The Annals of Probability
1987, Vol. 15, No. 4, 1419–1440

STRONG INVARIANCE PRINCIPLES FOR PARTIAL SUMS OF INDEPENDENT RANDOM VECTORS[1]

BY UWE EINMAHL

University of Cologne

An estimate in the multidimensional central limit theorem is obtained, which is used together with the Strassen–Dudley theorem to prove a strong approximation theorem for partial sums of independent, identically distributed d-dimensional random vectors. This theorem implies immediately multidimensional versions of the strong invariance principles of Strassen and Major as well as a new d-dimensional strong invariance principle which improves the known results for the 1-dimensional case. In particular, we are able to weaken the assumption in Major's strong invariance principle. At the same time, it is shown that the assumptions of our theorem are nearly necessary.

1. Introduction. The first one to prove strong invariance principles was Strassen (1964). He showed that given a p-measure $Q|\mathbb{B}$ with zero mean and variance 1, one can construct a p-space $(\Omega, \mathscr{A}, P)$ and two sequences of i.i.d.r.v.'s $\{X_n\}, \{Y_n\}$ with $P \circ X_1 = Q$, $P \circ Y_1 = N(0,1)$ such that the partial sums $S_n := \sum_1^n X_k$, $T_n := \sum_1^n Y_k$ satisfy

$$S_n - T_n = o\left(\sqrt{n \log\log n}\right) \quad \text{a.s.} \tag{1.1}$$

It is known that the convergence rate in (1.1) cannot be improved in general [cf. Major (1976b)].

The situation changes if $Q|\mathbb{B}$ fulfills additional integrability assumptions. Let $\mathscr{H}$ denote the set of all continuous, nonnegative functions H on $[0, \infty)$ such that $t^{-2}H(t)$ is nondecreasing and $t^{-3}H(t)$ is nonincreasing.

Assume that

$$\int H(|x|)Q(dx) < \infty, \quad \text{for some } H \in \mathscr{H}. \tag{1.2}$$

Breiman (1967) showed that under this condition a construction is possible such that

$$S_n - T_n = o\left(H^{-1}(n)\sqrt{\log f_n}\right) \quad \text{a.s.,} \tag{1.3}$$

where $f_n \uparrow \infty$ with $\sum_n (1/H^{-1}(n)^2 f_n^m) < \infty$ for some $m > 0$. Moreover he was able to show that no better convergence rate than $o(H^{-1}(n))$ can be reached in general. The question remained open, whether (1.3) is the best possible result.

Using an entirely different method, Major (1976a) obtained a construction for which

$$S_n - T_n = o\left(H^{-1}(n)\right) \quad \text{a.s.,} \tag{1.4}$$

Received August 1985; revised August 1986.

[1]This paper is part of the author's doctoral thesis at the University of Cologne.

AMS 1980 *subject classifications*. Primary 60F17; secondary 60F15.

Key words and phrases. Strong invariance principles, partial sums of independent random vectors, strong approximations.

if the function $H \in \mathscr{H}$ fulfills the condition

$$t^{-2-r}H(t) \text{ is nondecreasing for some } r > 0. \tag{1.5}$$

In the paper of Komlós, Major and Tusnády (1976) it was shown that (1.4) remains valid for a large class of functions H such that $t^{-3}H(t)$ is nondecreasing.

The main purpose of this paper is to find out, what convergence rates in the strong invariance principle can be reached, if condition (1.5) is not fulfilled—in particular, if $H(t) = t^2 h(t)$, where $h(t) \uparrow \infty$ is a slowly varying function.

Since (1.4) cannot be valid for all functions $H \in \mathscr{H}$ [take $H(t) = t^2$], we formulate a somewhat different strong approximation theorem.

(1.6) One can construct two sequences of independent r.v.'s $\{X_n\}, \{Y_n\}$ with $P \circ X_n = Q$, $P \circ Y_n = N(0, \sigma_n^2)$, $\sigma_n^2 \uparrow 1$, $n \in \mathbb{N}$, such that $S_n - T_n = o(H^{-1}(n))$ a.s.

According to Major's strong invariance principle, (1.6) is true for functions H satisfying (1.5). (Set $\sigma_n^2 = 1$.) Moreover Major (1979) has shown (1.6) for $H(t) = t^2$. We will show that (1.6) is true for all functions $H \in \mathscr{H}$. The sequence σ_n^2 can be chosen in a way such that

$$1 - \sigma_n^2 = o\left(H^{-1}(n)^2/n\right), \quad \text{as } n \to \infty. \tag{1.7}$$

From (1.6) and (1.7) we obtain immediately the strong invariance principles of Strassen and Major—the last one even under a condition less restrictive than (1.5)—as well as a new strong invariance principle, which "interpolates" between the strong invariance principles of Strassen and Major.

Philipp (1979) extended Strassen's strong invariance principle to the multidimensional case, whereas his convergence rate in Major's strong invariance principle is slightly worse than that in (1.4). Using a different method, Berger (1982) obtained the convergence rate (1.4) in the multidimensional case.

Our proof of (1.6) and (1.7) uses a variant of Philipp's method. Whereas Philipp (1979) uses the Strassen–Dudley theorem and an estimate of the Prohorov distance in the multidimensional central limit theorem, we use the Strassen–Dudley theorem in connection with a different estimate in the multidimensional central limit theorem, which is obtained in Section 3.

In this way we are able to prove (1.6) and (1.7) for arbitrary functions $H \in \mathscr{H}$ and for p-measures $Q|\mathbb{B}^d$.

2. The results. Throughout the rest of the paper $Q|\mathbb{B}^d$ denotes a p-measure on the Borel sets of the d-dimensional euclidean space $\mathbb{R}^d$ with zero mean and covariance matrix Σ. Log t stands for $\log(\max(t, e))$. $|\cdot|$ denotes the euclidean norm in $\mathbb{R}^d$, $\|\cdot\|$ the euclidean matrix norm [i.e., $\|A\| := \max\{|A \cdot x|: |x| = 1\}$, if A is a $d \times d$-matrix].

Let us now state our main result.

THEOREM 1. *Let $H \in \mathscr{H}$. Assume $\int H(|x|)Q(dx) < \infty$. Then one can construct two sequences of independent random vectors $\{X_n\}, \{Y_n\}$ such that $P \circ X_n = Q$, $P \circ Y_n = N(0, \Sigma_n)$, $\|\Sigma_n - \Sigma\| = o(H^{-1}(n)^2/n)$, $n \in \mathbb{N}$, and $S_n - T_n = o(H^{-1}(n))$ a.s.*

Theorem 1 enables us now to prove strong invariance principles, where the function $\overline{H}(t) := t^2 \operatorname{Log}\operatorname{Log} t$, $t \geq 0$, plays an important role. If $H \in \mathscr{H}$ is a function such that $H(t)/\overline{H}(t)$ is nondecreasing, we obtain the convergence rate $o(H^{-1}(n))$. On the other hand, we can reach only worse convergence rates, if $H(t)/\overline{H}(t) \downarrow 0$ as $t \uparrow \infty$.

THEOREM 2. *Let $H \in \mathscr{H}$ be such that $t^{-2}(\operatorname{Log}\operatorname{Log} t)^{-1}H(t)$ is nondecreasing. Assume $\int H(|x|)Q(dx) < \infty$. Then one can construct two sequences of independent random vectors $\{X_n\}, \{Y_n\}$ with $P \circ X_n = Q$, $P \circ Y_n = N(0, \Sigma)$ such that $S_n - T_n = o(H^{-1}(n))$ a.s.*

THEOREM 3. *Let $H \in \mathscr{H}$ be such that $t^{-2}(\operatorname{Log}\operatorname{Log} t)^{-1}H(t)$ is nonincreasing. Assume $\int H(|x|)Q(dx) < \infty$. Then one can construct two sequences of independent random vectors $\{X_n\}, \{Y_n\}$ with $P \circ X_n = Q$, $P \circ Y_n = N(0, \Sigma)$ such that*

$$\text{(a)} \quad S_n - T_n = o\left(H^{-1}(n)\sqrt{\frac{\operatorname{Log}\operatorname{Log} n}{h(n)}}\right) \quad a.s., \text{ where } h(t) := t^{-2}H(t)$$

and

$$\text{(b)} \qquad \frac{1}{H^{-1}(n)} \max_{1 \leq k \leq n} |S_k - T_k| \to_P 0.$$

Theorem 3(b), which we obtain as a byproduct of the proof of Theorem 3(a), is a weak invariance principle [cf. Philipp (1980)].

Let us now consider the functions $H_\alpha(t) := t^2(\operatorname{Log}\operatorname{Log} t)^\alpha$, $\alpha \in [0, 1]$. If the p-measure $Q|\mathbb{B}^d$ fulfills $\int H_\alpha(|x|)Q(dx) < \infty$, we obtain from Theorem 3(a)

$$S_n - T_n = o\left(\sqrt{n}\,(\operatorname{Log}\operatorname{Log} n)^{1/2-\alpha}\right) \quad \text{a.s.} \tag{2.1}$$

In the case $\alpha = 0$, (2.1) is just the multidimensional version of Strassen's invariance principle. If $\alpha = 1$, (2.1) is a special case of Theorem 2, which generalizes the strong invariance principle of Major. If $\alpha \in (0, 1)$, we obtain the exponent in the logarithmic term by linear interpolation between 0 and 1. Theorem 4 shows that this choice is reasonable.

THEOREM 4. *Let $\alpha \in (0, 1)$. There exists a p-measure $Q_\alpha|\mathbb{B}$ with zero mean, variance 1 and $\int H_\alpha(|x|)Q_\alpha(dx) < \infty$, such that for all sequences of i.i.d.r.v.'s $\{X_n\}, \{Y_n\}$ with $P \circ X_1 = Q_\alpha$, $P \circ Y_1 = N(0, 1)$ the following is true:*

$$\limsup_n \frac{|S_n - T_n|}{\sqrt{n}\,(\operatorname{Log}\operatorname{Log} n)^\beta} = \infty \quad a.s., \text{ if } \beta < \frac{1}{2} - \alpha.$$

Theorem 4 also shows that the assumption "$t^{-2}(\operatorname{Log}\operatorname{Log} t)^{-1}H(t)$ nondecreasing" in Theorem 2 cannot be replaced by "$t^{-2}(\operatorname{Log}\operatorname{Log} t)^{-\alpha}H(t)$ nondecreasing" with some $\alpha < 1$.

