BERRY–ESSEEN BOUNDS FOR STATISTICS OF WEAKLY DEPENDENT SAMPLES

V. Bentkus[1] F. Götze[1] A. Tikhomirov[1,2]

University of Bielefeld
Syktyvkar University and the Mathematical Department of Komi RC of the Russian Academy of Sciences

April 1995

Abstract. We prove Berry–Esseen bounds for a general class of asymptotically normal statistics which are functions of N weakly dependent random variables under easily verifiable conditions. In particular, we show the validity of the bound $\mathcal{O}(N^{-1/2}\log^{\delta} N)$ with some $\delta > 0$, for U-statistics, Studentized means, functions of sample means, functionals of empirical distribution functions and linear combinations of order statistics

1. Introduction and results

Let $X_1, X_2, \dots$ be a sequence of random variables taking values in arbitrary measurable space $(\mathcal{X}, \mathcal{A})$ which is *stationary in the strong sense.* We shall assume that the sequence satisfies an absolute regularity condition with coefficients $\beta(m) \to 0$, as $m \to \infty$,

$$\beta(m) \overset{\text{def}}{=} \sup_{k \geqslant 1} \mathbf{E} \sup \Big\{ \big| \mathbf{P}\{A \,\big|\, \sigma[1,k]\} - \mathbf{P}\{A\} \big| : \ A \in \sigma[k+m, \infty) \Big\},$$

where $\sigma[a,b]$ denotes the σ-algebra generated by the random variables X_l with $l \in [a,b]$.

The aim of this paper is to prove Berry–Esseen bounds for a sufficiently large class of statistics of weakly dependent random variables. Let $t = t_N$ be a real-valued

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

Key words and phrases. weakly dependent random variables, mixing, absolute regularity, Berry-Esseen bounds, asymptotically normal statistics, Studentized means, U-statistics, functions of sample means, functionals of the empirical distribution functions, linear combinations of order statistics.

[1] Research supported by the SFB 343 in Bielefeld.

[2] Research supported by Russian Foundation for Fundamental Research. Grant N93–011–1454.

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

function of N variables. We shall consider statistics $T = t(X_1, \dots, X_N)$ which can be represented as

$$T = S + R, \qquad \text{where} \quad S = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} g(X_j), \tag{1.1}$$

with some function $g : \mathcal{X} \to \mathbf{R}$ such that $\mathbf{E}\, g(X_1) = 0$ and some remainder term $R = R_N(X_1, \dots, X_N)$. The function g may depend on N; throughout we allow such dependence without explicitly mentioning it.

Let

$$\sigma_N^2 \overset{\text{def}}{=} \mathbf{E}\, S^2 = \mathbf{E}\, g^2(X_1) + 2 \sum_{j=1}^{N-1} (1 - jN^{-1})\, \mathbf{E}\, g(X_1)\, g(X_{1+j})$$

denote the variance of S, and let $\rho_s = \mathbf{E}\, |g(X_1)|^s$.

In the case of independent identically distributed (i.i.d.) random variables X_1, $X_2, \dots$ the linear part S of the statistic T is asymptotically normal as $N \to \infty$,

$$\mathbf{P}\{S < \sigma_N x\} \to \Phi(x), \qquad \text{where} \quad \Phi(x) \overset{\text{def}}{=} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\{-\frac{t^2}{2}\} dt,$$

provided that

$$\sigma_N^2 \geq \Sigma^2 > 0, \qquad \text{for all} \quad N, \tag{1.2}$$

and $\sup_N \rho_{2+\delta} < \infty$, for some fixed $\Sigma > 0$ and $\delta > 0$. In the weakly dependent case the additional condition $\sup_m m^{(1+\varepsilon)(1+1/\delta)} \beta(m) < \infty$, for some $\varepsilon > 0$, ensures the same result (see, Ibragimov and Linnik (1971), Ibragimov (1975), Eberlein (1984)). Thus the statistic T will be asymptotically normal provided that in addition $R \to 0$ in probability.

In order to prove Berry-Esseen bounds some stronger conditions are necessary. It is known (Tikhomirov (1980)) that the conditions

$$\sup_N \rho_3 \leq \rho < \infty \tag{1.3}$$

and

$$\beta(m) \leqslant K \exp\{-\beta m\}, \qquad \text{for all} \quad m \geqslant 1, \tag{1.4}$$

with some $K < \infty$ and $\beta > 0$, together imply

$$\sup_x \left| \mathbf{P}\{S < \sigma_N x\} - \Phi(x) \right| \leqslant A N^{-1/2} \log^2 N$$

with a constant A depending on K, β, Σ and ρ only. We shall extend this estimate for general nonlinear statistics.

