*The Annals of Probability*
1991, Vol. 19, No. 2, 724–739

# ON THE RATE OF CONVERGENCE IN THE MULTIVARIATE CLT

BY F. GÖTZE

*University of Bielefeld*

Berry–Esseen theorems are proved in the multidimensional central limit theorem without using Fourier methods. An effective and simple estimate of the error in the CLT for sums and convex sets using Stein's method and induction is derived. Furthermore, the error in the CLT for multivariate functions of independent random elements is estimated extending results of van Zwet and Friedrich to the multivariate case.

## 1. Introduction and results.

A. *The Berry–Esseen theorem for sums of independent random vectors.* Let $X_1, X_2, \dots, X_n$ denote independent random vectors in $\mathbb{R}^k$. Let $\|\cdot\|$ denote the Euclidean norm in $\mathbb{R}^k$ and define $S_n \triangleq X_1 + \cdots + X_n$. Assume that

$$\beta_3 \triangleq \sum_{j=1}^{n} E\|X_j\|^3 < \infty. \tag{1.1}$$

For reasons of simplicity consider normed random vectors such that

$$EX_j = 0, j = 1, \dots, n \quad \text{and} \quad \operatorname{Cov}(S_n) = k \times k\text{-identity}. \tag{1.2}$$

Let $\Phi$ denote the standard Gaussian distribution in $\mathbb{R}^k$ and let $Q_{(n)}$ denote the distribution of $S_n$.

For indicator functions of measurable convex sets in $\mathbb{R}^k$, say $f$, and i.i.d. vectors $X_j$, the estimate $|\int f\,d(Q_{(n)} - \Phi| \le c(k)\beta_3$ has been proved by Sazonov (1968) using the so-called method of compositions due to Lindeberg (1922) and Bergström (1944). Using truncation techniques and more elaborate Fourier methods, Bhattacharya (1975) obtained this estimate uniformly over more general translation classes of (unbounded) functions $f$. Using a Fourier related approach, Sweeting (1977) provided estimates for individual functions $f$ extending these results. The dependence of $c(k)$ on $k$ for convex sets has been shown to be of order $O(k)$ by Nagaev (1976) which has been improved to $O(k^{1/2})$ by Bentkus (1986) for the i.i.d. case. First, we will provide a short inductive proof of the Berry–Esseen theorem in $\mathbb{R}^k$ with explicit constant $c(k)$ for a class of functions including indicators of convex sets by the method of Stein (1972, 1987) adapted for $\mathbb{R}^k$ in order to demonstrate the method.

Let $\mathscr{A}$ denote a class of measurable functions uniformly bounded by some $K > 0$. For $f \in \mathscr{A}$ and $\delta > 0$, define $f_\delta^+(x) \triangleq \sup\{f(x+y)\colon \|y\| \le \delta\}$ and

Received December 1988; revised December 1989.
*AMS* 1980 *subject classifications*. Primary 60F05; secondary 62H10.
*Key words and phrases*. Central limit theorem, Berry–Esseen theorem, Stein's method, multivariate statistics.

$f_\delta^-(x) \triangleq -(-f)_\delta^+(x)$ and $\omega(f;\delta) \triangleq \int (f_\delta^+ - f_\delta^-)\,d\Phi$ as the average local variation of $f$. Furthermore, let $|f|_0 \triangleq \sup\{|f(x) - f(0)|: x \in \mathbb{R}^k\}$. We shall prove:

THEOREM 1.3. *Let $\mathscr{A}$ be closed under the supremum and affine transformations, that is, $f \in \mathscr{A}$ implies $f_\delta^{+/-}$ and $x \mapsto f(Mx)$ are again in $\mathscr{A}$ for every $\delta > 0$ and affine transformation $M$ of $\mathbb{R}^k$. Assume that* (1.1) *holds and*

$$\sup\{\omega(f;\varepsilon): f \in \mathscr{A}\} \le \Delta\varepsilon \tag{1.4}$$

*for every $\varepsilon > 0$ and some $\Delta > 0$ not depending on $n$. Then*

$$\sup\left\{\left|\int f\,d(Q_{(n)} - \Phi)\right|: f \in \mathscr{A}\right\} \le C_k \beta_3 \sup\{|f|_0: f \in \mathscr{A}\},$$

*where $C_k \triangleq 62.2 a_k \Delta + 10.7$ and $a_k^2$ denotes the* 0.875 *quantile of the $\chi^2$-distribution with k degrees of freedom, that is, $a_k = 2.04, 2.4, 2.69, 2.94$ for $k = 2$, 3, 4, 5 and $a_k \le 1.27k^{1/2}$ for $k \ge 6$.*

In particular we obtain for the class $\mathscr{C}$ of all measurable convex sets in $\mathbb{R}^k$, using the bound $\Delta \le 2k^{1/2}$ which follows from Corollary 32, page 24 of Bhattacharya and Ranga Rao (1986) (using Stirling's formula),

$$\sup_{C \in \mathscr{C}} |Q_{(n)}(C) - \Phi(C)| \le C_k \beta_3 n^{-1/2}, \tag{1.5}$$

where $C_k = 124.4 a_k k^{1/2} + 10.7 < 157.85k + 10$ for $k \ge 6$, which shows that these bounds are still of limited use for small sample sizes.

In the proof of Theorem 1.3 we shall use a multidimensional version of Stein's idea of defining differential equations in $\mathbb{R}^1$. In order to motivate the choice of our equation, let us sketch one way of obtaining an equation which is useful for convergence rate estimates.

Assume that $T_n$, $n \in \mathbb{N}$ denotes a sequence of $\mathbb{R}^k$-valued statistics of $n$ random elements $X_1, \dots, X_n$ with limit distribution $\pi$ on $\mathbb{R}^k$. Furthermore, let $V_{t,x}$, $t \ge 0$, denote a continuous time Markov process starting in $x \in \mathbb{R}^k$ with stationary distribution $\pi$ and infinitesimal generator $A$. For a bounded measurable function $h$, we have the following formal identity using the diffusion equation for $V_t$:

$$\begin{aligned} h(x) - \pi(h) &= Eh(V_{0,x}) - Eh(V_{\infty,x}) \\ &= -\int_0^\infty \frac{\partial}{\partial t} Eh(V_{t,x})\,dt \\ &= (A\psi)(x), \end{aligned} \tag{1.6}$$

$$\text{where } \psi(x) \triangleq -\int_0^\infty Eh(V_{t,x})\,dt.$$

For some generators $A$, the r.h.s. of $Eh(T_n) - \pi(h) = EA\psi(T_n)$ can in fact now be estimated by means of Taylor expansions and partial integration. This describes our view of Stein's method.

In our case we may choose $\pi = \Phi$ and $V_{t,x}$ as the Ornstein–Uhlenbeck process

$$dV_{t,x} = -\tfrac{1}{2}V_{t,x}\,dt + dw_t, \qquad V_{0,x} = x,$$

where $w_t$ denotes the standard Brownian motion in $\mathbb{R}^k$ and obtain as infinitesimal generator $A\psi(x) = \Delta\psi(x) - \psi^{(1)}(x)x$. Here $\Delta$ denotes the Laplace operator and $\psi^{(j)}v_1 \cdots v_j$ denotes the $j$th derivative of $\psi$ in the directions $\vartheta_1, \dots, \vartheta_j \in \mathbb{R}^k$.