By a slight modification of the proof of Theorem 2 it is possible to prove

THEOREM 5. *Let H be a continuous nonnegative function on $[0, \infty)$ such that $t^{-3}H(t)$ is nondecreasing and $t^{-4+r}H(t)$ is nonincreasing for some $r > 0$. Assume $\int H(|x|)Q(dx) < \infty$. Then one can construct two sequences of independent random vectors $\{X_n\}, \{Y_n\}$ with $P \circ X_n = Q$, $P \circ Y_n = N(0, \Sigma)$ such that*

$$S_n - T_n = o(H^{-1}(n)) \quad a.s.$$

Thus we obtain by means of the present method a partial extension of the above-mentioned result of Komlós, Major and Tusnády (1976) to the multidimensional case.

3. Estimates of the convergence rate in the multidimensional central limit theorem. Let $\xi_1, \dots, \xi_n: \Omega \to \mathbb{R}^d$ be independent random vectors with zero means and finite third moments. Let $\Gamma_n := \operatorname{cov}(\xi_1 + \cdots + \xi_n)$ be positive definite. Denote by $\lambda_n(\Lambda_n)$ the smallest (largest) eigenvalue of Γ_n. Set

$$\rho_n := \lambda_n^{-3/2}\sum_1^n E\big[|\xi_k|^3\big], \qquad P_n := P \circ \Gamma_n^{-1/2}\sum_1^n \xi_k.$$

From a well-known theorem of Yurinskii (1975) we obtain an estimate of the Prohorov distance in the multidimensional central limit theorem,

$$\rho(P_n, N(0, I)) \le C\rho_n, \tag{3.1}$$

where C is a positive constant, which depends only on d. I denotes the d-dimensional unit matrix. Recall that the Prohorov distance of two p-measures $Q_i|\mathbb{B}^d$, $i = 1, 2$, is defined by $\rho(Q_1, Q_2) = \inf\{\varepsilon > 0: Q_1(A) \le Q_2(A^\varepsilon) + \varepsilon$ for all closed sets $A \subseteq \mathbb{R}^d\}$, where $A^\varepsilon := \{x: \inf_{y \in A}|x - y| < \varepsilon\}$.

We set for $Q_i|\mathbb{B}^d$, $i = 1, 2$, $\delta > 0$: $\lambda(Q_1, Q_2, \delta) := \sup\{Q_1(A) - Q_2(A^\delta): A \subseteq \mathbb{R}^d$ closed$\}$. Then (3.1) can be rewritten

$$\lambda(P_n, N(0, I), C\rho_n) \le C\rho_n. \tag{3.2}$$

The purpose of this section is to find conditions, which guarantee that for given $\varepsilon \in (0, 1)$ and $\alpha > 1$ the following holds true:

$$\lambda(P_n, N(0, I), \rho_n^{1-\varepsilon}) = O(\rho_n^\alpha).$$

Our main result in this direction is

THEOREM 6. *Let $\xi_1, \dots, \xi_n: \Omega \to \mathbb{R}^d$ be independent random vectors with zero means. Assume $|\xi_k| \le K\sqrt{\lambda_n \operatorname{Log} 1/\rho_n}$ a.s., $k = 1, \dots, n$. Then we have for $\varepsilon \in (0, 1)$*

$$\lambda(P_n, N(0, I), C_{\varepsilon, K}\rho_n^{1-\varepsilon}) \le C'_{\varepsilon, K}\rho_n^{(1/8d)(\varepsilon/K)^2},$$

where $C_{\varepsilon, K}, C'_{\varepsilon, K}$ are positive constants depending on ε, K and d only.

To prove Theorem 6, we use a similar method to Yurinskii (1977). We denote by φ_Σ the density of $N(0, \Sigma)$. The main tool of the proof is

PROPOSITION 1. *Let the assumptions of Theorem 6 be fulfilled and let $\varepsilon \in (0, 1)$. Set $P_n^* := P_n * N(0, \sigma_n^2)^d$, where $1 \ge \sigma_n \ge \rho_n^{1-4\varepsilon/5}$. Let $x \to p_n^*(x)$ be* *the continuous Lebesgue density of* P_n^*. *Then*:

(a) *If* $\rho_n \le 1/e$, *we have uniformly for* $|x| \le \frac{1}{2}(\varepsilon/K)\sqrt{1+\sigma_n^2}\sqrt{\operatorname{Log} 1/\rho_n}$

$$|p_n^*(x)/\varphi_{(1+\sigma_n^2)I}(x) - 1| = O(\rho_n^{1-\varepsilon}),$$

where the constant in $O(\cdot)$ *depends only on* ε, K *and* d.

(b) *There exists a positive constant* $D_{\varepsilon, K}$ *depending only on* ε, K *and* d *such that, if* $\rho_n \le D_{\varepsilon, K}$, *we have for* $|x| \le \frac{1}{2}(\varepsilon/K)\sqrt{1+\sigma_n^2}\sqrt{\operatorname{Log} 1/\rho_n}$, $y \in \mathbb{R}^d$

$$p_n^*(y) \le 3p_n^*(x)\exp(-\langle h, \Gamma_n^{1/2}(y-x)\rangle).$$

h_x *is defined by* (3.20) *and satisfies*

$$\left| h_x - \frac{1}{1+\sigma_n^2}\Gamma_n^{-1/2}x \right| \le \frac{1}{2(1+\sigma_n^2)}|\Gamma_n^{-1/2}x|.$$

The lemmas needed in the proof of Proposition 1 and Theorem 6 are stated and proved in Section 8.

PROOF OF PROPOSITION 1. We denote by C_i, $i = 1, \dots, 11$, constants, which depend only on ε, K and d.

(i) Let η_k, $k = 1, \dots, n$, be independent random vectors such that $P \circ \eta_k = N(0, \sigma_n^2 \operatorname{cov}(\xi_k))$, $k = 1, \dots, n$. Assume furthermore that the ξ_k's are independent of the η_k's.

Setting $Z_k := \xi_k + \eta_k$, $k = 1, \dots, n$, we obtain

$$P_n^* = P \circ \Gamma_n^{-1/2} \sum_1^n Z_k. \tag{3.3}$$

We apply the technique of conjugated random vectors. Let for $h \in \mathbb{R}^d$ Z_k^h, $k = 1, \dots, n$, be independent random vectors such that we have for $A \in \mathbb{B}^d$

$$P(Z_k^h \in A) = \frac{1}{E[\exp(\langle h, Z_k \rangle)]} \int_A \exp(\langle h, x\rangle)(P \circ Z_k)(dx), \qquad k = 1, \dots, n. \tag{3.4}$$

Then it is easy to see that

$$E[Z_k^h] = \nabla L_k(h), \qquad \operatorname{cov}(Z_k^h) = L_k''(h), \tag{3.5}$$

where $L_k(h) := \log R_k(h)$, $R_k(h) := E[\exp(\langle h, Z_k \rangle)]$, $k = 1, \dots, n$. (L_k'' denotes the matrix of the second partial derivatives of L_k.)

From the definition of the Z_k^h's we infer

$$P \circ \sum_1^n Z_k \ll P \circ \sum_1^n Z_k^h, \quad \text{with density } x \to \exp\left(\sum_1^n L_k(h) - \langle h, x \rangle \right) \tag{3.6}$$

and

$$P \circ (Z_k^h - E[Z_k^h]) = P \circ (\xi_k^h - E[\xi_k^h]) * P \circ \eta_k, \qquad k = 1, \dots, n, \tag{3.7}$$

where ξ_k^h is defined by (3.4) with Z_k replaced by ξ_k.

(ii) Set $\Gamma_n^*(h) := \text{cov}(Z_1^h + \cdots + Z_n^h)$, $\Gamma_n^* := \text{cov}(Z_1 + \cdots + Z_n)$. By (3.7) we have

$$\Gamma_n^*(h) - \Gamma_n^* = \text{cov}\left(\xi_1^h + \cdots + \xi_n^h\right) - \text{cov}(\xi_1 + \cdots + \xi_n)$$
$$= \sum_1^n \left(L''_{1,k}(h) - L''_{1,k}(0)\right),$$

where $L_{1,k}(h) := \log R_{1,k}(h)$, $R_{1,k}(h) := E[\exp(\langle h, \xi_k\rangle)]$, $k = 1,\dots,n$.

Applying the Hölder inequality, we easily obtain for $1 \le \alpha, \beta, \gamma \le d$,

$$\left|\frac{\partial^3 L_{1,k}}{\partial x_\alpha\, \partial x_\beta\, \partial x_\gamma}(h)\right| \le 6E\left[|\xi_k^h|^3\right], \qquad k = 1,\dots,n. \tag{3.8}$$

Since $E[\xi_k] = 0$, we infer from Jensen's inequality that $R_{1,k}(h) \ge 1$. Hence

$$E\left[|\xi_k^h|^3\right] \le E\left[|\xi_k|^3 \exp(|h| \cdot |\xi_k|)\right], \qquad k = 1,\dots,n. \tag{3.9}$$

Set

$$H_n := \frac{3}{4}\frac{\varepsilon}{K}\left(\frac{\operatorname{Log} 1/\rho_n}{\lambda_n}\right)^{1/2}.$$

Since $|\xi_k| \le K\sqrt{\lambda_n \operatorname{Log} 1/\rho_n}$ a.s., (3.9) implies

$$E\left[|\xi_k^h|^3\right] \le E\left[|\xi_k|^3\right]\rho_n^{-3/4\varepsilon}, \quad \text{if } |h| \le H_n. \tag{3.10}$$

Using (3.8) and (3.10), it is easy to see that

$$\|\Gamma_n^*(h) - \Gamma_n^*\| \le C_1 \rho_n^{1-3/4\varepsilon}\sqrt{\operatorname{Log} 1/\rho_n}\,\lambda_n, \quad \text{if } |h| \le H_n. \tag{3.11}$$

Hence, if $\rho_n \le C_2$, say, we have

$$\|\Gamma_n^*(h) - \Gamma_n^*\| \le \lambda_n/3, \quad \text{for } |h| \le H_n. \tag{3.12}$$

(iii) (3.12) shows in particular that $\Gamma_n^*(h)$ is positive definite, hence $\Gamma_n^*(h)^{-1/2}$ is well defined for $|h| \le H_n$, if $\rho_n \le C_2$.