Let $\sigma^c[j,k]$ denote the σ-algebra generated by X_l such that $l \notin [j,k]$ and $1 \leq l \leq N$. In the case $k < j$, set $\sigma^c[j,k] = \sigma(X_1, \dots, X_N)$.

Theorem 1.1. *Assume that (1.2) – (1.4) hold. Let*

$$R_{j,k} = R_{j,k}(X_1, \dots, X_{j-1}, X_{k+1}, \dots X_N)$$

denote any $\sigma^c[j,k]$-measurable random variables such that $R_{j,j-1} = R$, for all $1 \leq j \leq N$, and let

$$\gamma \stackrel{\mathrm{def}}{=} \max\{\, \mathbf{E}^{\,2/3} \big| R_{j,k} - R_{j,k-1} \big|^{3/2} : \ \ |j-k| \leq \log^3 N \ \ and \ \ 1 \leqslant j \leqslant k \leqslant N\}.$$

Then

$$\sup_x \ \Big| \mathbf{P}\,\{T < \sigma_N x\} - \Phi(x) \Big| \leqslant A\,N^{-1/2}\log^2 N + A\,\mathbf{E}\,|R| + A\gamma\sqrt{N}\log^3 N$$

with a constant A depending only on K, β, Σ, and ρ. If the function g is independent of N and

$$\sigma^2 \stackrel{\mathrm{def}}{=} \lim_{N\to\infty} \sigma_N^2 = \mathbf{E}\,g^2(X_1) + 2\sum_{j=1}^{\infty} \mathbf{E}\,g(X_1)\,g(X_{1+j}) > 0$$

then $\sup_x \Big| \mathbf{P}\,\{T < \sigma x\} - \Phi(x) \Big|$ satisfies the same estimate with a constant A depending only on K, β, σ, and ρ.

For the proof of the result we apply a modification of the method used by Stein (1972) and Tikhomirov (1980) and others. This result seems to be the first Berry–Esseen bound for a general class of statistics of dependent samples. In Section 2 we apply Theorem 1.1 for functions of sample means, functionals of the empirical distribution functions, for Studentized means, for U-statistics and for linear combinations of order statistics. In all these applications the estimation of $\mathbf{E}\,|R|$ and γ in Theorem 1.1 is quite simple and reduces to the estimation of certain moments. The random variables $R_{j,k}$ may be obtained by the simple rule: "remove all terms of R involving random variables X_l with l such that $j \leq l \leq k$".

In the literature only special classes of (nonlinear) statistics of weakly dependent samples were considered so far: Yoshihara (1976) proved the asymptotic normality for a class of U-statistics; Denker (1982), Denker and Keller (1983) proved the asymptotic normality and functional limit theorems for a classes of U-statistics and von Mises statistics and obtained Berry-Esseen bounds; Yoshihara (1984) obtained a Berry–Esseen bound for U-statistics; for results concerning sums see, for example, a review of Sunklodas (1991). Edgeworth expansions for statistics of dependent samples were considered by Götze and Hipp (1983, 1992).

Our result is similar to a Berry–Esseen bound of van Zwet (1984) for symmetric statistics of independent samples. In the independent case more precise estimates are known (see, for example, Friedrich (1989), Bolthausen and Götze (1993), Bentkus and Götze (1994), for lower estimates see Bentkus, Götze and Zitikis (1994). We could improve the moment conditions for the non-linear part of the statistic combining methods developed for independent random variables and those of the present paper, but detailed proofs would require large amount of routine work. Furthermore, the dependence on $\log N$ in Theorem 1.1 may be improved.

The mixing condition (1.4) is relatively weak. For example, solutions of Ito stochastic equations in Euclidean spaces satisfy this condition and do not satisfy conditions with other stronger (e.g., uniformly) mixing coefficients, see Veretennikov (1987). Heinrich (1992) found a clear and simple condition yielding (1.4) for stationary renewal processes. The condition (1.4) is fulfilled as well when it is is fulfilled for ψ-mixing coefficients $\psi(m)$ or for φ-mixing coefficients $\varphi(m)$ since $\beta(m) \le \varphi(m) \le \psi(m)$.

Notation. By $A, A_1, A_2, \dots$ we shall denote generic constants which may depend on parameters of interest, like K, β, ρ, α, ... By m we shall denote a natural number such that for a sufficiently large generic constant A, $m \approx A \log N$, for example, $m = [A \log N]$.

By $\tau, \tau_1, \tau_2, \dots$ we shall denote a sequence of i.i.d. r.v. uniformly distributed on $[0,1]$, and independent of all other r.v. By $\mathbf{E}_\tau$ we shall denote the conditional expectation given all random variables but τ.