Replacing $w_t$ by a Poisson process and $\pi$ by a Poisson distribution, one obtains the difference equations of Barbour (1988) who treated the multivariate Poisson case. The defining equation

$$\Delta\psi(x) - \psi^{(1)}(x)x = h(x) - \Phi(h) \tag{1.7}$$

has to be interpreted carefully.

It is valid only for points $x$, where $h(x)$ is locally Hölder continuous. Applying this equation we need to evaluate expressions of type $E(\Delta\psi(S_n + x) - \Delta\psi(S_n + x'))$ which is made easier by introducing a smoothed version of $h - \Phi(h)$, say $-\chi_t(\cdot|h)$, $t > 0$ small, which is *adapted* to the equation (1.7). Define for $1 \geq t \geq 0$,

$$\chi_t(x|h) \triangleq \int \Big(h(y) - h\big(t^{1/2}y + (1-t)^{1/2}x\big)\Big)\,\Phi(dy). \tag{1.8}$$

Notice that in terms of the Ornstein–Uhlenbeck process $V_{t,x}$, we may write $-\chi_t(x|h) = E(h(V_{s,x}) - h(V_{\infty,x}))$, where $1 - e^{-s} \triangleq t$, since $\mathscr{D}(V_{s,x}) = \mathscr{D}(xe^{-s/2} + (1-e^{-s})^{1/2}w_1)$. Then $-\chi_0(\cdot|h) = h - \Phi(h)$ and $\chi_1(\cdot|h) \equiv 0$.

Assuming that

$$\chi_t(x|h) = O_x(|\log(|1-t)|^{-\gamma}) \quad \text{holds for } t \uparrow 1 \text{ and some } \gamma > 1, \tag{1.9}$$

the functions

$$\psi_t(x) \triangleq \frac{1}{2}\int_t^1 \chi_s(x|h)\frac{ds}{1-s} \tag{1.10}$$

are well-defined solutions of

$$\Delta\psi_t(x) - \psi_t^{(1)}(x)x = -\chi_t(x|h) \tag{1.11}$$

for every $x \in \mathbb{R}^k$ and $0 < t \leq 1$ [cf. Lemma 2.1(i)].

Of course we have to estimate the error in replacing $h - \Phi(h)$ by $-\chi_t(\cdot|h)$ by an appropriate smoothing inequality (see Lemma 2.11).

The equation (1.11) leads us to estimate

$$|Eh(S_n) - \Phi(h)| = |E\Delta\psi_t(S_n) - E\psi_t^{(1)}(S_n)S_n|$$

applying Stein's method similarly as in Bolthausen (1984) (in his proof of the Berry–Esseen theorem for general linear permutation statistics) combining the equation (1.11) with an induction on $n$.

Notice that in one dimension the connection with Stein's equation $f^{(1)}(x) - f(x)x = h(x) - \Phi(h)$ is given by $f(x) = \psi^{(1)}(x)$.

B. *A Berry–Esseen result for functions of independent random elements.* Let $X_1, \dots, X_n$ denote independent random elements taking values in a measurable space $(\mathscr{X}, \mathscr{B})$. Let $t_n\colon \mathscr{X}^n \to \mathbb{R}^k$ denote a measurable function. We shall consider statistics of the type

$$T = t_n(X_1, \dots, X_n).$$

For convenience we may assume

$$ET = 0. \tag{1.12}$$

Define $T_j \triangleq E(T|X_j)$ and let $\lambda_j \triangleq$ [smallest eigenvalue of $\operatorname{Cov}(T_1 + \cdots + T_j)$]. Assume that (perhaps after rearrangement of $X_1, \dots, X_n$)

$$\operatorname{Cov}(T_1 + \cdots + T_n) = k \times k\text{-identity}, \qquad \lambda_j \geq \frac{j}{2n}, \qquad 1 \leq j \leq n. \tag{1.13}$$

Let $\Delta \triangleq T - \sum_{j=1}^n T_j$ denote the remainder in Hoeffding's projection. Define $\check{X}_j \triangleq (X_l\colon l \neq j)$ and assume that the following moments of first and second order differences measuring the stochastic influence of single observations and pairs of observations on $\Delta$ are well defined.

Let $\Delta_j \triangleq \Delta - E(\Delta|\check{X}_j)$, $\Delta_{jl} \triangleq \Delta_j - E(\Delta_j|\check{X}_l)$, $1 \leq j < l \leq n$ and define

$$\begin{gathered} \beta_3 \triangleq \sum_{j=1}^n E\|T_j\|^3, \qquad \beta_{3,0} \triangleq \max_j E\|T_j\|^3, \qquad \beta_{p,1} \triangleq \max_{1 \leq j \leq n} E\|\Delta_j\|^p, \\ \beta_{p,2} \triangleq \max_{1 \leq j < l \leq n} E\|\Delta_{jl}\|^p \quad \text{for } 2 \geq p \geq 1. \end{gathered} \tag{1.14}$$

THEOREM 1.15. *Assume that conditions* (1.12)–(1.14) *hold. Let $f$ denote a measurable function such that $|f|_0 = \sup_x |f(x) - f(0)| < \infty$ and let $S$ denote a random vector with standard normal distribution $\Phi$. When $\frac{3}{2} \leq p < 2$, there exists a constant $c$ depending on $k$ and $p$ only such that*

$$\begin{aligned} |Ef(T) - Ef(S)| \leq c|f|_0 &\left( \beta_3 + \frac{1}{2-p} n^{3/2} \beta_{3,0}^{1/3} \beta_{p,1} + n^2 \beta_{3,0}^{2/3} \beta_{3/2,2}^{2/3} \right) \\ &+ c\omega(f; c_1 \beta_3), \end{aligned} \tag{1.16}$$

*provided that $\beta_3 < c_2 (\log n)^{-3}/k$, where $c_1$ and $c_2$ denote absolute constants.*

The estimate extends results of van Zwet (1984) and Friedrich (1989) to the case of multivariate statistics $T$. Using a similar martingale decomposition for $\Delta$ as in the latter paper while avoiding Fourier inversion techniques, we arrive at a comparable error bound.

The proof is based on Theorem A.3 of the Appendix which might be of independent interest. This result can be also proved via Stein's method and the equation (1.6) and induction but for vectors $T_j$ with nearly the same covariance and third moment behavior. Therefore we decided to give the shorter proof based on standard techniques.

A Berry–Esseen type estimate can be proved by the Stein method of Theorem 1.3 for multivariate functions of dependent random vectors like general rank and sampling statistics. The estimates depend on slightly higher moments of $\Delta_j$ as in (1.11) but not on moments of $\Delta_{ij}$. Details will appear somewhere else.