Let $x \to p_n^h(x)$ be the continuous Lebesgue density of

$$P_n^h := P \circ \Gamma_n^*(h)^{-1/2} \sum_1^n \left(Z_k^h - E\left[Z_k^h\right]\right) \quad [\text{notice } (3.7)].$$

The purpose of this part is to show that

$$\sup_x |p_n^h(x) - \varphi_I(x)| \le C_3 \rho_n^{1-3/4\varepsilon}, \quad \text{if } |h| \le H_n,\ \rho_n \le C_2. \tag{3.13}$$

Set

$$l_{3,n}(h) := \max_{|t|=1} \frac{\Sigma_1^n E\left[|\langle Z_k^h - E\left[Z_k^h\right], t\rangle|^3\right]}{\langle t, \Gamma_n^*(h) t\rangle^{3/2}}.$$

From (3.7), (3.10) and (3.12) we easily obtain

$$l_{3,n}(h) \le C_4 \rho_n^{1-3/4\varepsilon}, \quad \text{if } |h| \le H_n,\ \rho_n \le C_2. \tag{3.14}$$

Denote by f_n^h the characteristic function of P_n^h. From Theorems 8.4 and 8.9, Bhattacharya and Rao (1976) and (3.14) we infer for $|h| \le H_n$,

$$(3.15)\quad \left| f_n^h(t) - \exp\left(-\frac{1}{2}|t|^2\right)\right| \le C_5\rho_n^{1-3/4\varepsilon}|t|^3\exp\left(-\frac{|t|^2}{3}\right), \quad \text{if } |t| \le C_6\rho_n^{3/4\varepsilon-1}.$$

By (3.7) we have for arbitrary $t \in \mathbb{R}^d$

$$|f_n^h(t)| \le \left| E\left[\exp\left(\left\langle it, \Gamma_n^*(h)^{-1/2}\sum_1^n \eta_k\right\rangle\right)\right]\right|$$

$$= \exp\left(-\tfrac{1}{2}\sigma_n^2\langle t, \Gamma_n^*(h)^{-1/2}\Gamma_n^*\Gamma_n^*(h)^{-1/2}t\rangle\right) \le \exp\left(-\tfrac{1}{4}\sigma_n^2|t|^2\right),$$

since $\|\Gamma_n^*(h)^{-1/2}\Gamma_n^*\Gamma_n^*(h)^{-1/2} - I\| \le \frac{1}{2}$ by (3.12).

Using the Fourier inversion formula, we obtain for $|h| \le H_n$, if $\rho_n \le C_2$,

$$\sup_x |p_n^h(x) - \varphi_I(x)| \le \frac{1}{(2\pi)^d}\int\left| f_n^h(t) - \exp\left(-\frac{|t|^2}{2}\right)\right| dt$$

$$\le C_7\rho_n^{1-3/4\varepsilon} + \frac{2}{(2\pi)^d}\int_{\{|t| > C_6\rho_n^{3/4\varepsilon-1}\}}\exp\left(-\sigma_n^2\frac{|t|^2}{4}\right) dt$$

$$\le C_3\rho_n^{1-3/4\varepsilon},$$

since $\sigma_n \ge \rho_n^{1-4/5\varepsilon}$.

Before we come to the next part of the proof, let us still remark that (3.12) and hence (3.13) remain true for $C_2 \le \rho_n \le 1/e$, if $h = 0$. Thus we have

$$(3.16)\quad \sup_x |p_n^*(x) - \varphi_{(1+\sigma_n^2)I}(x)| \le C_3\rho_n^{1-\varepsilon}.$$

(iv) Using (3.6), it is straightforward to check that

$$(3.17)\quad p_n^*(x) = \varphi_{(1+\sigma_n^2)I}(x)\exp\left(\sum_1^n L_k(h) - \langle h, \Gamma_n^{1/2}x\rangle + \frac{1}{2}\frac{|x|^2}{1+\sigma_n^2}\right) r_n^h(x),$$

with

$$(3.18)\quad r_n^h(x) = \sqrt{\det \Gamma_n^*\Gamma_n^*(h)^{-1}}\,\sqrt{2\pi}^{\,d}\, p_n^h\left(\Gamma_n^*(h)^{-1/2}\left(\Gamma_n^{1/2}x - \nabla\left(\sum_1^n L_k\right)(h)\right)\right).$$

We set $g_n(h) := \Gamma_n^{*-1}\nabla(\sum_1^n L_k)(h)$, $h \in \mathbb{R}^d$. Let $J_h g_n$ be the Jacobi matrix of g_n, i.e., $J_h g_n = \Gamma_n^{*-1}\Gamma_n^*(h)$. By relation (3.12) we have

$$(3.19)\quad \|J_h g_n - I\| \le \tfrac{1}{3}, \quad \text{if } |h| \le H_n,\ \rho_n \le C_2.$$

From Lemma 1 we obtain for $|\Gamma_n^{-1/2}x| \le \frac{2}{3}(1+\sigma_n^2)H_n$ a unique h_x such that $|h_x| \le H_n$ and

$$(3.20)\quad g_n(h_x) = \frac{1}{1+\sigma_n^2}\Gamma_n^{-1/2}x.$$

From (3.11) and (3.20) it follows easily that

$$\left| h_x - \frac{1}{1+\sigma_n^2}\Gamma_n^{-1/2}x \right| \le C_1\rho_n^{1-3/4\varepsilon}\sqrt{\operatorname{Log} 1/\rho_n}\,|h_x| \tag{3.21}$$

and

$$\left| h_x - \frac{1}{1+\sigma_n^2}\Gamma_n^{-1/2}x \right| \le \frac{1}{3}|h_x|, \quad \text{if } \rho_n \le C_2. \tag{3.22}$$

(v) Let now

$$|x| \le \frac{1}{2}\frac{\varepsilon}{K}\sqrt{1+\sigma_n^2}\sqrt{\operatorname{Log} 1/\rho_n}$$

be fixed. (Then trivially $|\Gamma_n^{-1/2}x| \le \frac{2}{3}(1+\sigma_n^2)H_n$.) Setting $h = h_x$ in (3.17) we obtain from Taylor's theorem and (3.20), if $\rho_n \le C_2$,

$$p_n^*(x) = \varphi_{(1+\sigma_n^2)I}(x)\exp(\gamma_n(x))\sqrt{\det \Gamma_n^*\Gamma_n^*(h_x)^{-1}}\,\sqrt{2\pi}^{\,d}\,p_n^{h_x}(0),$$

with

$$\gamma_n(x) = \frac{1}{2}\left(\frac{|x|^2}{1+\sigma_n^2} - \langle h_x, \Gamma_n^*(h'_x)h_x\rangle\right) \quad \text{and} \quad |h'_x| \le H_n.$$

Using (3.11) and (3.21), one can easily show that

$$|\gamma_n(x)| \le C_8\rho_n^{1-\varepsilon}. \tag{3.23}$$

Since we have

$$\left|\sqrt{\det \Gamma_n^*\Gamma_n^*(h_x)^{-1}} - 1\right| \le C_9\|\Gamma_n^*\Gamma_n^*(h_x)^{-1} - I\| \le C_{10}\rho_n^{1-\varepsilon},$$

by (3.11), we obtain from (3.13) and (3.23)

$$|p_n^*(x)/\varphi_{(1+\sigma_n^2)I}(x) - 1| \le C_{11}\rho_n^{1-\varepsilon}, \quad \text{if } \rho_n \le C_2.$$

If $C_2 \le \rho_n \le 1/e$, (a) follows immediately from (3.16). Thus we have shown (a).

(vi) In order to prove (b), we apply formula (3.17) with x and y. Setting $h = h_x$, we obtain from (3.13)

$$\begin{aligned} p_n^*(y) &= p_n^*(x)\exp(-\langle h_x, \Gamma_n^{1/2}(y-x)\rangle)\frac{p_n^{h_x}\left(\Gamma_n^*(h_x)^{-1/2}\Gamma_n^{1/2}(y-x)\right)}{p_n^{h_x}(0)} \\ &\le 3p_n^*(x)\exp(-\langle h_x, \Gamma_n^{1/2}(y-x)\rangle), \end{aligned}$$

if we have chosen $D_{\varepsilon,K}$ so small that $C_3\rho_n^{1-3/4\varepsilon} \le 1/2\sqrt{2\pi}^{\,d}$, if $\rho_n \le D_{\varepsilon,K}$. From (3.22) we infer

$$|h_x| - \frac{1}{1+\sigma_n^2}|\Gamma_n^{-1/2}x| \le \frac{1}{3}|h_x|.$$

Hence

$$\frac{2}{3}|h_x| \le \frac{1}{1+\sigma_n^2}|\Gamma_n^{-1/2}x|.$$

This together with (3.22) implies part (b). □

Proof of Theorem 6. The proof is based on Proposition 1 and Lemma 2, Section 8. We denote by $C_{12}, \dots, C_{19}$ positive constants, which depend only on ε, K and d.

First we observe that it suffices to show

$$\lambda\left(P_n^*, N\left(0, (1+\sigma_n^2)I\right), C_{12}\rho_n^{1-\varepsilon}\right) \le C_{13}\rho_n^{(1/8d)(\varepsilon/K)^2}, \quad \text{if } 1 \ge \sigma_n \ge \rho_n^{1-4/5\varepsilon}. \tag{3.24}$$

Setting $\sigma_n = \rho_n^{1-4/5\varepsilon}$, the assertion follows immediately from (3.24) and the simple smoothing inequality

$$\begin{aligned}\lambda(P_n, N(0,I), \delta) \le \lambda\left(P_n^*, N\left(0,(1+\sigma_n^2)I\right), \delta/2\right)\\ +2N\left(0,\sigma_n^2\right)^d\{|x| \ge \delta/4\}.\end{aligned} \tag{3.25}$$

To prove (3.24), we show

$$\begin{aligned}&P_n^*(A) \le N\left(0,(1+\sigma_n^2)I\right)(A^{\delta_n'}) + C_{14}\rho_n^{(1/8d)(\varepsilon/K)^2}\\ &\text{for all closed sets } A \subseteq \left\{|x|_+ \le \frac{1}{2\sqrt{d}}\frac{\varepsilon}{K}\sqrt{1+\sigma_n^2}\sqrt{\operatorname{Log} 1/\rho_n}\right\},\end{aligned} \tag{3.26}$$

where $\delta_n' = C_{12}\rho_n^{1-\varepsilon}$ and

$$P_n^*\left\{|x|_+ > \frac{1}{2\sqrt{d}}\frac{\varepsilon}{K}\sqrt{1+\sigma_n^2}\sqrt{\operatorname{Log} 1/\rho_n}\right\} \le C_{15}\rho_n^{(1/8d)(\varepsilon/K)^2}. \tag{3.27}$$

$|\cdot|_+$ is defined by $|x|_+ := \max\{|x_i|: i = 1, \dots, d\}$ for $x = (x_1, \dots, x_d)$.

