Probab. Theory Relat. Fields 109, 367–416 (1997)

Probability Theory and Related Fields

Uniform rates of convergence in the CLT for quadratic forms in multidimensional spaces

V. Bentkus★, F. Götze★

Fakultät für Mathematik, Universität Bielefeld, Postfach 100131, D-33501 Bielefeld 1, Germany
E-mail addresses: bentkus@mathematik.uni-bielefeld.de; goetze@mathematik.uni-bielefeld.de

Received: 9 January 1997 / In revised form: 15 May 1997

Summary. Let $X, X_1, X_2, \ldots$ be a sequence of i.i.d. random vectors taking values in a d-dimensional real linear space $\mathbb{R}^d$. Assume that $\mathbf{E}\,X = 0$ and that X is not concentrated in a proper subspace of $\mathbb{R}^d$. Let G denote 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 $S_N = N^{-1/2}(X_1 + \cdots + X_N)$ and show that

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

provided that $d \ge 9$ and the fourth moment of X exists. The bound $\mathcal{O}(N^{-1})$ is optimal and improves, e.g., the well-known bound $\mathcal{O}(N^{-d/(d+1)})$ due to Esseen (1945). The result extends to the case of random vectors taking values in a Hilbert space. Furthermore, we provide explicit bounds for Δ_N and for the concentration function of the random variable $\mathbb{Q}[S_N]$.

AMS Subject Classification (1991): Primary 60F05; secondary 62E20

1. Introduction

Let $\mathbb{R}^d$ denote the d-dimensional space of real vectors $x = (x_1, \ldots, x_d)$ with scalar product $\langle \cdot, \cdot \rangle$ and norm $|x|^2 = \langle x, x \rangle = x_1^2 + \cdots + x_d^2$. Since our results are independent of the dimension (provided $d \ge 9$), it will be convenient to

Key words and phrases. Central Limit Theorem, concentration inequalities, convergence rates, Berry–Esseen bounds, Edgeworth expansions, multidimensional spaces, Hilbert spaces, quadratic forms, ellipsoids, hyperboloids, lattice point problem

★ Research supported by the SFB 343 in Bielefeld

denote by $\mathbb{R}^\infty$ a real separable Hilbert space. Thus, $\mathbb{R}^\infty$ consists of all real sequences $x = (x_1, x_2, \ldots)$ such that $|x|^2 = x_1^2 + x_2^2 + \cdots < \infty$.

Let $X, X_1, X_2, \ldots$ be a sequence of i.i.d. random vectors taking values in $\mathbb{R}^d$. If $\mathbf{E}\,X = 0$ and $\mathbf{E}\,|X|^2 < \infty$, then the sums

$$S_N = N^{-1/2}(X_1 + \cdots + X_N)$$

converge weakly to a mean zero Gaussian random vector, say G, such that its covariance operator $\mathbb{C} = \operatorname{cov} G \colon \mathbb{R}^d \to \mathbb{R}^d$ is equal to $\operatorname{cov} X$.

For a linear symmetric and bounded operator $\mathbb{Q} : x \mapsto \mathbb{Q}x$ mapping $\mathbb{R}^d$ into $\mathbb{R}^d$, define the quadratic form $\mathbb{Q}[x] = \langle \mathbb{Q}x, x \rangle$. We call $\mathbb{Q}$ non-degenerate if $\ker \mathbb{Q} = \{0\}$, or equivalently, if $\mathbb{Q}$ is injective. If $d < \infty$, non-degeneracy means that $\mathbb{Q}$ is invertible.

Notice that the distribution of the quadratic form $\mathbb{Q}[G]$ may be represented up to a shift as the distribution of a finite (resp. eventually infinite) weighted sum of squares of i.i.d. standard Gaussian variables, for $d < \infty$ (resp., $d = \infty$).

Write

$$\sigma^2 = \beta_2, \quad \beta_q = \mathbf{E}\,|X|^q, \quad \text{for } q \geq 0 \ .$$

Theorem 1.1. *Let* $\mathbf{E}\,X = 0$. *Assume that* $\mathbb{Q}$ *and* $\mathbb{C}$ *are non-degenerate and that* $d \geq 9$ *or* $d = \infty$. *Then*

$$\sup_x \Big| \mathbf{P}\big\{\mathbb{Q}[S_N] \leq x\big\} - \mathbf{P}\big\{\mathbb{Q}[G] \leq x\big\} \Big| \leq c(\mathbb{Q}, \mathbb{C})\beta_4/N \ .$$

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

Remark. A rather straightforward inspection of the proofs shows that Theorem 1.1 holds for $d = 8$ with $\mathcal{O}(N^{-1} \log^\delta N)$ instead of $\mathcal{O}(N^{-1})$, for some $\delta > 0$. It is likely that the results of the paper remain valid for $d \geq 5$. We need the assumption $d \geq 9$ for estimation of integrals over the characteristic function of the quadratic form for Fourier frequencies $N^{3/5} \ll |t| \ll N$.

An earlier version of results of this paper was published in Bentkus and Götze (1995b).

The bound of Theorem 1.1 is optimal since the distribution function of $|S_N|^2$ (for bounded $X \in \mathbb{R}^d$) may have jumps of order $\mathcal{O}(N^{-1})$, for all $1 \leq d \leq \infty$. See, for example, Bentkus and Götze (1996). In that paper a similar bound $\mathcal{O}(N^{-1})$ as in Theorem 1.1 is proved even for $d \geq 5$ assuming that $\mathbb{Q}$ is diagonal on the subspace spanned by five coordinates of X, which have to be stochastically independent and independent of other coordinates. Both results are based on discretization (i.e., a reduction to lattice valued random vectors) and symmetrization techniques. The independence assumption in Bentkus and Götze (1996) allowed to apply an adaption of the Hardy–Littlewood circle method. For the general case described in Theorem 1.1, we had to develop a new tool – a multiplicative inequality for characteristic functions.

Theorem 1.1 and the method of its proof are related to the well known lattice point problem for conic sections in number theory. Assume that $\mathbb{R}^d$ is finite dimensional and that $\langle \mathbb{Q}x, x\rangle > 0$, for $x \neq 0$. Write $\operatorname{vol} E_s$ for the volume of the ellipsoid

$$E_s = \{x \in \mathbb{R}^d : \mathbb{Q}[x] \le s\}, \quad \text{for } s \ge 0 \ .$$

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

The following result corresponds to Theorem 1.1.

Theorem 1.2 (Bentkus and Götze 1995a, 1997). *For* $d \ge 9$,

$$\sup_{a \in \mathbb{R}^d} \left| \frac{\operatorname{vol}_{\mathbb{Z}}(E_s + a) - \operatorname{vol} E_s}{\operatorname{vol} E_s} \right| = \mathcal{O}\left(\frac{1}{s}\right), \quad \text{for } s \ge 1 \ ,$$

where the constant in $\mathcal{O}(s^{-1})$ *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 $d \ge 9$, and it improves the classical estimate $\mathcal{O}(s^{-d/(d+1)})$ 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. For Hilbert spaces the optimal order of error under the conditions of Theorem 1.1 had been investigated intensively. For a more detailed discussion of the literature on error bounds in probability theory for finite and infinite dimensional spaces and the lattice point problem in number theory, see Bentkus and Götze (1996, 1995a, 1997). Under somewhat more restrictive moment and dimension conditions the estimate $\mathcal{O}(N^{-1+\varepsilon})$, $\varepsilon \downarrow 0$ as $d \uparrow \infty$, was proved in Götze (1979), by a result for bivariate U-statistics. Assuming special smoothness properties, which are satisfied, e.g., by L_p-type functionals of uniform empirical processes, error bounds $\mathcal{O}(N^{-1})$ (and even Edgeworth type expansions) are established in Götze (1979), Bentkus (1984), Bentkus, Götze and Zitikis (1993). Since Theorem 1.1 and more detailed Theorems 1.3–1.5 give a rather complete and explicit solution to the problem for $d < \infty$ and $d = \infty$, it may be helpful to add a few comments on differences between both cases. Error bounds of order $\mathcal{O}(N^{-1/2})$ and better in Theorem 1.1 for general ellipsoids could not be proved via an extension of Esseen's (1945) techniques in $\mathbb{R}^d$ since there is no Lebesgue measure in Hilbert spaces. The symmetrization inequality for characteristic functions introduced in Götze (1979), which is related to Weyl's inequality for trigonometric sums, provided a sufficiently general tool to analyze the infinite dimensional case. An extension of this inequality together with some other new ideas (see below) supplies the basic techniques to prove sharp results *both* in the finite and infinite dimensional cases.

It is likely that the dimensional dependence of our results is not optimal. In order to prove the rate $\mathcal{O}(N^{-1})$ we required that $d \ge 9$. Assumptions like the diagonality of $\mathbb{Q}$, $\mathbb{C}$ and the independence of coordinates allow to reduce the dimension requirement to $d \ge 5$, see Bentkus and Götze (1996). Some yet unpublished results of Götze (1994) indicate that for sums of two independent *arbitrary* quadratic forms (each of rank $d \geq 3$) the rate $\mathcal{O}(N^{-1})$ holds as well. In view of lower bounds of order $\mathcal{O}(N^{-1} \log N)$ for dimension $d = 4$ in the corresponding lattice point problem, an optimal condition would be the assumption $d \geq 5$. To prove (or disprove) that $d \geq 5$ is sufficient for the rate $\mathcal{O}(N^{-1})$ seems to be a difficult problem, since its solution implies a solution of the corresponding unsolved problem for the lattice point remainder for *arbitrary* ellipsoids. The question of precise convergence rates in lower dimensions $2 \leq d \leq 4$ still remains completely open (even in the case $\mathbb{Q} = \mathbb{I}$ and for random vectors with coordinates which are independent Rademacher variables). For instance, in the case $d = 2$ a precise convergence rate would imply a solution of the so called circle problem. Known lower bounds in the circle problem correspond to $\mathcal{O}(N^{-3/4} \log^{\delta} N)$ in our setup. A famous conjecture by Hardy (1916) says that up to logarithmic factors this is the true order.

Introduce the distribution functions

$$F(x) = \mathbf{P}\{\mathbb{Q}[S_N - a] \leq x\}, \quad F_0(x) = \mathbf{P}\{\mathbb{Q}[G - a] \leq x\}, \quad a \in \mathbb{R}^d \ . \tag{1.1}$$

Furthermore, define the Edgeworth correction $F_1(x) = F_1(x; \mathcal{L}(X), \mathcal{L}(G))$ as a function of bounded variation (for $d \geq 9$; see Lemma 5.7) such that $F_1(-\infty) = 0$ and its Fourier–Stieltjes transform is given by

$$\widehat{F}_1(t) = \frac{2(it)^2}{3\sqrt{N}} \mathbf{E}\, \mathrm{e}\{t\mathbb{Q}[Y]\} \Big(3\langle \mathbb{Q}X, Y\rangle \langle \mathbb{Q}X, X\rangle + 2it\langle \mathbb{Q}X, Y\rangle^3 \Big), \quad Y \stackrel{\text{def}}{=} G - a \ . \tag{1.2}$$

In (1.2) and throughout we write $\mathrm{e}\{x\} = \exp\{ix\}$ and assume that all random vectors and variables are independent in aggregate, if the contrary is not clear from the context. Notice that $F_1 = 0$ if $a = 0$ or if $\mathbf{E}\,\langle X, y\rangle^3 = 0$, for all $y \in \mathbb{R}^d$. In particular, $F_1 = 0$ if X is symmetric. For a representation of F_1 as an integral (as in Bhattacharya and Rao 1986) over finite dimensional $\mathbb{R}^d$, see (1.21).

Write

$$\Delta_N = \sup_x |F(x) - F_0(x) - F_1(x)| \ .$$

We shall provide explicit bounds for Δ_N. These bounds yield Theorem 1.1. To formulate the results we need more notation.

We denote by $\sigma_1^2 \geq \sigma_2^2 \geq \cdots$ the eigenvalues of $\mathbb{C}$, counting their multiplicities. We have $\sigma^2 = \sigma_1^2 + \sigma_2^2 + \cdots$ We write $\theta_1^4 \geq \theta_2^4 \geq \cdots$ for the eigenvalues of $(\mathbb{C}\mathbb{Q})^2$.

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

Let $p > 0$ and $\delta \geq 0$. For a random vector $Y \in \mathbb{R}^d$, we introduce the following condition

$$\mathcal{N}(p, \delta, \mathcal{S}, Y) : \mathbf{P}\big\{|Y - e| \leq \delta\big\} \geq p, \quad \text{for all } e \in \mathcal{S} \cup \mathbb{Q}\mathcal{S} \ . \tag{1.3}$$

We shall refer to condition (1.3) as condition $\mathscr{N}(p,\delta,\mathscr{S},Y)=\mathscr{N}(p,\delta,\mathscr{S},Y;\mathbb{Q})$. Notice that, for any Gaussian random vector G with non-degenerate covariance, the condition $\mathscr{N}(p,\delta,\mathscr{S},G)$ holds for any $\mathscr{S}$ and $\delta>0$, with *some* $p>0$. Thus, condition (1.3) for G is just a reformulation of non-degeneracy.

A particular case where explicit lower bounds for p in (1.3) can be given in terms of eigenvalues of $\mathbb{C}$ and $\mathbb{Q}$, is the following: there exists an orthonormal system $\mathscr{S}_o=\{e_1,\dots,e_s\}$ of eigenvectors of $\mathbb{C}$ such that $\mathbb{Q}\mathscr{S}_o$ is again a system of eigenvectors of $\mathbb{C}$. We shall refer to this assumption saying that condition $\mathscr{B}(\mathscr{S}_o,\mathbb{C})=\mathscr{B}(\mathscr{S}_o,\mathbb{C};\mathbb{Q})$ is fulfilled, and we shall write

$$\mathscr{B}(\mathscr{S}_o,\mathbb{C}):\ \lambda_s^2=\min_{e\in\mathscr{S}_o\cup\mathbb{Q}\mathscr{S}_o}\sigma_e^2\ , \tag{1.4}$$

where σ_e^2 denotes the eigenvalue of $\mathbb{C}$ corresponding to the eigenvector e. In particular, such a system $\mathscr{S}_o$ exists provided that $\mathbb{Q}$ and $\mathbb{C}$ are diagonal in a common orthonormal basis, and, if $\mathbb{Q}$ is isometric, we can choose $\mathscr{S}_o$ such that $\lambda_s^2=\sigma_s^2$.

See Lemma 5.5 below for some properties of σ_j, θ_j and λ_j.

Let us introduce truncated random vectors

$$X^\diamond=X\mathbf{I}\{|X|\le\sigma\sqrt{N}\},\quad X_\diamond=X\mathbf{I}\{|X|>\sigma\sqrt{N}\},\quad X^\diamond+X_\diamond=X\ ,$$

and their moments

$$\Lambda_4=\frac{N}{(\sigma\sqrt{N})^4}\mathbf{E}|X^\diamond|^4,\quad \Pi_q=\frac{N}{(\sigma\sqrt{N})^q}\mathbf{E}|X_\diamond|^q\ . \tag{1.5}$$

Define $F_1^\diamond(x)=F_1(x;\mathscr{L}(X^\diamond),\mathscr{L}(G))$ just replacing X by $X^\diamond$ in (1.2), and

$$\Delta_N^\diamond=\sup_x|F(x)-F_0(x)-F_1^\diamond(x)|\ .$$

By c we shall denote generic absolute positive constants. If a constant depends on, say s, then we shall write c_s or $c(s)$.

In Theorems 1.3–1.5 we assume that $\sigma<\infty$ and $\delta=1/300$. Furthermore, in Theorems 1.3–1.6 we assume without loss of generality (see Remark 1.7) that the symmetric operator $\mathbb{Q}$ is isometric, that is, that $\mathbb{Q}^2$ is the identity operator $\mathbb{I}$. Furthermore, we denote c_0 a positive absolute constant, for example one may choose $c_0=1$.

Theorem 1.3. *Let* $\mathbf{E}X=0$, $s=13$ *and* $13\le d\le\infty$. *Then we have*:

(i) *Assume that condition* $\mathscr{N}(p,\delta,\mathscr{S}_o,c_0G/\sigma)$ *holds. Then*

$$\Delta_N^\diamond\le C(\Pi_2+\Lambda_4)(1+|a/\sigma|^6),\quad \Delta_N\le C(\Pi_3+\Lambda_4)(1+|a/\sigma|^6) \tag{1.6}$$

with $C=cp^{-6}+c(\sigma/\theta_8)^8$;

(ii) *Assume that condition* $\mathscr{B}(\mathscr{S}_o,\mathbb{C})$ *is fulfilled. Then the constant in* (1.6) *satisfies* $C\le\exp\{c\sigma^2\lambda_{13}^{-2}\}$.

Theorem 1.4. *Let* X *be symmetric*, $s=9$ *and* $9\le d\le\infty$. *Then* $\Delta_N=\Delta_N^\diamond=\sup_x|F(x)-F_0(x)|$ *and we have*:

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

$$\Delta_N \le C(\Pi_2 + \Lambda_4)(1 + |a/\sigma|^4), \quad C = cp^{-4} \ ; \tag{1.7}$$

(ii) *Assume that condition $\mathcal{B}(\mathcal{S}_o, \mathbb{C})$ is fulfilled. Then the constant in* (1.7) *allows the estimate $C \le \exp\{c\sigma^2/\lambda_9^2\}$.*

Unfortunately, the bounds of Theorem 1.3 are not applicable in dimensions $d = 9, 10, 11, 12$. The following Theorem 1.5 holds for finite-dimensional spaces with $d \ge 9$ only. Compared with Theorem 1.3, the bounds of Theorem 1.5 have a more natural dependence on $|a|$. However, they depend on the smallest eigenvalue σ_d, which makes them unstable in cases where one of coordinates of X degenerates.

Theorem 1.5. *Let $\mathbf{E}X = 0$, $s = 9$ and $9 \le d < \infty$. Then we have:*
(i) *Assume that condition $\mathcal{N}(p, \delta, \mathcal{S}_o, c_0 G/\sigma)$ holds. Then*

$$\Delta_N^\diamond \le C(\Pi_2 + \Lambda_4)(1 + |a/\sigma|^3), \quad \Delta_N \le C(\Pi_3 + \Lambda_4)(1 + |a/\sigma|^3) \ , \tag{1.8}$$

with $C = cp^{-3}(\sigma/\sigma_d)^4$;
(ii) *Assume that condition $\mathcal{B}(\mathcal{S}_o, \mathbb{C})$ is fulfilled. Then the constant in* (1.8) *allows the bound $C \le \frac{\sigma^4}{\sigma_d^4} \exp\left\{c\frac{\sigma^2}{\lambda_9^2}\right\}$.*

Theorems 1.3 and 1.5 yield Theorem 1.1, choosing $a = 0$ and using the bound $\Pi_3 + \Lambda_4 \le \beta_4/(\sigma^4 N)$.

Define the symmetrization $\tilde{X}$ of a random vector X as a random vector such that $\mathcal{L}(\tilde{X}) = \mathcal{L}(X_1 - X_2)$.

Introduce the concentration function

$$Q(X; \lambda) = Q(X; \lambda; \mathbb{Q}) = \sup_{a,x} \mathbf{P}\{x \le \mathbb{Q}[X + a] \le x + \lambda\}, \quad \text{for } \lambda \ge 0 \ .$$

Theorem 1.6. *Assume that $9 \le s \le d \le \infty$ and $0 \le \delta \le 1/(5s)$. Let $Z_N = X_1 + \cdots + X_N$. For any random vector X we have:*
(i) *If condition $\mathcal{N}(p, \delta, \mathcal{S}_o, \tilde{X})$ is fulfilled with some $p > 0$ then*

$$Q(Z_N; \lambda) \le c_s(pN)^{-1} \max\{1; \lambda\}, \quad \lambda \ge 0 \ ; \tag{1.9}$$

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

$$Q(Z_N; \lambda) \le c_s(pN)^{-1} \max\{m; \lambda\}, \quad \lambda \ge 0 \ ; \tag{1.10}$$

(iii) *Assume that $\tilde{X}$ is not concentrated in a proper closed linear subspace of $\mathbb{R}^d$. Then, for any $\delta > 0$ and $\mathcal{S}$ there exists a natural number m such that the condition*

$$\mathcal{N}(p, \delta, \mathcal{S}, m^{-1/2}\tilde{Z}_m) \text{ holds with some } p > 0 \ . \tag{1.11}$$

We say that a random vector Y is concentrated in $\mathbb{L} \subset \mathbb{R}^d$ if $\mathbf{P}\{Y \in \mathbb{L}\} = 1$.

It is interesting to compare the bounds of Theorem 1.6 and of Theorem 2.1 below with the classical bound $\mathcal{O}(N^{-1/2} \max\{1; \lambda\})$ for the concentration of sums (see, e.g., Th. 9 of Ch. 3 in Petrov 1975). Theorem 1.6 and 2.1 are useful for investigations of infinitely divisible approximations, see Bentkus, Götze and Zaitsev (1997).

Remark 1.7. The assumption that the symmetric operator $\mathbb{Q}$ is isometric, i.e., that $\mathbb{Q}^2$ is the identity operator $\mathbb{I}$, simplifies the notation and does not restrict generality. Indeed, any symmetric operator $\mathbb{Q}$ may be decomposed as $\mathbb{Q} = \mathbb{Q}_1\mathbb{Q}_0\mathbb{Q}_1$, where $\mathbb{Q}_0$ is symmetric and isometric and $\mathbb{Q}_1$ is symmetric bounded and non-negative, that is, $\langle \mathbb{Q}_1 x, x\rangle \geq 0$, for all $x \in \mathbb{R}^d$. Thus, for any symmetric $\mathbb{Q}$, we can apply all our bounds replacing the random vector X by $\mathbb{Q}_1 X$, the Gaussian random vector G by $\mathbb{Q}_1 G$,, the shift a by $\mathbb{Q}_1 a$,, etc. In the case of concentration functions, $Q(X; \lambda; \mathbb{Q}) = Q(\mathbb{Q}_1 X; \lambda; \mathbb{Q}_0)$, and we may apply Theorem 1.6 provided $\mathbb{Q}_1 X$ (instead of X) satisfies the conditions.

We conclude the Introduction by a brief description of the basic elements of the proof – a discretization procedure, a double large sieve and multiplicative inequalities.

Let $\varepsilon_1, \varepsilon_2, \ldots$ denote i.i.d. symmetric Rademacher random variables.

Let $\delta > 0$ and $\mathscr{S} = \{e_1, \ldots, e_s\} \subset \mathbb{R}^d$. We say that a discrete random vector $Y \in \mathbb{R}^d$ (or its distribution $\mathscr{L}(Y)$) belongs to the class $\Gamma(\delta; \mathscr{S})$ (briefly $\mathscr{L}(Y) \in \Gamma(\delta; \mathscr{S})$) if 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 $|z_j - e_j| \leq \delta$, for all $1 \leq j \leq s$.

For a bounded and measurable function $H \colon \mathbb{R}^d \to \mathbf{B}$ taking values in a Banach space $(\mathbf{B}, |\cdot|_{\mathbf{B}})$, define the norm $|H|_\infty = \sup_x |H(x)|_{\mathbf{B}}$.

Discretization. Assume that a random vector $W \in \mathbb{R}^d$ is independent of the sum $Z_N = X_1 + \cdots + X_N$ and that the symmetrization $\tilde{X}$ of X satisfies $\mathbf{P}\{|\tilde{X} - e| \leq \delta\} \geq p > 0$, for all $e \in \mathscr{S}$. Then, for any $\gamma \geq 0$ and natural k, with $0 \leq k \leq pN/(4s)$,

$$\left|\mathbf{E} H(2Z_N + W)\right| \leq c_\gamma (pN)^{-\gamma} |H|_\infty + \sup_\Gamma \sup_{b \in \mathbb{R}^d} \left|\mathbf{E} H(Y_1 + \cdots + Y_k + b)\right|$$

(cf. Lemmas 6.1 and 6.2 of Section 6), where $\sup_\Gamma$ is taken over all independent random vectors $Y_1, \ldots, Y_k$ of the class $\Gamma(\delta, \mathscr{S})$.