REMARKS 1.17. (a) The estimate (1.16) remains true for $p = 2$ with $1/(2-p)$ replaced by $\log n$.

(b) We have $\beta_{p,1} \le n\beta_{p,2}$ [see (2.26)] and $\beta_{p,2}^{1/p} \ge \beta_{q,2}^{1/q}$, $p \ge q \ge 1$.

(c) Therefore, to achieve an error bound of order $O(n^{-1/2} + \omega(f; n^{-1/2}))$ in (1.16), we may assume $\beta_{3,0} = O(n^{-3/2})$ which implies $\beta_3 = O(n^{-1/2})$, together with

$$\beta_{3/2,2}^{2/3} = O(n^{-3/2}) \tag{1.18}$$

and

$$\beta_{p,1} = O(n^{-3/2}) \quad \text{or} \quad \beta_{p,2} = O(n^{-5/2}) \quad \text{for } \tfrac{3}{2} \le p < 2. \tag{1.19}$$

The assumptions (1.18) and (1.19) follow from

$$\beta_{5/3,2}^{3/5} = O(n^{-3/2}). \tag{1.20}$$

(d) In the case of a $U$-statistic $U \triangleq \sum_{i \le i_1 < \cdots < i_r \le n} h_{i_1,\dots,i_r}$, where $h_{i_1,\dots,i_r} \triangleq h_{i_1,\dots,i_r}(X_{i_1},\dots,X_{i_r})$ and $h_{i_1,\dots,i_r}(\cdot)$ denotes a measurable function of $\mathscr{X}^r$ to $\mathbb{R}$, it follows by Minkowsky's inequality that (1.20) holds provided that

$$\max_{1 \le j_1 < \cdots > j_r \le n} \left(E\|h_{j_1,\dots,j_r}\|^{5/3}\right)^{3/5} \le Dn^{-r+1/2}, \qquad D > 0. \tag{1.21}$$

For further applications to $L$- and $R$-statistics, see Friederich (1988) and van Zwet (1984).

These moment conditions sharpen results on the error in the CLT for multivariate $U$-statistics of degree two by Götze (1987) which were obtaincd by using characteristic function methods.

**2. Lemmas and proofs of the results.** For a measure $Q$ on $\mathbb{R}^k$ and $\varepsilon > 0$, define $Q_\varepsilon(A) \triangleq Q(\varepsilon^{-1}A)$. Let $h$ denote a bounded measurable function.

LEMMA 2.1. (i) *For every $x \in \mathbb{R}^k$, $0 < t \le 1$, we have*

$$\Delta\psi_t(x) - \psi_t^{(1)}(x)x = -\chi_t(x|h). \tag{2.2}$$

(ii) *There exist absolute constants $c_j > 0$ such that for $t > 0$,*

$$\sup_x |\psi_t^{(j)}(x)| \le c_j |h|_0 A_j(t), \tag{2.3}$$

*where $A_j(t) \triangleq t^{-(j-2)/2}$, $j > 2$, $A_2(t) \triangleq \log(t^{-1})$ and $A_1(t) \triangleq A_0(t) \triangleq 1$. For $k = 1$, we have $A_2(t) \triangleq 1$.*

(iii) Let Q denote a finite (signed) Borel measure on $\mathbb{R}^k$. Then

$$\left|\int \psi_t^{(j)}(x)v^j Q(dx)\right| \le c_j A_j(t)\sup\left\{\left|\int h(xs+a)Q(dx)\right| : 0 \le s \le 1,\ a \in \mathbb{R}^k\right\}\|v\|^j \tag{2.4}$$

for every $1 > t > 0$ and $j \ge 1$. Note for later use that we may choose $c_3 \triangleq (1-t)^{1/2}\alpha_3$, with $\alpha_3 \triangleq \int|\varphi^{(3)}(y)v^3|\, d^k y \le 1.602\|v\|^3$.

Proof. (i) In order to validate the formal calculations of (1.6), notice that by change of variables $y \to t^{-1/2}(y-(1-t)^{1/2}x)$,

$$\chi_t(x|h) = \int h\, d\Phi - \int h(y)\varphi_t\big(y-(1-t)^{1/2}x\big)\, d^k y, \tag{2.5}$$

where $\varphi_t(z) \triangleq t^{-k/2}\varphi(t^{-1/2}z)$ and $\varphi(z)$ denotes the Lebesgue density of the $k$-variate standard normal distribution $\Phi$.

Now it is straightforward to verify that the equation

$$2(1-t)\frac{d}{dt}u_t(x) = \Delta u_t(x) - u_t^{(1)}(x)x, \qquad t > 0, \tag{2.6}$$

is satisfied for the functions $u_t(x) \triangleq \varphi_t(y-(1-t)^{1/2}x)$ with $y$ arbitrary, $u_t(x) \triangleq \Phi(h)$ and thus for $u_t(x) \triangleq \chi_t(x|h)$ by interchanging differentiations and integration in $y \in \mathbb{R}^k$. This is obviously allowed since $h$ is uniformly bounded. Similar interchanges of differentiation and integration prove for $j \ge 1$,

$$\begin{aligned} |\chi_s^{(j)}(x|h)v^j| &\le \sup\left\{|h\big(y+x(1-s)^{1/2}\big)-h(0)| : y \in \mathbb{R}^k\right\} \\ &\quad \times(1-s)^{j/2}\int|\varphi_s^{(j)}(y)v^j|\, d^k y \\ &\le |h|_0(1-s)^{j/2}s^{-j/2}\alpha_j\|v\|^j, \end{aligned} \tag{2.7}$$

where $\alpha_j$ denotes an absolute constant *independent of* $k$.

Now (2.6) implies, after division by $2(1-t)$ with $u_t(x) \triangleq -\chi_t(x|h)$ and $t > 0$,

$$\begin{aligned} \chi_1(x|h) - \chi_t(x|h) &= \int_t^1 ds\frac{d}{ds}\chi_s(x|h) \\ &= \frac{1}{2}\int_t^1 \Delta\chi_s(x|h)\frac{ds}{1-s} - \frac{1}{2}\int_t^1 \chi_s^{(1)}(x|h)x\frac{ds}{1-s}. \end{aligned} \tag{2.8}$$

Thus $\chi_1(x|h) \equiv 0$ and interchanging integration and differentiation in (2.8) which is justifiable because of (2.7) yields the result (2.2).

(ii) The estimates follow immediately from (2.7) by interchanging differentiation and integration in $s$. The estimates for $k = 1$ and $A_2(t)$ follow from the properties of the function $\psi_t^{(1)}$ being a solution of Stein's equation [see Stein (1972)].