Proof of (3.26). It suffices to prove (3.26) under the condition that ρ_n is sufficiently small. The general case follows then by a possible enlargement of the constant C_{14}. We use induction on the dimension d.

If $d = 1$, the assertion follows easily from Lemma 2 and Proposition 1.

Assume now that the assertion holds true for dimension $(d-1)$. We set $\hat{P}_n := P_n^* \circ T$, where $T(x) = (x_1, \dots, x_{d-1})$ for $x = (x_1, \dots, x_d)$. Then we have $\hat{P}_n := (P \circ \Sigma_1^n \hat{\xi}_k) * N(0, \sigma_n^2)^{d-1}$, where $\hat{\xi}_k := T(\Gamma_n^{-1/2}\xi_k)$: $\Omega \to \mathbb{R}^{d-1}$, $k = 1, \dots, n$, are independent random vectors. Since $\operatorname{cov}(\hat{\xi}_1 + \cdots + \hat{\xi}_n) = I_{d-1}$ [$= (d-1)$-dimensional unit matrix], $|\hat{\xi}_k| \le |\Gamma_n^{-1/2}\xi_k| \le K\sqrt{\operatorname{Log} 1/\rho_n}$ a.s., $k = 1, \dots, n$, $\hat{\rho}_n := \Sigma_1^n E[|\hat{\xi}_k|^3] \le \rho_n$, the $\hat{\xi}_k$'s fulfill the assumptions of Theorem 6.

By the induction hypothesis we have

$$\begin{aligned}&\hat{P}_n(A) \le N\left(0, 1+\sigma_n^2\right)^{d-1}(A^{\hat{\delta}_n}) + C_{14}(d-1)\hat{\rho}_n^{(1/8(d-1))(\varepsilon/K)^2}\\ &\text{for all closed sets } A \subseteq \left\{u \in \mathbb{R}^{d-1}: |u|_+ \le \frac{1}{2\sqrt{d-1}}\frac{\varepsilon}{K}\sqrt{1+\sigma_n^2}\sqrt{\operatorname{Log} 1/\rho_n}\right\},\end{aligned} \tag{3.28}$$

where $\hat{\delta}_n = C_{12}(d-1)\hat{\rho}_n \le C_{12}(d-1)\rho_n$. Let now for $u \in \mathbb{R}^{d-1}$, $v \to p_n^*(v|u)$ be the conditional density of $P_n^* \circ \pi_d$ given $(\pi_1, \dots, \pi_{d-1}) = u$, where π_i denotes the projection on the ith coordinate, $i = 1, \dots, d$.

If $u \to \hat{p}_n(u)$ denotes the density of $\hat{P}_n$, we have

$$p_n^*(v|u) = p_n^*(u, v)/\hat{p}_n(u). \tag{3.29}$$

Applying Proposition 1 for $\xi_1, \dots, \xi_n$ and $\hat{\xi}_1, \dots, \hat{\xi}_n$, we obtain positive constants C_{16}, C_{17} depending only on ε, K and d such that we have for

$$\max(|u|_+, |v|) \leq \frac{1}{2\sqrt{d}} \frac{\varepsilon}{k} \sqrt{1 + \sigma_n^2} \sqrt{\operatorname{Log} 1/\rho_n},$$

$$\exp\left(-C_{16}\rho_n^{1-\varepsilon}\right) \leq p_n^*(v|u) \leq \exp\left(C_{16}\rho_n^{1-\varepsilon}\right), \quad \text{if } \rho_n \leq C_{17}. \tag{3.30}$$

Let now

$$A \subseteq \left\{|x|_+ \leq \frac{1}{2\sqrt{d}} \frac{\varepsilon}{K} \sqrt{1 + \sigma_n^2} \sqrt{\operatorname{Log} 1/\rho_n}\right\}$$

be a closed set. Then trivially

$$P_n^*(A) = \int\int 1_{A_u}(v) p_n^*(v|u)\, dv \hat{p}_n(u)\, du, \tag{3.31}$$

where $A_u := \{v \in \mathbb{R}: (u, v) \in A\}$. We assume w.l.o.g. that C_{17} is small enough such that

$$C_{16}\rho_n^{1-\varepsilon} \leq \frac{1}{\sqrt{2\pi}\, e^4} \quad \text{and} \quad \frac{1}{2\sqrt{d}} \frac{\varepsilon}{K} \sqrt{\operatorname{Log} 1/\rho_n} \geq 2, \quad \text{if } \rho_n \leq C_{17}.$$

Using (3.30) and Lemma 2, we obtain for $u \in \mathbb{R}^{d-1}$,

$$\int 1_{A_u}(v) p_n^*(v|u)\, dv \leq N\left(0, 1 + \sigma_n^2\right)\left((A_u)^{\tilde{\delta}_n}\right) + 2\rho_n^{(1/8d)(\varepsilon/K)^2}, \tag{3.32}$$

where $\tilde{\delta}_n = C_{18}\rho_n^{1-\varepsilon}$. [Notice that $A_u = \varnothing$, if

$$|u|_+ > \frac{1}{2\sqrt{d}} \frac{\varepsilon}{K} \sqrt{1 + \sigma_n^2} \sqrt{\operatorname{Log} 1/\rho_n}.]$$

Since $(A_u)^{\tilde{\delta}_n} \subseteq (A^{\tilde{\delta}_n})_u$, we infer from (3.31) and (3.32)

$$\begin{aligned} P_n^*(A) &\leq \int N\left(0, 1 + \sigma_n^2\right)\left((A^{\tilde{\delta}_n})_u\right)\hat{p}_n(u)\, du + 2\rho_n^{(1/8d)(\varepsilon/K)^2} \\ &= \int \hat{P}_n\left((A^{\tilde{\delta}_n})_v\right) N\left(0, 1 + \sigma_n^2\right)(dv) + 2\rho_n^{(1/8d)(\varepsilon/K)^2}, \quad \text{if } \rho_n \leq C_{17}. \end{aligned}$$

From (3.28) we obtain finally, if $\rho_n \leq C_{17}$,

$$\begin{aligned} P_n^*(A) &\leq \int N\left(0, 1 + \sigma_n^2\right)^{d-1}\left((A^{\tilde{\delta}_n + \hat{\delta}_n})_v\right) N\left(0, 1 + \sigma_n^2\right)(dv) \\ &\qquad + (2 + C_{14}(d-1))\rho_n^{(1/8d)(\varepsilon/K)^2} \\ &= N\left(0, (1 + \sigma_n^2)I\right)\left(A^{\tilde{\delta}_n + \hat{\delta}_n}\right) + \left(2 + C_{14}(d-1)\right)\rho_n^{(1/8d)(\varepsilon/K)^2}. \end{aligned}$$

[Notice that w.l.o.g.

$$(A^{\tilde{\delta}_n})_v \subseteq \left\{u \in \mathbb{R}^{d-1}: |u|_+ \leq \frac{1}{2\sqrt{d-1}} \frac{\varepsilon}{K} \sqrt{1 + \sigma_n^2} \sqrt{\operatorname{Log} 1/\rho_n}\right\},$$

if $\rho_n \leq C_{17}$.]

PROOF OF (3.27). We prove the inequality

$$
(3.33)\qquad \begin{aligned} &P_n^*\{|x|_+ > r\} \le C_{15}\exp\left(-\frac{r^2}{2(1+\sigma_n^2)}\right),\\ &0 \le r \le \frac{1}{2}\frac{\varepsilon}{K}\sqrt{1+\sigma_n^2}\sqrt{\operatorname{Log} 1/\rho_n}\,. \end{aligned}
$$

Setting $P_{n,i}^* := P_n^* \circ \pi_i$, $i = 1,\dots,d$, we obtain

$$
(3.34)\qquad P_n^*\{|x|_+ > r\} \le \sum_{i=1}^{d} P_{n,i}^*\{t \in \mathbb{R}: |t| > r\}.
$$

Let $i \in \{1,\dots,d\}$ be fixed. We have

$$
P_{n,i}^* = \left(P \circ \sum_{k=1}^{n} \hat{\xi}_{k,i}\right) * N(0,\sigma_n^2),
$$

where $\hat{\xi}_{k,i} = \pi_i \circ \Gamma_n^{-1/2}\xi_k$, $k = 1,\dots,n$.

Using the same arguments as in the proof of (3.26), one can show that the $\hat{\xi}_{k,i}$'s fulfill the assumptions of the proposition. Thus we obtain, for the density $p_{n,i}^*$ of $P_{n,i}^*$, if $\rho_n \le C_{19}$, say,

$$
(3.35)\qquad p_{n,i}^*(r) \le 2\varphi_{1+\sigma_n^2}(r)
$$

and

$$
(3.36)\qquad p_{n,i}^*(t) \le 3p_{n,i}^*(r)\exp(-h_{n,i}(t-r)), \qquad t \ge r,
$$

where

$$
h_{n,i} \ge \frac{1}{2}\frac{r}{1+\sigma_n^2} \ge \frac{r}{4}.
$$

Integration of (3.36) yields

$$
P_{n,i}^*(r,\infty) \le \frac{24}{r}\exp\left(-\frac{r^2}{2(1+\sigma_n^2)}\right).
$$

Similarly, we obtain

$$
P_{n,i}^*(-\infty,-r) \le \frac{24}{r}\exp\left(-\frac{r^2}{2(1+\sigma_n^2)}\right).
$$

From (3.34) we finally obtain

$$
P_n^*\{|x|_+ > r\} \le 48d\exp\left(-\frac{r^2}{2(1+\sigma_n^2)}\right), \quad \text{if } \rho_n \le C_{19}.
$$

[Notice that $48d\exp(-(r^2/2(1+\sigma_n^2)) \ge 1$, if $0 \le r \le 1$.] It is easy now to see that (3.33) holds, if we choose $C_{15} \ge 48d$ large enough. □

4. Proof of Theorem 1. The main tool of the proof is the following.

THEOREM 7. *Let $\xi_1, \dots, \xi_n$: $\Omega \to \mathbb{R}^d$ be independent random vectors with zero means. Assume*

$$|\xi_k| \le \frac{\varepsilon}{14\sqrt{d}} \sqrt{\lambda_n \operatorname{Log} 1/\rho_n} \quad a.s., \qquad k = 1, \dots, n,$$

where $\lambda_n(\Lambda_n)$ is the smallest (largest) eigenvalue of $\Gamma_n := \operatorname{cov}(\xi_1 + \cdots + \xi_n)$, $\rho_n := \lambda_n^{-3/2}\sum_1^n E[|\xi_k|^3]$, $\varepsilon \in (0,2)$. One can construct a p-space $(\Omega_0, \mathscr{A}_0, P_0)$ and random vectors S_n, T_n: $\Omega_0 \to \mathbb{R}^d$ such that $P_0 \circ S_n = P \circ \sum_1^n \xi_k$, $P_0 \circ T_n = N(0, \Gamma_n)$ and $E[|S_n - T_n|^2] \le D_\varepsilon \Lambda_n \rho_n^{2-\varepsilon}$, where D_ε is a positive constant depending on ε and d only.