This discretization allows to reduce the estimation of the characteristic function $\widehat{F}(t) = \mathbf{E}\, e\{t\mathbb{Q}[S_N - a]\}$, $t \in \mathbb{R}$, to the estimation of functions like $\varphi(t) = \mathbf{E}\, e\{t\mathbb{Q}[S'_N]\}$, where $S'_N = N^{-1/2}(Y_1 + \cdots + Y_N)$ is a sum of independent (non-identically distributed!) random vectors of class $\Gamma(\delta, \mathscr{S})$. Notice that norms of the random vectors Y_j are bounded from above by a constant independent of N.

The symmetrization inequality (see Lemma 5.1) reduces the estimation of the function φ to the estimation of the characteristic function

$$\mathbf{E}\, e\{t\langle T_N, T'_N\rangle\} \;, \tag{1.12}$$

where T_N and T'_N denote independent random vectors, which are normalized sums of i.i.d. random vectors of class Γ, taking values in a s-dimensional subspace of $\mathbb{R}^d$.

Let us choose $\delta = 1/(5s)$ and $\mathscr{S} = \mathscr{S}_o$ to be an orthonormal system in $\mathbb{R}^d$. The double large sieve (see Lemma 4.7) allows to estimate (1.12) and to obtain:

$$\left|\varphi(t)\right| = \mathcal{O}(|t|^{-s/2}), \text{ for } |t| \le \sqrt{N}, \quad \left|\varphi(t)\right| = \mathcal{O}\Big(\left|t/N\right|^{s/2}\Big), \text{ for } |t| \ge \sqrt{N} \ .$$

These relations imply that

$$\left|\widehat{F}(t)\right| = \mathcal{O}(|t|^{-s/2}), \text{ for } |t| \le \sqrt{N}, \quad \left|\widehat{F}(t)\right| = \mathcal{O}\Big(\left|t/N\right|^{s/2}\Big), \text{ for } |t| \ge \sqrt{N} \ . \tag{1.13}$$

Since for Gaussian X,

$$\left|\widehat{F}(t)\right| \sim |t|^{-s/2}, \quad \text{for } |t| \ge 1 \ ,$$

the first bound in (1.13) is precise (up to a constant).

The inequalities (1.13) allow to prove in Theorem 1.1 only error bounds $\mathcal{O}(N^{-\alpha})$, for some $\alpha < 1$. This is due to possible oscillations of $|\widehat{F}(t)|$ between 0 and 1, as $|t| \sim N$. The oscillations are restricted by the following *multiplicative inequality* (cf. Lemma 8.1): for any $t \in \mathbb{R}$,

$$\left|\varphi(t)\varphi(t+\delta)\right| = \mathcal{O}(\delta^{-s/2}), \quad \text{for } \ 0 < \delta \le \sqrt{N} \ , \tag{1.14}$$

$$\left|\varphi(t)\varphi(t+\delta)\right| = \mathcal{O}\Big((\delta/N)^{s/2}\Big), \quad \text{for } \delta \ge \sqrt{N} \ . \tag{1.15}$$

Notice that the right-hand sides of (1.14) and (1.15) are independent of t. In other words, (1.14) and (1.15) reflect a certain stationarity in the behavior of the characteristic function φ (and $\widehat{F}$ as well): if $|\varphi(t_0)|$ is sufficiently large, then $|\varphi(t)|$ is bounded from above (near t_0) similarly as it is bounded near $t = 0$. The inequalities (1.14) and (1.15) guarantee that the distance between maxima of $|\varphi(t)|$ has to be sufficiently large, and that the integral of $|\varphi(t)/t|$ around its maxima is $\mathcal{O}(N^{-1})$ provided that $s \ge 9$.

We conclude the Section by introducing the notation used throughout the proofs. In Section 2 we prove bounds for concentration functions. The proofs, being technically simpler as those of Theorems 1.3–1.5, already contain all the principal ideas. In Section 3 Theorems 1.3–1.5 are proved. In Section 4 we extend the well-known double large sieve used in number theory to arbitrary (unbounded) probability distributions. In Section 5 we have collected some simple but useful auxiliary Lemmas. Section 6 contains a description of the discretization of expectations, using random selections. In Section 7 we prove estimates for characteristic functions. The proofs of these estimates are based on conditioning, discretization, as well as on the double large sieve. Section 8 is devoted to the study of the crucial multiplicative inequality for characteristic functions. This is an extension of an inequality for trigonometric sums introduced in Bentkus and Götze (1995a, 1997). Section 9 deals with expansions of characteristic functions.

Notation. We write $\mathrm{e}\{x\} = \exp\{ix\}$.

By $\lceil \alpha \rceil$ we shall denote the integer part of a number α.

By $c, c_1, \dots$ we shall denote generic absolute positive constants. If a constant depends on, say s, then we shall point out the dependence as c_s or $c(s)$. We shall write $A \ll B$, if there exists an absolute constant c such that $A \le cB$. Similarly, $A \ll_s B$, if $A \le c(s)B$.

By $\bar{X}$ and $X_1, X_2, \dots$ we shall denote independent copies of a random vector X, and $\mathscr{L}(X)$ shall denote the distribution of X.

By $\tilde{X}$ we shall denote a symmetrization of X, for example $\tilde{X} = X - \bar{X}$.

For the sake of brevity we shall write throughout

$$\beta = \beta_4, \quad \Pi = \Pi_2, \quad \Lambda = \Lambda_4 \ .$$

We write $Z_N = X_1 + \cdots + X_N$ and $S_N = N^{-1/2} Z_N$.

By $\mathbf{I}\{A\}$ we denote the indicator of an event A.

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.

By $\widehat{F}$ 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 function

$$\mathscr{M}(t; N) = 1\big/\sqrt{|t| N}, \text{ for } |t| \le N^{-1/2}, \quad \mathscr{M}(t; N) = \sqrt{|t|}, \text{ for } |t| \ge N^{-1/2} \ . \tag{1.16}$$

Notice that, for $s > 0$,

$$2^{-1}\Big(|t/N|^{-s/2} + |t|^{s/2}\Big) \le \mathscr{M}^s(t; N) \le |t/N|^{-s/2} + |t|^{s/2} \ . \tag{1.17}$$

Instead of normalized sums S_N, it is more convenient to consider the sums Z_N. Introduce the distribution functions

$$\Psi(x) = \mathbf{P}\{\mathbb{Q}[Z_N - b] \le x\}, \quad \Psi_0(x) = \mathbf{P}\{\mathbb{Q}[\sqrt{N}\, G - b] \le x\} \tag{1.18}$$

with $b = \sqrt{N}\, a$. Define the Edgeworth correction $\Psi_1(x; \mathscr{L}(X), \mathscr{L}(G))$ as a function of bounded variation (for $d \ge 9$; see Lemma 5.7) such that $\Psi_1(-\infty) = 0$. Its Fourier–Stieltjes transform is equal to

$$\widehat{\Psi}_1(t) = -\frac{1}{\sqrt{N}} \mathbf{E}\bigg(\frac{4i}{3} \langle N\mathbb{D}Y, X\rangle^3 + 2\langle N\mathbb{D}Y, X\rangle N\mathbb{D}[X]\bigg)\, \mathrm{e}\{N\mathbb{D}[Y]\} \ , \tag{1.19}$$

where $Y = G - a$ and $\mathbb{D} = t\mathbb{Q}$. Define as well $\Psi_1^\diamond(x) = \Psi_1(x; \mathscr{L}(X^\diamond), \mathscr{L}(G))$ just replacing in (1.19) the random vector X by $X^\diamond$. Recall that the truncated random vectors $X^\diamond$, $X_\diamond$ and their moments are defined in (1.5). In Sections 2, 3 and 9 we shall denote

$$X' = X^\diamond - \mathbf{E} X^\diamond + W \ , \tag{1.20}$$

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 random vector W exists by Lemma 2.4. Finally by $Z_N^\diamond$ (resp. Z_N') we shall denote sums of N independent copies of $X^\diamond$ (resp. X').

In finite dimensional spaces we have the following representation of the Edgeworth correction. Let ϕ 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 we have

$$F_1(x/N) = \Psi_1(x) = \frac{1}{6\sqrt{N}} \chi(A_x/\sqrt{N}), \quad A_x = \{u \in \mathbb{R}^d : \mathbb{Q}[u-b] \le x\} \ , \tag{1.21}$$

with the signed measure

$$\chi(A) = \int_A \mathbf{E} p'''(x) X^3 \, dx, \qquad \text{for measurable } A \subset \mathbb{R}^d \ , \tag{1.22}$$

and where

$$p'''(x)u^3 = p(x)\Big(3\langle \mathbb{C}^{-1}u, u\rangle \langle \mathbb{C}^{-1}x, u\rangle - \langle \mathbb{C}^{-1}x, u\rangle^3\Big) \tag{1.23}$$

denotes the third Frechet derivative of p in the direction u. We can write a similar representation for $\Psi_1^\diamond(x) = \Psi_1(x; \mathscr{L}(X^\diamond), (G))$ just replacing X by $X^\diamond$ in (1.21).

We shall often use the following Rosenthal type inequality. Let $\xi_1, \dots, \xi_N$ denote independent random vectors which have mean zero and assume values in $\mathbb{R}^\infty$. Then

$$\mathbf{E}\Big|\sum_{j=1}^N \xi_j\Big|^q \ll_q \sum_{j=1}^N \mathbf{E}|\xi_j|^q + \Big(\sum_{j=1}^N \mathbf{E}|\xi_j|^2\Big)^{q/2}, \quad 0 < q < \infty \ . \tag{1.24}$$

This inequality easily follows from a result of Acosta (1981) and the fact that Hilbert spaces are type 2 spaces, that is, $\mathbf{E}\big|\sum_{j=1}^N \xi_j\big|^2 \ll \sum_{j=1}^N \mathbf{E}|\xi_j|^2$.

Acknowledgement. We would like to thank A. Zaitsev for helpful discussions.

2. Proofs of bounds for concentration functions; truncation

We start the Section with Theorem 2.1 which (under additional restrictions) provides more explicit bounds for the concentration than those of Theorem 1.6. In the next Theorem we assume that c_0 is an arbitrary positive absolute constant, for example, one can choose $c_0 = 1$. Recall as well that we write $\beta = \beta_4$, $\Pi = \Pi_2$ and $\Lambda = \Lambda_4$.

Theorem 2.1. *Assume that* $9 \le d \le \infty$ *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* $\mathscr{N}(p, \delta, \mathscr{S}_o, c_0 G/\sigma)$ *with* $s = 9$ *and* $\delta = 1/200$. *Then*

$$Q(Z_N; \lambda) \ll p^{-2} \max\{\Pi + \Lambda; \lambda\sigma^{-2}N^{-1}\}, \quad \lambda \ge 0 \ . \tag{2.1}$$

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

(ii) *Let the condition* $\mathscr{B}(\mathscr{S}_o, \mathbb{C})$ (*see* (1.4)) *be fulfilled with* $s = 9$. *Then*

$$Q(Z_N;\lambda) \le \exp\{c\sigma^2/\lambda_9^2\} \max\{\Pi + \Lambda; \lambda\sigma^{-2}N^{-1}\}, \quad \lambda \ge 0 \ , \tag{2.2}$$

where c denotes a sufficiently large absolute constant.

Proof of Theorems 1.6 and 2.1. Below we shall prove the assertions (1.9); (1.9)$\Longrightarrow$(1.10); (1.11); (1.10)$\Longrightarrow$(2.1) and (2.1)$\Longrightarrow$(2.2). □

For the integration of characteristic functions we shall use the following Theorem 2.2. This Theorem slightly extends Lemma 6.1 in Bentkus and Götze (1997). Its proof repeats in essence the proof of the Lemma mentioned, and will be published elsewhere.

Theorem 2.2. *Let* $\varphi(t)$, $t \ge 0$, *denote a continuous function such that* $\varphi(0) = 1$ *and* $0 \le \varphi \le 1$. *Assume that*

$$\varphi(t)\varphi(t+\tau) \le \Theta \mathscr{M}^{2s}(\tau; N), \quad \textit{for all } t \ge 0 \textit{ and } \tau \ge 0 \ ,$$

with some $\Theta \ge 1$ *independent of t and* τ. *Then, for any* $T \ge 1$, $0 < B \le 1$ *and* $N \ge 1$,

$$\int_{B/\sqrt{N}}^{T} \sqrt{\varphi(t)}\, \frac{dt}{t} \ll_s \frac{\Theta(1+\log T)}{N} + \Theta B^{-s/2} N^{-s/4}, \quad \textit{for } s > 8 \ .$$

For $T \ge t_0$, $t_1 \ge 0$, define the integrals

$$I_0 = \int_{-t_1}^{t_1} \big|\widehat{\Psi}(t)\big|\, dt, \quad I_1 = \int_{t_0 \le |t| \le T} |\widehat{\Psi}(t)|\, \frac{dt}{|t|} \ ,$$

where $\widehat{\Psi}$ denotes the Fourier–Stieltjes transform of the distribution function Ψ of $\mathbb{Q}[Z_N - b]$ (see (1.18)).

Lemma 2.3. *Assume the condition* $\mathscr{N}(p, \delta, \mathscr{S}_o, \tilde{X})$ *with some* $0 \le \delta \le 1/(5s)$ *and* $s \ge 9$. *Let*

$$t_0 = c_0(s)(pN)^{-1+2/s}, \quad t_1 = c_1(s)(pN)^{-1/2}, \quad c_2(s) \le T \le c_3(s)$$

with some positive constants $c_j(s)$, $0 \le j \le 3$. *Then*

$$I_0 \ll_s (pN)^{-1}, \quad I_1 \ll_s (pN)^{-1} \ . \tag{2.3}$$

Proof. In the proof we shall denote $k = pN$. Without loss of generality we shall assume that $k \ge c_s$, for a sufficiently large constant c_s. Indeed, if $k \le c_s$, then we can derive (2.3) using $|\widehat{\Psi}| \le 1$. Another consequence of $k \ge c_s$ is that $1/k \le t_0 \le t_1 \le T$.

Let us prove (2.3) for I_0. By Theorem 7.1 we have

$$|\widehat{\Psi}(t)| \ll_s \mathscr{M}^s(t;k), \quad k = pN \ .$$

Since $|\widehat{\Psi}| \le 1$, we have $\big|\widehat{\Psi}(t)\big| \ll_s \min\{1; \mathscr{M}^s(t;k)\}$. Furthermore, denoting $t_2 = k^{-1/2}\max\{1; c_1(s)\}$ and using the definition of the function $\mathscr{M}$, we obtain

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

thus proving (2.3) for I_0.

It remains to estimate I_1. We shall use Theorem 2.2. By Corollary 6.3 we have

$$I_1 = \int_{t_1 \le |t| \le T} |\widehat{\Psi}(t)| \frac{dt}{|t|} \ll_{\gamma,s} \sup_{\Gamma} \sup_{b \in \mathbb{R}^d} I + k^{-\gamma} \log \frac{T}{t_1} \ , \tag{2.4}$$

for any $\gamma > 0$, with

$$I = \int_{t_1 \le |t| \le T} \sqrt{\varphi_n(t/4)}\, \frac{dt}{|t|} = \int_{t_1/4 \le |t| \le T/4} \sqrt{\varphi_n(t)}\, \frac{dt}{|t|}$$

and

$$\varphi_n(t) = \left| \mathbf{E}\, e\{t \mathbb{Q}[Y_n + \mathbb{Q}\mathbb{Y}_n' + b]\} \right|^2, \quad n = \lceil k/(5s) \rceil \ ,$$

where $Y_n = U_1 + \cdots + U_n$ and $Y_n' = U_1' + \cdots + U_n'$ denote sums of independent vectors, and $\sup_\Gamma$ is taken over all $\{\mathscr{L}(U_j), \mathscr{L}(U_j') \colon 1 \le j \ \le n\} \subset \Gamma(\delta; \mathscr{S}_o)$.

Put $\gamma = 2$ in (2.4). Then it remains to show that $I \ll_s 1/k$. By Lemma 8.1,

$$\varphi_n(t)\varphi_n(t - \tau) \ll_s \mathscr{M}^{2s}(\tau; n), \qquad \text{for any } t, \tau \in \mathbb{R} \ .$$

Hence, replacing N by n in Theorem 2.2, we obtain $I \ll_s 1/n \ll_s 1/k$. $\square$

Proof of (1.9). Using a well-known inequality (see, for example Petrov 1975, Lemma 3 of Ch. 3), we have

$$Q(Z_N; \lambda) \le 2 \sup_{a \in \mathbb{R}^d} \max\left\{\lambda; \frac{1}{T}\right\} \int_{-T}^{T} |\widehat{\Psi}(t)|\, dt \ , \tag{2.5}$$

for any $T > 0$. To estimate the integral in (2.5) we shall apply Lemma 2.3. Let us choose $T = 1$. Then using $1 \le 1/|t|$, for $|t| \le 1$, we have

$$\int_{-T}^{T} |\widehat{\Psi}(t)|\, dt \le \int_{|t| \le (pN)^{-1/2}} |\widehat{\Psi}(t)|\, dt + \int_{(pN)^{-1/2} \le |t| \le 1} |\widehat{\Psi}(t)| \frac{dt}{|t|} \overset{\text{def}}{=} I_0 + I_1 \ .$$

Lemma 2.3 implies $I_0 \ll_s 1/(pN)$ and $I_1 \ll_s 1/(pN)$. $\square$

Proof of (1.9) $\Longrightarrow$ (1.10). Without loss of generality we can assume that $N/m \ge 2$. Let $Y_1, Y_2, \ldots$ denote independent copies of $m^{-1/2} Z_m$. Write $W_k = Y_1 + \cdots + Y_k$. Then $\mathscr{L}(Z_N) = \mathscr{L}(\sqrt{m} W_k + b)$ with $k = \lceil N/m \rceil$ and with some b independent of W_k. Consequently, we have $Q(Z_N; \lambda) \le Q(W_k; \lambda/m)$. In order to estimate $Q(W_k; \lambda/m)$ we can apply (1.9) replacing Z_N by W_k. We obtain

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

Proof of (1.11). The existence of the desired $m = m(\mathscr{L}(X))$ can be proved as follows. Let $\tau > 0$. Split the distribution $\mu = \mathscr{L}(\tilde{X}) = uv + v\vartheta$, with the conditional distributions

$$v(A) = \mathbf{P}\{\tilde{X} \in A \big| |\tilde{X}| \le \tau\}, \quad \vartheta(A) = \mathbf{P}\{\tilde{X} \in A \big| |\tilde{X}| > \tau\} \ ,$$

and parameters $u = \mathbf{P}\{|\tilde{X}| \le \tau\}$ and $v = 1 - u$. Denote the minimal closed linear subspace supporting v by $\mathbb{L}_\tau$. For any $e \in \mathbb{R}^d$, the distance $\rho(e, \mathbb{L}_\tau) = \inf\{|x - e| : x \in \mathbb{L}_\tau\} \to 0$ as $\tau \to \infty$, since $\tilde{X}$ is not concentrated in a subspace. Choose and fix $\tau = \tau_0$ such that $\dim L_{\tau_0} \ge 1$ and $\rho(e, \mathbb{L}_\tau) < \delta/2$, for all $e \in \mathscr{S} \cup \mathbb{Q}\mathscr{S}$. The measure v in $\mathbb{L}_{\tau_0}$ has bounded support, and therefore it satisfies the Central Limit Theorem with a limiting Gaussian measure, say γ. Any ball in $\mathbb{L}_{\tau_0}$ with positive radius has positive measure γ since the covariance operator $\operatorname{cov} v = \operatorname{cov} \gamma$ is non-degenerate as an operator in $\mathbb{L}_{\tau_0}$, due to the definition of $\mathbb{L}_{\tau_0}$. Consider the balls $B_e = \{x \in \mathbb{R}^d : |x - e| < \delta\}$ and $B'_e = B_e \cap \mathbb{L}_{\tau_0}$. Writing μ^m for the m-fold convolution of the measure μ, we have

$$\mathscr{L}(m^{-1/2}\tilde{S}_m)(A) = (uv + v\vartheta)^m (m^{1/2}A) \ge u^m v^m (m^{1/2}A) \ ,$$

for any measurable $A \subset \mathbb{R}^d$. Therefore, for sufficiently large $m = m(v, \mathscr{S})$, we obtain

$$\mathscr{L}(m^{-1/2}\tilde{S}_m)(B_e) \ge u^m v^m (m^{1/2}B'_e) \ge 2^{-1}u^m \gamma(B'_e) > 0, \quad \text{for all } e \in \mathscr{S} \cup \mathbb{Q}\mathscr{S} \ ,$$

since the Gaussian measure γ of balls with positive radius is positive. $\square$

Recall that truncated random vectors and moments were defined by (1.5) and (1.20), and that $\mathbb{C} = \operatorname{cov} X = \operatorname{cov} G$. We omit the simple proof of the following Lemma.

Lemma 2.4. *The random vectors $X^\diamond$, $X_\diamond$ satisfy*

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

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

$$2 \operatorname{cov} G_* = 2 \operatorname{cov} X^\diamond = \operatorname{cov} \tilde{X^\diamond}, \quad \langle \operatorname{cov} Wx, x\rangle = \mathbf{E}\langle X_\diamond, x\rangle^2 + \langle \mathbf{E}X^\diamond, x\rangle^2 \ .$$

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

Recall, that $Z_N^\diamond$ denotes a sum of N independent copies of $X^\diamond$.

Lemma 2.5. *Let $\varepsilon > 0$. There exist absolute positive constants c and c_1 such that the condition $\Pi \le c_1 p\delta^2/(\varepsilon^2\sigma^2)$ implies that*

$$\mathscr{N}(p, \delta, \mathscr{S}, \varepsilon G) \Longrightarrow \mathscr{N}(p/4, 4\delta, \mathscr{S}, \varepsilon(2m)^{-1/2}\tilde{Z}_m^\diamond) \ ,$$

for $m \ge c\varepsilon^4\sigma^4 N\Lambda/(p\delta^4)$.

Proof. The result of the Lemma follows from the following relations (2.6)–(2.7), since p, δ, ε in these relations are arbitrary and $\mathbf{E}|\tilde{X^\diamond}|^4 \le 16\mathbf{E}|X^\diamond|^4 = 16N\sigma^4\Lambda$.

For the Gaussian random vector G_* defined in Lemma 2.4, we have

$$\mathscr{N}(2p,\delta,\mathscr{S},\varepsilon G) \Longrightarrow \mathscr{N}(p,2\delta,\mathscr{S},\varepsilon G_*), \qquad \text{provided that } \Pi \le p\delta^2/(2\varepsilon^2\sigma^2) \ . \tag{2.6}$$

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

$$\mathscr{N}(2p,\delta,\mathscr{S},\varepsilon G_*) \Longrightarrow \mathscr{N}(p,2\delta,\mathscr{S},\varepsilon(2m)^{-1/2}\tilde{Z}_m^\diamond) \ . \tag{2.7}$$

Let us prove (2.6). For $e \in \mathbb{R}^d$ define p_e by $2p_e = \mathbf{P}\{|\varepsilon G - e| < \delta\}$. Assuming that $\Pi \le p_e\delta^2/(2\varepsilon^2\sigma^2)$, it suffices to prove that $\mathbf{P}\{|\varepsilon G_* - e| < 2\delta\} \ge p_e$. Replacing δ by δ/ε and e by e/ε, we see that we can assume that $\varepsilon = 1$. Applying Lemma 2.4, we obtain

$$\begin{aligned}\mathbf{P}\{|G_* - e| < 2\delta\} &\ge \mathbf{P}\{|W| + |G - e| < 2\delta\} \ge \mathbf{P}\{|W| < \delta,\ |W| + |G-e| < 2\delta\}\\ &\ge \mathbf{P}\{|W| < \delta \text{ and } |G - e| < \delta\} \ge 2p_e - \mathbf{P}\{|W| \ge \delta\}\\ &\ge 2p_e - \delta^{-2}\mathbf{E}|W|^2 \ge 2p_e - 2\delta^{-2}\sigma^2\Pi \ge p_e \ ,\end{aligned}$$

and (2.6) follows.