(iii) We have by Fubini's theorem and interchanging differentiation and integration,

$$\begin{aligned}\left|\int \psi_t^{(j)}(x) v^j Q(dx)\right| &= \left|\int_t^1 ds (1-s)^{-1+j/2}\right. \\ &\qquad \left.\times \int h(y) \varphi_s^{(j)}\big(y-(1-s)^{1/2}x\big) v^j\, d^k y\, Q(dx)\right| \\ &= \left|\int_t^1 ds(1-s)^{-1+j/2}\right. \\ &\qquad \left.\times \int \left(\int h\big(y-(1-s)^{1/2}x\big)\, Q(dx)\right) \varphi_s^{(j)}(y) v^j\, d^k y\right|,\end{aligned}$$

which yields the desired estimate (2.4) using the second step of (2.7). The value of $c_3$ follows by inspection of these arguments using $\frac{1}{2}\int_t^1 ds(1-s)^{1/2}s^{-3/2} \le (1-t)^{1/2}$ and the fact that $\langle u, v\rangle$ has standard normal distribution under $\Phi(du)$ when $\|v\| = 1$. □

LEMMA 2.9 (Stein's method). *Let $\overline{X}_j$ denote an independent copy of $X_j$, $1 \le j \le n$. Furthermore, define $S_{n,j} \triangleq N_j(S_n - X_j)$, where $N_j$ denotes a positive definite matrix such that* Cov$(S_{n,j}) = k \times k$*-identity. Let $\tau$ be uniformly distributed in* $[0,1]$ *and independent of $X_1, \dots, X_n$. Finally, let*

$$H_{j,t}(x|u,v) \triangleq E\psi_t^{(3)}\big(N_j^{-1}x + \tau v\big)\Big(\tfrac{1}{3}(u+v)^3 - \big(\tfrac{4}{3} - \tau\big)v^3 - \tfrac{1}{3}u^3\Big).$$

*Then*

$$|E\chi_t(S_n|h)| = \left|\sum_{j=1}^n H_{j,t}\big(S_{n,j}|\overline{X}_j, X_j\big)\right|. \tag{2.10}$$

PROOF. Using (2.2), we have

$$\begin{aligned}-E\chi_t(S_n|h) &= E\big(\Delta\psi_t(S_n) - \psi_t^{(1)}(S_n)S_n\big) \\ &= \sum_{j=1}^n E\big(\psi_t^{(2)}(S_n)\overline{X}_j^2 - \psi_t^{(1)}(S_n)X_j\big) \\ &= \sum_{j=1}^n \Big[E\psi_t^{(2)}(S_n - X_j)\overline{X}_j^2 + E\psi_t^{(3)}\big(S_n - (1-\tau)X_j\big)\overline{X}_j^2 X_j \\ &\qquad - E\psi_t^{(1)}(S_n - X_j)X_j - E\psi_t^{(2)}(S_n - X_j)X_j^2 \\ &\qquad - E\psi_t^{(3)}\big(S_n - (1-\tau)X_j\big)X_j^3(1-\tau)\Big].\end{aligned}$$

Since $X_l$ and $\overline{X}_j$ are independent for every $l, j$ and $\overline{X}_j$ and $X_j$ have the same distribution, the first and fourth term in the square bracket cancel and the third term vanishes by independence of $S_n - X_j$ and $X_j$.

Since $\psi_t^{(3)}(\cdot)(u^2v - v^3(1-\tau)) = \psi_t^{(3)}(\cdot)(\frac{1}{3}(u+v)^3 - (\frac{1}{3}+1-\tau)v^3 - \frac{1}{3}u^3)$, the second and fifth term together yield the function $H_{j,t}$. This proves (2.10). □

LEMMA 2.11. *Assume that* (1.3) *holds. For* $\varepsilon > 0$, *let* $b \triangleq (1-\varepsilon^2)$ *and* $\delta \triangleq \varepsilon/b$. *Then*

$$\sup\left\{\left|\int f\,d(Q_{(n)} - \Phi)\right| : f \in A\right\}$$

$$(2.12) \qquad \leq \tfrac{4}{3}\sup\left\{\left|\int \chi_{\varepsilon^2}(\cdot\,|h)\,dQ_{(n)}\right| : h \in \mathscr{A}\right\} + \tfrac{5}{2}a_k\,\delta\Delta,$$

*where* $a_k^2$ *is the* $\frac{7}{8}$-*quantile of the* $\chi^2$-*distribution with* $k$ *degrees of freedom.*

PROOF. From standard smoothing inequalities [see Bhattacharya and Ranga Rao (1986), Lemma 11.14, page 95], we obtain the following upper bound for the l.h.s. of (2.12) using $\mathscr{A} = \{f(b\cdot^{-1}): f \in \mathscr{A}\}$:

$$\left[\sup\left\{\left|\int h\,d(Q_{(n)} - \Phi)_b * \Phi_\varepsilon\right| : h = f_{a\delta}^{\pm}(b\cdot^{-1}),\ f \in \mathscr{A}\right\}\right.$$

$$\left. + (\alpha+1)\alpha\,\delta\Delta\right]\Big/(2\alpha - 1),$$

where $\alpha = \Phi(\|x\| \leq a_k) = \frac{7}{8}$ and $\Phi_\varepsilon(A)$ is defined at the start of Section 2. This together with $\int h\,d(Q_{(n)} - \Phi)_b * \Phi_\varepsilon = -\int \chi_{\varepsilon^2}(\cdot\,|h)\,dQ_{(n)}$ yields (2.12). □

PROOF OF THEOREM 1.3. We shall show by induction on $n$ that

$$(2.13) \qquad \sup\left\{\left|\int f\,d(Q_{(n)} - \Phi)\right| : f \in \mathscr{A}\right\} \leq C_k\beta_3,$$

where $C_k = 62.2a_k\Delta + 10.7$.

Notice that w.l.o.g. $\sup\{|f|_0 : f \in \mathscr{A}\} = 1$. Since $\beta_3 \geq n^{-1/2}\beta_2^{3/2} = n^{-1/2}k^{3/2}$ by Hölder's inequality, the estimate (2.10) obviously holds for $\beta_3 > C_k^{-1}$ or $n \leq n_0 \triangleq C_k^2k^3$.

Assume that (2.13) has been proved for $n - 1 \geq n_0$. From Lemma 2.11, we conclude that it is sufficient to use (2.2) with $t = \varepsilon^2 > 0$ to be determined later. Notice that (1.9) holds since for $h \in \mathscr{A}$,

$$|\chi_t(x|h)| \leq \left|\int \Big(h(y) - h\big(yt^{1/2} + (1-t)^{1/2}x\big)\Big) I\big(\|y\| > (1-t)^{-1/2}\big)\,\Phi(dy)\right|$$

$$+ \Phi\big(\|y\| > (1-t)^{1/2}\big)$$

$$\leq c\Delta\big((1-t)^{1/2}\|x\| + (1-t)^{1/2}\big) + O\big((1-t)^{k/2}\big), \qquad k > 2,\ t\uparrow 1.$$

Let $Q_{n,j}$ denote the distribution of $S_{n,j}$. By Lemma 2.9, we may write

$$
\begin{aligned}
|E\chi_t(S_n|h)| &\le \sum_{j=1}^{n} E\left|\int H_{j,\varepsilon^2}\left(\cdot|\overline{X}_j, X_j\right) dQ_{n,j}\right| \\
&\le \sum_{j=1}^{n} E\left|\int H_{j,\varepsilon^2}\left(\cdot|\overline{X}_j, X_j\right) d(Q_{n,j} - \Phi)\right| \qquad (2.14)\\
&\quad + \sum_{j=1}^{n} E\left|\int H_{j,\varepsilon^2}\left(\cdot|\overline{X}_j, X_j\right) d\Phi\right| \\
&\triangleq I_1 + I_2, \quad \text{say.}
\end{aligned}
$$