To prove Theorem 7 we apply the well-known Strassen–Dudley theorem [cf. Dudley (1968), Theorem 2], Theorem 6 and a moment inequality which is proved in Section 8.

PROOF OF THEOREM 7. W.l.o.g. we assume $\rho_n \le \frac{1}{2}$. Let $P_n := P \circ \Gamma_n^{-1/2}\sum_1^n \xi_k$. From Theorem 6 (applied with $\varepsilon/2$ and $K = \varepsilon/14\sqrt{d}$) we infer that $\lambda(P_n, N(0, I), C_\varepsilon \rho_n^{1-\varepsilon/2}) \le C_\varepsilon' \rho_n^6$, where C_ε and C_ε' are positive constants depending on ε and d only.

Using the Strassen–Dudley theorem, we obtain a p-space $(\Omega_0, \mathscr{A}_0, P_0)$ and random vectors S_n, T_n: $\Omega_0 \to \mathbb{R}^d$ with the above distributions such that

$$P_0\left\{|S_n - T_n| > C_\varepsilon \sqrt{\Lambda_n}\, \rho_n^{1-\varepsilon/2}\right\} \le C_\varepsilon' \rho_n^6. \tag{4.1}$$

Since we have by the Hölder inequality for $\delta > 0$,

$$\begin{aligned} E\left[|S_n - T_n|^2\right] &\le \delta^2 + E\left[|S_n - T_n|^2 1_{\{|S_n - T_n| > \delta\}}\right] \\ &\le \delta^2 + E\left[|S_n - T_n|^3\right]^{2/3} (P\{|S_n - T_n| > \delta\})^{1/3}, \end{aligned}$$

we obtain

$$E\left[|S_n - T_n|^2\right] \le C_\varepsilon^2 \Lambda_n \rho_n^{2-\varepsilon} + 2C_\varepsilon'^{1/3}\left(E\left[|S_n|^3\right]^{2/3} + E\left[|T_n|^3\right]^{2/3}\right)\rho_n^2. \tag{4.2}$$

[Notice that $E[|S_n - T_n|^3]^{2/3} \le 2(E[|S_n|^3]^{2/3} + E[|T_n|^3]^{2/3})$ by the Minkowski inequality.]

Using the obvious inequality

$$E\left[|S_n|^3\right] \le \sqrt{d} \sum_{i=1}^{d} E\left[|S_{n,i}|^3\right], \quad \text{if } S_n = (S_n, \dots, S_{n,d}),$$

we infer from Lemma 3, Section 8, since $\rho_n \le \frac{1}{2}$,

$$E\left[|S_n|^3\right]^{2/3} \le C_{20} \Lambda_n \log 1/\rho_n.$$

Furthermore, we have $E[|T_n|^3]^{2/3} \le C_{21}\Lambda_n$, where C_{20}, C_{21} are positive constants depending on d only. This together with (4.2) implies the assertion. □

We prove Theorem 1 for all functions $H \in \mathscr{H}$, which satisfy

$$(4.3) \qquad t^{-3}H(t) \to 0, \quad \text{as } t \to \infty.$$

If (4.3) is not fulfilled, we have $\int |x|^3 Q(dx) < \infty$ and $a_1 n^{1/3} \le H^{-1}(n) \le a_2 n^{1/3}$ for some positive constants a_i, $i = 1, 2$. The assertion follows then from Theorem 5 [applied with $H(t) = t^3$].

Using the same argument as Major (1976a) [cf. [10] or Csörgő and Révész (1981), Lemmas 2.6.1 and 2.6.2], one can easily see that it suffices to obtain a construction such that

$$(4.4) \qquad S_n - T_n = O(H^{-1}(n)) \quad \text{a.s.}$$

Let now $\{\xi_k\}$ be a sequence of i.i.d. random vectors with $P \circ \xi_1 = Q$. W.l.o.g. we assume $\Sigma = I$. We set

$$\bar{\xi}_k = \xi_k 1_{\{|\xi_k| < H^{-1}(k)\}}, \qquad \tilde{\xi}_k = \bar{\xi}_k - E[\bar{\xi}_k], \qquad k \in \mathbb{N}.$$

Then we have

$$(4.5) \qquad \sum_1^n \xi_k - \sum_1^n \tilde{\xi}_k = o(H^{-1}(n)) \quad \text{a.s.}$$

The proof of (4.5) is similar to the proof of relations (2.3) and (2.4) in Major (1979) and will therefore be omitted. We set $\Sigma_n := \operatorname{cov}(\tilde{\xi}_n)$, $n \in \mathbb{N}$. Then it is easy to see that

$$\|\Sigma_n - I\| = o\left(\frac{H^{-1}(n)^2}{n}\right) \quad \text{as } n \to \infty.$$

We construct now a p-space $(\hat{\Omega}, \hat{\mathscr{A}}, \hat{P})$ and two sequences of independent random vectors $\{\hat{X}_n\}, \{\hat{Y}_n\}$ with $\hat{P} \circ \hat{X}_n = P \circ \tilde{\xi}_n$, $\hat{P} \circ \hat{Y}_n = N(0, \Sigma_n)$ such that the partial sums $\hat{S}_n := \sum_1^n \hat{X}_k$, $\hat{T}_n := \sum_1^n \hat{Y}_k$ fulfill

$$(4.6) \qquad \hat{S}_n - \hat{T}_n = O(H^{-1}(n)) \quad \text{a.s.}$$

From (4.5), (4.6) and Lemma A.1 in Berkes and Philipp (1979) we infer (4.4), hence the assertion.

To prove (4.6), we show

$$(4.7) \qquad \sum_{m=1}^{\infty} P\left\{ \max_{1 \le k \le 2^{m-1}} |\hat{S}_k^{(m)} - \hat{T}_k^{(m)}| \ge C_{22} H^{-1}(2^m) \right\} < \infty,$$

where C_{22} is a positive constant depending on d only,

$$\hat{S}_k^{(m)} := \hat{S}_{k-1+2^{m-1}}, \qquad \hat{T}_k^{(m)} := \hat{T}_{k-1+2^{m-1}}, \qquad k = 1, \dots, 2^{m-1},\ m \ge 1.$$

[Using the Borel–Cantelli lemma, we obtain from (4.7) for $2^{m-1} \le k < 2^m$,

$$\begin{aligned} |\hat{S}_k(\omega) - \hat{T}_k(\omega)| &\le K(\omega) + C_{22} \sum_1^m H^{-1}(2^s) \\ &\le K(\omega) + C_{22} \left(\sum_1^m (2^{s-m})^{1/4} \right) H^{-1}(2^m) \\ &\le K(\omega) + C_{22} \frac{\sqrt{2}}{1 - 2^{-1/4}} H^{-1}(k) \quad \text{a.s.,} \end{aligned}$$

since $t^{-4}H(t)$ is nonincreasing and $t^{-2}H(t)$ is nondecreasing.] For the following proof of (4.7) we need Theorem 7, and Lemma 4 of Section 8.

PROOF OF (4.7). (i) We denote by $C_{22},\dots,C_{26}$ constants depending on d only. We set $\beta_m := E[|\bar{\xi}_{2^m}|^3]$, $m \geq 1$. Lemma 4 shows in particular that $\beta_m/H^{-1}(2^m)$ is a null-sequence. Thus, we can define $i_m \in \{0,\dots,m-1\}$ for $m \geq m_0$, say, by

$$(4.8)\quad C_{23}H^{-1}(2^m)^2\left(\log\frac{H^{-1}(2^m)}{\beta_m}\right)^{-1} \leq 2^{i_m} < 2C_{23}H^{-1}(2^m)^2\left(\log\frac{H^{-1}(2^m)}{\beta_m}\right)^{-1},$$

where C_{23} is a sufficiently large chosen positive constant, which is determined by (4.11). Furthermore, we assume w.l.o.g. that $\|\Sigma_k - I\| \leq \frac{1}{2}$ for $k \geq 2^{m_0-1}$. Let now $m \geq m_0$ be fixed. We set

$$n := 2^{i_m},\ l := 2^{m-1}/n,\ \rho_j := \lambda_j^{-3/2}\sum_{(j-1)n+1}^{jn} E\left[|\tilde{\xi}_k^{(m)}|^3\right],$$

where λ_j (Λ_j) is the smallest (largest) eigenvalue of $\Gamma_j = \sum_{(j-1)n+1}^{jn}\operatorname{cov}(\tilde{\xi}_k^{(m)})$, $j = 1,\dots,l$, $\bar{\xi}_k^{(m)} := \bar{\xi}_{k-1+2^{m-1}}$, $1 \leq k \leq 2^{m-1}$. Then we have for $1 \leq j \leq l$,

$$(4.9)\qquad \frac{1}{2}C_{23}H^{-1}(2^m)^2\left(\log\frac{H^{-1}(2^m)}{\beta_m}\right)^{-1} \leq \frac{1}{2}n \leq \lambda_j \leq \Lambda_j \leq \frac{3}{2}n.$$