Let us prove (2.7). Notice that $\operatorname{cov}(\varepsilon\tilde{X}^\diamond/\sqrt{2}) = \operatorname{cov}(\varepsilon G_*)$. Therefore, to prove (2.7) it suffices to apply Lemma 5.3, replacing in that Lemma X by $\varepsilon\tilde{X}^\diamond/\sqrt{2}$. □

Proof (1.10) $\Longrightarrow$ (2.1). The proof is based on truncation of random vectors.

Recall that we assumed that $s = 9$ and $\delta = 1/200$. By a well known truncation argument, we have

$$\left|\mathbf{P}\{Z_N \in A\} - \mathbf{P}\{Z_N^\diamond \in A\}\right| \le N\mathbf{P}\{|X| > \sigma\sqrt{N}\} \le \Pi \ , \tag{2.8}$$

for any measurable set A, and

$$Q(Z_N,\lambda) \le \Pi + Q(Z_N^\diamond,\lambda) \ . \tag{2.9}$$

Write $K = \varepsilon/\sqrt{2m}$ with $\varepsilon = c_0/\sigma$. Then, by Lemma 2.5, we have

$$\mathscr{N}(p,\delta,\mathscr{S}_o,\varepsilon G) \Longrightarrow \mathscr{N}(p/4,4\delta,\mathscr{S}_o,K\tilde{Z}_m^\diamond) \ , \tag{2.10}$$

provided that

$$\Pi \le c_1 p, \quad m \ge cN\Lambda/p \ . \tag{2.11}$$

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

The non-degeneracy condition (2.10) for $K\tilde{Z}_m^\diamond$ allows to apply (1.10) of Theorem 1.6, and we obtain

$$Q(Z_N^\diamond,\lambda) = Q(K Z_N^\diamond, K^2\lambda) \ll (pN)^{-1}\max\{m; K^2\lambda\} \ , \tag{2.12}$$

for any m such that (2.11) is fulfilled. Choosing the minimal m in (2.11), we obtain

$$Q(Z_N^\diamond,\lambda) \ll p^{-2}\max\{\Lambda; \lambda/(\sigma^2 N)\} \ . \tag{2.13}$$

Combining the estimates (2.9) and (2.13), we conclude the proof. □

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 5.4 and the assumption $\mathcal{B}(\mathcal{S}_o, \mathbb{C})$, the effective lower bound $p \geq \exp\{c\sigma^2/\lambda_9^2\}$ follows. $\square$

3. Proofs of Theorems 1.3–1.5

The proof of Theorem 1.3 is rather complicated. It starts with a truncation of random vectors and an application of the Fourier transform to the functions Ψ and Ψ_j. We shall estimate integrals over the Fourier transforms using results of Sections 2, 5–9. The proof of Theorem 1.4 essentially repeats with certain simplifications the proof of Theorem 1.3. For the proof of Theorem 1.5 we shall apply in addition some elements of the standard techniques used in the case of the CLT in multidimensional spaces (cf. e.g., Bhattacharya and Rao 1986).

We shall use the following approximate and precise formulas for the Fourier inversion. A smoothing inequality of Prawitz (1972) implies (see Bentkus and Götze 1996, Section 4) that

$$F(x) = \frac{1}{2} + \frac{i}{2\pi} \text{ V.P.} \int_{-K}^{K} \mathrm{e}\{-xt\}\widehat{F}(t)\, \frac{dt}{t} + R \ , \tag{3.1}$$

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

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

Here 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 $F\colon \mathbb{R} \to \mathbb{R}$ 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 holds (see, e.g., Chung 1974)

$$F(x) = \frac{1}{2}F(\infty) + \frac{i}{2\pi} \lim_{M\to\infty} \text{ V.P.} \int_{|t|\leq M} \mathrm{e}\{-xt\}\widehat{F}(t)\, \frac{dt}{t} \ . \tag{3.2}$$

The formula is well-known for distribution functions. For functions of bounded variation, it extends by linearity arguments.

In this Section we shall assume that the following conditions are fulfilled

$$\mathbb{Q}^2 = \mathbb{I}, \quad \sigma^2 = 1, \quad 9 \leq s \leq 13, \quad \delta = 1/300, \quad \mathcal{N}(p, \delta, \mathcal{S}_o, c_0 G) \ . \tag{3.3}$$

Notice that the assumption $\sigma^2 = 1$ does not restrict generality since from Theorems 1.3–1.5 with $\sigma = 1$ we can derive the general result replacing X by X/σ, G/σ, etc. Other assumptions in (3.3) are included as conditions in Theorems 1.3–1.5. The assumption $\sigma^2 = 1$ yields (recall that we write $\Pi = \Pi_2$ and $\Lambda = \Lambda_4$)

$$N^{-1} \le 2(\Pi + \Lambda), \quad \Pi + \Lambda \le 1, \quad \sigma_j \le 1, \quad \lambda_j \le 1, \quad \theta_j \le 1 \ . \tag{3.4}$$

Recall that functions Ψ and Ψ_j are defined by (1.18) and (1.19). In that notation we have

$$\Delta_N = \sup_{x\in\mathbb{R}} |\Delta_N(x)|, \quad \Delta_N^\diamond = \sup_{x\in\mathbb{R}} |\Delta_N^\diamond(x)| \ , \tag{3.5}$$

where

$$\Delta_N(x) = \Psi(x) - \Psi_0(x) - \Psi_1(x), \quad \Delta_N^\diamond(x) = \Psi(x) - \Psi_0(x) - \Psi_1^\diamond(x) \tag{3.6}$$

since the functions F, F_0 and F_1 defined by (1.1)–(1.2) satisfy

$$F(x) = \Psi(xN), \quad F_j(x) = \Psi_j(xN), \quad \widehat{F}(tN) = \widehat{\Psi}(t), \quad \widehat{F}_j(tN) = \widehat{\Psi}_j(t) \ . \tag{3.7}$$

Reduction of Theorems 1.3–1.5 *to the proof that*

$$\Delta_N^\diamond \ll (p^{-6} + \theta_8^{-8})(\Pi + \Lambda)(1 + |a|^6), \quad s = 13 \ , \tag{3.8}$$

$$\Delta_N^\diamond \ll p^{-4}(\Pi + \Lambda)(1 + |a|^4), \quad s = 9 \ , \tag{3.9}$$

$$\Delta_N^\diamond \le p^{-3}\sigma_d^{-4}(\Pi + \Lambda)(1 + |a|^3), \quad s = 9 \ , \tag{3.10}$$

respectively.

We have to prove that the bounds (3.8)–(3.10) imply the desired bounds for Δ_N, and to show that the assertions (i) of Theorems 1.3–1.5 imply (ii).

To derive the bounds for Δ_N it suffices to note that $\Pi \le \Pi_3$ and to verify that

$$\sup_x |\Psi_1(x) - \Psi_1^\diamond(x)| \ll \theta_8^{-6}(1 + |a|^3)\Pi_3 \tag{3.11}$$

in the case of Theorem 1.3, and that

$$\sup_x |\Psi_1(x) - \Psi_1^\diamond(x)| \ll \sigma_d^{-3}\Pi_3 \tag{3.12}$$

in the case of Theorem 1.5. But the bound (3.11) is implied by (5.5) of Lemma 5.7 with $s = 8$. The bound (3.12) one can easily prove using the representation (1.21) of the Edgeworth correction as a signed measure in finite dimensional $\mathbb{R}^d$ and estimating the variation of that measure. Indeed, using (1.21), we have

$$\sup_x |\Psi_1(x) - \Psi_1^\diamond(x)| \ll N^{-1/2} I, \quad I \stackrel{\text{def}}{=} \int_{\mathbb{R}^d} \left| \mathbf{E} p'''(x) X^3 - \mathbf{E} p'''(x) X^{\diamond 3} \right| dx \ .$$

By the explicit formula (1.23), the function $u \mapsto p'''(x)u^3$ is a 3-linear form in the variable u. Therefore, using $X = X^\diamond + X_\diamond$ and $|X^\diamond||X_\diamond| = 0$, we have $p'''(x)X^3 - p'''(x)X^{\diamond 3} = p'''(x){X_\diamond}^3$, and

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

whence (3.12) follows.

Let us prove (i)$\Longrightarrow$(ii). This follows from $p \ge \exp\{c\lambda_{13}^{-2}\}$ and the obvious inequalities $\theta_8 \ge \theta_{13} \ge \lambda_{13}$. To obtain $p \ge \exp\{c\lambda_{13}^{-2}\}$ we can use (i) in the case when the condition $\mathscr{N}(p, \delta, \mathscr{S}_o, c_0 G)$ is fulfilled with $c_0 = \delta/4 = 1/1200$. Indeed, the condition $\mathscr{B}(\mathscr{S}_o, \mathbb{C})$ guarantees that $e \in \mathscr{S}_o \cup \mathbb{Q}\mathscr{S}_o$ are eigenvectors of the covariance operator $\mathbb{C}$, and we can get the lower bound for p by an application of Lemma 5.4 using $c_0 = \delta/4$. $\square$

While proving (3.8)–(3.10) we can and shall assume that

$$\Pi \le cp, \quad \Lambda \le cp \ , \tag{3.13}$$

and in the case of (3.10),

$$\Pi \le cp\sigma_d^4, \quad \Lambda \le cp\sigma_d^4 \ , \tag{3.14}$$

with a sufficiently small positive absolute constant c. These assumptions do not restrict generality. Indeed, in the case of (3.9) the symmetry of X implies that $\Psi_1^\diamond = 0$, and we have $\Delta_N^\diamond \le 1$. Thus, if at least one of the assumptions (3.13) is not fulfilled, we obviously obtain (3.9). In the case of (3.8) we can estimate $|\Psi_1^\diamond|$ using (5.4), and we get

$$\Delta_N^\diamond \le 1 + |\Psi_1^\diamond| \ll (1 + |a|^3)\theta_8^{-6} \ ,$$

which again allows to assume (3.13) (notice that $\theta_8^{-6} p^{-1} \le p^{-4} + \theta_8^{-8}$).

If the assumption (3.14) does not hold, the estimate

$$|\Psi_1^\diamond| \ll_d N^{-1/2}\mathbf{E}|\mathbb{C}^{-1/2}X^\diamond|^3 \ll_d \sigma_d^{-2}\Lambda^{1/2} \ . \tag{3.15}$$

immediately implies (3.10). For a proof of (3.15) we can use the representation (1.21) of the Edgeworth correction as a signed measure. Estimating the variation of that measure and using

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

we obtain (3.15).

Recall that the random vectors $X^\diamond$, X' and sums $Z_N^\diamond$, Z_N' of their independent copies are defined in (1.5) and (1.20). Write $\Psi^\diamond$ for the distribution function of $\mathbb{Q}[Z_N^\diamond - b]$. For $0 \le k \le N$ introduce the distribution function

$$\Psi^{(k)}(x) = \mathbf{P}\Big\{\mathbb{Q}[G_1 + \cdots + G_k + X_{k+1}' + \cdots + X_N' - b] \le x\Big\} \ . \tag{3.16}$$

Notice that $\Psi^{(0)} = \Psi'$, $\Psi^{(N)} = \Psi_0$.

Lemma 3.1. *Assume that $\Pi \le c_1 p$ and that a number $1 \le m \le N$ satisfies $m \ge c_2 N\Lambda/p$, with some sufficiently small (resp. large) positive absolute constant c_1 (resp. c_2). Let c_3 be an absolute constant. Write*

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

Let F denote any of the functions $\Psi^\diamond$, $\Psi^{(k)}$ or Ψ_0. Then we have

$$F(x) = \frac{1}{2} + \frac{i}{2\pi}\text{V.P.} \int_{|t|\le t_1} \mathrm{e}\{-xtK\}\widehat{F}(tK)\,\frac{dt}{t} + R_1 \ , \tag{3.17}$$

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

$$\Psi_1^\diamond(x) = \frac{i}{2\pi}\int_{|t|\le t_1} \mathrm{e}\{-xtK\}\widehat{\Psi}_1^\diamond(tK)\,\frac{dt}{t} + R_2 \tag{3.18}$$

with $|R_2| \ll \theta_8^{-8}(\Lambda + pm/N)(1+|a|^3)$.

Proof. Let us prove (3.17). For the proof we shall combine (3.1) and Lemma 2.3. Changing the variable $t = \tau K$ in the approximate Fourier inversion formula (3.1), we obtain

$$F(x) = \frac{1}{2} + \frac{i}{2\pi}\ \mathrm{V.P.} \int_{|t|\le 1} \mathrm{e}\{-xtK\}\widehat{F}(tK)\,\frac{dt}{t} + R\ , \tag{3.19}$$

where

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

Notice all functions $\Psi^\diamond$, $\Psi^{(k)}$, Ψ_0 are distribution functions of the following type of random variables:

$$\mathbb{Q}[U+T], \quad U \stackrel{\mathrm{def}}{=} G_1 + \cdots + G_k + X_{k+1}^\diamond + \cdots + X_N^\diamond\ ,$$

with some $0 \le k \le N$, where the random vector T is independent of $X_j^\diamond$ and G_j, for all j. Let us consider two alternative cases: $k \ge N/2$ and $k < N/2$.

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

$$K^{1/2}U \stackrel{\mathscr{D}}{=} Y_1 + \cdots + Y_l + T_1 \tag{3.20}$$

with $l = \lceil N/(2m) \rceil$ and some random T_1 independent of $Y_1, \ldots, Y_l$. By Lemma 2.5 we have

$$\mathscr{N}(p, \delta, \mathscr{S}, c_0 G) \Longrightarrow \mathscr{N}(p/4, 4\delta, \mathscr{S}, \tilde{Y}) \tag{3.21}$$

provided that

$$\Pi \le c_1 p \quad \text{and} \quad m \ge c_2 N\Lambda/p\ . \tag{3.22}$$

The inequalities in (3.22) are just conditions of the Lemma. Due to (3.20) and (3.21), we can use Lemma 2.3 in order to estimate the integrals in (3.19). Replacing in that Lemma X by Y and N by l, we obtain (3.17) in the case $k < N/2$.

The case $k \ge N/2$. We can proceed as in the previous case defining however Y as a sum of m independent copies of $K^{1/2}G$. The condition (3.21) is fulfilled since now $\mathscr{L}(\tilde{Y}) = \mathscr{L}(c_0 G/\sigma)$, and (3.17) follows.

To prove (3.18) we can apply the Fourier inversion (3.2) to the function $\Psi_1^\diamond$. Using the estimate (5.6) of Lemma 5.7 with $s = 8$ and $t_* = t_1 K$, we obtain

$$\int_{|t|\ge t_1} |\Psi_1^\diamond(tK)|\,\frac{dt}{|t|} \ll \Lambda^{1/2}(t_1 KN)^{-1}(1+|a|^3)\theta_8^{-8}\ .$$

Since $(t_1 KN)^{-1} \ll (pm/N)^{1/2}$, we conclude the proof of (3.18) by an application of the arithmetic-geometric mean inequality. $\square$

Let us introduce the following upper bound $\varkappa = \varkappa(t) = \varkappa(t; N, \mathscr{L}(X), \mathscr{L}(G))$ of for the characteristic function of quadratic forms (cf. Bentkus 1984, Bentkus, Götze and Zitikis 1993). We define $\varkappa = \varkappa(t; N, \mathscr{L}(X)) + \varkappa(t; N, \mathscr{L}(G))$, where

$$\varkappa(t; N, \mathscr{L}(X)) = \sup_{a \in \mathbb{R}^d} \left| \mathbf{E}\, \mathrm{e}\{t\mathbb{Q}[Z_k] + \langle a, Z_k \rangle\} \right|, \quad Z_k = X_1 + \cdots + X_k \ , \tag{3.23}$$

with $k = [(N-2)/14]$. Notice that if $\mathbb{Q}^2 = \mathbb{I}$, then we can replace $t\mathbb{Q}[Z_k] + \langle a, Z_k \rangle$ by $t\mathbb{Q}[Z_k - a]$ in the definition (3.23).

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

$$\int_{|t| \le t_1} (|t|K)^\alpha \varkappa(tK; N, \mathscr{L}(X^\diamond), \mathscr{L}(G)) \frac{dt}{|t|} \ll_\alpha (pN)^{-\alpha}, \qquad \text{for } \ 0 \le \alpha < s/2 \ .$$

Proof. By (3.21), the condition $\mathscr{N}(p/4, 4\delta, \mathscr{S}, K^{1/2}\tilde{Z}_m^\diamond)$ is fulfilled. Therefore, collecting independent copies of $K^{1/2}X^\diamond$ in groups as in (3.20), we can apply Theorem 7.1. We obtain

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

A similar upper bound holds for $\varkappa(tK; N, \mathscr{L}(G))$ (cf. the proof of (3.17) in the case of $k > N/2$. Using the definition of the function $\mathscr{M}(\cdot,\cdot)$ and (1.16) in order to get rid of absolute constants, we get

$$\varkappa(tK; N, \mathscr{L}(X^\diamond), \mathscr{L}(G)) \ll_s \min\left\{1; (m/(tpN))^{s/2}\right\}, \qquad \text{for } |t| \le t_1 \ .$$

Integrating that bound (cf. the estimation of I_1 in Lemma 2.3), we conclude the proof of the Lemma. $\square$

Reduction of (3.8)–(3.10) *to an estimation of*

$$\Delta_N' = \sup_x |\Psi'(x) - \Psi_0(x) - \Psi_1^\diamond(x)| \ , \tag{3.24}$$

where Ψ' is the distribution function of $\mathbb{Q}[Z_N' - b]$. It suffices to prove that the quantity Δ_N' satisfies inequalities of type (3.8)–(3.10). Indeed, let us prove that

$$\sup_x |\Psi(x) - \Psi'(x)| \ll p^{-2}(\Pi + \Lambda)(1 + |a|^2) \ . \tag{3.25}$$

Using truncation (cf. (2.8)), we have $|\Psi - \Psi^\diamond| \le \Pi$, and

$$\sup_x |\Psi(x) - \Psi'(x)| \le \Pi + \sup_x |\Psi^\diamond(x) - \Psi'(x)| \ . \tag{3.26}$$

In order to estimate $|\Psi^\diamond - \Psi'|$, we shall apply Lemmas 3.1 and 3.2. The number m in these Lemmas exists, as it follows from the second inequality in (3.13). Let us choose the minimal m, that is, $m \sim N\Lambda/p$. Then

$(pN)^{-1}m \ll \Lambda/p^2$ and $m/N \ll \Lambda/p$. Therefore, using (3.5), (3.6) and Lemma 3.1 we have

$$\sup_x \left|\Psi^\diamond(x) - \Psi'(x)\right| \ll p^{-2}\Lambda + \int_{|t|\le t_1} \left|\widehat{\Psi}^\diamond(\tau) - \widehat{\Psi}'(\tau)\right| \frac{dt}{|t|}, \quad \tau = tK \ . \tag{3.27}$$

We shall prove that

$$|\widehat{\Psi}^\diamond(\tau) - \widehat{\Psi}'(\tau)| \ll \varkappa\Pi|\tau|N(1+|\tau|N)(1+|a|^2) \ , \tag{3.28}$$

with $\varkappa = \varkappa(\tau; N, \mathscr{L}(X^\diamond))$. Combining (3.26)–(3.28), using $\tau = tK$ and integrating the inequality (3.28) with help of Lemma 3.2, we derive (3.25).

Let us prove (3.28). Recall that $X' = X^\diamond - \mathbf{E}X^\diamond + W$, 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 2.4). Writing $D = Z_N^\diamond - E\,Z_N^\diamond - b$, we have

$$Z_N^\diamond - b = D + \mathbf{E}\,Z_N^\diamond, \quad Z_N' \overset{\mathscr{D}}{=} D + \sqrt{N}W \ ,$$

and $|\widehat{\Psi}^\diamond(\tau) - \widehat{\Psi}'(\tau)| \le |f_1(\tau)| + |f_2(\tau)|$ with

$$\begin{aligned} f_1(\tau) &= \mathbf{E}\,e\{\tau\mathbb{Q}[D+\sqrt{N}W]\} - \mathbf{E}\,e\{\tau\mathbb{Q}[D]\}, \\ f_2(t) &= \mathbf{E}\,e\{\tau\mathbb{Q}[D+\mathbf{E}\,Z_N^\diamond]\} - \mathbf{E}\,e\{\tau\mathbb{Q}[D]\} \ . \end{aligned} \tag{3.29}$$

We have to show that both $|f_1(t)|$ and $|f_2(t)|$ are bounded from above by the right hand side of (3.28). Let us consider f_1. We can write $\mathbb{Q}[D+\sqrt{N}W] = \mathbb{Q}[D] + A + B$ with $A = 2\sqrt{N}\langle\mathbb{Q}D, W\rangle$ and $B = N\mathbb{Q}[W]\}$. Taylor expansions of the exponent in (3.29) 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 (notice that $\mathbf{E}\,W = 0$)

$$|f_1(\tau)| \ll \varkappa|\tau|N\,\mathbf{E}|W|^2 + \varkappa\tau^2 N\,\mathbf{E}|W|^2\mathbf{E}|D|^2 \ , \tag{3.30}$$

where $\varkappa = \varkappa(\tau; N, \mathscr{L}(X^\diamond))$. The estimation of the remainders of these expansions is based on the splitting and conditioning techniques described in Section 9. Using $\sigma = 1$, $\mathbf{E}|W|^2 \ll \Pi$ and $\mathbf{E}|D|^2 \ll N(1+|a|^2)$, we derive from (3.30) that

$$|f_1(\tau)| \ll \varkappa\Pi|\tau|N(1+|\tau|N))(1+|a|^2) \ . \tag{3.31}$$

Expanding in powers of $\mathbf{E}\,Z_N^\diamond = N\mathbf{E}X^\diamond$ and proceeding similarly to the proof of (3.31), we obtain

$$|f_2(t)| \ll \varkappa\Pi|\tau|N(1+|a|) \ ,$$

which concludes the proof of (3.28). $\square$

Proof of Theorems 1.3 and 1.4. Relations (3.8), (3.24) and (3.25) reduce the proof of Theorem 1.3 to showing that Δ_N' is bounded by the right hand side of (3.8), assuming that $s = 13$. Similarly, for the proof of Theorem 1.4 we have to show that Δ_N' is bounded by the right hand side of (3.9), assuming $s = 9$ and symmetry of X.