Put $Q \triangleq Q_{n,j} - \Phi$ in Lemma 2.1(iii). Then (2.4) yields by the definition of $H_{j,\varepsilon^2}$,

$$
\begin{aligned}
&\left|\int H_{j,\varepsilon^2}\left(x|\overline{X}_j, X_j\right)(Q_{n,j} - \Phi)(dx)\right| \\
&\quad \le \alpha_3(1-\varepsilon^2)^{1/2}\varepsilon^{-1}\sup\left\{\left|\int f\, d(Q_{n,j} - \Phi)\right| : f \in \mathscr{A}\right\} \qquad (2.15)\\
&\quad\quad \times\left(\tfrac{1}{3}\|\overline{X}_j + X_j\|^3 + \left(\tfrac{4}{3} - \tau\right)\|X_j\|^3 + \tfrac{1}{3}\|\overline{X}_j\|^3\right).
\end{aligned}
$$

Hence we obtain by the induction hypothesis (2.13) taking expectations over $\overline{X}_j$, $X_j$ and $\tau$ in (2.14) and using $\|\overline{X}_j + X_j\|^3 \le 4(\|\overline{X}_j\|^3 + \|X_j\|^3)$ the inequality

$$
\begin{aligned}
I_1 &\le \alpha_3(1-\varepsilon^2)\varepsilon^{-1}\sum_{j=1}^{n}\left(\tfrac{4}{3} + \tfrac{4}{3} + \tfrac{4}{3} - E\tau + \tfrac{1}{3}\right)E\|X_j\|^3 C_k \qquad (2.16)\\
&\quad \times \sum_{\substack{1\le l\le n\\ l\ne j}} E\|N_j X_l\|^3.
\end{aligned}
$$

Notice that in order to apply the induction hypothesis we have to use the fact that $\mathscr{A}$ is invariant under $x \to N_j^{-1}x$.

Since $N_j^{-2} = \operatorname{Cov}(S_n - X_j)$ and $\langle x, (Id - N_j^{-2})x\rangle = E\langle X_j, x\rangle^2 \le \beta_3^{2/3}\|x\|^2$, the inequality $\beta_3 < C_k^{-1}$ implies $\|N_j x\| \le f\|x\|$, where $f \le (1 - C_k^{-2/3})^{1/2}$. Thus

$$
I_1 \le \beta_3\left(\tfrac{23}{6}\alpha_3 f^3(1-\varepsilon^2)^{1/2}\varepsilon^{-1}\beta_3 C_k\right). \tag{2.17}
$$

By (1.8), we obtain with $M_{j,t} \triangleq Id + (1-t)(N_j^{-2} - Id)$,

$$
\left|\int \chi_s\left(x + N_j^{-1}z\right)v^3\, d\Phi(z)\right| = \int h\, d\Phi - \int h(y)\varphi_{M_{j,s}}\left(y - (1-s)^{1/2}x\right) d^k y,
$$

where $\varphi_\Sigma$ denotes the $k$-variate normal density with covariance matrix $\Sigma$ and mean zero. Thus with $\alpha_3$ defined in Lemma 2.1(iii),

$$\left|\int \psi_t^{(3)}\left(x + N_j^{-1}z\right)v^3\,\Phi(dz)\right| \le \tfrac{1}{2}\int_t^1 ds(1-s)^{-1+3/2}$$
$$\times\left|\int h(y)\varphi_{M_{j,s}}^{(3)}\left(y - (1-s)^{1/2}x\right)v^3\,d^ky\right|$$
$$\le \tfrac{1}{2}\tfrac{2}{3}|h|_0 \sup_s \int\left|\varphi_{M_{j,s}}^{(3)}(y)v^3\right|d^ky$$
$$\le \tfrac{1}{3}\sup_s\left\langle vM_{j,s}^{-1}v\right\rangle^{3/2}\alpha_3 \le \tfrac{1}{3}f^3\|v\|^3\alpha_3.$$

This inequality together with the definition of $H_{j,\varepsilon^2}$ implies similarly as in (2.15)–(2.16),

$$(2.18)\qquad I_2 \le \frac{\alpha_3}{3}f^3\sum_{j=1}^n\left(\tfrac{4}{3} + \tfrac{4}{3} + \tfrac{4}{3} - E\tau + \tfrac{1}{3}\right)E\|X_j\|^3 \le \alpha_3\beta_3\tfrac{23}{18}f^3.$$

Hence (2.17)–(2.18) together with (2.14) and Lemma 2.11 yields

$$D \triangleq \sup\left\{\left|\int f\,d(Q_{(n)} - \Phi)\right| : f \in \mathscr{A}\right\}$$
$$\le \beta_3\tfrac{23}{6}\alpha_3 f^3\left((1-\varepsilon^2)^{1/2}\varepsilon^{-1}\beta_3 C_k + \tfrac{1}{3}\right) + \tfrac{5}{2}a_k\varepsilon(1-\varepsilon^2)^{-1/2}\Delta.$$

Choosing $\varepsilon$ optimally, that is, such that the first and third terms are approximately equal, we get $\varepsilon(1-\varepsilon^2)^{-1/2} \triangleq (\tfrac{23}{15}C_k\alpha_3 f^3/a_k\Delta)^{1/2}\beta_3$ and therefore

$$D \le C_k\beta_3\left[\tfrac{23}{18}\alpha_3 f^3 C_k^{-1} + C_k^{-1/2}2\left(\tfrac{115}{12}\alpha_3 a_k\Delta f^3\right)^{1/2}\right].$$

Choosing $C_k$ and hence $n_0$ sufficiently large, we obtain that the factor in square brackets is smaller than 1. More precisely: Since $C_k\beta_3 \le 1$, we have $f \le (1 - C_k^{-2/3})^{-1/2}$. Some tedious but straightforward numerical estimates now lead to the choice (2.10) for $C_k$ thus completing the induction. □

