation in Banach spaces* Preprint 90-100 SFB 343, Universität Bielefeld, 1990 (In Russian: Itogi Nauki i Tehniki, ser. Sovr. Probl. Matem., Moskva, VINITI, **81**, 39–139, 1991)

[8] Bentkus, V., Götze, F. and Zaitsev, A.Yu., *Approximations of quadratic forms of independent random vectors by accompanying laws*, Theory Probab. Appl., **42**, 2, 189–212, 1997

[9] Bentkus, V., Götze, F. and Zitikis, R., *Asymptotic expansions in the integral and local limit theorems in Banach spaces with applications to $\omega$-statistics*, J. Theor. Probab., **6**, 4, 727–780, 1993

[10] Bhattacharya, R.N. and Ranga Rao, R., *Normal Approximation and Asymptotic Expansions*, Wiley New York, 1986

[11] Bogatyrev, S.A., Götze, F. and Ulyanov, V.V., *Non-uniform bounds for short asymptotic expansions in the CLT for balls in a Hilbert space*, J. Multivariate Anal. **97**, 9, 2041–2056, 2006

[12] Cassels, J.W.S., *An introduction to the geometry of numbers*, Springer, Berlin–Göttingen–Heidelberg, 1959

[13] Chung, K.L., *A course in probability theory*, Academic Press, New York and London, 1974

[14] Davenport, H., *Indefinite quadratic forms in many variables* (II), Proc. London Math. Soc., **8**, 3, 109–126, 1958

[15] Esseen, C.-G., *Fourier analysis of distribution functions*, Acta Math., **77**, 1–125, 1945

[16] Fricker, F., *Einführung in die Gitterpunktlehre*, Birkhäuser, Basel–Boston–Stuttgart, 1982

[17] Götze, F., *Asymptotic expansions for bivariate von Mises functionals*, Z. Wahrsch. verw. Geb., **50**, 333–355, 1979

[18] Götze, F., *Lattice point problem and values of quadratic forms*, Invent. Math., **157**, 1, 195–226, 2004

[19] Götze, F. and Margulis G.A., *Distribution of values of quadratic forms*, Preprint 09-001 SFB 701, Universität Bielefeld, Bielefeld, 2009

[20] Götze, F. and Ulyanov, V., *Uniform approximations in the CLT for balls in Euclidian spaces*, Preprint 00-034 SFB 343, Universität Bielefeld, Bielefeld, 2000

[21] Götze, F. and Ulyanov, V., *Asymptotic disrtribution of $\chi^2$-type statistics*, Preprint 03-033, Research group "Spectral analysis, asymptotic distributions and stochastic dynamics", 2003

[22] Götze, F. and Zaitsev, A.Yu., *Uniform rates of convergence in the CLT for quadratic forms*, Preprint 08-119 SFB 701, Universität Bielefeld, Bielefeld, 2008

[23] Hardy, G.H., *The average order of the arithmetical functions $P(x)$ and $\Delta(x)$*, Proc. London Math. Soc. **15**, 2, 192–213, 1916

[24] Landau, E., *Zur analytischen Zahlentheorie der definiten quadratischen Formen*, Sitzber. Preuss. Akad. Wiss. **31**, 458–476, 1915

[25] Lenstra, A.K., Lenstra, H.W.Jr. and Lovász, L., *Factoring polynomials with rational coefficients*, Math. Ann. **261**, 4, 515–534, 1982

[26] Mumford, D., *Tata Lectures on Theta I*, Birkhäuser, Boston–Basel–Stuttgart, 1983

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

[28] Nagaev, S.V. and Chebotarev, V.I., *On the accuracy of Gaussian approximation in Hilbert space*, Limit theorems of probability theory (Vilnius, 1999). Acta Appl. Math., **58**, 1-3, 189–215, 1999

[29] Nagaev, S.V. and Chebotarev, V.I., *On the accuracy of Gaussian approximation in Hilbert space*, Siberian Adv. Math., **15**, 1, 11–73, 2005

[30] Petrov, V.V., *Sums of independent random variables*, Springer, Berlin–Heidelberg–New York, 1975

[31] Prawitz, H. *Limits for a distribution, if the characteristic function is given in a finite domain*, Skand. AktuarTidskr., 138–154, 1972

[32] Senatov, V.V. *Normal approximation: new results, methods and problems*, Modern Probability and Statistics. VSP, Utrecht, 1998. viii+363 pp.

[33] Weyl, H., *Über die Gleichverteilung der Zahlen mod-Eins*, Math. Ann., **77**, 313–352, 1915/16

[34] Zalesskiĭ, B.A., Sazonov, V.V. and Ulyanov, V. V., *Normal approximation in a Hilbert space.* I; II, Theory Probab. Appl. **33**, 2, 207–227; 3, 473–487, 1988

*E-mail address*: goetze@math.uni-bielefeld.de

Fakultät für Mathematik,Universität Bielefeld, Postfach 100131, D-33501 Bielefeld, Germany

*E-mail address*: zaitsevpdmi.ras.ru

St. Petersburg Department of Steklov Mathematical Institute,, Fontanka 27, St. Petersburg 191023, Russia