We shall apply Lemmas 3.1 and 3.2. Choosing $m \sim N\Lambda/p$ as in the proof of (3.25), and using (3.5), (3.6), (3.16), (3.24) and Lemma 3.1 we have

$$\Delta'_N \ll I + (p^{-2} + \theta_8^{-8})\Lambda(1 + |a|^3) \tag{3.32}$$

with

$$I = \int_{|t| \le t_1} \big|\widehat{\Delta}'_N(\tau)\big| \frac{dt}{|t|}, \quad \tau = tK \ .$$

Define the Edgeworth correction $\Psi'_1(x) = \Psi_1(x; \mathscr{L}(X'), (G))$ by replacing X by X' in the definition (1.19). We have

$$\big|\widehat{\Delta}'_N\big| \le |\widehat{\Psi}' - \widehat{\Psi}_0(\tau) - \widehat{\Psi}'_1| + |\widehat{\Psi}'_1 - \widehat{\Psi}^\diamond_1| \ .$$

Below we shall prove that

$$|\widehat{\Psi}'(\tau) - \widehat{\Psi}_0 - \widehat{\Psi}'_1(\tau)| \ll \varkappa(\Pi + \Lambda)\tau^2 N^2 (1 + \tau^4 N^4)(1 + |a|^6) \ , \tag{3.33}$$

$$|\widehat{\Psi}'_1(\tau) - \widehat{\Psi}^\diamond_1(\tau)| \ll \varkappa(\Pi + \Lambda)\tau^2 N^2 (1 + |\tau| N)(1 + |a|^3) \ , \tag{3.34}$$

with $\varkappa = \varkappa(\tau; N, \mathscr{L}(X^\diamond), \mathscr{L}(G))$. Using $\tau = tK$ and integrating the inequalities (3.33)–(3.34) with the help of Lemma 3.2, we derive

$$I \ll p^{-6}(1 + |a|^6)(\Pi + \Lambda) \ ,$$

which combined with (3.32) shows that Δ'_N is bounded from above by the right hand side of (3.8), thus proving Theorem 1.3. Notice that the requirement $s = 13$ was needed for the integration of (3.33) only since the highest power of τ in (3.33) is 6; in all other parts of the proof the requirement $s = 9$ suffices.

In the symmetric case we have $\Psi^\diamond_1 = \Psi'_1 = 0$, and we can repeat the previous proof. Instead of (3.33)–(3.34) we shall prove that

$$|\widehat{\Psi}'(\tau) - \widehat{\Psi}_0(\tau)| \ll \varkappa(\Pi + \Lambda)\tau^2 N^2 (1 + \tau^2 N^2)(1 + |a|^4) \ . \tag{3.35}$$

An integration of this bound shows that Δ'_N is bounded from above by the right hand side of (3.9), thus proving Theorem 1.4. The highest power of τ in (3.35) is 4, and for the integration the assumption $s = 9$ is sufficient.

Thus it remains to prove (3.33)–(3.35).

Let us prove (3.33). Recall that $\sigma = 1$. Write $\beta'_q = \mathbf{E}|X'|^q$ and $\beta' = \beta'_4$. The covariances of X' and G are equal, and we can apply Lemma 9.2. Replacing in that Lemma X by X', we have

$$\big|\widehat{\Psi}'(\tau) - \widehat{\Psi}_0(\tau) - \Psi'_1(\tau)\big| \ll \varkappa\tau^2 N \beta' (1 + \tau^4 N^4)(1 + \beta'_6/N^2)(1 + |a|^6) \ , \tag{3.36}$$

where $\varkappa = \varkappa(\tau; \mathscr{L}(X'), \mathscr{L}(G))$. Since $X' = X^\diamond - \mathbf{E}X^\diamond + W$, and W is independent of $X^\diamond$, we obtain

$$\varkappa(\tau; N, \mathscr{L}(X'), \mathscr{L}(G)) \le \varkappa(\tau; N, \mathscr{L}(X^\diamond), \mathscr{L}(G)) \ .$$

Furthermore, for $q \ge 2$ we have

$$\beta'_q \ll_q \mathbf{E}|X^\diamond|^q + \mathbf{E}|W|^q \ll_q N^{(q-2)/2}$$

since $\sigma = 1$. Similarly, $\beta' \ll (\Pi + \Lambda)N$, and (3.36) yields (3.33).

The proof of (3.35) repeats the proof of (3.33), replacing however Lemma 9.2 by Lemma 9.1.

Let us prove (3.34). The proof is based on the observation that $X' = X^\diamond - \mathbf{E}X^\diamond + W$ and that the Fourier–Stieltjes transforms $\widehat{\Psi}_1^\diamond$ and $\widehat{\Psi}_1'$ are 3-linear forms in X and X' respectively. For the proof it suffices to use $\sqrt{N}G \overset{\mathscr{D}}{=} G_1 + \cdots + G_N$, standard splitting-conditioning and symmetrization techniques, the bounds $|\mathbf{E}X^\diamond| \le N^{-1/2}\Pi$ and $\mathbf{E}|W|^2 \ll \Pi$. We omit related technicalities, and refer to Lemmas 5.7 and 9.1, where similar proofs are carried out in detail. $\square$

Proof of Theorem 1.5. Again, due to (3.10) and (3.24) we have to verify that Δ_N' is bounded by the right hand side of (3.10), that is, that

$$\Delta_N' \ll p^{-3}\sigma_d^{-4}(\Pi + \Lambda)(1 + |a|^3) \ . \tag{3.37}$$

Recall that the distribution function $\Psi^{(k)}$ is defined in (3.16). For any $1 \le k \le N$, we have

$$\Delta_N' \le I_1 + I_2 + I_3, \qquad I_1 = \sup_x |\Psi'(x) - \Psi^{(k)}(x)| \ ,$$

$$I_2 = \sup_x |\Psi^{(k)}(x) - \Psi_0(x) - \Psi_1'(x)|, \qquad I_3 = \sup_x |\Psi_1'(x) - \Psi_1^\diamond(x)| \ . \tag{3.38}$$

Below we shall prove that

$$I_1 \ll p^{-2}\Lambda + p^{-3}kN^{-2}\left(\beta' + \sqrt{N\beta'}\right)(1 + |a|^3) \ , \tag{3.39}$$

$$I_2 \ll_d \frac{k\sqrt{\beta'}}{N^{3/2}\sigma_d^2} + \frac{\beta'}{\sigma_d^4 N} + \frac{\beta'^{(d+7)/2}N^{(d+3)/2}}{k^{d+5}\sigma_d^{2d+14}} \tag{3.40}$$

$$I_3 \ll \sigma_d^{-3}(\Pi + \Lambda) \tag{3.41}$$

with $\beta' = \mathbf{E}|X'|^4$. Let us choose $k \sim \sigma_d^{-2}\sqrt{N\beta'}$. Such $k \le N$ exists since we assumed (3.13), (3.14) and $\beta' \ll N(\Pi + \Lambda)$. Thus, (3.38)–(3.41) yield

$$\Delta_N' \ll p^{-3}\sigma_d^{-4}\left(\Pi + \Lambda + \frac{\beta'}{N} + \frac{\beta'^{3/2}}{N^{3/2}}\right)(1 + |a|^3) \quad ,$$

and (3.37) follows by an application of $\beta' \ll N(\Pi + \Lambda)$ and $\Pi + \Lambda \le 1$.

Let us prove (3.39). As in the proof of Theorem 1.3, applying Lemma 3.1 we obtain

$$I_1 \ll p^{-2}\Lambda + \int_{|t| \le t_1} |\widehat{\Psi}'(\tau) - \widehat{\Psi}^{(k)}(\tau)| \frac{dt}{|t|}, \quad \tau = tK \ .$$

Applying Lemma 9.3 and replacing in that Lemma X by X', t by τ and β by $\beta' = \mathbf{E}|X'|^4$, we have

$$|\widehat{\Psi}'(\tau) - \widehat{\Psi}^{(k)}(\tau)| \ll \varkappa\tau^2 k\left(\beta' + |\tau|N\beta' + |\tau|N\sqrt{N\beta'}\right)(1 + |a|^3) \ .$$

Integrating with the help of Lemma 3.2 we obtain (3.39).

Let us prove (3.40). Define the measure χ' by replacing X by X' in (1.22). Using representations of type (1.21) for Ψ_1' and $\Psi_1^\diamond$, we have

$$I_2 \le I_4 + I_5$$

with

$$I_4 = \sup_x |\Psi^{(k)}(x) - \Psi_0(x) - \frac{N-k}{6N^{3/2}} \chi'(A_x/\sqrt{N})|, \quad I_5 = \frac{k}{6N^{3/2}} \sup_{A \subset \mathbb{R}^d} |\chi'(A)| ,$$

where $A_x = \{u \in \mathbb{R}^d : \mathbb{Q}[u-b] \le x\}$.

Write $Z'_{kN} = \sum_{j=1}^k G_j + \sum_{j=k+1}^N X'_j$. A re-normalization of random vectors implies

$$I_4 \le \delta'_N \stackrel{\text{def}}{=} \sup_{A \subset \mathbb{R}^d} \left| \mathbf{P}\{Z'_{kN} \in \sqrt{N}A\} - \mathbf{P}\{Z'_{NN} \in \sqrt{N}A\} - \frac{N-k}{6N^{3/2}} \chi'(A) \right| .$$

To estimate δ'_N we can apply Lemma 9.4 with X_j replaced by X'_j. We get

$$\delta'_N \ll_d \frac{\beta'}{\sigma_d^4 N} + \frac{\beta'^{(d+7)/2} N^{(d+3)/2}}{k^{d+5} \sigma_d^{2d+14}} .$$

Using a representation of type (1.22) for χ' and estimating the variation of the signed measure, we obtain

$$I_5 \ll_d kN^{-3/2} \sigma_d^{-2} \sqrt{\beta'} .$$

Collecting these bounds, we obtain (3.40).

It remains to verify (3.41). Using representations of type (1.21) for the Edgeworth corrections, we have

$$I_3 = \sup_x |\Psi_1'(x) - \Psi_1^\diamond(x)| \le N^{-1/2} \sup_{A \subset \mathbb{R}^d} |\chi'(A) - \chi^\diamond(A)| .$$

Both χ' and $\chi^\diamond$ are 3-linear forms in X' and $X^\diamond$ respectively. Using $X' = X^\diamond - \mathbf{E}X^\diamond + W$ and estimating the variations of the signed measures, we arrive at (3.41). $\square$

4. An extension of the double large sieve to unbounded distributions

The large sieves of Linnik (1941) (see also Graham and Kolesnik 1991) play a key role in some problems of analytic number theory. The main result of the Section–Lemma 4.1 extends the double large sieve to (unbounded) random vectors. Lemma 4.7 presents an application of this bound to sums of i.i.d. vectors in $\mathbb{R}^d$. We need this lemma for the proof of the main result of this paper.

In this section we shall assume that $\mathbb{R}^d$ is finite dimensional, i. e., $d < \infty$.

Let $|x|_\infty$ denote the max-norm $|x|_\infty = \max_{1 \le j \le d} |x_j|$.

Lemma 4.1. *Let m denote an integer such that $m > d/2$. Assume that the independent random vectors $U, V \in \mathbb{R}^d$ are sums of independent random vectors U_j and V_j with non-negative characteristic functions,*

$$U = U_1 + \cdots + U_{2m+2} \text{ and } V = V_1 + \cdots + V_{2m+2} \ .$$

Let $R > 0$ and $T > 0$ denote some positive real numbers. Write

$$A = 1 + \mathbf{E}|U/T|^{2m}, \quad B = 1 + \mathbf{E}|V/R|^{2m} \ ,$$

and

$$C = 1 + (TR)^{-2m}, \quad D = (TR)^d \ .$$

Then we have

$$\mathbf{E}\, \mathrm{e}\{\langle U, V\rangle\} \ll_{m,d} A\, B\, C\, D \max_{1\le a\le 2m+1} \mathbf{P}\{|RU_a| \le 1\} \max_{1\le b\le 2m+1} \mathbf{P}\{|TV_b| \le 1\} \ .$$

Note that the bound of Lemma 4.1 depends on U_{2m+2} and V_{2m+2} through moments A and B only.

To compare our results with the corresponding results for trigonometric sums, we include a special probabilistic version of the double large sieve bound as Proposition 4.2. We omit the proof since it differs mainly in notation from similar proofs in Graham and Kolesnik (1991).

Proposition 4.2. *Let $U, V \in \mathbb{R}^d$ be independent random vectors with non-negative characteristic functions. Assume that $\mathbf{P}\{|U|_\infty \le T\} = 1$ and $\mathbf{P}\{|V|_\infty \le R\} = 1$ with some positive constants such that $TR > 1$. Then*

$$\mathbf{E}\, \mathrm{e}\{\langle U, V\rangle\} \ll_d (RT)^d \mathbf{P}\{|TV|_\infty \le 1\} \mathbf{P}\{|RU|_\infty \le 1\} \ .$$

Corollary 4.3. *For arbitrary positive numbers R_0 and T_0 write $\bar{R} = \min\{R; R_0\}$ and $\bar{T} = \min\{T; T_0\}$. Then, under the assumptions of Lemma 4.1, we have*

$$\mathbf{E}\, \mathrm{e}\{\langle U, V\rangle\} \ll_{m,d} A\, B\, C\, D \max_{1\le a\le 2m+1} I_a \max_{1\le b\le 2m+1} \bar{I}_b \ ,$$

where

$$I_a = \int_{|s|\le 1} \mathbf{E}\, \mathrm{e}\{\langle \bar{R}^2 U_a/R, s\rangle\}\, ds, \quad \bar{I}_b = \int_{|s|\le 1} \mathbf{E}\, \mathrm{e}\{\langle \bar{T}^2 V_b/T, s\rangle\}\, ds \ .$$

Remark. Lemma 4.1 can be extended to vectors

$$T = (T_1, \ldots, T_d) \quad \text{and} \quad R = (R_1, \ldots, R_d)$$

with positive coordinates. Define

$$T^{-1} = (T_1^{-1}, \ldots, T_d^{-1}) \quad \text{and} \quad Ts = (T_1 s_1, \ldots, T_d s_d) \ ,$$

for $s = (s_1, \ldots, s_d) \in \mathbb{R}^d$. Then Lemma 4.1 still holds with

$$A = 1 + \mathbf{E}\big|T^{-1}U\big|^{2m}, \quad B = 1 + \mathbf{E}\big|R^{-1}V\big|^{2m} \ ,$$

and

$$C = 1 + \sum_{j=1}^{d} (T_j R_j)^{-2m}, \quad D = \prod_{j=1}^{d} T_j R_j \ .$$

Similar extensions hold for other results of the section.

To prove Lemma 4.1 and its Corollary 4.3 we need several lemmas which may be of separate interest.

Lemma 4.4. *Let U and V be independent random vectors taking values in $\mathbb{R}^d$. Let $R > 0$ and $T > 0$. Let g and h denote positive functions. Assume that g as well as its inverse Fourier transform, say $\widehat{g}$, are Lebesgue integrable and* $\mathbf{E} 1/g(V/R) < \infty$. *Then*

$$|\mathbf{E}\{\langle U, V\rangle\}| \ll_d J_1 J_2 \ ,$$

where

$$J_1 = \int_{\mathbb{R}^d} \left| h(s/T) \mathbf{E} \frac{\mathrm{e}\{\langle V, s\rangle\}}{g(V/R)} \right| ds \ ,$$

$$J_2 = \sup_{s \in \mathbb{R}^d} \left| \int_{\mathbb{R}^d} \frac{\mathrm{e}\{-\langle \tau, s\rangle\}}{h(s/T)} g(\tau/R) \mathbf{E}\, \mathrm{e}\{\langle U, \tau\rangle\} \, d\tau \right| \ .$$

Proof of Lemma 4.4. Since $g(x) = c_d \int_{\mathbb{R}^d} \mathrm{e}\{\langle x, y\rangle\} \widehat{g}(y)\, dy$, we get

$$\begin{aligned}
\mathbf{E}\, \mathrm{e}\{\langle U, V\rangle\} &= \mathbf{E}\, \mathrm{e}\{\langle U, V\rangle\} \frac{g(V/R)}{g(V/R)} \\
&= c_d \mathbf{E} \int_{\mathbb{R}^d} \frac{\mathrm{e}\{\langle V, U + \tau/R\rangle\}}{g(V/R)} \widehat{g}(\tau) \, d\tau \\
&= c_d \int_{\mathbb{R}^d} \mathbf{E} \frac{\mathrm{e}\{\langle V, s\rangle\}}{g(V/R)} \widehat{g}(Rs - RU) R^d \, ds \ ,
\end{aligned}$$

and

$$|\mathbf{E}\, \mathrm{e}\{\langle U, V\rangle\}| \ll_d \int_{\mathbb{R}^d} \left| \mathbf{E} \frac{h(s/T) \mathrm{e}\{\langle V, s\rangle\}}{g(V/R)} \right| \left| \frac{\mathbf{E} \widehat{g}(Rs - RU) R^d}{h(s/T)} \right| ds \ .$$

Consequently,

$$|\mathbf{E}\, \mathrm{e}\{\langle U, V\rangle\}| \ll_d J_1 J_3, \quad \text{where} \quad J_3 = \sup_{s \in \mathbb{R}^d} \left| \frac{\mathbf{E} \widehat{g}(Rs - RU)}{h(s/T)} R^d \right| \ .$$

Representing $\widehat{g}$ as the inverse Fourier transform of g, we obtain $J_3 \ll_d J_2$, and the result of the Lemma follows. □

Lemma 4.4 yields the following

Corollary 4.5. *Let* $\mathbf{E}|U|^{2m} < \infty$ *and* $\mathbf{E}|V|^{2m} < \infty$, *for some* $m > d/2$. *Then, for* $g(s) = (1 + |s|^2)^{-m}$, *we have*

$$|\mathbf{E}\, \mathrm{e}\{\langle U, V\rangle\}| \ll_d J_1 \bar{J}_2 \ ,$$

where

$$J_1 = \int_{\mathbb{R}^d} g(s/T)\Big|\mathbf{E}\Big(1+|V/R|^2\Big)^m \mathrm{e}\{\langle V,s\rangle\}\Big|\,ds\ ,$$

$$\bar J_2 = \int_{\mathbb{R}^d} \big|(1-T^{-2}\Delta_\tau)^m (g(\tau/R)\mathbf{E}\mathrm{e}\{\langle U,\tau\rangle\})\big|\,d\tau\ ,$$

and where the Laplace operator $\Delta_\tau = \partial_1^2 + \cdots + \partial_d^2$ *acts on the variables* τ.

Proof. Notice that $(1+|s/T|^2)^m \mathrm{e}\{-\langle\tau,s\rangle\} = (1-T^{-2}\Delta_\tau)^m \mathrm{e}\{-\langle\tau,s\rangle\}$. Thus, assuming that $g=h$, the integral J_2 from Lemma 4.4 we can represent as

$$J_2 = \sup_{s\in\mathbb{R}^d}\left|\int_{\mathbb{R}^d}\big((1-T^{-2}\Delta_\tau)^m \mathrm{e}\{-\langle\tau,s\rangle\}\big) g(\tau/R)\mathbf{E}\mathrm{e}\{\langle U,\tau\rangle\}\,d\tau\right|\ .$$

Integrating by parts we derive $J_2 \le \bar J_2$. $\square$

Lemma 4.6. *Assume all conditions of Lemma* 4.1 *except positivity of the characteristic functions which is replaced by the assumption that the random vectors have mean zero*. *Write*

$$\mathscr{I}_a = \int_{\mathbb{R}^d} g(s/R)|\mathbf{E}\,\mathrm{e}\{\langle U_a,s\rangle\}|\,ds, \quad \bar{\mathscr{I}}_b = \int_{\mathbb{R}^d} g(s/T)|\mathbf{E}\,\mathrm{e}\{\langle V_b,s\rangle\}|\,ds\ ,$$

where $g(s) = (1+|s|^2)^{-m}$. *Then*

$$|\mathbf{E}\,\mathrm{e}\{\langle U,V\rangle\}| \ll_{m,d} A\,B\,C \max_{1\le a\le 2m+1}\mathscr{I}_a \max_{1\le b\le 2m+1}\bar{\mathscr{I}}_b\ .$$

Proof. We shall derive the result from Corollary 4.5. It is sufficient to show that

$$\Big|\mathbf{E}(1+|V/R|^2)^m \mathrm{e}\{\langle V,s\rangle\}\Big| \ll_{m,d} B \sum_{1\le b\le 2m+1} |\mathbf{E}\,\mathrm{e}\{\langle V_b,s\rangle\}|\ , \tag{4.1}$$

$$\big|(1-T^{-2}\Delta_\tau)^m (u(\tau)g(\tau/R))\big| \ll_{m,d} A\,C \sum_{1\le a\le 2m+1} |\mathbf{E}\,\mathrm{e}\{\langle U_a,\tau\rangle\}| g(\tau/R)\ , \tag{4.2}$$

where we denote $u(\tau) = \mathbf{E}\,\mathrm{e}\{\langle U,\tau\rangle\}$.

Let us prove (4.1). We have

$$\Big|\mathbf{E}(1+|V/R|^2)^m \mathrm{e}\{\langle V,s\rangle\}\Big| \ll_{m,d} \sum_{j=0}^m R^{-2j} I_j\ ,$$

where

$$I_j = \big|\mathbf{E}|V|^{2j}\,\mathrm{e}\{\langle V,s\rangle\}\big|\ .$$

In order to estimate I_j recall that $V = V_1 + \cdots + V_{2m+2}$. Thus

$$|V|^{2j} = \left(\sum_{l=1}^{2m+2}\sum_{r=1}^{2m+2}\langle V_l,V_r\rangle\right)^j = \sum_{1\le l_1,\dots,l_j\le 2m+2}\ \sum_{1\le r_1,\dots,r_j\le 2m+2}\ \prod_{q=1}^j \langle V_{l_q},V_{r_q}\rangle\ ,$$

and instead of I_j we have to bound

$$J \stackrel{\text{def}}{=} \left| \mathbf{E} \prod_{q=1}^{j} \langle V_{l_q}, V_{m_q} \rangle \mathrm{e}\{\langle V, s \rangle\} \right| .$$

The number of (different) V_j in the product $\prod_{q=1}^{j} \langle V_{l_q}, V_{r_q} \rangle$ does not exceed $2j \le 2m$. Thus, this product is independent of at least one of vectors $V_1, \dots, V_{2m+1}$, say V_b. Therefore

$$J \le \left(\mathbf{E} \prod_{q=1}^{j} |V_{l_q}| \; |V_{m_q}| \right) |\mathbf{E}\, \mathrm{e}\{\langle V_b, s \rangle\}| .$$

Applying the geometric-arithmetic mean inequality we get

$$\mathbf{E} \prod_{q=1}^{j} |V_{l_q}| \, |V_{r_q}| \ll_{m,d} \sum_{q=1}^{j} \left(\mathbf{E}|V_{l_q}|^{2j} + \mathbf{E}|V_{r_q}|^{2j} \right) \ll_{m,d} \mathbf{E}|V|^{2j}$$

since by Jensen's inequality

$$\mathbf{E}|X|^{2j} = \mathbf{E}|X + \mathbf{E}Y|^{2j} = \mathbf{E}|\mathbf{E}_Y(X+Y)|^{2j} \le \mathbf{E}|X+Y|^{2j}$$

provided that X and Y are independent and $\mathbf{E}Y = 0$. Collecting these estimates, we obtain (4.1).