PROOF OF THEOREM 1.10. Let $S$ and $Y$ be random vectors independent of $X_1, \dots, X_n$ with standard normal distribution and distribution $K$ [defined in (A1)–(A2) of the Appendix] on $\mathbb{R}^k$. By the smoothing lemma of Sweeting (1977), Lemma 5, we have

$$D \triangleq |Ef(T) - Ef(S)|$$
$$\le \frac{1}{2\alpha - 1}\left[\max_{+,-}|Ef_{b\varepsilon}^{\pm}(T + \varepsilon Y) - Ef_{b\varepsilon}^{\pm}(S + \varepsilon Y)|\right.$$
$$+2|f|_0 P(\|Y\| > bl) + e^1\omega(f; 2b\varepsilon)$$
$$\left.+2|f|_0\varepsilon''^{-(k-2)}\exp\left[(\varepsilon''^{-2}/2)\right] + 2|f|_0\left(\frac{1-\alpha}{\alpha}\right)^l\right],$$

where $\alpha \triangleq P(\|Y\| > b) > \frac{1}{2}$, $0 < \varepsilon < \varepsilon' < 1/b$, $l \in \mathbb{N}$ s.t. $\varepsilon'' \triangleq lb\varepsilon' < 1$. Choosing $\varepsilon \triangleq c\beta_3$, $\varepsilon' \triangleq \varepsilon^{1/2}$ and $l = [\varepsilon^{-1/2}/b]$, we obtain for $\beta_3$ sufficiently small,

$$D \le c \max_{+,-} |Ef_{b\varepsilon}^{\pm}(T + \varepsilon Y) - Ef_{b,\varepsilon}^{\pm}(S + \varepsilon Y)| + c|f|_0\beta_3 + c\omega(f; c\beta_3). \tag{2.19}$$

Write $g(x)$ for $Ef_{b\varepsilon}^{\pm}(x + \varepsilon Y)$. Then we have

$$\begin{aligned} |E(g(T) - g(s))| &\le |E(g(S_n) - g(S))| \\ &\quad + \sum_{j=1}^{n} |E\big(g(S_{j-1}) - g(S_j)\big)| \triangleq I + J_1 + \cdots + J_n, \end{aligned} \tag{2.20}$$

say, where

$$S_j \triangleq T_1 + \cdots + T_j + E\big(T|X_{j+1}, \dots, X_n\big); \qquad S_0 = T, \qquad S_n = T_1 + \cdots + T_n.$$

By interchanging integration and differentiation, Theorem A.3 immediately leads (with $h$ replaced by $g$) to

$$|E\psi_0^{(j)}(S_n - X_j + x)v^j| \le c|h|_0\|v\|^j,$$

which together with Lemma 2.9 implies

$$I \le c\beta_3. \tag{2.21}$$

In order to estimate $J_l$ write

$$\begin{aligned} J_l &\le |E(g(S_{l-1}) - g(S_l) - g'(S_l)\delta_l)| + |Eg'(S_l)\delta_l| \\ &\triangleq J_{l,1} + J_{l,2}, \end{aligned} \tag{2.22}$$

say, where $\delta_l \triangleq E(\Delta|X_l, \dots, X_n) - E(\Delta|X_{l+1}, \dots, X_n)$. From Theorem A.3, we obtain for $1 < p \le 2$, $0 \le t \le 1$, $\|a\| \le \lambda_l^{1/2}$,

$$E|E\big(g^{(2)}(S_l + ta)|X_{l+1}, \dots, X_n\big)a^2| \le c\lambda_l^{-p/2}\|a\|^p \tag{2.23}$$

and therefore

$$\begin{aligned} &E\big|E\big(g'(S_l + ta) - g'(S_l)|X_{l+1}, \dots, X_n\big)a\big| \\ &\quad \le c\lambda_l^{-p/2}\|a\|^p + c\lambda_l^{-1/2}\|a\|I\big(\|a\| > \lambda_l^{1/2}\big) \\ &\quad \le c\lambda_l^{-p/2}\|a\|^p. \end{aligned} \tag{2.24}$$

By assumption (1.13), $\lambda_j \ge j/2n$ and therefore (2.24) implies for $p < 2$,

$$\begin{aligned} \sum_{l=1}^{n} J_{l,1} &\le cn^{p/2} \sum_{l=1}^{n} l^{-p/2}E\|\delta_l\|^p \\ &\le \frac{c}{2-p} n^1 \max_l E\|\Delta_l\|^p = \frac{c}{2-p} n\beta_{p,1}, \end{aligned} \tag{2.25}$$

since by Hölder's inequality $E\|\delta_l\|^p \le E\|\Delta_l\|^p$.

For $p = 2$, the factor $1/(2 - p)$ may be replaced by $\log n$.

Furthermore, define $\delta_{l|m} \triangleq E(\delta_l|X_l, X_{m+1}, \dots, X_n)$ and $\delta_{l,m} \triangleq \delta_{l|m-1} - \delta_{l|m}$. Let $\delta_m = 0$, $\delta_{l|m} = 0$ and $\delta_{l,m} = 0$ for $l = 1, \dots, n-1$, $m > n$.

By Lemma 1 of Chatterji (1969), the martingale structure of $S_m$ and $\delta_{j|l}$ yields for $1 \le p \le 2$,

$$(2.26)\qquad \begin{aligned} &E|S_m - S_l|^p \le 2^{2-p} \sum_{i=m+1}^{l} E|\delta_i|^p, \qquad 0 \le m < l \le n,\\ &E|\delta_{j|l} - \delta_{j|m}|^p \le 2^{2-p} \sum_{i=m+1}^{l} E|\delta_{j,i}|^p, \qquad 1 \le j < l < m \le n. \end{aligned}$$

Rewriting $J_{l,2}$ we have

$$(2.27)\qquad \begin{aligned} J_{l,2} &\le \sum_{j=1}^{\lfloor n/l \rfloor} |Eg'(S_{jl})\delta_{l|jl} - Eg'(S_{(j+1)l})\delta_{l|(j+1)l}|\\ &\le \sum_{j=1}^{\lfloor n/l \rfloor} (J_{l,j,3} + J_{l,j,4}), \end{aligned}$$

say, where

$$J_{l,j,3} \triangleq |E(g'(S_{jl}) - g'(S_{(j+1)l}))\delta_{l|jl}|,$$

$$J_{l,j,4} \triangleq \left| Eg'(S_{(j+1)l}) \sum_{m=1}^{l} \delta_{l,jl+m} \right|.$$

Since by construction,

$$J_{l,j,3} = |E[(g'(S_{jl}) - g'(S_{(j+1)l})) - (g'(S_{jl} - T_l) - g'(S_{(j+1)l} - T_l))]\delta_{l|jl}|,$$

we obtain by similar arguments as in (2.24) using Hölder's inequality and (1.13),

$$\begin{aligned} J_{l,j,3} &\le c\lambda_{jl}^{-(p+1)/2} E\|\delta_{l|jl}\|\,\|T_l\| \left\| \sum_{m=1}^{l} \delta_{jl+m} \right\|^{p-1}\\ &\le c(jl)^{(p+1)/2} n^{(p+1)/2} E^{1/q} \left( \left\| \sum_{m=1}^{l} \delta_{jl+m} \right\|^{(p-1)q} \|T_l\|^q \right) E^{1/p} \|\delta_{l|jl}\|^p, \end{aligned}$$

where $1/p + 1/q = 1$. By independence of $\delta_{jl+m}$ and $T_l$, $p = (p-1)q$, together with $E\|\delta_{l|jl}\|^p \le E\|\delta_l\|^p \le E\|\Delta_l\|^p$ (Hölder's inequality) and $E\|\sum_{m=1}^{l} \delta_{jl+m}\|^p \le l \max_m E\|\delta_{jl+m}\|^p$ [see (2.26)], we conclude