Since $E[|\tilde{\xi}_k^{(m)}|^3] \leq 8\mathrm{E}[|\bar{\xi}_k^{(m)}|^3] \leq 8\beta_m$, $1 \leq k \leq 2^{m-1}$, by the Hölder inequality, we infer from (4.9)

$$(4.10)\qquad \rho_j \leq 16\sqrt{2}\,C_{23}^{-1/2}\frac{\beta_m}{H^{-1}(2^m)}\left(\log\frac{H^{-1}(2^m)}{\beta_m}\right)^{1/2},\qquad j = 1,\dots,l.$$

If we have chosen C_{23} large enough, we obtain from (4.9) and (4.10)

$$(4.11)\qquad 2H^{-1}(2^m) \leq \frac{1}{22\sqrt{d}}\sqrt{\lambda_j \operatorname{Log} 1/\rho_j},\qquad j = 1,\dots,l.$$

Hence

$$(4.12)\qquad |\tilde{\xi}_k^{(m)}| \leq \frac{1}{22\sqrt{d}}\sqrt{\lambda_j \operatorname{Log} 1/\rho_j}\ \text{ a.s.},\qquad (j-1)n < k \leq jn.$$

(ii) Using Theorem 7, we obtain a p-space $(\Omega_0, \mathscr{A}_0, P_0)$ and independent random vectors $(U_j^{(m)}, V_j^{(m)})$, $j = 1,\dots,l$, such that

$$P_0 \circ U_j^{(m)} = P \circ \sum_{(j-1)n+1}^{jn} \tilde{\xi}_k^{(m)},\qquad P_0 \circ V_j^{(m)} = N\left(0, \operatorname{cov}\left(U_j^{(m)}\right)\right)$$

and

$$(4.13)\qquad E\left[|U_j^{(m)} - V_j^{(m)}|^2\right] \leq C_{24}n\left(\frac{\beta_m}{H^{-1}(2^m)}\right)^{4/3},\qquad j = 1,\dots,l.$$

[Use (4.9), (4.10) and (4.12).]

Using Lemma A.1 of Berkes and Philipp (1979) we obtain a p-space $(\hat{\Omega}, \hat{\mathscr{A}}, \hat{P})$ and two sequences of independent random vectors $\hat{X}_k^{(m)}$, $k = 1,\dots,2^{m-1}$, and $\hat{Y}_k^{(m)}$, $k = 1,\dots,2^{m-1}$, such that

$$\hat{P} \circ \hat{X}_k^{(m)} = P \circ \tilde{\xi}_k^{(m)}, \qquad \hat{P} \circ \hat{Y}_k^{(m)} = N\big(0, \operatorname{cov}\big(\tilde{\xi}_k^{(m)}\big)\big), \qquad k = 1,\dots,2^{m-1}$$

and

$$P_0 \circ U_j^{(m)} = \hat{P} \circ \sum_{(j-1)n+1}^{jn} \hat{X}_k^{(m)}, \qquad P_0 \circ V_j^{(m)} = \hat{P} \circ \sum_{(j-1)n+1}^{jn} \hat{Y}_k^{(m)}, \qquad j = 1,\dots,l. \tag{4.14}$$

(iii) If $\hat{S}_k^{(m)} = (\hat{S}_{k,1}^{(m)},\dots,\hat{S}_{k,d}^{(m)})$, $\hat{T}_k^{(m)} = (\hat{T}_{k,1}^{(m)},\dots,\hat{T}_{k,d}^{(m)})$, we have by Kolmogorov's inequality

$$\begin{aligned}\hat{P}\Big\{\max_{1\le j\le l}|\hat{S}_{jn}^{(m)} - \hat{T}_{jn}^{(m)}| \ge \sqrt{d}\,H^{-1}(2^m)\Big\} &\le \sum_{i=1}^{d}\hat{P}\Big\{\max_{1\le j\le l}|\hat{S}_{jn,i}^{(m)} - \hat{T}_{jn,i}^{(m)}| \ge H^{-1}(2^m)\Big\}\\ &\underset{(4.14)}{\le} \frac{1}{H^{-1}(2^m)^2}\sum_{1}^{l} E\big[|U_j^{(m)} - V_j^{(m)}|^2\big]\\ &\underset{(4.13)}{\le} C_{24}\frac{2^{m-1}}{H^{-1}(2^m)^2}\left(\frac{\beta_m}{H^{-1}(2^m)}\right)^{4/3}.\end{aligned}$$

(iv) Let now C_{25} be a sufficiently large chosen constant. From the first Lévy inequality, (4.8), (4.9) and (4.14) we infer for $1 \le j \le l$, $1 \le i \le d$,

$$\begin{aligned}&\hat{P}\Big\{\max_{(j-1)n<k\le jn}|\hat{S}_{k,i}^{(m)} - \hat{S}_{(j-1)n,i}^{(m)}| \ge C_{25}H^{-1}(2^m)\Big\}\\ &\quad\le 2\hat{P}\Big\{|\hat{S}_{jn,i}^{(m)} - \hat{S}_{(j-1)n,i}^{(m)}| \ge \frac{C_{25}}{2}H^{-1}(2^m)\Big\}\\ &\quad\le 2P_0\Big\{|U_{j,i}^{(m)} - V_{j,i}^{(m)}| \ge \frac{C_{25}}{4}H^{-1}(2^m)\Big\} + 2P_0\Big\{|V_{j,i}^{(m)}| \ge \frac{C_{25}}{4}H^{-1}(2^m)\Big\}.\end{aligned}$$

Thus we obtain from (4.13)

$$\begin{aligned}&\hat{P}\Big\{\max_{1\le j\le l}\max_{(j-1)n\le k\le jn}|\hat{S}_k^{(m)} - \hat{S}_{(j-1)n}^{(m)}| \ge C_{25}\sqrt{d}\,H^{-1}(2^m)\Big\}\\ &\quad\le 2C_{24}\frac{2^{m-1}}{H^{-1}(2^m)^2}\left(\frac{\beta_m}{H^{-1}(2^m)}\right)^{4/3} + 2\sum_{j=1}^{l}\sum_{i=1}^{d}P_0\Big\{|V_{j,i}^{(m)}| \ge \frac{C_{25}}{4}H^{-1}(2^m)\Big\}\\ &\quad\le C_{26}\frac{2^m}{H^{-1}(2^m)^2}\left(\frac{\beta_m}{H^{-1}(2^m)}\right)^{4/3}.\end{aligned} \tag{4.15}$$

Similarly, we obtain

$$\begin{aligned}&\hat{P}\Big\{\max_{1\le j\le l}\max_{(j-1)n\le k\le jn}|\hat{T}_k^{(m)} - \hat{T}_{(j-1)n}^{(m)}| \ge C_{25}\sqrt{d}\,H^{-1}(2^m)\Big\}\\ &\quad\le C_{26}\frac{2^m}{H^{-1}(2^m)^2}\left(\frac{\beta_m}{H^{-1}(2^m)}\right)^{4/3}.\end{aligned} \tag{4.16}$$

Using Lemma 4, Section 8, we easily get (4.7) from (iii), (4.15) and (4.16). □

5. Proof of Theorems 2 and 3. W.l.o.g. we assume $\Sigma = I$. From Theorem 1 we obtain a p-space $(\Omega, \mathscr{A}, P)$ and two sequences of independent random vectors $\{X_n\}, \{\tilde{Y}_n\}$ such that $P \circ X_n = Q$, $P \circ \tilde{Y}_n = N(0, \Sigma_n)$, $\|\Sigma_n - I\| = o(H^{-1}(n)^2/n)$, $n \in \mathbb{N}$, and

$$S_n - \tilde{T}_n = o\big(H^{-1}(n)\big) \quad \text{a.s.} \tag{5.1}$$

Since $\Sigma_n \to I$, we may w.l.o.g. assume that Σ_n is positive definite for all $n \in \mathbb{N}$.

We set $Y_n := \Sigma_n^{-1/2}\tilde{Y}_n$, $n \in \mathbb{N}$. It is obvious that $\{Y_n\}$ is a sequence of independent $N(0, I)$ distributed random vectors. Furthermore, $\{Y_n - \tilde{Y}_n\}$ is a sequence of independent random vectors such that

$$P \circ \big(Y_n - \tilde{Y}_n\big) = N\Big(0, \big(\Sigma_n^{1/2} - I\big)^2\Big), \qquad n \in \mathbb{N}. \tag{5.2}$$

Since $\|(\Sigma_n^{1/2} - I)^2\| \le \|\Sigma_n - I\|^2$, we obtain

$$\Big\|\big(\Sigma_n^{1/2} - I\big)^2\Big\| = o\left(\frac{H^{-1}(n)^4}{n^2}\right). \tag{5.3}$$

We denote by $T_{n,i}(\tilde{T}_{n,i})$ the ith component of $T_n(\tilde{T}_n)$, $i = 1, \dots, d$. Set $\hat{B}_{n,i} := E[(T_{n,i} - \tilde{T}_{n,i})^2]$, $i = 1, \dots, d$. From (5.3) it follows that

$$\hat{B}_{n,i} = o\left(\sum_1^n \frac{H^{-1}(k)^4}{k^2}\right) = o\left(\frac{H^{-1}(n)^4}{n}\right), \qquad i = 1, \dots, d. \tag{5.4}$$

[Hint: If we assume that $t^{-4+r}H(t)$ is nonincreasing for some $r > 0$, we have

$$\sum_{k=1}^n \frac{H^{-1}(k)^4}{k^2} \le \frac{H^{-1}(n)^4}{n^{4/(4-r)}} \sum_{k=1}^n k^{(2r-4)/(4-r)} = O\left(\frac{H^{-1}(n)^4}{n}\right).]$$

Using (5.4), we obtain from the a.s. stability criterion [cf. Loève (1977), 18.2.II.A]

$$\begin{aligned} T_{n,i} - \tilde{T}_{n,i} &= o\left(\frac{H^{-1}(n)^2}{\sqrt{n}}\sqrt{\log\log n}\right) \\ &= o\left(H^{-1}(n)\sqrt{\frac{\log\log n}{h(H^{-1}(n))}}\right) \quad \text{a.s.}, \qquad i = 1, \dots, d, \end{aligned} \tag{5.5}$$

where $h(t) := t^{-2}H(t)$, $t \ge 0$.

Theorem 2 follows immediately from (5.1) and (5.5).

To prove Theorem 3(a), we remark that $h(H^{-1}(n))/h(n) \to 1$ as $n \to \infty$, since $(\operatorname{Log}\operatorname{Log} t)^{-1}h(t)$ is nonincreasing.