The proof of (4.2) is a little bit more involved. We have

$$\left| (1 - T^{-2}\Delta_\tau)^m (u(\tau) g(\tau/R)) \right| \ll_{m,d} \sum_{s=0}^{m} T^{-2s} I_s ,$$

where

$$I_s = \left| \Delta_\tau^s (u(\tau) g(\tau/R)) \right| .$$

Differentiating the product we get

$$I_s \ll_{m,d} \sum_{|\alpha|+|\beta|=2s} \left| \partial_1^{\alpha_1} \dots \partial_d^{\alpha_d} u(\tau) \right| R^{-|\beta|} \left| (\partial_1^{\beta_1} \dots \partial_d^{\beta_d} g)(\tau/R) \right| .$$

Here $\alpha = (\alpha_1, \dots, \alpha_d)$ and $\beta = (\beta_1, \dots, \beta_d)$ denote non-negative integer multi-indices with $|\alpha| = \alpha_1 + \dots + \alpha_d$. The function g satisfies

$$\left| (\partial_1^{\beta_1} \dots \partial_d^{\beta_d} g)(\tau) \right| \ll_{m,d} g(\tau) .$$

Furthermore, writing the random vector $U = (U_{(1)}, \dots, U_{(d)})$ in coordinates of $\mathbb{R}^d$, we have

$$\left| \partial_1^{\alpha_1} \dots \partial_d^{\alpha_d} u(\tau) \right| = \left| \mathbf{E} U_{(1)}^{\alpha_1} \dots U_{(d)}^{\alpha_d} \mathrm{e}\{\langle U, \tau \rangle\} \right| .$$

The product $U_{(1)}^{\alpha_1} \dots U_{(d)}^{\alpha_d}$ is a polynomial of order $|\alpha| \le 2s \le 2m$ in U. Thus arguing similarly as in the proof of (4.1) we obtain

$$\left| \mathbf{E} U_{(1)}^{\alpha_1} \dots U_{(d)}^{\alpha_d} \mathrm{e}\{\langle U, \tau \rangle\} \right| \ll_{m,d} \mathbf{E}|U|^{|\alpha|} \sum_{1 \le a \le 2m+1} \left| \mathbf{E}\, \mathrm{e}\{\langle U_a, \tau \rangle\} \right| ,$$

and (4.2) follows since

$$\sum_{s=0}^{m} \sum_{|\alpha|+|\beta|=2s} T^{-2s} R^{-|\beta|} \mathbf{E}|U|^{|\alpha|} \ll_{m,d} A\,C \ . \quad \square$$

Proof of Lemma 4.1. We shall derive the result from Lemma 4.6. Let $X \in \mathbb{R}^d$ denote a random vector with non-negative characteristic function. It is sufficient to show that

$$J \stackrel{\text{def}}{=} \int_{\mathbb{R}^d} g(s/R) \mathbf{E}\,\mathrm{e}\{\langle X, s\rangle\}\, ds \ll_{m,d} R^d \mathbf{P}\{|RX| \le 1\} \ . \tag{4.3}$$

Introduce the concentration function

$$\mathcal{Q}(X;\lambda) = \sup_{s\in\mathbb{R}^d} \mathbf{P}\{|X-s|_\infty \le \lambda\}, \quad \lambda \ge 0 \ .$$

Assume that a function $p(s) = p_1(s_1)\dots p_d(s_d)$ can be represented as a product such that the functions $p_j \colon \mathbb{R} \to \mathbb{R}$ are even, non-negative, non-increasing on $[0,\infty)$ and $p_j(t) = p_j(0)$, for $|t| \le 1$, for all $1 \le j \le d$. It is known that (see Zaitsev 1988a, Lemma 5.3, or 1988b, Lemma 2.6)

$$\int_{\mathbb{R}^d} p(s) \mathbf{E}\mathrm{e}\{\langle Z, s\rangle\}\, ds \ll_d \mathcal{Q}(Z;1) \int_{\mathbb{R}^d} p(s)\, ds \ , \tag{4.4}$$

for any random vector Z with non-negative characteristic function.

Let us apply the bound (4.4) to the integral in (4.3). Define $p(s) = \theta(s_1)\dots\theta(s_d)$, where the function $\theta \colon \mathbb{R} \to \mathbb{R}$ satisfies

$$\theta(t) = 1, \text{ for } |t| \le 1, \text{ and } \theta(t) = (2/(1+t^2))^{m/d}, \text{ for } |t| \ge 1 \ .$$

Then $g(s) \ll_{m,d} p(s)$ and a change of variable of integration together with (4.4) implies $J \ll_{m,d} R^d \mathcal{Q}(RX;1)$. To finish the proof of (4.3) we should replace the norm $|\cdot|_\infty$ by $|\cdot|$. To this end notice that $\mathcal{Q}(X;1) \ll_d \delta^{-d}\mathcal{Q}(X;\delta)$, for any $0 < \delta \le 1$ and random vector X. Furthermore, the inequality

$$\sup_{s\in\mathbb{R}^d} \mathbf{P}\{|X-s| \le \lambda\} \ll_d \mathbf{P}\{|X| \le \lambda\}, \quad \lambda \ge 0 \ ,$$

holds for any random vector X with non-negative characteristic function. A proof is based on an application of Parseval's equality, cf. the proof of Lemma 5.1 in Zaitsev (1988a). $\square$

Proof of Corollary 4.3. Obviously

$$\mathbf{P}\{|RU_a| \le 1\} \le \mathbf{P}\{\bar{R}^2 |U_a/R| \le 1\}$$

and the result follows from Lemma 4.1 by an application of

$$\mathbf{P}\{|X| \le 1\} \ll_d \int_{|s|\le 1} |\mathbf{E}\,\mathrm{e}\{\langle X, s\rangle\}|\, ds \ ,$$

see Esseen (1968), Lemma 6.1. $\square$

We shall denote $|\mathbb{A}| = \sup_{|x|=1} |\mathbb{A}x|$.

Lemma 4.7. *Let $\mathbb{A}$ be a $d \times d$ matrix. Let $X \in \mathbb{R}^d$ denote a random vector with covariance $\mathbb{C}$. Assume that there exists a constant c_d such that*

$$\mathbf{P}\{|X| \le c_d\} = 1, \quad |\mathbb{A}| \le c_d, \quad |\mathbb{A}^{-1}| \le c_d, \quad |\mathbb{C}^{-1}| \le c_d \ . \tag{4.5}$$

Let U and V denote independent random vectors which are sums of N independent copies of X. Then (*the function $\mathscr{M}$ is defined by* (1.16))

$$|\mathbf{E}\, e\{t\langle \mathbb{A}U, V\rangle\}| \ll_d \mathscr{M}^{2d}(t; N), \quad \textit{for } t \in \mathbf{R} \ . \tag{4.6}$$

Proof. Introducing sums U' and V' of $\lceil N/2 \rceil$ independent copies of the symmetrization $\tilde{X}$ of X, we have

$$|\mathbf{E}\, e\{t\langle \mathbb{A}U, V\rangle\}| \le \mathbf{E}\, e\{t\langle \mathbb{A}U', V'\rangle\} \ . \tag{4.7}$$

To prove (4.7), one should proceed as follows: split $V = V_1 + V_2 + V_3$ into three independent sums such that each of V_1 and V_2 is a sum of $\lceil N/2 \rceil$ independent copies of X; condition on U and V_3 and to apply the equality $|\mathbf{E}\, e\{\langle x, V_1 + V_2\rangle\}| = |\mathbf{E}\, e\{\langle x, V_1\rangle\}|^2 = \mathbf{E}\, e\{\langle x, \tilde{V}_1\rangle\}$, which is valid for any $x \in \mathbb{R}^d$ and any i.i.d. random vectors V_1, V_2; repeat the procedure with U instead of V.

In the proof of the Lemma we can assume that N is sufficiently large, that is, that $N \ge c(d)$, with a sufficiently large constant $c(d)$. Otherwise the result follows from the trivial bound $|\mathbf{E}\, e\{t\langle \mathbb{A}U, V\rangle\}| \le 1$ and the estimate $1 \ll_d \mathscr{M}^{2d}(t; N)$, valid for $N < c(d)$. Furthermore, without loss of generality we shall assume that U and V are sums of $4dN$ independent copies of $\tilde{X}$. To see this, use (4.7) and replace $\lceil N/2 \rceil$ by $4dN$. Since N is arbitrary and (1.17) is fulfilled, such a replacement can change in (4.6) only constants depending on d. Thus we can write

$$U = U_1 + \cdots + U_{4d}, \quad V = V_1 + \cdots + V_{4d} \tag{4.8}$$

with i.i.d. random vectors $U_1, \ldots, U_{4d}, V_1, \ldots, V_{4d}$ such that U_1 is a sum of N independent copies of $\tilde{X}$. Thus, assuming (4.8) we have to estimate $\mathbf{E}\, e\{t\langle \mathbb{A}U, V\rangle\}$. Due to the symmetry of random vectors, we may assume as well that $t > 0$.

Let us apply Corollary 4.3 replacing U and V by $\sqrt{t}\mathbb{A}U$ and $\sqrt{t}V$, and choosing $m = 2d - 1$, $T^2 = R^2 = \varepsilon^2 tN$, $T_0^4 = R_0^4 = \varepsilon^4 N$, where we write $\varepsilon = 1/(2c_d^2 + c_d)$. By the Rosenthal's inequality and (4.5), the moments $\mathbf{E}|\mathbb{A}U_j|^{4d-2}$ and $\mathbf{E}|V_j|^{4d-2}$ are bounded from above by $c(d)N^{2d-1}$. Furthermore, we can assume that $T^2 = \varepsilon^2 tN \ge 1$. Indeed, otherwise the result of the Lemma is obvious since $1 \ll_d \mathscr{M}(t; N)$, for $0 < tN \ll_d 1$. Thus, Corollary 4.3 implies

$$\mathbf{E}\, e\{t\langle \mathbb{A}U, V\rangle\} \ll_d T^{2d} I \bar{I} \ , \tag{4.9}$$

where

$$I = \int_{|x| \le 1} \mathbf{E}\, e\{\sqrt{t}\langle \bar{T}^2 \mathbb{A}U_1/T, x\rangle\}\, dx, \quad \bar{I} = \int_{|x| \le 1} \mathbf{E}\, e\{\sqrt{t}\langle \bar{T}^2 V_1/T, x\rangle\}\, dx \ .$$

The bound (4.9) implies the Lemma provided that we can show that

$$I \ll_d (tN)^{-d/2} T^d \bar{T}^{-2d}, \quad \bar{I} \ll_d (tN)^{-d/2} T^d \bar{T}^{-2d} \quad . \tag{4.10}$$

Thus it remains to prove (4.10). We shall estimate I. The estimate of $\bar{I}$ is similar.

Write $\xi = \sqrt{t}\langle \bar{T}^2 \tilde{X}/T, \mathbb{A}'x\rangle$, where $\mathbb{A}'$ is the transposed matrix $\mathbb{A}$. The i.i.d. property implies that

$$\mathbf{E}\, \mathrm{e}\{\sqrt{t}\langle \bar{T}^2 \mathbb{A}\mathbb{U}_1/T, x\rangle\} = (\mathbf{E}\, \mathrm{e}\{\xi\})^N \quad . \tag{4.11}$$

Using (4.5) we have $|\tilde{X}| \le 2c_d$, $|\mathbb{A}'x| \le c_d$, for $|x| \le 1$. Therefore

$$|\xi| \le 1, \text{ provided that } 2c_d^2 \sqrt{t}\, \bar{T}^2/T \le 1 \quad .$$

But the inequality $2c_d^2\sqrt{t}\,\bar{T}^2/T \le 1$ is clearly fulfilled due to our choice of ε and T, T_0. Hence the symmetry of $\tilde{X}$ and the inequality $\cos u \le 1 - u^2/4$, for $|u| \le 1$ together imply

$$\mathbf{E}\, \mathrm{e}\{\xi\} = \mathbf{E} \cos \xi \le 1 - \frac{1}{4}\mathbf{E}\xi^2 \le \exp\left\{-\frac{1}{4}\mathbf{E}\xi^2\right\} \quad . \tag{4.12}$$

We have

$$\mathbf{E}\xi^2 = t\bar{T}^4 T^{-2} \mathbf{E}\langle \tilde{X}, \mathbb{A}'x\rangle^2 = 2t\bar{T}^4 T^{-2} \langle \mathbb{C}\mathbb{A}'x, \mathbb{A}'x\rangle \quad .$$

Using $|\mathbb{C}^{-1}| \le c_d$ and $|\mathbb{A}^{-1}| \le c_d$, we have $\langle \mathbb{C}z, z\rangle \ge |z|^2/c_d$ and $|\mathbb{A}'x| \ge |x|/c_d$. Consequently, $\mathbf{E}\xi^2 \ge 2c_d^{-3} t\bar{T}^4 T^{-2}|x|^2$, and, in view of (4.11) and (4.12), we obtain

$$I \le \int_{|x|\le 1} \mathbf{E} \exp\{-Nc_d^{-3} t\bar{T}^4 T^{-2}|x|^2/2\}\, dx \ll_d (tN)^{-d/2} \bar{T}^{-2d} T^d \quad ,$$

hence proving (4.10) and the Lemma. $\square$

5. Technical Lemmas

The following symmetrization inequality (see Bentkus and Götze 1996, Lemma 3.1) improves an inequality due to Götze (1979).

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*

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

a real-valued polynomial of second order. Then

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

We shall need the following auxiliary bound for the convergence rate in the CLT in finite dimensional and Hilbert spaces for expectations of smooth functions.

Lemma 5.2. *Assume that X is symmetric. Let a function $\varphi : \mathbb{R}^d \to \mathbb{R}$ be four times Frechet differentiable. Then there exists an absolute constant c such that*

$$|\mathbf{E}\varphi(S_N) - \mathbf{E}\varphi(G)| \le c\beta N^{-1} \sup_{x\in\mathbb{R}^d} \left\|\varphi^{(4)}(x)\right\| \ .$$

Proof. For the sake of completeness we include a sketch of the proof. Write

$$W_j = N^{-1/2}(X_1 + \cdots + X_{j-1} + G_{j+1} + \cdots + G_N) \ .$$

Then

$$|\mathbf{E}\varphi(S_N) - \mathbf{E}\varphi(G)| \le \delta_1 + \cdots + \delta_N \ ,$$

where, by Taylor expansions in powers of $X_j/\sqrt{N}$ and $G_j/\sqrt{N}$,

$$\delta_j = \left|\mathbf{E}\varphi(W_j + X_j/\sqrt{N}) - \mathbf{E}\varphi(W_j + G_j/\sqrt{N})\right| \le cN^{-2}\mathbf{E}|X|^4 \sup_{x\in\mathbb{R}^d} \left\|\varphi^{(4)}(x)\right\| \ .$$

□

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 $m \ge c\beta/(p\delta^4)$.*

Proof. We shall apply Lemma 5.2. Let $e \in \mathbb{R}^d$. Write $p_e = \mathbf{P}\{|G - e| \le \delta\}$. It is sufficient to show that there exists an absolute constant c such that

$$2\mathbf{P}\{|S_m - e| \le 2\delta\} \ge p_e, \quad \text{for } m \ge c\beta/(p_e\delta^4) \ . \tag{5.1}$$

Consider a function $\varphi : \mathbb{R}^d \to \mathbb{R}$ with infinitely many bounded derivatives such that

$$0 \le \varphi \le 1, \quad \varphi(x) = 1, \text{ for } |x| \le 1, \quad \varphi(x) = 0, \text{ for } |x| \ge 2 \ .$$

Applying Lemma 5.2 we have

$$\begin{aligned} 2\mathbf{P}\{|S_m - e| \le 2\delta\} &\ge 2\mathbf{E}\varphi\left(\frac{S_m - e}{\delta}\right) \ge 2\mathbf{E}\varphi\left(\frac{G - e}{\delta}\right) - cm^{-1}\delta^{-4}\beta \\ &\ge 2p_e - p_e = p_e \ , \end{aligned}$$

for $m \ge c\beta/(p_e\delta^4)$, thus proving (5.1). □

Lemma 5.4. *Assume that $0 < 4\varepsilon \le \delta \le 1$. 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 $p_e = \mathbf{P}\{|\varepsilon\sigma^{-1}G - e| \le \delta\}$ satisfies $p_e \ge \exp\{-c\sigma^2\varepsilon^{-2}\sigma_e^{-2}\}$ with some positive absolute constant c.*

Proof. Introduce the Gaussian random vectors $G - \langle G, e\rangle e$ and $\langle G, e\rangle e$. These vectors are independent since e is an eigenvector of $\mathbb{C} = \operatorname{cov} G$. We have

$$p_e \ge \mathbf{P}\left\{\varepsilon\sigma^{-1}|G - \langle G, e\rangle e| + \left|\varepsilon\sigma^{-1}\langle G, e\rangle e - e\right| \le \delta\right\} \ge p_1 p_2 \ ,$$

where

$$p_1 = \mathbf{P}\big\{\varepsilon\sigma^{-1}|G - \langle G, e\rangle e| \le \delta/2\big\} \quad \text{and} \quad p_2 = \mathbf{P}\big\{\big|\varepsilon\sigma^{-1}\langle G, e\rangle e - e\big| \le \delta/2\big\} \ .$$

Using Chebyshev's inequality and $\sigma_e^2 = \mathbf{E}\langle G, e\rangle^2$, we have

$$\begin{aligned} p_1 &= 1 - \mathbf{P}\big\{\varepsilon\sigma^{-1}|G - \langle G, e\rangle e| > \delta/2\big\} \ge 1 - 4\varepsilon^2\delta^{-2}\sigma^{-2}\mathbf{E}|G - \langle G, e\rangle e|^2 \\ &\ge 1 - 4\varepsilon^2\delta^{-2}\sigma^{-2}\mathbf{E}|G|^2 \ge 1 - 4\varepsilon^2\delta^{-2} \ \ge 1/2 \end{aligned}$$

provided that $4\varepsilon \le \delta$.

We can write $p_2 = \mathbf{P}\big\{\big|\varepsilon\sigma^{-1}\sigma_e\eta - 1\big| \le \delta/2\big\}$, where η is a standard Gaussian random variable. Using $\sigma \ge \sigma_e$, $\delta \ge 4\varepsilon$ and $\delta \le 1$, and replacing the standard normal density by its minimal value in the interval $[0, \sigma/(\varepsilon\sigma_e)]$, we obtain

$$\begin{aligned} p_2 \ge \mathbf{P}\big\{1 - \delta/2 \le \varepsilon\sigma^{-1}\sigma_e\eta \le 1\big\} &\ge c_1 \exp\big\{-\sigma^2/(2\varepsilon^2\sigma_e^2)\big\} \\ &\ge \exp\big\{-c\sigma^2/(\varepsilon^2\sigma_e^2)\big\} \ . \quad \square \end{aligned}$$

Lemma 5.5. *Assume that* $\mathbb{Q}^2 = \mathbb{I}$. *The collections of eigenvalues* (*counting their multiplicities*) *of* $(\mathbb{Q}\mathbb{C})^2$, $(\mathbb{C}\mathbb{Q})^2$ *and* $\mathbb{D} \overset{\text{def}}{=} \sqrt{\mathbb{Q}\mathbb{C}\mathbb{Q}}\,\mathbb{C}\sqrt{\mathbb{Q}\mathbb{C}\mathbb{Q}}$ *are equal. Furthermore, the eigenvalues* $\theta_1^4 \ge \theta_2^4 \ge \dots$ *of* $(\mathbb{C}\mathbb{Q})^2$ *satisfy*

$$\sum_{j\ge 1} \theta_j^4 \le \sigma_1^2\sigma^2, \quad \theta_j \le \sigma \ . \tag{5.2}$$

If the condition $\mathscr{B}(\mathscr{S}_o, \mathbb{C})$ (see (1.4)) *is fulfilled then* $\lambda_s \le \theta_s$ *and* $\lambda_s \le \sigma_s$.

Proof. For any bounded linear operators $\mathbb{A}, \mathbb{B}\colon \mathbb{R}^d \to \mathbb{R}^d$ such that $\mathbb{A} \ge 0$, the collections of non-zero eigenvalues of $\mathbb{A}\mathbb{B}$, $\mathbb{B}\mathbb{A}$ and $\sqrt{\mathbb{A}}\mathbb{B}\sqrt{\mathbb{A}}$ coincide (for a proof of this simple fact see Vakhania 1981, or Lemma 2.3 in Bentkus 1984). Hence, all operators mentioned in the Lemma have the same sets of eigenvalues.

Let us prove (5.2). For any $e \in \mathbb{R}^d$ we have

$$\langle \mathbb{D}e, e\rangle \le \sigma_1^2\langle\sqrt{\mathbb{Q}\mathbb{C}\mathbb{Q}}e, \sqrt{\mathbb{Q}\mathbb{C}\mathbb{Q}}e\rangle = \sigma_1^2\langle\mathbb{C}\mathbb{Q}e, \mathbb{Q}e\rangle \ .$$

Therefore, for any orthonormal basis $\{e_j\}$ of $\mathbb{R}^d$ we obtain

$$\sum_{j\ge 1}\theta_j^4 = \sum_{j\ge 1}\langle\mathbb{D}e_j, e_j\rangle \le \sigma_1^2\sum_{j\ge 1}\langle\mathbb{C}\mathbb{Q}e_j, \mathbb{Q}e_j\rangle = \sigma_1^2\sigma^2$$

since $\{\mathbb{Q}e_j\}$ is again an orthonormal basis of $\mathbb{R}^d$. $\square$

Lemma 5.6. *Assume that* $\mathbb{Q}^2 = \mathbb{I}$. *The characteristic functions*

$$u(t) = \mathbf{E}\,\mathrm{e}\{t\langle\mathbb{Q}G_1, G_2\rangle\}, \quad \widehat{F}_0(t) = \mathbf{E}\mathrm{e}\{t\mathbb{Q}[G - a]\}$$

satisfy

$$|\widehat{F}_0(t)|^2 \le u(4t/3), \quad u(t) = \prod_{j=1}^{d}(1 + \theta_j^4 t^2)^{-1/2} \ . \tag{5.3}$$

Proof. Writing $G \overset{\mathcal{D}}{=} (G_1 + G_2 + G_3)/\sqrt{3}$ and applying the symmetrization Lemma 5.1, we derive the inequality in (5.3). Clearly

$$u(t) = \mathbf{E}\exp\{-t^2\langle \mathbb{C}\mathbb{Q}G, \mathbb{Q}G\rangle/2\} = \mathbf{E}\exp\Big\{-t^2|W|^2/2\rangle\Big\} \ ,$$

where W denotes a centered Gaussian vector such that $\operatorname{cov} W = \sqrt{\mathbb{Q}\mathbb{C}\mathbb{Q}}\mathbb{C}\sqrt{\mathbb{Q}\mathbb{C}\mathbb{Q}}$. By Lemma 5.5, the operators $\operatorname{cov} W$ and $(\mathbb{C}\mathbb{Q})^2$ have the same eigenvalues θ_j^4, and we can write $W \overset{\mathcal{D}}{=} \sum_{j=1}^d \theta_j^2 g_j e_j$, where $\{g_j\}$ denotes a sequence of i.i.d. standard normal variables and $\{e_j\}$ is an orthonormal basis of $\mathbb{R}^d$ corresponding to the eigenvalues θ_j^4 of $\operatorname{cov} W$. Consequently, $u(t) = \prod_{j=1}^d \mathbf{E}\exp\Big\{-t^2\theta_j^4 g_1^2/2\Big\}$, and simple calculations complete the proof of (5.3). $\square$

Lemma 5.7. *Let* $\mathbb{Q}^2 = \mathbb{I}$ *and* $\sigma < \infty$. *Then the Edgeworth corrections* $F_1^\diamond$ *and* $\Psi_1^\diamond$ (see (1.2) *and* (1.19)) *are functions of bounded variation provided that* $s \geq 9$. *If* $\beta_3 < \infty$ *and* $s \geq 9$ *then* F_1 *and* Ψ_1 *are functions of bounded variation as well. Furthermore, assuming* $\sigma = 1$ *and* $s \geq 7$, *we have*

$$\sup_x |\Psi_1^\diamond(x)| \ll_s \theta_s^{-6}(1 + |a|^3) \ , \tag{5.4}$$

$$\sup_x |\Psi_1(x) - \Psi_1^\diamond(x)| \ll_s \theta_s^{-6}(1 + |a|^3)\Pi_3 \ , \tag{5.5}$$

and, for any $t_* \geq c/N$, *where c is an absolute positive constant*,

$$\int_{|t| \geq t_*} |\widehat{\Psi}_1^\diamond(t)| \, \frac{dt}{|t|} \ll_s \Lambda^{1/2}(t_* N)^{3-s/2}(1 + |a|^3)\theta_s^{-s} \ . \tag{5.6}$$

Proof. We shall consider the case of functions Ψ_1 and $\Psi_1^\diamond$ only since $F_1(x) = \Psi_1(xN)$ and $F_1^\diamond(x) = \Psi_1^\diamond(xN)$ (see (3.7)). Assuming $\sigma = 1$, using splittings of G into independent components and the conditioning and symmetrization techniques of Section 9, we obtain

$$|\widehat{\Psi}_1(t)| \ll (tN)^2 \mathbf{E}|X|^3 N^{-1/2}\sqrt{u(ctN)}(1 + |tN|)(1 + |a|^3) \tag{5.7}$$

with the function u defined in Lemma 5.6. A similar bound with $\mathbf{E}|X^\diamond|^3$ instead of $\mathbf{E}|X|^3$ holds for $|\widehat{\Psi}_1^\diamond(t)|$. Lemma 5.6 implies $u(t) \ll_s \theta_s^{-2s}|t|^{-s}$ since $1 \geq \theta_1 \geq \theta_2 \geq \ldots$ Thus, both $\widehat{\Psi}_1^\diamond$ and $\widehat{\Psi}_1$ are integrable provided that $s \geq 9$, and $\Psi_1^\diamond$ and Ψ_1 have bounded variations.