$$\begin{aligned} J_{l,j,3} &\le cn^{(p+1)/2} j^{(p+1)/2} l^{(p+1)/2+1/q} E^{1/q} \|T_l\|^q \max_m (E^{1/q}\|\delta_{jl+m}\|^p)(E^{1/p}\|\Delta_l\|^p)\\ &\le cn^{(p+1)/2} j^{(p+1)/2} l^{(p+1)/2+1/q} \beta_3^{1/3} \beta_{p,1}^{1/p+1/q}, \end{aligned}$$

by Hölder's inequality. Since $\sum_{j:\, jl \le n} j^{(p+1)/2} \le c(n/l)^{1+(p+1)/2}$, we have

$$\begin{aligned}(2.28)\qquad \sum_{l=1}^{n} \sum_{jl \le n} J_{l,j,3} &\le c\frac{n^{(p+1)/2}}{2-p}\left(\sum_{l=1}\left(\frac{n}{l}\right)^{1+(p+1)/2} l^{(p+1)/2+1-1/p}\beta_{3,0}^{1/3}\beta_{p,1}\right)\\ &\le c\frac{n^{2-1/p}}{2-p}\beta_{3,0}^{1/3}\beta_{p,1}\end{aligned}$$

for $1 \le p < 2$. For $p = 2$, the factor $1/(2-p)$ may be replaced by $\log n$. Furthermore, write

$$\begin{aligned}J_{l,j,4} \le \sum_{m=1}^{l} \Big| E\big[\big(g'(S_{(j+1)l}) - g'(S_{(j+1)l} - T_l)\big)\\ -\big(g'(S_{(j+1)l} - T_{jl+m}) - g'(S_{(j+1)l} - T_l - T_{jl+m})\big)\big]\delta_{l,jl+m}\Big|.\end{aligned}$$

By similar arguments as in (2.24), we obtain

$$J_{l,j,4} \le c\sum_{m=1}^{l} (jl+m)^{-3/2} n^{3/2} E\|\delta_{l,jl+m}\|\,\|T_l\|\,\|T_{jl+m}\|.$$

Hence, by

$$E\big(\|\delta_{j,jl+m}\|\,\|T_{jl+m}\|\,\|T_l\|\big) \le E^{2/3}\|\delta_{j,jl+m}\|^{3/2}\|T_{jl+m}\|^{3/2}E^{1/3}\|T_l\|^3$$

and independence of $T_{jl+m}$ and $\delta_{l,jl+m}$, we have

$$(2.29)\qquad \sum_{l=1}^{n}\sum_{lj\le n} J_{l,j,4} \le cn^{3/2}\sum_{l=1}^{n}\sum_{t=l+1}^{n}(t-1)^{-3/2}\beta_{3,0}^{2/3}\beta_{3/2,2}^{2/3} \le cn^2\beta_{3,0}^{2/3}\beta_{3/2,2}^{2/3}.$$

Summarizing, we obtain by (2.19), (2.20), (2.25), (2.28) and (2.29)

$$(2.30)\qquad D \le c\big(\beta_3 + \omega(f; c\beta_3) + n^{3/2}\beta_{3,0}^{1/3}\beta_{p,1}/(2-p) + n^2\beta_{3,0}^{2/3}\beta_{3/2,2}^{2/3}\big),$$

using

$$n\beta_{p,1} \le cn^{3/2}\beta_{p,1}\beta_{3,0}^{1/3}. \qquad \square$$

## APPENDIX

Let $Y$ denote a random vector independent of $X_1, \dots, X_n$ with distribution $K$ in $\mathbb{R}^k$ and characteristic function $\hat{K}(t)$ satisfying

$$(\text{A.1})\qquad \int \|x\|^p K(dx) < \infty \quad \text{for every } p \in \mathbb{N},$$

$$(\text{A.2})\qquad |\hat{K}(t)| = 0 \quad \text{for every } t \text{ such that } \|t\| > 1.$$

For an example, see Bhattacharya and Ranga Rao [(1986), Theorem 10.1, page 85].

The following result is a useful tool for deriving asymptotic expansions in the CLT for multivariate functions of independent random elements [see Götze (1985)], in so far as it does not require truncation.

THEOREM A.3. *Assume that $X_1, \dots, X_n$ denote independent random vectors in $\mathbb{R}^k$ such that*

$$EX_j = 0, \qquad 1 \le j \le n, \qquad \beta_s \triangleq \sum_{j=1}^{n} E\|X_j\|^s < \infty, \qquad s \ge 3, \tag{A.4}$$

*and*

$$\lambda_n \triangleq \textit{minimal eigenvalue of } \operatorname{Cov}(X_1 + \cdots + X_n) > 0. \tag{A.5}$$

*Let f denote a measurable function $f\colon \mathbb{R}^k \to \mathbb{R}$ and assume*

$$\begin{aligned} &M_s(f) \triangleq \sup\{f(x)/(1+\|x\|^s) : x \in \mathbb{R}^k\} < \infty,\\ &\beta_s \lambda_n^{-s/2} < (\log n)^{-s}/16k. \end{aligned} \tag{A.6}$$

*Then there is an absolute constant $c_1$ such that with $\varepsilon \triangleq c_1 \beta_3/\lambda_n$,*

$$F(x) \triangleq Ef(X_1 + \cdots + X_n + \varepsilon Y + x)$$

*satisfies*

$$|F^{(r)}(x)a^r| \le c(r,k) M_s(f)(1+\|x\|^s)\big(1+\beta_2^{s/2}\big)\big(1+\lambda_n^{-r/2}\big)\|a\|^r$$

*for every $x, a \in \mathbb{R}^k$, $r \in \mathbb{N} \cup \{0\}$.*

PROOF. For $I \subset \{1,\dots,n\}$, let $R_I \triangleq \sum_{j\in I} X_j$ and $R_n \triangleq X_1 + \cdots + X_n$. Let $k_\varepsilon$ denote the Lebesgue density of $K_\varepsilon(A) \triangleq K(\varepsilon^{-1}A)$. Let $C_n$ denote a positive definite matrix such that $C_n R_n$ has covariance equal to the $k \times k$-identity matrix. Replacing $f(x)$ by $\bar f(x) = f(C_n^{-1}x)$, $\varepsilon$ by $\bar\varepsilon \triangleq \varepsilon/\lambda_n^{1/2}$ and $K_\varepsilon$ by $\bar K_{\bar\varepsilon}(A) \triangleq K_\varepsilon(C_n^{-1}A)$, we obtain $M_s(\bar f) \le M_s(f)(1 + \|C_n^{-1}\|^s)$ and $\hat{\bar K}_{\bar\varepsilon}(t) = 0$ for $\varepsilon\|C_n\|\,\|t\| = \bar\varepsilon\|t\| > 1$. Thus (A.1)–(A.2) hold for $\bar K$ as well. Let $c$ denote in the following a generic constant depending on $k$ only. Writing $F(x) = E\bar f(C_n R_n + \varepsilon C_n Y + C_n x)$, we conclude that it is sufficient to prove (A.6) for the normalized quantities above. In order to simplify notations we shall write again $f$ instead of $\bar f$, $R_n$ instead of $C_n R_n$, etc.