To prove Theorem 3(b), it suffices to show for $\varepsilon > 0$,

$$P\Big\{\max_{1 \le k \le n} |T_{k,i} - \tilde{T}_{k,i}| \ge \varepsilon H^{-1}(n)\Big\} \to 0, \qquad i = 1, \dots, d. \tag{5.6}$$

[Notice (5.1).] Using Kolmogorov's inequality, this follows immediately from (5.4). □

6. Proof of Theorem 5. First we show that Theorem 1 remains valid, if H is a continuous, nonnegative function on $[0, \infty)$ such that $t^{-3}H(t)$ is nondecreasing and $t^{-4+r}H(t)$ is nonincreasing for some $r > 0$. We can use nearly the same arguments as in Section 4. Instead of Lemma 4 we use the relation

$$\sum_m \frac{2^m}{H^{-1}(2^m)^2}\left(\frac{\beta_m}{H^{-1}(2^m)}\right)^{2-r/2} < \infty,$$

which holds, since β_m remains bounded.

We replace the constant C_{23} by a positive constant $\hat{C}_1$ depending on d and r only. It is easy to see that (iii) can be estimated by

$$\hat{C}_2\frac{2^{m-1}}{H^{-1}(2^m)^2}\left(\frac{\beta_m}{H^{-1}(2^m)}\right)^{2-r/2},$$

if we have chosen $\hat{C}_1$ large enough. Similarly, (4.15) and (4.16) can be estimated by

$$\hat{C}_4\frac{2^{m}}{H^{-1}(2^m)^2}\left(\frac{\beta_m}{H^{-1}(2^m)}\right)^{2-r/2},$$

if we replace C_{25} by a sufficiently large constant $\hat{C}_3$ depending on r and d only.

Using the same arguments as in Section 5, we infer Theorem 5 from the modified Theorem 1. □

7. Proof of Theorem 4. The main tools of the proof are Theorem 1 and

PROPOSITION 2. *Let $\{Y_k\}$ and $\{\tilde{Y}_k\}$ be sequences of independent random variables such that $P \circ Y_k = N(0,1)$, $P \circ \tilde{Y}_k = N(0, \sigma_k^2)$, $\sigma_k^2 \uparrow 1$, $k \in \mathbb{N}$. Assume $(1 - \sigma_n^2)\log\log n \to \infty$. Then we have for the partial sums $T_n := \Sigma_1^n Y_k$, $\tilde{T}_n := \Sigma_1^n \tilde{Y}_k$, $n \in \mathbb{N}$,*

$$\limsup_n \frac{|\tilde{T}_{2^n} - T_{2^n}|}{\left(1 - \sigma_{2^{n+1}}^2\right)\sqrt{2^n \log n}} \geq \frac{1}{8} \quad a.s.$$

PROOF. Let $t \to \alpha(t)$, $t > 0$, be a continuous differentiable map such that

(7.1) $$\alpha(t) \to \infty, \quad \text{as } t \to \infty,$$

(7.2) $$t \to \alpha'(t), \qquad t > 0, \text{ is nonincreasing},$$

(7.3) $$0 \leq \alpha'(t) \leq 1/t, \qquad t > 0,$$

(7.4) $$\left(1 - \sigma_{2^n}^2\right)\log n/\alpha(n) \to \infty, \quad \text{as } n \to \infty.$$

We set $x_n := 2^{n/2}\sigma_{2^n}\bar{x}_n$, $y_n := 2^{n/2}\bar{y}_n$, where

$$\bar{x}_n := \left(\operatorname{Log} 1/\alpha'(n) + \alpha(n) - \tfrac{1}{2}\operatorname{Log}\operatorname{Log} 1/\alpha'(n)\right)^{1/2}$$

and

$$\bar{y}_n := \left(\operatorname{Log} 1/\alpha'(n) - \tfrac{1}{2}\operatorname{Log}\operatorname{Log} 1/\alpha'(n)\right)^{1/2}.$$

Then it is easy to see that

$$\sum_{n=1}^{\infty} P\big(|\tilde{T}_{2^n} - \tilde{T}_{2^{n-1}}| \ge x_n\big) < \infty \tag{7.5}$$

and

$$\sum_{n=1}^{\infty} P\big(|T_{2^n} - T_{2^{n-1}}| \ge y_n\big) = \infty. \tag{7.6}$$

Using the Borel–Cantelli lemma, we obtain from (7.5) and (7.6) that almost surely

$$\big|\big(T_{2^n} - \tilde{T}_{2^n}\big) - \big(T_{2^{n-1}} - \tilde{T}_{2^{n-1}}\big)\big| \ge y_n - x_n \quad \text{infinitely often.} \tag{7.7}$$

Since

$$\begin{aligned} y_n - x_n &= 2^{n/2}\bar{y}_n(1 - \sigma_{2^n}) - 2^{n/2}\sigma_{2^n}(\bar{x}_n - \bar{y}_n) \\ &\ge 2^{n/2}\bar{y}_n\left[\frac{1}{2}\big(1 - \sigma_{2^n}^2\big) - \frac{\bar{x}_n^2 - \bar{y}_n^2}{\bar{y}_n^2}\right] \\ &\underset{(7.3)}{\ge} 2^{(n-1)/2}\sqrt{\log n}\left[\frac{1}{2}\big(1 - \sigma_{2^n}^2\big) - \frac{2\alpha(n)}{\log n}\right] \\ &\ge \frac{1}{4}\big(1 - \sigma_{2^n}^2\big)\sqrt{2^n \log n}, \end{aligned}$$

for sufficiently large n by (7.4), we easily obtain the assertion from (7.7). □

To simplify our notation, we set $h_\alpha(t) := (\operatorname{Log}\operatorname{Log} t)^\alpha$, $t \ge 0$. Let now a p-measure $Q_\alpha|\mathbb{B}$ be given such that

$$\int x Q_\alpha(dx) = 0, \qquad \int x^2 Q_\alpha(dx) = 1, \qquad \int H_\alpha(|x|) Q_\alpha(dx) < \infty \tag{7.8}$$

and

$$\int_{\{|x| \ge H_\alpha^{-1}(n)\}} x^2 Q_\alpha(dx) = (\operatorname{Log}\operatorname{Log}\operatorname{Log} n)^{-2} h_\alpha(n)^{-1} := \gamma_{\alpha, n}, \qquad n \ge 0. \tag{7.9}$$

[We set $p_{\alpha, n} := (\gamma_{\alpha, n} - \gamma_{\alpha, n+1})/H_\alpha^{-1}(n)^2$, $n \ge 1$. It is easy to see that $\sum_{n=1}^{\infty} p_{\alpha, n} \le 1$. Thus we can define a p-measure $Q_\alpha|\mathbb{B}$ such that $Q_\alpha\{H_\alpha^{-1}(n)\} = Q_\alpha\{-H_\alpha^{-1}(n)\} = \frac{1}{2}p_{\alpha, n}$, $n \in \mathbb{N}$, and $Q_\alpha\{0\} = 1 - \sum_{n=1}^{\infty} p_{\alpha, n}$. From the definition it follows immediately that $\int x Q_\alpha(dx) = 0$, $\int x^2 Q_\alpha(dx) = 1$ and $\int_{\{|x| \ge H_\alpha^{-1}(n)\}} x^2 Q_\alpha(dx) = \gamma_{\alpha, n}$, $n \ge 1$. Furthermore, we have

$$\int H_\alpha(|x|) Q_\alpha(dx) = \sum_{n=1}^{\infty} p_{\alpha, n} n = \sum_{n=2}^{\infty} \gamma_{\alpha, n}\big(h_\alpha\big(H_\alpha^{-1}(n)\big) - h_\alpha\big(H_\alpha^{-1}(n-1)\big)\big).$$

It is easy to see that the last series is finite.]

Let now $\{X_n\}$ and $\{Y_n\}$ be sequences of i.i.d.r.v.'s such that $P \circ X_1 = Q_\alpha$, $P \circ Y_1 = N(0,1)$. From Theorem 1 [cf. (4.6)] we obtain a p-space $(\Omega', \mathscr{A}', P')$ and two sequences of independent random variables $\{X_n'\}$ and $\{\tilde{Y}_n\}$ such that

$$P' \circ X_n' = P \circ X_n, \qquad P' \circ \tilde{Y}_n = N\big(0, \sigma_n^2\big), \qquad n \in \mathbb{N}, \tag{7.10}$$

where

$$\sigma_n^2 = \int_{\{|x|<H_\alpha^{-1}(n)\}} x^2 Q_\alpha(dx) - \left(\int_{\{|x|<H_\alpha^{-1}(n)\}} x Q_\alpha(dx)\right)^2$$

and

$$S_n' - \tilde{T}_n = O\big(H_\alpha^{-1}(n)\big) \quad \text{a.s.} \tag{7.11}$$

Using Lemma A.1 in Berkes and Philipp (1979) we may w.l.o.g. assume that $(\Omega', \mathscr{A}', P') = (\Omega, \mathscr{A}, P)$ and $X_n' = X_n$, $n \in \mathbb{N}$. Thus we can infer from (7.11)

$$\limsup_n \frac{|S_n - T_n|}{\sqrt{n}\,(\operatorname{Log}\operatorname{Log} n)^\beta} = \limsup_n \frac{|T_n - \tilde{T}_n|}{\sqrt{n}\,(\operatorname{Log}\operatorname{Log} n)^\beta} \quad \text{a.s.,} \qquad \text{if } \beta < \frac{1}{2} - \alpha. \tag{7.12}$$

Using (7.9) and (7.12), we can conclude the proof by an application of Proposition 2. □

8. Lemmas.

LEMMA 1. *Let* $g: \mathbb{R}^d \to \mathbb{R}^d$ *be a continuous differentiable map with* $g(0) = 0$. *Assume that the Jacobi matrix* $J_h g$ *fulfills*

$$\|J_h g - I\| \le \tfrac{1}{3}, \quad \text{for } h \in U_r(0) := \{h: |h| \le r\}.$$

Then

(a) $g|_{U_r(0)}$ *is injective,*
(b) $g(U_r(0)) \supseteq U_{2/3r}(0)$.

PROOF. Cf. Edwards (1973), Lemma 3.2, Chapter 3. □

The following lemma refines Lemma 10 of Yurinskii (1977).