Estimating $\mathbf{E}|X^\diamond|^3 \leq \sqrt{N}$ and using a bound of type (5.7) for $|\widehat{\Psi}_1^\diamond(t)|$, we obtain

$$|\widehat{\Psi}_1^\diamond(t)| \ll_s (tN)^2(1 + |tN|)(1 + |a|^3)(1 + \theta_s^4 t^2 N^2)^{-s/4} \ , \tag{5.8}$$

whence, using the Fourier inversion formula (3.2), we obtain (5.4).

To prove (5.5) and (5.6) we use again the Fourier inversion formula (3.2), estimates of type (5.7), (5.8), $\sigma = 1$ and

$$X = X^\diamond + X_\diamond, \quad \mathbf{E}|X_\diamond|^3 = \sqrt{N}\Pi_3, \quad N^{-1/2}\mathbf{E}|X^\diamond|^3 \leq \Lambda^{1/2} \ .$$

6. Discretization of expectations

Assume that the symmetrization $\tilde{X}$ of a random vector $X \in \mathbb{R}^d$ satisfies

$$\mathbf{P}\{|\tilde{X} - e| \leq \delta\} \geq p \ , \tag{6.1}$$

for some $\delta \geq 0$, $0 < p \leq 1$ and $e \in \mathbb{R}^d$.

Recall that $\varepsilon_1, \varepsilon_2, \ldots$ denote i.i.d symmetric Rademacher random variables.

Throughout this Section we assume that $\mathscr{S} = \{e_1, \ldots, e_s\} \subset \mathbb{R}^d$ is an arbitrary finite subset such that card $\mathscr{S} = s$. Recall that a discrete random vector $U \in \mathbb{R}^d$ belongs to the class $\Gamma(\delta, \mathscr{S})$, $\delta > 0$, if U is distributed as $\varepsilon_1 z_1 + \cdots + \varepsilon_s z_s$, with some (non-random) $z_1, \ldots, z_s \in \mathbb{R}^d$ such that $|z_j - e_j| \leq \delta$, for all $1 \leq j \leq s$.

For a bounded and measurable function $H\colon \mathbb{R}^d \to \mathbf{B}$ with values in a Banach space $(\mathbf{B}, |\cdot|_{\mathbf{B}})$, define the norm $|H|_\infty = \sup_x |H(x)|_{\mathbf{B}}$.

Let $H_t\colon \mathbb{R}^d \to \mathbf{B}$ denote a family of bounded functions indexed by $t \in \mathbb{R}$ such that the functions $(t, x) \mapsto H_t(x)$ and $t \mapsto |H_t|_\infty$ are measurable. In Lemmas 6.1 and 6.2 we shall apply the discretization procedure to $|\mathbf{E}H(Z_N)|$ and $\int |\mathbf{E}H_t(Z_N)|dt$, where we write $\int = \int_{\mathbb{R}}$.

In Lemma 6.1 we write $n = \lceil pN/2 \rceil$ and $T = U_1 + \cdots + U_n$, where U_j are independent random vectors of class $\Gamma(\delta, \{e\})$, that is $U_j = \varepsilon_j z_j$, for some non-random z_j such that $|z_j - e| \leq \delta$.

Lemma 6.1. *Assume that X satisfies* (6.1). *Let $W \in \mathbb{R}^d$ denote a random vector independent of Z_N. Then, for any $\gamma \geq 0$, we have*

$$\int |\mathbf{E}H_t(2Z_N + W)|\, dt \leq I + c_\gamma (pN)^{-\gamma} \int |H_t|_\infty \, dt \tag{6.2}$$

with

$$I = \sup_{\Gamma} \sup_{b \in \mathbb{R}^d} \int |\mathbf{E}H_t(T + W + b)|\, dt \ ,$$

where $\sup_\Gamma$ *denotes the supremum over all $\mathscr{L}(U_1), \ldots, \mathscr{L}(U_n) \in \Gamma(\delta, \{e\})$ such that T is independent of W.*

Proof. Replacing H_t by the function $x \mapsto \mathbf{E}_W H_t(x + W)$, we can assume that $W = 0$. Furthermore, we can assume that $pN \geq c$ with a sufficiently large absolute constant c since otherwise (6.2) follows from the trivial estimate $|\mathbf{E}H_t(2Z_N + W)| \leq |H_t|_\infty$.

Let $\alpha_1, \alpha_2, \ldots$ be a sequence of i.i.d random variables, independent of all other random variables and vectors and such that

$$\mathbf{P}\{\alpha_1 = 0\} = \mathbf{P}\{\alpha_1 = 1\} = 1/2 \ .$$

Then the sum

$$V^\alpha = \sum_{j=1}^{N} (\alpha_j X_j + (1 - \alpha_j)\bar{X}_j)$$

has the same distribution as Z_N (recall that $\bar{X}$ denotes an independent copy of X). Notice that the random variables $\varepsilon_1, \varepsilon_2, \dots$ defined by $\varepsilon_j = 2\alpha_j - 1$ are i.i.d symmetric Rademacher random variables. Introduce the sum $V^\varepsilon = \varepsilon_1 \tilde{X}_1 + \cdots + \varepsilon_N \tilde{X}_N$. Using the relation $\alpha_j = \frac{\varepsilon_j}{2} + \frac{1}{2}$, we may write $V^\alpha = \frac{1}{2} V^\varepsilon + b$, for some $b = b_N(X_1, \dots, X_N)$, which is independent of $\varepsilon_1, \dots, \varepsilon_N$. Conditioning we have (recall that we assumed that $W = 0$)

$$\int |\mathbf{E} H_t(2Z_N)|\, dt \le \mathbf{E} \int |\psi|\, dt, \quad \text{where} \quad \psi = \mathbf{E}_\varepsilon H_t(V^\varepsilon + b) \ , \tag{6.3}$$

and where $\mathbf{E}_\varepsilon$ denotes the partial integration with respect to the distribution of $\varepsilon_1, \varepsilon_2, \dots$

Introduce the independent events

$$A(l) = \{|\tilde{X}_l - e| \le \delta\}, \quad l = 1, 2, \dots \ .$$

Notice that $p' \overset{\text{def}}{=} \mathbf{P}\{A(l)\} \ge p$ by the assumption (see (6.1)). Consider the event

$$B_N = \left\{ \text{at least } \frac{p'N}{2} \text{ of events } A(1), \dots, A(N) \text{ occur} \right\} \ ,$$

and introduce Bernoulli random variables $\xi_l = \mathbf{I}\{A(l)\}$. For the complement B_N^c of the event B_N, we have

$$\mathbf{P}\{B_N^c\} = 1 - \mathbf{P}\{B_N\} \ll_\gamma (pN)^{-\gamma}, \quad \text{for any } \gamma > 0 \ . \tag{6.4}$$

Let us prove (6.4). We can assume that $\gamma \ge 1$. Write $\eta_j = \xi_j - \mathbf{E}\xi_j$ and notice that $\mathbf{E}|\eta_1|^{2\gamma} \ll_\gamma \mathbf{E}|\xi_1|^{2\gamma} = p'$. Using $p'N \ge pN \ge c$, Chebyshev's inequality and Rosenthal's type inequality (1.24), we obtain

$$\begin{aligned} \mathbf{P}\{B_N^c\} &= \mathbf{P}\left\{\xi_1 + \cdots + \xi_N < \frac{p'N}{2}\right\} \ll_\gamma (p'N)^{-2\gamma} \mathbf{E}|\eta_1 + \cdots + \eta_N|^{2\gamma} \\ &\ll_\gamma (p'N)^{-2\gamma}\Big(N\mathbf{E}|\eta_1|^{2\gamma} + \big(N\mathbf{E}\eta_1^2\big)^\gamma\Big) \ll_\gamma (p'N)^{-\gamma} \ , \end{aligned}$$

whence (6.4) follows since $p' \ge p$.

Consequently, (6.4) yields

$$\mathbf{E} \int |\psi|\, dt \le c_\gamma (pN)^{-\gamma} \int |H_t|_\infty\, dt + \mathbf{E}\mathbf{I}\{B_N\} \int |\psi|\, dt \ ,$$

and it remains to show that

$$\mathbf{I}\{B_N\} \int |\psi|\, dt \le \sup_\Gamma \sup_{b \in \mathbb{R}^d} \int |\mathbf{E} H_t(T + b)|\, dt \tag{6.5}$$

since (6.3) holds. If the event B_N does not occur, then $\mathbf{I}\{B_N\} = 0$, and (6.5) is fulfilled. If B_N occurs then at least $pN/2$ of events $A(l)$ occur, say $A(l_1), \dots, A(l_a)$ with some $a \ge pN/2 \ge n$. Reorder the random variables $\varepsilon_l \tilde{X}_l$, $1 \le l \le N$, so that $A(1), \dots, A(a)$ occur. Then we may write

$$\mathbf{I}\{B_N\} \int |\psi|\, dt = \mathbf{I}\{B_N\} \int \big|\mathbf{E}_\varepsilon H_t(\varepsilon_1 \tilde{X}_1 + \cdots + \varepsilon_a \tilde{X}_a + b + b_1)\big|\, dt \tag{6.6}$$

with some b_1 independent of $\varepsilon_1, \dots, \varepsilon_a$. Conditioning in (6.6) on ε_l such that $n < l \le a$, we obtain

$$\mathbf{I}\{B_N\} \int |\psi|\, dt \le \sup_{b\in\mathbb{R}^d} \mathbf{I}\{B_N\} \int \left|\mathbf{E}_{\varepsilon_1,\dots,\varepsilon_n} H_t(\varepsilon_1 \tilde{X}_1 + \cdots + \varepsilon_n \tilde{X}_n + b)\right| dt \ ,$$

whence (6.5) follows. $\square$

The next Lemma 6.2 allows to replace sums of independent random vectors by sums of discrete bounded random vectors of a very special type.

Lemma 6.2. *Let* $\mathbb{Q}^2 = \mathbb{I}$. *Assume* $\mathscr{N}(p, \delta, \mathscr{S}, \tilde{X})$. *Let* $n = \lceil pN/(5s) \rceil$. *Then, for any* $\gamma \ge 0$, *we have*

$$\int |\mathbf{E} H_t(2Z_N + b)|\, dt \le I + c_\gamma(s)(pN)^{-\gamma} \int |H_t|_\infty\, dt \tag{6.7}$$

with

$$I = \sup_{\Gamma} \sup_{b\in\mathbb{R}^d} \int |\mathbf{E} H_t(Y + \mathbb{Q}Y' + b)|\, dt \ ,$$

and

$$|\mathbf{E} H(2Z_N + b)| \le c_\gamma(s)(pN)^{-\gamma} |H|_\infty + \sup_{\Gamma} \sup_{b\in\mathbb{R}^d} |\mathbf{E} H(Y + \mathbb{Q}Y' + b)| \ , \tag{6.8}$$

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

Proof. To prove (6.8), it suffices to set $H_t = H\mathbf{I}\{0 \le t \le 1\}$ in (6.7).

It remains to prove (6.7). We shall apply $2s$ times Lemma 6.1. As in the proof of Lemma 6.1 we may assume that $pN \ge c_s$, for a sufficiently large constant c_s.

Let $r = \lceil N/(2s) \rceil$ and $m = \lceil pr/2 \rceil$. Notice that $m \ge n$ since $pN \ge c_s$ is sufficiently large.

Introducing i.i.d random vectors V_j such that their common distribution is equal to $\mathscr{L}(X_1 + \cdots + X_r)$, and collecting summands in Z_N in groups, we may write

$$Z_N + b = V_1 + \cdots + V_{2s} + b_1 \ ,$$

where b_1 is independent of V_j, $1 \le j \le 2s$. Conditioning on b_1 we obtain

$$J \overset{\text{def}}{=} \int |\mathbf{E} H_t(2Z_N + b)|\, dt \le \sup_{b\in\mathbb{R}^d} \int |\mathbf{E} H_t(2V_1 + \cdots + 2V_{2s} + b)|\, dt \ . \tag{6.9}$$

Write $W = 2V_2 + \cdots + 2V_{2s} + b$. Then $2V_1 + \cdots + 2V_{2s} + b = 2V_1 + W$. To estimate the right hand side of (6.9) we can apply Lemma 6.1 replacing in that Lemma Z_N by V_1, N by r, n by m and e by e_1. The condition $\mathscr{N}(p, \delta, \mathscr{S}, \tilde{X})$ with $\mathscr{S} = \{e_1, \dots, e_s\}$ guarantees that $\mathbf{P}\{|\tilde{X} - e_1| \le \delta\} \ge p$. Thus, using $N \ll_s r \ll_s N$ we get

$$J \le c_\gamma(s)(pN)^{-\gamma} \int |H_t|_\infty\, dt + \sup_{\Gamma} \sup_{b\in\mathbb{R}^d} \int \left|\mathbf{E} H_t(T^{(1)} + W + b)\right| dt \ , \tag{6.10}$$

where $T^{(1)}$ is a sum of $m = \lceil pr/2 \rceil$ independent random vectors of the class $\Gamma(\delta, \{e_1\})$. Introducing i.i.d. Rademacher random variables ε_{kj}, we can write

$$T^{(1)} = \sum_{j=1}^{m} \varepsilon_{1j} z_{1j}, \quad \text{with } |z_{1j} - e_1| \le \delta \ .$$

Splitting $W = V_2 + W'$ with $W' = V_3 + \cdots + V_{2s}$, to the last integral in (6.10) we can apply the procedure leading from (6.9) to (6.10) with V_2 instead of V_1. Now we can repeat the procedure with $V_3, \cdots, V_{2s}$. We obtain

$$J \le c_\gamma(s)(pN)^{-\gamma} |H_t|_\infty + \sup_{\Gamma} \sup_{b \in \mathbb{R}^d} \int \left| \mathbf{E} H_t(T^{(1)} + \cdots + T^{(2s)} + b) \right| dt \qquad (6.11)$$

with $T^{(k)} = \sum_{j=1}^{m} \varepsilon_{kj} z_{kj}$ such that

$$\left| z_{kj} - e_k \right| \le \delta, \quad \text{for } 1 \le k \le s, \quad \left| z_{kj} - \mathbb{Q}\dot{e}_k \right| \le \delta, \quad \text{for } s+1 \le k \le 2s \ .$$

Reordering the summands we can write

$$T^{(1)} + \cdots + T^{(s)} = \sum_{k=1}^{s} \sum_{j=1}^{m} \varepsilon_{kj} z_{kj} = \sum_{j=1}^{m} U_j, \quad \text{with } U_j = \sum_{k=1}^{s} \varepsilon_{kj} z_{kj}. \qquad (6.12)$$

Define $z'_{kj} = \mathbb{Q} z_{kj}$. Since $\mathbb{Q}^2 = \mathbb{I}$, we have $z_{kj} = \mathbb{Q} z'_{kj}$. Furthermore, the inequality $\left| z_{kj} - \mathbb{Q} e_k \right| \le \delta$ is equivalent to $|z'_{kj} - e_k| \le \delta$. Thus, we can write

$$T^{(s+1)} + \cdots + T^{(2s)} = \sum_{k=1}^{s} \sum_{j=1}^{m} \varepsilon_{kj} \mathbb{Q} z'_{kj} = \mathbb{Q} \sum_{j=1}^{m} U'_j, \quad \text{with } U'_j = \sum_{k=1}^{s} \varepsilon_{kj} z'_{kj} \ . \qquad (6.13)$$

The relations (6.11)–(6.13) together imply (6.7) with $m \ge n$ instead of n. But we can remove all summands U_j, U'_j with $n < j \le m$ conditioning on them. □

Lemma 6.2 allows to bound the following integrals over the characteristic functions.

Corollary 6.3. *Let* $\mathbb{Q}^2 = \mathbb{I}$. *Assume* $\mathcal{N}(p, \delta, \mathcal{S}, \tilde{X})$. *Write* $n = \lceil pN/(5s) \rceil$. *Then, for any* $0 < A \le B$ *and* $\gamma \ge 0$, *we have*

$$\int_{A \le |t| \le B} |\mathbf{E}\, \mathrm{e}\{t\mathbb{Q}[Z_N - a]\}| \, \frac{dt}{|t|} \le I + c_\gamma(s)(pN)^{-\gamma} \log \frac{B}{A} \ ,$$

with

$$I = \sup_{\Gamma} \sup_{b \in \mathbb{R}^d} \int_{A \le |t| \le B} \sqrt{\varphi(t/4)} \, \frac{dt}{|t|}, \quad \varphi(t) \overset{\text{def}}{=} \left| \mathbf{E}\, \mathrm{e}\{t\mathbb{Q}[Y + \mathbb{Q}Y' + b]\} \right|^2 \ ,$$

where $Y = U_1 + \cdots + U_n$ *and* $Y' = U'_1 + \cdots + U'_n$ *denote sums of independent (non-i.i.d.!) vectors, and* $\sup_\Gamma$ *is taken over all* $\{\mathcal{L}(U_j), \mathcal{L}(U'_j) : 1 \le j \le n\} \subset \Gamma(\delta; \mathcal{S})$.

Proof. It is sufficient to choose $H_t(x) = |t|^{-1} \mathbf{I}\{A \le |t| \le B\} \mathbf{E}\, \mathrm{e}\{t\mathbb{Q}[x]/4\}$ in (6.7) of Lemma 6.2. □

7. Bounds for characteristic functions

The main result of the Section is the following Theorem 7.1, which is valid without any moment assumptions.

Theorem 7.1. *Assume that* $\mathbb{Q}^2 = \mathbb{I}$ *and that the condition* $\mathscr{N}(p, \delta, \mathscr{S}_o, \tilde{X})$ *holds with some* $0 < p \le 1$ *and* $0 \le \delta \le 1/(5s)$. *Then*

$$|\mathbf{E}\ e\{t\mathbb{Q}[Z_N - a]\}| \ll_s \mathscr{M}^s(t; pN) \ ,$$

where the function $\mathscr{M}$ *is defined by* (1.16).

Corollary 7.2. *Assume that* $\mathbb{Q}^2 = \mathbb{I}$ *and that the condition* $\mathscr{N}(p, \delta, \mathscr{S}_o, G)$ *holds with some* $0 < p \le 1$ *and* $0 < \delta \le 1/(10s)$. *Then*

$$|\mathbf{E}\ e\{t\mathbb{Q}[Z_N - a]\}| \ll_s (1 + \delta^s)\mathscr{M}^s(tm_0; pN/m_0), \quad m_0 = \beta/p \ .$$

Proof of Corollary 7.2. The proof can be reduced to collecting summands in the sum Z_N in groups of size, say m. By Lemma 5.3, a normalized sum, say Y, of m independent copies of X satisfies the condition $\mathscr{N}(p/2, 2\delta, \mathscr{S}_o, \tilde{Y})$ provided that $m \sim \beta/(p\delta^4)$, and therefore we can apply Theorem 7.1. $\square$

To prove Theorem 7.1 we need the auxiliary Lemmas 7.3 and 7.4. Lemma 7.3 is a initial step for an application of the double large sieve in Lemma 7.4. In Section 8 we shall extend the methods of this Section for the proof of the multiplicative inequality.

In the proof of the next Lemma we shall combine discretization techniques (see Lemma 6.2) with symmetrization arguments (see Lemma 5.1). An application of the geometric-arithmetic mean inequality will then reduce the problem to the i.i.d case.

Lemma 7.3. *Assume that* $\mathbb{Q}^2 = \mathbb{I}$ *and that the condition* $\mathscr{N}(p, \delta, \mathscr{S}, \tilde{X})$ *holds with some* $0 < p \le 1$ *and* $\delta > 0$. *Write* $n = \lceil pN/(11s) \rceil$. *Then*

$$|\mathbf{E}\, e\{t\mathbb{Q}[Z_N - a]\}| \le c_s(\gamma)(pN)^{-\gamma} + \sup_{\Gamma} \sqrt{\mathbf{E}\ e\{t\langle \tilde{W}, \tilde{W}'\rangle/2\}} \ , \tag{7.1}$$

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

Proof. While proving (7.1) we can assume that $pN \ge c_s$ with a sufficiently large constant c_s, since otherwise (7.1) is trivially fulfilled.

Write $H(x) = e\{t\mathbb{Q}[x]/4\}$ and $b = -2a$. Then

$$|\mathbf{E}\, e\{t\mathbb{Q}[Z_N - a]\}| = |\mathbf{E} H(2Z_N + b)| \ ,$$

and the inequality (6.8) of Lemma 6.2 implies

$$\begin{aligned} |\mathbf{E}\ e\{t\mathbb{Q}[Z_N - a]\}| &\le c_\gamma(s)(pN)^{-\gamma}|H|_\infty + \sup_\Gamma \sup_{b\in\mathbb{R}^d} |\mathbf{E} H(Y + \mathbb{Q}\mathbb{Y}' + b)| \\ &= c_\gamma(s)(pN)^{-\gamma} + \sup_\Gamma \sup_{b\in\mathbb{R}^d} |\varphi| \end{aligned} \tag{7.2}$$

with $\varphi =: \mathbf{E}\ \mathrm{e}\{t\mathbb{Q}[Y + \mathbb{Q}\mathbb{Y}' + b]/4\}$, where

$$Y = U_1 + \cdots + U_k \quad \text{and} \quad Y' = U_1' + \cdots + U_k', \quad k = \lceil pN/(5s) \rceil \ ,$$

denote independent sums of independent non-i.i.d. vectors, and $\sup_\Gamma$ is taken over all $\{\mathscr{L}(U_j), \mathscr{L}(U_j') : 1 \le j \le k\} \subset \Gamma(\delta\,; \mathscr{S})$.