Define $E_I' f \triangleq Ef \prod_{j\in I} I(\|X_j\| \le 1)$. Then

$$\begin{aligned} D &\triangleq \int Ef(R_n + y) k_\varepsilon^{(r)}(y) a^r \prod_j \big(I(\|X_j\| \le 1) + I(\|X_j\| > 1)\big)\, d^k y\\ &= \sum_I EF_I(R_{I^c}) \prod_{j\in I^c} I(\|X_j\| > 1), \end{aligned} \tag{A.7}$$

where $F_I(z) \triangleq \int E_I' f(R_I + z + y) k_\varepsilon^{(r)}(y) a^r\, d^k y$. We have

$$\begin{aligned} |F_I(z)| &\le c M_s(f)(1+\|z\|^s) \sup\Big\{(1+\|z\|^{k+1+s}) E_I' k_\varepsilon^{(r)}(R_I + z) a^r : z \in \mathbb{R}^k\Big\}\\ &\le c M_s(f)(1+\|z\|^s) \max_{|\beta| \le k+1} \int \big|D^\beta\big(E_I' \exp[itR_I] \hat K(t\varepsilon)(ta)^r\big)\big|\, d^k t\\ &\le c M_s(f)(1+\|z\|^s), \end{aligned} \tag{A.8}$$

by Lemma 11.6 and Corollary 14.4 of Bhattacharya and Ranga Rao (1986), provided that $\beta_s \le (16k(\log n)^s)^{-1}$ and $|I| > n - \log n$, $n$ sufficiently large.

Here we have used $\|\mathrm{Cov}\, R_I - Id\| \le \beta_s^{2/s}(\log n)^{1-2/s} < \frac{1}{8}$ for $n$ sufficiently large. From (A.7) and (A.8), we obtain

$$
\begin{aligned}
D \le cM_s(f) \sum_{l=0}^{\log n} \sum_{|I|=n-l} E \prod_{j\notin I} I(\|X_j\| > 1)\left(1 + l^{s-1} \sum_{j\notin I} E\|X_j\|^s I(\|X_j\| > 1)\right) \\
+ cM_s(f) \sum_{l>\log n} \varepsilon^{-r} \sum_{|I|=n-l} E_I'\left(1 + \left\|\sum_{j\in I} X_j\right\|^s\right) \prod_{j\notin I} E\|X_j\|^s.
\end{aligned} \tag{A.9}
$$

The estimate of the second summand of (A.9) follows from $\int |k_\varepsilon^{(r)}(y)|\, d^k y = O(\varepsilon^{-r})$ and $EI(\|X_j\| > 1) \le E\|X_j\|^s$. Since

$$
\sum_{|I|=l} \sum_{j\in I} E\|X_j\|^s I(\|X_j\| > 1) \prod_{l\in I,\, l\ne j} E\|X_l\|^s = o(\beta_s^l) l!^{-1},
$$

we obtain from (A.9),

$$
\begin{aligned}
D &\le cM_s(f)\left\{1 + \sum_{l=1}^{\log n} l^{s-1}\beta_s^l l!^{-1} + \left(\sum_{l>\log n} c\beta_s^l l!^{-1}\right)\varepsilon^{-r}\right\} \\
&\le cM_s(f),
\end{aligned} \tag{A.10}
$$

by our assumption on the size of $\beta_s$.

For the estimation of the second part of (A.9), we have used $E_I'\|R_I\|^p \le c_p$ for every $p$ [see, e.g., Götze and Hipp (1978), Lemma 4.5, page 75]. □

**Acknowledgment.** I would like to thank the referee for valuable suggestions which helped to improve the manuscript.

## REFERENCES

BARBOUR, A. D. (1988). Stein's method and Poisson process convergence. *J. Appl. Probab.* **25A** 175–184.

BENTKUS, V. (1986). Dependence of the Berry–Esseen estimate on the dimension. *Litovsk. Mat. Sb.* **26** 205–210. (In Russian.)

BERGSTRÖM, H. (1944). On the central limit theorem. *Skand. Aktuarietidskr.* **27** 139–153.

BERGSTRÖM, H. (1945). On the central limit theorem in the space $\mathbb{R}^k$, $k > 1$. *Skand. Aktuarietidskr.* **28** 106–127.

BHATTACHARYA, R. N. (1975). On the errors of normal approximation. *Ann. Probab.* **3** 815–828.

BHATTACHARYA, R. N. and RANGA RAO, R. (1986). *Normal Approximation and Asymptotic Expansion*. Krieger, Melbourne, Fla.

BOLTHAUSEN, E. (1984). An estimate of the remainder in combinatorial central limit theorem. *Z. Wahrsch. Verw. Gebiete* **66** 379–386.

CHATTERJI, S. D. (1969). An $L^p$-convergence theorem. *Ann. Math. Statist.* **40** 1068–1070.

FRIEDRICH, K. O. (1985). Berry–Esseen Abschätzungen für nichtparametrische Statistiken. Ph.D. dissertation, Albert-Ludwigs Universität, Freiburg im Breisgau.

FRIEDRICH, K. O. (1989). A Berry–Esseen bound for functions of independent random variables. *Ann. Statist.* **17** 170–183.

GÖTZE, F. (1985). Asymptotic expansions in functional limit theorems. *J. Multivariate Anal.* **16** 1–20.

GÖTZE, F. (1987). Approximations for multivariate $U$-statistics. *J. Multivariate Anal.* **22** 212–229.

GÖTZE, F. and HIPP, C. (1978). Asymptotic expansions in the CLT under moment conditions. *Z. Wahrsch. Verw. Gebiete* **42** 67–87.

LINDEBERG, Y. W. (1922). Eine neue Herleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechnung. *Math. Z.* **15** 221–225.

NAGAEV, S. V. (1976). An estimate of the remainder term in the multidimensional central limit theorem. *Proc. Third Japan–USSR Symp. Probab. Theory. Lecture Notes in Math.* **550** 419–438. Springer, Berlin.

SAZONOV, V. V. (1968). On the multi-dimensional central limit theorem. *Sankhyā Ser. A* **30** 181–204.

STEIN, C. (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. *Proc. Sixth Berkeley Symp. Math. Statist. Probab.* **2** 583–602. Univ. California Press, Berkeley.

STEIN, C. (1987). *Approximate Computation of Expectations. IMS Lecture Notes—Monograph Ser.* **7**. IMS, Hayward, Calif.

SWEETING, T. J. (1977). Speed of convergence for the multidimensional central limit theorem. *Ann. Probab.* **5** 28–41.

VAN ZWET, W. R. (1984). A Berry–Esseen bound for symmetric statistics. *Z. Wahrsch. Verw. Gebiete* **66** 425–440.

FAKULTÄT FÜR MATHEMATIK
UNIVERSITÄT BIELEFELD
POSTFACH 8640
4800 BIELEFELD 1
GERMANY