LEMMA 2. *Let* $G_n|\mathbb{B}$ *be a p-measure with Lebesgue density* $x \to g_n(x)$. *Assume* $\exp(-\alpha_n) \le g_n(x)/\varphi_{\sigma_0^2}(x) \le \exp(\alpha_n)$ *for* $|x| \le b_n$, *where* $\alpha_n \le 1/\sqrt{2\pi}\,e^4$ *and* $b_n \ge 2\sigma_0$. *Then we have for all closed sets* $A \subseteq [-b_n, b_n]$,

$$G_n(A) \le N(0, \sigma_0^2)(A^{\bar{\alpha}_n}) + 2\left(1 - \Phi\left(\frac{b_n}{\sigma_0}\right)\right),$$

where $\bar{\alpha}_n := C_{27}\sigma_0\alpha_n$ *with an absolute constant* C_{27}.

PROOF. We use arguments similar to those in Yurinskii (1977). W.l.o.g. we assume $\sigma_0 = 1$. Let first $A = [a, b]$ with $1 + \alpha_n/2 \le a < b \le b_n$. Then we have for $x \in [a, b]$,

$$g_n(x) \le \varphi(x - \alpha_n) \quad \text{and} \quad \varphi(x) \le g(x - \alpha_n).$$

Integrating these inequalities over $[a, b]$, we obtain

(8.1) $G_n[a, b] \le N(0,1)[a - \alpha_n, b]$ and $N(0,1)[a, b] \le G_n[a - \alpha_n, b]$.

If $0 \le a \le 1 + \alpha_n/2$, $b \le b_n$, we have

$$G_n\left[a - \sqrt{2\pi}\, e^4\alpha_n, a\right] \ge \sqrt{2\pi}\, e^4\alpha_n \min_{|x|\le 2} g_n(x) \ge \alpha_n e.$$

Since

$$N(0,1)[a, b] \le G_n[a, b]\exp(\alpha_n) \le G_n[a, b] + \alpha_n e,$$

we infer

$$N(0,1)[a, b] \le G_n\left[a - \sqrt{2\pi}\, e^4\alpha_n, b\right]. \tag{8.2}$$

Similarly, we obtain

$$G_n[a, b] \le N(0,1)\left[a - \sqrt{2\pi}\, e^4\alpha_n, b\right]. \tag{8.3}$$

From (8.1)–(8.3) it follows that

$$\begin{aligned} &G_n[a, b] \le N(0,1)\left[a - \sqrt{2\pi}\, e^4\alpha_n, b\right] \quad \text{and} \\ &N(0,1)[a, b] \le G_n\left[1 - \sqrt{2\pi}\, e^4\alpha_n, b\right], \end{aligned} \tag{8.4}$$

for all intervals $[a, b] \subseteq [0, b_n]$. Similarly,

$$\begin{aligned} &G_n[-b, -a] \le N(0,1)\left[-b, -a + \sqrt{2\pi}\, e^4\alpha_n\right] \quad \text{and} \\ &N(0,1)[-b, -a] \le G_n\left[-b, -a + \sqrt{2\pi}\, e^4\alpha_n\right], \end{aligned} \tag{8.5}$$

for all intervals $[-b, -a] \subseteq [-b_n, 0]$.

Using the same arguments as in Lemma 10, Yurinskii (1977), we first infer from (8.4) and (8.5) that for all intervals $A \subseteq [-b_n, b_n]$ with $0 \in A$,

$$G_n(A) \le N(0,1)\left(A^{\sqrt{2\pi}e^4\alpha_n}\right) + 2(1 - \Phi(b_n)), \tag{8.6}$$

and obtain the assertion finally from (8.4)–(8.6). □

LEMMA 3. *Let ξ_k, $k = 1, \dots, n$, be independent random variables with zero means. Assume $|\xi_k| \le K_n$ a.s., $k = 1, \dots, n$. Then we have*

$$E\left[\left|\sum_1^n \xi_k\right|^3\right] \le 12\sqrt{\pi}\, B_n^{3/2} + 384 K_n^3,$$

where $B_n := \sum_1^n E[\xi_k^2]$.

PROOF. By means of partial integration we obtain

$$E\left[\left|\sum_1^n \xi_k\right|^3\right] = \int_0^\infty 3x^2 P\left(\left|\sum_1^n \xi_k\right| \ge x\right) dx.$$

Using the exponential bound 19.1.A, Loève (1977), we infer

$$E\left[\left|\sum_1^n \xi_k\right|^3\right] \le 6\int_0^\infty x^2\left(\exp\left(-\frac{x^2}{4B_n}\right) + \exp\left(-\frac{x}{4K_n}\right)\right)dx$$
$$= 12\sqrt{\pi}\,B_n^{3/2} + 384K_n^3. \qquad \square$$

Lemma 4. *Let $\xi\colon \Omega \to \mathbb{R}^d$ be a random vector such that $E[H(|\xi|)] < \infty$ for some $H \in \mathscr{H}$. Then we have for all $\eta > 0$,*

$$\sum_{m=1}^\infty \frac{2^m}{H^{-1}(2^m)^2}\left(\frac{\beta_m}{H^{-1}(2^m)}\right)^{1+\eta} < \infty, \quad \textit{where } \beta_m = E\left[|\xi|^3 1_{\{|\xi| \le H^{-1}(2^m)\}}\right],$$
$$m \ge 1.$$

Proof. It suffices to prove the lemma for $\eta < 1$. Since $t^{-3}H(t)$ is nonincreasing, we have

$$\beta_m \le \frac{H^{-1}(2^m)^3}{2^m} E[H(|\xi|)].$$

Thus it suffices to show that

$$\sum_m \left(\frac{2^m}{H^{-1}(2^m)^2}\right)^{1-\eta}\frac{\beta_m}{H^{-1}(2^m)} < \infty. \tag{8.7}$$

We set

$$p_k := E\left[H(|\xi|)1_{\{H^{-1}(2^{k-1}) \le |\xi| < H^{-1}(2^k)\}}\right], \qquad k \ge 2,$$
$$p_1 := E\left[H(|\xi|)1_{\{|\xi| < H^{-1}(2)\}}\right].$$

Since $H \in \mathscr{H}$, we obtain

$$\sum_{m=1}^\infty \left(\frac{2^m}{H^{-1}(2^m)^2}\right)^{1-\eta}\frac{\beta_m}{H^{-1}(2^m)}$$
$$\le \sum_{m=1}^\infty \left(\frac{2^m}{H^{-1}(2^m)^2}\right)^{1-\eta}\frac{1}{H^{-1}(2^m)}\sum_{k=1}^m \frac{H^{-1}(2^k)^3}{2^k}p_k$$
$$= \sum_{k=1}^\infty \left(\sum_{m=k}^\infty \left(\frac{2^m}{H^{-1}(2^m)^3}\right)^{1-2/3\eta} 2^{-m\eta/3}\right)\frac{H^{-1}(2^k)^3}{2^k}p_k$$
$$\le \sum_{k=1}^\infty \left(\sum_{m=k}^\infty 2^{-m\eta/3}\right)\left(\frac{H^{-1}(2^k)^3}{2^k}\right)^{2/3\eta}p_k$$
$$\le \sum_{k=1}^\infty \left(\sum_{m=k}^\infty 2^{-m\eta/3}\right)H^{-1}(1)^{2\eta}2^{k\eta/3}p_k = H^{-1}(1)^{2\eta}\frac{E[H(|\xi|)]}{1-2^{-\eta/3}} < \infty. \qquad \square$$

Acknowledgments. I would like to thank my thesis advisor, D. Landers, for encouraging me to work in this area and for supporting me during the preparation. I also want to thank N. Herrndorf for many useful discussions.

REFERENCES

[1] BERGER, E. (1982). Fast sichere Approximation von Partialsummen unabhängiger und stationärer ergodischer Folgen von Zufallsvektoren. Dissertation, Universität Göttingen.
[2] BERKES, I. and PHILIPP W. (1979). Approximation theorems for independent and weakly dependent random vectors. *Ann. Probab.* **7** 29–54.
[3] BHATTACHARYA, R. N. and RAO, R. R. (1976). *Normal Approximation and Asymptotic Expansions*. Wiley, New York.
[4] BREIMAN, L. (1967). On the tail behaviour of sums of independent random variables. *Z. Wahrsch. verw. Gebiete* **9** 20–24.
[5] CSÖRGŐ, M. and RÉVÉSZ, P. (1981). *Strong Approximations in Probability and Statistics*. Academic, New York.
[6] DUDLEY, R. M. (1968). Distances of probability measures and random variables. *Ann. Math. Statist.* **39** 1563–1572.
[7] EDWARDS, C. H., JR. (1973). *Advanced Calculus of Several Variables*. Academic, New York.
[8] KOMLÓS, J., MAJOR, P. and TUSNÁDY, G. (1976). An approximation of partial sums of independent rv's and the sample df. II. *Z. Wahrsch. verw. Gebiete* **34** 33–58.
[9] LOÈVE, M. (1977). *Probability Theory* **1**, 4th ed. Springer, New York.
[10] MAJOR, P. (1976a). The approximation of partial sums of independent r.v.'s. *Z. Wahrsch. verw. Gebiete* **35** 213–220.
[11] MAJOR, P. (1976b). Approximation of partial sums of i.i.d.r.v.'s when the summands have only two moments. *Z. Wahrsch. verw. Gebiete* **35** 221–229.
[12] MAJOR, P. (1979). An improvement of Strassen's invariance principle. *Ann. Probab.* **7** 55–61.
[13] PHILIPP, W. (1979). Almost sure invariance principles for sums of B-valued random variables. *Probability in Banach spaces II. Lecture Notes in Math.* **709** 171–193. Springer, Berlin.
[14] PHILIPP, W. (1980). Weak and L^p-invariance principles for sums of B-valued random variables. *Ann. Probab.* **8** 68–82; correction **14** (1986) 1095–1101.
[15] STRASSEN, V. (1964). An invariance principle for the law of the iterated logarithm. *Z. Wahrsch. verw. Gebiete* **3** 211–226.
[16] YURINSKII, V. V. (1975). A smoothing inequality for the estimates of the Lévy–Prokhorov distance. *Theory Probab. Appl.* **20** 1–10.
[17] YURINSKII, V. V. (1977). On the approximation of convolutions by normal laws. *Theory Probab. Appl.* **22** 653–667.

MATHEMATICAL INSTITUTE
UNIVERSITY OF COLOGNE
WEYERTAL 86-90
D-5000 COLOGNE 41
WEST GERMANY