In the view of (7.2), it remains to show that

$$|\varphi|^2 \le \sup_\Gamma \mathbf{E}\ \mathrm{e}\Big\{t\langle \tilde{W}, \tilde{W}'\rangle/2\Big\} \ . \tag{7.3}$$

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 $n = \lceil pN/(11s) \rceil$ independent summands U_j and U_j' respectively. Such an n exists since $pN \ge c_s$ with a sufficiently large c_s. The symmetrization Lemma 5.1 and symmetry of $\mathbb{Q}$ imply

$$2|\varphi|^2 \le \mathbf{E}\,\mathrm{e}\big\{t\langle \tilde{T}, \mathbb{Q}^2\tilde{R}\rangle/2\big\} + \mathbf{E}\ \mathrm{e}\big\{t\langle \tilde{T}, \mathbb{Q}^2\tilde{R}_1\rangle/2\big\} \ .$$

Recall that $\mathbb{Q}^2 = \mathbb{I}$. Furthermore,

$$\sup_\Gamma \mathbf{E}\ \mathrm{e}\big\{t\langle \tilde{T}, \tilde{R}\rangle\big\} = \sup_\Gamma \mathbf{E}\ \ \mathrm{e}\big\{t\langle \tilde{T}, \tilde{R}_1\rangle\big\} \ .$$

Thus,

$$|\varphi|^2 \le \sup_\Gamma \mathbf{E}\ \mathrm{e}\big\{t\langle \tilde{T}, \tilde{R}\rangle/2\big\} \ .$$

Applying the geometric-arithmetic mean inequality we have

$$\begin{aligned}
\mathbf{E}\,\mathrm{e}\big\{t\langle \tilde{T}, \tilde{R}\rangle/2\big\} &= \mathbf{E}\prod_{j=1}^n \mathbf{E}_{\tilde{U}_j}\ \mathrm{e}\big\{t\langle \tilde{U}_j, \tilde{R}\rangle/2\big\} \\
&\le \frac{1}{n}\sum_{j=1}^n \mathbf{E}\Big(\mathbf{E}_{\tilde{U}_j}\mathrm{e}\big\{t\langle \tilde{U}_j, \tilde{R}\rangle/2\big\}\Big)^n \\
&\le \sup_{\mathscr{L}(V)\in\Gamma(\delta,\mathscr{S})} \mathbf{E}\big(\mathbf{E}_{\tilde{V}}\ \mathrm{e}\big\{t\langle \tilde{V}, \tilde{R}\rangle/2\big\}\big)^n \\
&= \sup_{\mathscr{L}(V)\in\Gamma(\delta,\mathscr{S})} \mathbf{E}\ \ \mathrm{e}\big\{t\langle \tilde{W}, \tilde{R}\rangle/2\big\} \ .
\end{aligned} \tag{7.4}$$

Arguing as in (7.4), we replace $\tilde{R}$ by $\tilde{W}'$ in the last expectation, which concludes the proof of (7.3). $\square$

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

Lemma 7.4. *Assume that* $\delta \le 1/(5s)$. *Let* $W = V_1 + \cdots + V_n$ *and* $W' = V_1' + \cdots + V_n'$ *denote independent sums of independent copies of some random vectors* V *and* V' *such that* $\mathscr{L}(V), \mathscr{L}(V') \in \Gamma(\delta\,; \mathscr{S}_o)$. *Then*

$$\mathbf{E}\,\mathrm{e}\Big\{t\langle \tilde{W}, \tilde{W}'\rangle\Big\} \ll_s \mathscr{M}^{2s}(t\,; n), \quad \textit{for } t \in \mathbb{R} \ . \tag{7.5}$$

Proof. The Lemma easily follows from Lemma 4.7. Indeed, we can write

$$V = \varepsilon_1 z_1 + \cdots + \varepsilon_s z_s, \quad V' = \bar{\varepsilon}_1 z_1' + \cdots + \bar{\varepsilon}_s z_s'$$

with some z_j, $z_j' \in \mathbb{R}^d$ such that

$$|z_j - e_j| \le \delta, \quad |z_j' - e_j| \le \delta, \quad \text{for } 1 \le j \le s \ . \tag{7.6}$$

Consider the random vector $Y = (\tilde{\varepsilon}_1, \ldots, \tilde{\varepsilon}_s) \in \mathbb{R}^s$ with coordinates which are symmetrizations of i.i.d. Rademacher random variables. Obviously, $\mathbf{P}\{|Y| \le 2\sqrt{s}\} = 1$, and cov $Y = 2\mathbb{I}$, where $\mathbb{I}$ is the unity matrix. Let R and T denote independent sums of n independent copies of Y. Introduce the matrix $\mathbb{A} = \{a_{ij} : 1 \le i, j \le s\}$ with $a_{ij} = \langle z_i, z_j' \rangle$. Then we can write $\langle \tilde{W}, \tilde{W}' \rangle = \langle \mathbb{A} R, T \rangle$.

In order to estimate the characteristic function of $\langle \mathbb{A} R, T \rangle$, we shall apply Lemma 4.7, replacing d by s, N by n, U by R, and V by T. Since $\mathscr{S}_o$ is an orthonormal system, the inequalities (7.6) imply that $\mathbb{A} = \mathbb{I} + \mathbb{B}$ with some matrix $\mathbb{B} = \{b_{ij}\}$ such that $|b_{ij}| \le 2\delta + \delta^2$. Thus we have $|\mathbb{B}| \le |\mathbb{B}|_2 \le 2s\delta + s\delta^2$, where $|\mathbb{B}|_2$ denotes the Hilbert–Schmidt norm of the matrix $\mathbb{B}$. Therefore the condition $\delta \le 1/(5s)$ implies $|\mathbb{B}| \le 1/2$. Consequently, $|\mathbb{A}^{-1}| \le 2$. Thus, all conditions of Lemma 4.7 are fulfilled, and (7.5) follows. □

Proof of Theorem 7.1. We shall assume that $pN \ge c_s$ with a sufficiently large constant c_s, since otherwise the trivial inequality $|\mathbf{E}\, e\{t\mathbb{Q}[Z_N - a]\}| \le 1$ combined with

$$\inf_t \mathscr{M}^s(t\,; pN) = (pN)^{-s/4} \tag{7.7}$$

implies the result.

Let us apply Lemma 7.3 and Lemma 7.4 with $n = \lceil pN/(11s) \rceil$. We have

$$|\mathbf{E}\, e\{t\mathbb{Q}[Z_N - a]\}| \ll_{s,\gamma} (pN)^{-\gamma} + \mathscr{M}^s(t/2\,; n) \ .$$

By (1.17), $\mathscr{M}^s(t/2\,; n) \ll_s \mathscr{M}^s(t\,; pN)$. Using (7.7) and choosing $\gamma = s/4$, we have $(pN)^{-\gamma} \le \mathscr{M}^s(t\,; pN)$, which completes the proof of the Theorem. □

8. The multiplicative inequality

The main result of this section is the multiplicative inequality of Lemma 8.1 for characteristic functions of discrete random vectors. Combined with the discretization technique described in Section 5, the multiplicative inequality can be applied to bound integrals over general characteristic functions.

Introduce the independent sums

$$Y_n = \sum_{j=1}^n U_j, \quad Y_n' = \sum_{j=1}^n U_j', \quad \mathscr{L}(U_j), \mathscr{L}(U_j') \in \Gamma(\delta, \mathscr{S}_o), \text{ for } 1 \le j \le n \ , \tag{8.1}$$

of independent (non-identically distributed!) random vectors. Write

$$\varphi_n(t) = \left|\mathbf{E}\,\mathrm{e}\{t\mathbb{Q}[Y_n + \mathbb{Q}Y_n' + b]\}\right|^2, \quad b \in \mathbb{R}^d \ .$$

Lemma 8.1. *Assume that* $0 \le \delta \le 1/(5s)$ *and* $\mathbb{Q}^2 = \mathbb{I}$. *Then we have*

$$\varphi_n(t)\varphi_n(t+\tau) \ll_s \mathscr{M}^{2s}(\tau; n), \quad \text{for } t, \tau \in \mathbb{R} \ . \tag{8.2}$$

Proof. For an alternative proof of (8.2) see Bentkus and Götze (1995b).

In the proof we may assume that $n \ge 2$ since otherwise (8.2) is trivially fulfilled. We shall prove that

$$\varphi_n(t-\tau)\varphi_n(t+\tau) \ll_s \mathscr{M}^{2s}(\tau; n), \quad \text{for } t, \tau \in \mathbb{R} \ . \tag{8.3}$$

This estimate implies (8.2). Indeed, it suffices to put $t = s + \tau$ in (8.3) and to use the estimate $\mathscr{M}^{2s}(\tau; n) \ll_s \mathscr{M}^{2s}(2\tau; n)$, which can be easily verified using (1.17).

Notice that

$$\mathbb{Q}[x] - \mathbb{Q}[y] = \langle \mathbb{Q}(x+y), x-y\rangle, \quad 2\mathbb{Q}[x] + 2\mathbb{Q}[y] = \mathbb{Q}[x+y] + \mathbb{Q}[x-y] \ . \tag{8.4}$$

For an arbitrary random vector ξ, let $\bar{\xi}$ denote an independent copy of ξ. Writing

$$\theta(t) = \left|\mathbf{E}\,\mathrm{e}\{t\mathbb{Q}[\xi + b]\}\right| \ ,$$

using $\theta(t) = \theta(-t)$ and applying (8.4), we have

$$\begin{aligned}
\theta(t+\tau)\theta(t-\tau) &= \left|\mathbf{E}\,\mathrm{e}\{(t+\tau)\mathbb{Q}[\xi+b] - (t-\tau)\mathbb{Q}[\bar{\xi}+b]\}\right| \\
&= \left|\mathbf{E}\,\mathrm{e}\{t\mathbb{Q}[\xi+b] - t\mathbb{Q}[\bar{\xi}+b] + \tau\mathbb{Q}[\xi+b] + \tau\mathbb{Q}[\bar{\xi}+b]\}\right| \\
&= \left|\mathbf{E}\mathrm{e}\big\{t\langle \mathbb{Q}\xi + \mathbb{Q}\bar{\xi} + 2\mathbb{Q}b, \xi - \bar{\xi}\rangle + \tfrac{\tau}{2}\mathbb{Q}[\xi + \bar{\xi} + 2b] + \tfrac{\tau}{2}\,\mathbb{Q}[\xi - \bar{\xi}]\big\}\right| \ .
\end{aligned}$$

We shall use (8.5) with $\xi = Y_n + \mathbb{Q}Y_n'$. Let ε_{jl}, $\bar{\varepsilon}_{jl}$, ε_{jl}^*, j, $l \ge 1$, denote i.i.d. Rademacher random variables, which are independent in aggregate. The random vectors Y_n and Y_n' are sums of U_j and U_j', and using (8.1) we can write

$$\xi = Y_n + \mathbb{Q}Y_n' = \sum_{j=1}^{n}\sum_{l=1}^{s} \varepsilon_{jl} z_{jl} + \mathbb{Q}\sum_{j=n+1}^{2n}\sum_{l=1}^{s} \varepsilon_{jl} z_{jl} \ ,$$

where z_{jl} denote non-random vectors in $\mathbb{R}^d$ such that $|z_{jl} - e_l| \le \delta$, for all possible values of j and l. Consequently, we can write

$$\xi + \bar{\xi} = \sum_{j=1}^{n}\sum_{l=1}^{s}(\varepsilon_{jl} + \bar{\varepsilon}_{jl}) z_{jl} + \mathbb{Q}\sum_{j=n+1}^{2n}\sum_{l=1}^{s}(\varepsilon_{jl} + \bar{\varepsilon}_{jl}) z_{jl} \ ,$$

$$\xi - \bar{\xi} = \sum_{j=1}^{n}\sum_{l=1}^{s}(\bar{}_{jl} - \bar{\varepsilon}_{jl}) z_{jl} + \mathbb{Q}\sum_{j=n+1}^{2n}\sum_{l=1}^{s}(\varepsilon_{jl} - \bar{\varepsilon}_{jl}) z_{jl} \ .$$

Note that the 2-dimensional random vectors

$$(\varepsilon + \bar{\varepsilon}, \varepsilon - \bar{\varepsilon}) \quad \text{and} \quad ((1 + \varepsilon\bar{\varepsilon})\varepsilon^*, \varepsilon - \bar{\varepsilon}) \tag{8.6}$$

have the same distribution provided that $\varepsilon, \bar{\varepsilon}, \varepsilon^*$ are i.i.d. Rademacher random variables.

By (8.6), the joint distribution of $\xi + \bar{\xi}$ and $\xi - \bar{\xi}$ does not change, if we replace $\xi + \bar{\xi}$ by

$$\eta \overset{\text{def}}{=} \sum_{j=1}^{n} \sum_{l=1}^{s} (1 + \varepsilon_{jl}\bar{\varepsilon}_{jl})\varepsilon^*_{jl} z_{jl} + \mathbb{Q} \sum_{j=n+1}^{2n} \sum_{l=1}^{s} (1 + \varepsilon_{jl}\bar{\varepsilon}_{jl})\varepsilon^*_{jl} z_{jl} \ .$$

Thus, denoting by $\mathbf{E}_*$ the partial integration with respect to the distributions of the random variables ε^*_{jl}, using (8.5) and an inequality of type $|\mathbf{E}X|^2 \le \mathbf{E}\,|X|^2$, we have

$$\begin{aligned} \varphi_n(t+\tau)\varphi_n(t-\tau) &\le \mathbf{E}\Big|\mathbf{E}_*\, \mathrm{e}\Big\{t\langle \mathbb{Q}\eta + 2\mathbb{Q}b, \xi - \bar{\xi}\rangle + \tfrac{\tau}{2}\mathbb{Q}[\eta + 2b] + \tfrac{\tau}{2}\mathbb{Q}[\xi - \bar{\xi}]\Big\}\Big|^2 \\ &= \mathbf{E}\Big|\mathbf{E}_*\, \mathrm{e}\big\{t\langle \mathbb{Q}\eta, \xi - \bar{\xi}\rangle + \tfrac{\tau}{2}\mathbb{Q}[\eta + 2b]\big\}\Big|^2 \end{aligned} \tag{8.7}$$

since $\xi - \bar{\xi}$ and ε^*_{jl}, $j, l \ge 1$, are independent. Write

$$\eta = P_n + \mathbb{Q}P'_n, \quad P_n = \sum_{j=1}^{n} V_j, \quad P'_n = \sum_{j=n+1}^{2n} V_j, \quad V_j = \sum_{l=1}^{s} (1 + \varepsilon_{jl}\bar{\bar{jl}})\varepsilon^*_{jl} z_{jl} \ .$$

Given ε_{jl}, $\bar{\varepsilon}_{jl}$, the random vectors V_j, $1 \le j \le 2n$, are independent. Split $P_n = T + T_1$ and $P'_n = R + R_1 + R_2$ so that each of the sums T, R, R_1 contains $k = \lceil n/2 \rceil$ summands V_j. By an application of the symmetrization Lemma 5.1 (cf. the proof of (7.3)), we obtain

$$2\Big|\mathbf{E}_*\, \mathrm{e}\big\{t\langle \mathbb{Q}\eta, \xi - \bar{\xi}\rangle + \tfrac{\tau}{2}\mathbb{Q}[\eta + 2b]\big\}\Big|^2 \le \mathbf{E}_*\, \mathrm{e}\{\tau\langle \tilde{T}, \tilde{R}\rangle\} + \mathbf{E}_*\, \mathrm{e}\{\tau\langle \tilde{T}, \tilde{R}_1\rangle\} \tag{8.8}$$

since $\mathbb{Q}^2 = \mathbb{I}$. Similarly to the proof of (7.3), we have

$$\sup_{\Gamma} \mathbf{E}\mathbf{E}_*\, \mathrm{e}\{\tau\langle \tilde{T}, \tilde{R}\rangle\} = \sup_{\Gamma} \mathbf{E}\mathbf{E}_*\, \mathrm{e}\{\tau\langle \tilde{T}, \tilde{R}_1\rangle\} \ ,$$

and (8.7) together with (8.8) yields

$$\varphi_n(t+\tau)\varphi_n(t-\tau) \le \sup_{\Gamma} \mathbf{E}\, \mathrm{e}\{\tau\langle \tilde{T}, \tilde{R}\rangle\} \ , \tag{8.9}$$

where $\tilde{T} = \tilde{V}_1 + \cdots + \tilde{V}_k$, $\tilde{R} = \tilde{V}_{k+1} + \cdots + \tilde{V}_{2k}$ with

$$\tilde{V}_j = \sum_{l=1}^{s} (1 + \varepsilon_{jl}\bar{\varepsilon}_{jl})\tilde{\varepsilon}^*_{jl} z_{jl} \ ,$$

and where $\sup_{\Gamma}$ is taken over all z_{jl} such that $|z_{jl} - e_l| \le \delta$, for all possible values of j and l. The random variable $\varepsilon\bar{\varepsilon}$ is distributed as ε. Thus, replacing the random variables $\varepsilon_{jl}\bar{\varepsilon}_{jl}$ and $\tilde{\varepsilon}^*_{jl}$ by jointly independent copies, say ε_{jl} and $\tilde{\varepsilon}_{jl}$, we may assume that $\tilde{V}_j = \sum_{l=1}^{s}(1 + \varepsilon_{jl})\tilde{\varepsilon}_{jl} z_{jl}$.

By an application of the geometric-arithmetic inequality (cf. (7.4) in the proof of (7.3)), we obtain

$$\sup_{\Gamma} \mathbf{E}\, e\{\tau\langle \tilde{T}, \tilde{R}\rangle\} \le \sup_{\Gamma} \mathbf{E}\, e\{\tau\langle \tilde{W}, \tilde{W}'\rangle\} \ , \tag{8.10}$$

where $\tilde{W}$ (resp., $\tilde{W}'$) is a sum of $k = \lceil n/2 \rceil$ independent copies of

$$\tilde{U} = \sum_{l=1}^{s} (1+\varepsilon_l)\tilde{\varepsilon}_l z_l \quad \left(\text{resp., of } \tilde{U}' = \sum_{l=1}^{s} (1+\varepsilon_l)\tilde{\varepsilon}_l z_l'\right) \ ,$$

and $\sup_\Gamma$ is taken over all z_l, z_l' such that

$$|z_l - e_l| \le \delta, \quad |z_l' - e_l| \le \delta, \qquad \text{for} \quad 1 \le j \le s \ .$$

In order to bound $\mathbf{E}\, e\{\tau\langle \tilde{W}, \tilde{W}'\rangle\}$, we can apply a modification of Lemma 7.4 (it suffices in the proof of that Lemma just to replace the random vector $Y = (\tilde{\varepsilon}_1, \ldots, \tilde{\varepsilon}_s)$ by the random vector $((1+\varepsilon_1)\tilde{\varepsilon}_1, \ldots, (1+\varepsilon_s)\tilde{\varepsilon}_s)$. We obtain (recall that $k = \lceil n/2 \rceil$)

$$\mathbf{E}\, e\{\tau\langle \tilde{W}, \tilde{W}'\rangle\} \ll_s \mathscr{M}^{2s}(\tau; k) \ll_s \mathscr{M}^{2s}(\tau; n) \ . \tag{8.11}$$

The estimates (8.9)–(8.11) together imply (8.3). $\square$

9. Expansions of characteristic functions

In this Section we shall obtain bounds for

$$\widehat{\Delta}_N(t) = \widehat{\Psi}(t) - \widehat{\Psi}_0(t) - \widehat{\Psi}_1(t) \ ,$$

where Ψ and Ψ_j are defined by (1.18) and (1.19).

Throughout we shall assume that $\mathbf{E}X = \mathbf{E}\,G = 0$ and $\operatorname{cov} X = \operatorname{cov} G$.

Recall that we write

$$b = N^{1/2}a, \quad \beta_s = \mathbf{E}|X|^s, \quad \beta = \beta_4, \quad \sigma^2 = \beta_2, \quad Z_N = X_1 + \cdots + X_N \ .$$

We shall denote as well $\mathbb{D} = t\mathbb{Q}$ and $U_N = G_1 + \cdots + G_N$. Since $\sqrt{N}G \overset{\mathscr{D}}{=} U_N$, we can write

$$\widehat{\Psi}(t) = \mathbf{E}\, e\{\mathbb{D}[Z_N - b]\}, \quad \widehat{\Psi}_0(t) = \mathbf{E}\, e\{\mathbb{D}[U_N - b]\} \ ,$$
$$\widehat{\Psi}_1(t) = -\frac{1}{\sqrt{N}} \mathbf{E}\left(\frac{4i}{3}\langle N\mathbb{D}Y, X\rangle^3 + 2\langle N\mathbb{D}Y, X\rangle N\mathbb{D}[X]\right) e\{N\mathbb{D}[Y]\} \ ,$$

where $Y = G - a$. We shall use the upper bound $\varkappa = \varkappa(t; N, \mathscr{L}(X)) + \varkappa(t; N, \mathscr{L}(G))$ (cf. (3.23)), where

$$\varkappa(t; N, \mathscr{L}(X)) = \sup_{a \in \mathbb{R}^d} |\mathbf{E}\, e\{\mathbb{D}[Z_k] + \langle a, Z_k\rangle\}|, \quad k = [(N-2)/14] \ .$$

We start with Lemma 9.1 since its proof is simpler than the proof of the main Lemma 9.2. Here we shall discuss standard technical steps which will be used in the proofs of Lemmas 9.2 and 9.3. In order to remove lower order terms, we shall use frequently without mentioning simple inequalities like $\beta_3^2 \le \beta\sigma^2$

(a consequence of Hölder's inequality), $x^\alpha y^\gamma \ll_{\alpha,\gamma} x^{\alpha+\gamma} + y^{\alpha+\gamma}$, for $x, y, \alpha, \gamma > 0$, as well as $1 + x^\alpha \ll 1 + x^\gamma$ with $\alpha \le \gamma$.

Lemma 9.1. *Assume that* $\sigma = 1$, $\mathbb{Q}^2 = \mathbb{I}$ *and* $\mathbf{E}\langle X, x\rangle^3 = 0$, *for all* $x \in \mathbb{R}^d$. *Then* $\widehat{\Delta}_N(t) = \widehat{\Psi}(t) - \widehat{\Psi}_0(t)$ *and we have*

$$|\widehat{\Delta}_N(t)| \ll \varkappa\beta t^2 N(1 + t^2N^2)(1 + \beta/N)(1 + |a|^4) \ .$$

Proof. For arbitrary $\mathbb{D} : \mathbb{R}^d \to \mathbb{R}^d$ we shall prove that

$$|\widehat{\Delta}_N(t)| \ll \varkappa N|\mathbb{D}|^4\beta(|b|^4 + N\beta + N^2\sigma^4) + \varkappa N|\mathbb{D}|^2\beta \tag{9.1}$$

with $\sigma^2 = \mathbf{E}|X|^2$. Setting

$$\sigma = 1, \quad |\mathbb{D}| = \sup_{|x|=1} |\mathbb{D}x| \le |t|, \quad b = \sqrt{N}a$$

in (9.1), we obtain the result of the Lemma.

We start with the following Bergström type identity:

$$\widehat{\Delta}_N(t) = \widehat{\Psi}(t) - \widehat{\Psi}_0(t) = \mathbf{E}\,\mathrm{e}\big\{\mathbb{D}[Z_N - b]\big\} - \mathbf{E}\,\mathrm{e}\big\{\mathbb{D}[U_N - b]\big\} = \sum_{k=1}^{N} J_k \ ,$$

where $J_k = \mathbf{E}\,\mathrm{e}\{\mathbb{D}[T + X]\} - \mathbf{E}\,\mathrm{e}\{\mathbb{D}[T + G]\}$. Here we used the notation:

$$T = G_2 + \cdots + G_k + X_{k+1} + \cdots + X_N - b \ .$$

In order to obtain (9.1), it suffices to prove that

$$|J_k| \ll \varkappa|\mathbb{D}|^4\beta(|b|^4 + N\beta + N^2\sigma^4) + \varkappa|\mathbb{D}|^2\beta \ . \tag{9.2}$$

Writing $\mathbb{D}[T + u] = \mathbb{D}[T] + 2\langle\mathbb{D}T, u\rangle + \mathbb{D}[u]$ and expanding in powers of $\mathbb{D}[u]$ with $u = X$ and $u = G$ respectively, we obtain $J_k = I_0 + I_1 + R_1$, where

$$I_r = \mathbf{E}(i\mathbb{D}[X])^r\,\mathrm{e}\{\mathbb{D}[T] + 2\langle\mathbb{D}T, X\rangle\} - \mathbf{E}(i\mathbb{D}[G])^r\,\mathrm{e}\{\mathbb{D}[T] + 2\langle\mathbb{D}T, G\rangle\} \ ,$$

and $|R_1| \le |\mathbb{D}|^2\beta\varkappa$. Thus, to conclude the proof of (9.2) it suffices to show that $|I_0|$ and $|I_1|$ are bounded from above by the right hand side of (9.2).

Let us estimate $|I_0|$. Let τ be a random variable uniformly distributed in $[0, 1]$ and independent of all other random variables and vectors. Expanding in powers of $2\langle\mathbb{D}T, u\rangle$ with $u = X$ and $u = G$, we obtain $I_0 = L_X - L_G$, where

$$L_Z = \tfrac{8}{3}\mathbf{E}(1 - \tau)^3\langle\mathbb{D}T, Z\rangle^4\,\mathrm{e}\{\mathbb{D}[T] + 2\tau\langle\mathbb{D}T, Z\rangle\} \ .$$

In this expansion lower order terms cancel since the expectation and the covariance of X are equal to those of G, and since $\mathbf{E}\langle X, x\rangle^3 = \mathbf{E}\langle G, x\rangle^3$, for all $x \in \mathbb{R}^d$. In order to estimate L_Z split the sum $T = \sum_{j=1}^5 T_j$ into five sums containing an approximately equal number of summands X_l or G_l, and include the shift b in one of these sums. Then

$$|L_Z| \le \sum_{j_1,j_2,j_3,j_4=1}^{5} L^*, \quad L^* = \left|\mathbf{E}(1 - \tau)^3\prod_{r=1}^{4}\langle\mathbb{D}T_{j_r}, Z\rangle\,\mathrm{e}\{\mathbb{D}[T] + 2\tau\langle\mathbb{D}T, Z\rangle\}\right| \ .$$

Note that $\{j_1, j_2, j_3, j_4\}$ is a proper subset of $\{1, \dots, 5\}$. Thus, without loss of generality we may assume that j_1, j_2, j_3, j_4 are different from $j = 5$. Conditioning on the random vectors with the indices present in L^*, we have

$$L^* \le \mathbf{E} \prod_{r=1}^{4} \big| \langle \mathbb{D} T_{j_r}, Z \rangle \big| \big| \mathbf{E}_{T_5} \, \mathrm{e}\{\mathbb{D}[T] + 2\tau \langle \mathbb{D} T, Z \rangle\} \big| \ .$$

The expectation $|\mathbf{E}_{T_5} \dots|$ is bounded from above by $\varkappa$. The product can be estimated using the arithmetic-geometric mean inequality, and it is bounded by a sum of terms $\mathbf{E} |\langle \mathbb{D} T_j, Z \rangle|^4$. Applying Rosenthal's inequality we have

$$\mathbf{E} \big| \langle \mathbb{D} T_j, Z \rangle \big|^4 \le |\mathbb{D}|^4 \mathbf{E} |T_j|^4 |Z|^4 \le c |\mathbb{D}|^4 \beta (|b|^4 + N\beta + N^2 \beta_2^2) \ .$$

Collecting the bounds we see that $|I_0|$ is bounded by the right-hand side of (9.2).

The estimation of $|I_1|$ is similar to that of $|I_0|$. Here we should expand in powers of $2\langle \mathbb{D} T, Z \rangle$ in a shorter series. We arrive at moments of type $\mathbf{E} |\mathbb{D}[Z]| \, |\langle \mathbb{D} T_j, Z \rangle|^2$. Using $ab \le a^2 + b^2$, we obtain that these moments are bounded from above by a sum of $\mathbf{E} |\mathbb{D}[Z]|^2$ and $\mathbf{E} |\langle \mathbb{D} T_j, Z \rangle|^4$. But these in turn are bounded from above by the right-hand side of (9.2), as already shown in the estimation of $|I_0|$. $\square$

Lemma 9.2. *Assume that* $\sigma = 1$ *and* $\mathbb{Q}^2 = \mathbb{I}$. *Then we have*

$$|\widehat{\Delta}_N(t)| \ll \varkappa \beta t^2 N (1 + t^4 N^4)(1 + \beta_6 / N^2)(1 + |a|^6) \ .$$

Proof. For arbitrary $\mathbb{D} : \mathbb{R}^d \to \mathbb{R}^d$ and σ we shall prove that

$$|\widehat{\Delta}_N(t)| \ll N \varkappa |\mathbb{D}|^2 \beta \Big(1 + |\mathbb{D}|^2 |b|^4 + |\mathbb{D}|^4 (N^2 \sigma^2 \beta_6 + N^4 \sigma^8 + N \sigma^2 |b|^6) \Big) \ . \tag{9.3}$$

Setting $\sigma = 1$, $|\mathbb{D}| \le |t|$ and $b = \sqrt{N} a$ in (9.3), we obtain the result of the Lemma.

It is easy to notice that the following Bergström type identity holds:

$$\begin{aligned} \widehat{\Delta}_N(t) &= \widehat{\Psi}(t) - \widehat{\Psi}_0(t) - \widehat{\Psi}_1(t) = \mathbf{E}\, \mathrm{e}\{\mathbb{D}[Z_N - b]\} - \mathbf{E}\, \mathrm{e}\{\mathbb{D}[U_N - b]\} - \widehat{\Psi}_1(t) \\ &= NJ - \widehat{\Psi}_1(t) + \sum_{k=2}^{N} (k-1)(J_1 - J_2 - J_3 + J_4) \ . \end{aligned} \tag{9.4}$$

Here we used the notation:

$$\begin{aligned} J = \mathbf{E}\, \mathrm{e}\{\mathbb{D}[S + X]\} - \mathbf{E}\, \mathrm{e}\{\mathbb{D}[S + G]\}, \quad S = G_2 + \dots + G_N - b, \\ J_1 = \mathbf{E}\, \mathrm{e}\{\mathbb{D}[T + X + \bar{X}]\}, \quad J_2 = \mathbf{E}\, \mathrm{e}\{\mathbb{D}[T + G + \bar{X}]\}, \\ J_3 = \mathbf{E}\, \mathrm{e}\{\mathbb{D}[T + X + \bar{G}]\}, \quad J_4 = \mathbf{E}\, \mathrm{e}\{\mathbb{D}[T + G + \bar{G}]\} \ , \end{aligned}$$

and $T = G_3 + \dots + G_k + X_{k+1} + \dots + X_N - b$.

In the view of (9.4), the relation (9.3) follows provided that we verify that $J = N^{-1} \widehat{\Psi}_1(t) + R_0$ with

$$|R_0| \ll \varkappa|\mathbb{D}|^2\beta\Big(1+N^2|\mathbb{D}|^2\sigma^4+|\mathbb{D}|^2|b|^4\Big) \ , \tag{9.5}$$

and that

$$\begin{aligned}|J_1-J_2-J_3+J_4| \\ &\ll \varkappa|\mathbb{D}|^3\beta_3^2+\varkappa|\mathbb{D}|^4\beta(\beta+N\sigma^4+|b|^2\sigma^2) \\ &\quad+\varkappa|\mathbb{D}|^6\beta_3^2(N\beta_6+N^3\sigma^6+|b|^6) \ .\end{aligned} \tag{9.6}$$

Let us prove (9.5). Taylor expansions in powers of $\mathbb{D}[u]$ and $2\langle\mathbb{D}S,u\rangle$ with $u=X$ and $u=G$ combined with the techniques used in the proof of Lemma 9.1 give

$$J=J_0+R_1, \quad J_0 \overset{\mathrm{def}}{=} -\frac{4i}{3}\mathbf{E}\langle\mathbb{D}S,X\rangle^3\,\mathrm{e}\{\mathbb{D}[S]\}-2\mathbf{E}\langle\mathbb{D}S,X\rangle\mathbb{D}[X]\mathrm{e}\{\mathbb{D}[S]\} \tag{9.7}$$

with some R_1 bounded similarly as R_0 in (9.5). To complete the proof of (9.5) it suffices to replace S in (9.7) by $S+G_1\overset{\mathscr{D}}{=}\sqrt{N}G-b$. This can be done using again Taylor expansions in powers of $\langle\mathbb{D}G_1,X\rangle$ and $\mathbb{D}[G_1]$. A remainder term, say R_2, of such a replacement is bounded similarly as R_0 in (9.5). Therefore the remark that $G_1+\cdots+G_N\overset{\mathscr{D}}{=}\sqrt{N}G$ concludes the proof of (9.5).

Thus, in order to complete the proof of the Lemma, it remains to prove (9.6). Writing

$$\mathbb{D}[T+u+v]=\mathbb{D}[T]+K(u)+K(v)+2\langle\mathbb{D}u,v\rangle, \quad K(u)\overset{\mathrm{def}}{=}\mathbb{D}[u]+2\langle\mathbb{D}T,u\rangle \ ,$$

expanding in powers of $2\langle\mathbb{D}u,v\rangle$ and applying the conditioning techniques used in the proof of Lemma 9.1, we obtain

$$J_s=J_{s0}+J_{s1}+2^{-1}J_{s2}+R_s, \quad 1\le s\le 4 \ ,$$

where, for example,

$$J_{1r}=\mathbf{E}(i\langle\mathbb{D}X,\bar{X}\rangle)^r\,\mathrm{e}\{\mathbb{D}[T]\}L(X)L(\bar{X}), \quad L(u)\overset{\mathrm{def}}{=}\mathrm{e}\{K(u)\} \ ,$$

for $0\le r\le 2$, and

$$|R_s|\le\varkappa|\mathbb{D}|^3\beta_3^2, \text{ for all } \quad 1\le s\le 5 \ .$$

Thus

$$|J_1-J_2-J_3+J_4|\le|I_0|+|I_1|+|I_2|+c\varkappa|\mathbb{D}|^3\beta_3^2 \ , \tag{9.8}$$

where

$$I_r=J_{1r}-J_{2r}-J_{3r}+J_{4r}, \quad 0\le r\le 2 \ .$$

The estimate (9.8) shows that in order to prove (9.6) it suffices to verify that, for $r=0,1,2$,

$$|I_r|\ll\varkappa|\mathbb{D}|^6\beta_3^2(|b|^6+N\beta_6+N^3\sigma^6)+\varkappa|\mathbb{D}|^4\beta_3^2(|b|^2+N\sigma^2)+\varkappa|\mathbb{D}|^4\beta^2 \ . \tag{9.9}$$

Let us prove (9.9) for I_0. It is easy to see that

$$I_0 = \mathbf{E}\,\mathrm{e}\{\mathbb{D}[T]\}(\mathbf{E}_X L(X) - \mathbf{E}_G L(G))(\mathbf{E}_{\bar{X}} L(\bar{X}) - \mathbf{E}_{\bar{G}} L(\bar{G})) \ , \tag{9.10}$$

where $L(u) = \mathrm{e}\{K(u)\}$. Using Taylor expansions of the function $x \to \mathrm{e}\{x\}$, let us represent the differences in (9.10) as sums of terms of third or fourth order in $u = X, \bar{X}, G, \bar{G}$. More precisely, expanding in powers of $y \stackrel{\text{def}}{=} \mathbb{D}[u]$, we have

$$L(u) = \mathrm{e}\{x\} + iy\,\mathrm{e}\{x\} - \mathbf{E}_\tau (1-\tau) y^2\, \mathrm{e}\{\tau y + x\}, \quad x \stackrel{\text{def}}{=} 2\langle \mathbb{D}T, u\rangle \ . \tag{9.11}$$

Expanding now $\mathrm{e}\{x\}$ in powers of x with a remainder $\mathcal{O}(x^3)$ for the first summand in the right hand side of (9.11), and with a remainder $\mathcal{O}(x)$ for the second summand, we obtain a representation of $L(u)$ as a sum of terms up to fourth order in u. Using this representation of $L(u)$ we obtain the desired representation for the difference $\mathbf{E}_X L(X) - \mathbf{E}_G L(G)$ in (9.10); notice that in this representation terms of order two and less cancel since X and G have the same covariances and expectations. A similar representation is valid for the second difference $\mathbf{E}_{\bar{X}} L(\bar{X}) - \mathbf{E}_{\bar{G}} L(\bar{G})$ in (9.10). Multiplying the representations of the differences term-wise, applying splitting techniques similar to the proof of Lemma 9.1, and using Rosenthal's inequality we derive (9.9) for I_0.

Let us prove (9.9) for I_1. It is easy to see that $I_1 = iJ(\bar{X}) - iJ(\bar{G})$, where

$$J(u) = \mathbf{E}\,\mathrm{e}\{\mathbb{D}[T]\}\mathbf{E}_u L(u)(\mathbf{E}_X \langle \mathbb{D}X, u\rangle L(X) - \mathbf{E}_G\langle \mathbb{D}G, u\rangle L(G)) \ . \tag{9.12}$$

Using the equality of means and covariances of $X, \bar{X}, G, \bar{G}$ we can replace $L(u)$ in (9.12) by $L(u) - 1 - 2i\langle \mathbb{D}T, u\rangle$. By Taylor expansions $L(u) - 1 - 2i\langle \mathbb{D}T, u\rangle = \mathcal{O}(\langle \mathbb{D}T, u\rangle^2 + \mathbb{D}[u])$. Similarly

$$\mathbf{E}_X\langle \mathbb{D}X, u\rangle L(X) - \mathbf{E}_G\langle \mathbb{D}G, u\rangle L(G) = \mathcal{O}(\langle \mathbb{D}X, u\rangle\langle \mathbb{D}T, u\rangle^2 + \mathbb{D}[u]) \ .$$

Thus we can proceed as in the estimation of I_0, and to obtain (9.9) for I_1.

The proof of (9.9) for I_2 is somewhat simpler than the proof for I_1, so we omit it. $\square$

For $0 \le k \le N$ define

$$\widehat{\Psi}^{(k)}(t) = \mathbf{E}\,\mathrm{e}\{t\mathbb{Q}[G_1 + \cdots + G_k + X_{k+1} + \cdots + X_N - a]\} \ .$$

Notice that $\widehat{\Psi}^{(0)}(t) = \widehat{\Psi}(t)$ and $\widehat{\Psi}^{(N)}(t) = \widehat{\Psi}_0(t)$.

Lemma 9.3. *Assume that $\sigma = 1$ and $\mathbb{Q}^2 = \mathbb{I}$. Then we have*

$$\left|\widehat{\Psi}(t) - \widehat{\Psi}^{(k)}(t)\right| \ll \varkappa t^2 k(\beta + |t| N\beta + |t| N\sqrt{N\beta})(1 + |a|^3) \ .$$

Proof. Obviously $|\widehat{\Psi}(t) - \widehat{\Psi}^{(k)}(t)| \le I_1 + \cdots + I_k$, where

$$I_j = \left|\mathbf{E}\,\mathrm{e}\{t\mathbb{Q}[S + X]\} - \mathbf{E}\,\mathrm{e}\{t\mathbb{Q}[S + G]\}\right| \ ,$$

$$S \stackrel{\text{def}}{=} G_1 + \cdots + G_{j-1} + X_{j+1} + \cdots + X_N - a \ .$$

An application of splitting and conditioning techniques, combined with Taylor expansions of the exponents with remainders $\mathcal{O}((t\mathbb{Q}[u])^2)$ and $\mathcal{O}(\langle t\mathbb{Q}S, u\rangle^3)$ with $u = X$ and $u = G$, conclude the proof. $\square$

Define the distributions

$$\mu(A) = \mathbf{P}\{U_k + \sum_{j=k+1}^{N} X_j \in \sqrt{N}A\}, \quad \mu_0(A) = \mathbf{P}\{U_N \in \sqrt{N}A\} \ .$$

For measurable sets $A \subset \mathbb{R}^d$ define the Edgeworth correction (to the distribution μ) as $\mu_1^{(k)}(A) = (N-k)N^{-3/2}\chi(A)$, where the signed measure χ is given by (1.22). Introduce the signed measure $v = \mu - \mu_0 - \mu_1^{(k)}$.

Lemma 9.4. *Assume that $d < \infty$ and $1 \le k \le N$. Then,*

$$\delta_N \stackrel{\text{def}}{=} \sup_{A \subset \mathbb{R}^d} |v(A)| \ll_d \frac{\beta}{\sigma_d^4 N} + \frac{\beta^{(d+7)/2} N^{(d+3)/2}}{k^{d+5}\sigma_d^{2d+14}} \ . \tag{9.13}$$

An outline of the proof. (cf. the proof of Lemma 2.5 in Bentkus and Götze 1996). Assuming that $\operatorname{cov} X = \operatorname{cov} G = \mathbb{I}$, we shall prove that

$$\delta_N \ll_d \frac{\beta}{N} + \frac{\beta^{(d+7)/2} N^{(d+3)/2}}{k^{d+5}} \ . \tag{9.14}$$

Applying (9.14) to $\mathbb{C}^{-1/2}X$ and $\mathbb{C}^{-1/2}G$ and estimating $|\mathbb{C}^{-1/2}| \le 1/\sigma_d$, we obtain (9.13).

While proving (9.14) we can assume that $\beta/N \le c_d$ and $N \ge 1/c_d$ with a sufficiently small positive constant c_d. Otherwise (9.14) follows from the trivial bound

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

To prove (9.14) we shall apply truncation of X_j, centering and a correction of the covariances of Gaussian summands G_j, for $k+1 \le j \le N$. Namely, in (9.14) we can replace X_j by $X_j^\diamond = X_j \mathbf{I}\{|X_j| \le \sqrt{N}\}$ up to an error β/N. The centering, that is, a replacement of $X_j^\diamond$ by $X_j' \stackrel{\text{def}}{=} X_j^\diamond - \mathbf{E}X^\diamond$, produces an error bounded by β/N. A correction of the covariances of the Gaussian random vectors yields a similar error. We shall denote the corrected Gaussian random vectors by G_j'.

After such a replacement of X_j and G_j by X_j' and G_j', for $k+1 \le j \le N$, we can assume that all eigenvalues of $\operatorname{cov} X_j'$ belong to the interval [1/2, 3/2]. Otherwise a trivial bound $\beta \ge cN$ implies the result. As a consequence of the truncation we have

$$\mathbf{E}|S_N|^s \ll_{s,d} 1, \quad s > 0 \ . \tag{9.15}$$

Denoting by Z_m' and U_m' sums of m independent copies of X' and G' respectively, introduce the multidimensional characteristic functions $g(t) = \mathbf{E}\,\mathrm{e}\{\langle t, G\rangle\}$,

$$f(t) = \mathbf{E}\,\mathrm{e}\Big\{\langle N^{-1/2}t, Z_{N-k}'\rangle\Big\}, \quad f_0(t) = \mathbf{E}\,\mathrm{e}\Big\{\langle N^{-1/2}t, U_{N-k}'\rangle\Big\} \ ,$$

$$f_1(t) = \frac{N-k}{6N^{3/2}}\,\mathbf{E}\langle it, X'\rangle^3 f_0(t), \quad \widehat{v}(t) = (f(t) - f_0(t) - f_1(t))g(\varepsilon t), \ \varepsilon^2 = k/N \ .$$

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 Bentkus and Götze (1996), we obtain

$$\delta_N \ll_d \max_{|\alpha|\le 2d} \int_{t\in\mathbb{R}^d} |\partial^\alpha \widehat{v}(t)|\, dt \ . \tag{9.16}$$

In order to derive (9.14) from (9.16), it suffices to prove that, for $|\alpha| \le 2d$,

$$|\partial^\alpha \widehat{v}(t)| \ll_d (1+|t|^3) g(\varepsilon t/\sqrt{2d+1}), \quad \varepsilon^2 = k/N \ , \tag{9.17}$$

$$|\partial^\alpha \widehat{v}(t)| \ll_d \beta N^{-1}(1+|t|^6)\exp\{-c_1(d)|t|^2\}, \quad \text{for } |t|^2 \le c_2(d)N/\beta \ . \tag{9.18}$$

Indeed, using (9.17) and denoting $M = 2d+10$, $T = c_d\sqrt{N}/\beta^{1/2}$, $c_d > 0$, we obtain

$$\int\limits_{|t|\ge T} |\partial^\alpha \widehat{v}(t)|\, dt \ll_d \int\limits_{|t|\ge T} |t|^3 g(\varepsilon t/\sqrt{2d+1})\, dt \ll_d \varepsilon^{-M} \int\limits_{|t|\ge T} |t|^{d+2-M} dt \ , \tag{9.19}$$

and it is easy to see that the last integral in (9.19) is bounded from above by the second summand in the right hand side of (9.14). In the proof of (9.19) we used $\sqrt{N}/\beta^{1/2} \ge c_d > 0$ and $g(t) \ll \exp\{-c|t|^2\} \ll_d |t|^{-M}$. Similarly, using (9.18), we can integrate $|\partial^\alpha \widehat{v}(t)|$ over $|t| \le T$, and the integral is bounded from above by $c_d\beta/N$.

To prove (9.17) we can write $g(\varepsilon t) = g^{2d+1}(\varepsilon t/\sqrt{2d+1})$ and differentiate the product. Using (9.15) we obtain (9.17).

One can prove (9.18) using a Bergström type identity similar to (9.4), the estimates (9.15) and a version of the standard techniques provided in Bhattacharya and Rao (1986). □

References

de Acosta, A.: Inequalities for B-valued random vectors with applications to the strong law of large numbers. Ann. Prob. **9**, 157–161 (1991)

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

Bentkus, V., Götze, F., Zitikis, R.: Asymptotic expansions in the integral and local limit theorems in Banach spaces with applications to ω-statistics. J. Theor. Probab. **6**, no. 4, 727–780 (1993)

Bentkus, V., Götze, F.: On the lattice point problem for ellipsoids. Russian Acad. Sc. Doklady **343**, no. 4, 439–440 (1995a)

Bentkus, V., Götze, F.: Optimal rates of convergence in functional limit theorems for quadratic forms. Preprint 95-091 SFB 343, Universität Bielefeld (1995b)

Bentkus, V., Götze, F.: Optimal rates of convergence in the CLT for quadratic forms. Ann. Probab. **24**, no. 1, 466–490 (1996)

Bentkus, V., Götze, F.: On the lattice point problem for ellipsoids. Acta Arithmetica **80**, no. 2, 101–125 (1997)

Bentkus, V., Götze, F., Zaitsev, A.Yu.: Approximations of quadratic forms of independent random vectors by accompanying laws. Theory Probab. appl. **42**, no. 2, 308–335 (1997)

Bhattacharya, R.N., Ranga Rao, R.: Normal Approximation and Asymptotic Expansions. Wiley, New York (1986)

Chung, K.L.: A course in probability theory. Academic Press, New York and London (1974)
Esseen, C.-G.: Fourier analysis of distribution functions. Acta Math. **77**, 1–125 (1945)
Esseen, C.-G.: On the concentration function of a sum of independent random variables. Z. Wahrsch. verw. Geb. **9**, 290–308 (1968)
Götze, F.: Asymptotic expansions for bivariate von Mises functionals. Z. Wahrsch. verw. Geb. **50**, 333–355 (1979)
Graham, S.W., Kolesnik, G.: Van der Corput's method of exponential sums, Cambridge University Press, London Mathem. Soc. Lecture Notes series. Cambridge New York Port Chester Melbourne Sydney (1991)
Hardy, G.H.: The average order of the arithmetical functions $P(x)$ and $\Delta(x)$. Proc. London Math. Soc. **15**, no. 2, 192–213 (1916)
Landau, E.: Zur analytischen Zahlentheorie der definiten quadratischen Formen. Sitzber. Preuss. Akad. Wiss. **31**, 458–476 (1915)
Linnik, Ju.V.: The large sieve. Dokl. Akad. Nauk SSSR **30**, 292–294 (1941)
Petrov, V.V.: Sums of independent random variables. Springer, Berlin Heidelberg New York, (1975)
Prawitz, H.: Limits for a distribution, if the characteristic function is given in a finite domain. Skand. AktuarTidskr 138–154 (1972)
Vakhania, N.N.: Probability distributions on linear spaces. North-Holland Publishing Co., New York-Amsterdam (1981)
Zaitsev, A.Yu.: Multidimensional generalized method of triangular functions. J. Soviet Math. **43**, no. 6, 2797–2810; (a translation from Zap. Nauch. Sem. LOMI) (1988a)
Zaitsev, A.Yu.: Estimates for the closeness of successive convolutions of multidimensional symmetric distributions. Probab. Th. Rel. Fields **79**, 175–200 (1988b)