Let $\hat{\xi}$ denote an independent copy of random variable ξ, which is also independent of all other random variables appearing in the specific context.

We shall often use the following simple inequality. Assume that ξ is $\sigma(-\infty, k]$-measurable and that η is $\sigma[k+m, \infty)$-measurable and that random variables ξ and η take values in a Polish space. If $\sup_{u,v} |\varphi(u,v)| \le D$ then

$$\Big| \mathbf{E}\, \varphi(\xi, \eta) - \mathbf{E}\, \varphi(\xi, \widehat{\eta}) \Big| \leqslant D\beta(m). \tag{1.5}$$

2. Applications

In this Section we shall apply Theorem 1.1 to various special statistics which shows that in fact the conditions can be easily verified. Similar examples were considered in the independent case by Bentkus, Götze and van Zwet (1994).

2.1 Functions of sample means. Assume for this subsection only that $X_1, X_2, \dots$ take values in a real separable Banach space $\mathcal{B}$. Consider the statistic

$$T = \sqrt{N}\,\big(H(\overline{X}) - H(0)\big)$$

for a function $H : \mathcal{B} \to \mathbf{R}$, where the sample mean $\overline{X} \overset{\text{def}}{=} N^{-1}(X_1 + \cdots + X_N)$.

We shall use Taylor's expansion for sufficiently smooth functions $f : \mathcal{B} \to \mathbf{R}$ (see, e.g. Cartan (1971)),

$$f(x+h) = f(0) + f'(x)h + \cdots + \frac{1}{k!} f^{(k)}(x)h^k + \frac{1}{k!} \mathbf{E}_\tau (1-\tau)^k f^{(k+1)}(x+\tau h)h^{k+1},$$

where $f^{(j)}(x)h^j$ denotes the j-th Frechet derivative of f at point x in the direction h.

Assuming that H is Frechet differentiable and using Taylor's expansion we may write $g(x) \stackrel{\text{def}}{=} H'(0)x$, and

$$R \stackrel{\text{def}}{=} \sqrt{N}\,(H(\overline{X}) - H(0) - H'(0)\overline{X}) = \sqrt{N}\,\mathbf{E}_\tau (1-\tau) H''(\tau \overline{X})\overline{X}^2.$$

The function g is independent of N, and

$$T = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} g(X_j) + R.$$

Denote

$$M_s \stackrel{\text{def}}{=} \sum_{j=1}^{s} \sup_{x \in B} \left\| H^{(j)}(x) \right\|.$$

Theorem 2.1. *Assume that the mixing condition (1.4) is fulfilled. Let*

$$\mathbf{E}\, X_1 = 0, \qquad \rho \stackrel{\text{def}}{=} \mathbf{E}\, \|X_1\|^3 < \infty.$$

Assume that $\sigma^2 > 0$. If $M_3 < \infty$ then there exists a constant A depending only on $\mathcal{B}$, K, β, ρ, M_3 and σ such that

$$\delta_N \stackrel{\text{def}}{=} \sup_x \left| \mathbf{P}\left\{ T < \sigma x \right\} - \Phi(x) \right| \leqslant C N^{-1/2} \log^4 N.$$

If the Banach space $\mathcal{B}$ is of type 2 then the smoothness condition $M_3 < \infty$ may be relaxed to $M_2 < \infty$.

A Banach space $\mathcal{B}$ is of type 2 if there exists a constant $C = C(\mathcal{B})$ such that $\mathbf{E}\, \| \sum_{i=1}^N Y_i \|^2 \le C \sum_{i=1}^N \mathbf{E}\, \|Y_i\|^2$, for any independent centered random variables Y_i. Finite dimensional spaces, Hilbert spaces and L_p, l_p, $2 \le p < \infty$, are of type 2.

Proof of Theorem 2.1. We shall derive the result from Theorem 1.1.

The verification of $|\frac{\sigma_N}{\sigma} - 1| \le \frac{A}{N}$ is easy. Therefore it is sufficient to prove that

$$\mathbf{E}\,|R| \le A m N^{-1/2} \tag{2.1}$$

and that

$$\mathbf{E}\,|R_{j,k} - R_{j,k-1}|^{3/2} \le A(m/N)^{3/2}, \tag{2.2}$$

where

$$R_{j,k} \overset{\text{def}}{=} \sqrt{N}\,(H(\overline{X}_{[j,k]}) - H(0) - H'(0)\overline{X}_{[j,k]}), \quad \text{and} \quad \overline{X}_{[j,k]} \overset{\text{def}}{=} \overline{X} - N^{-1} \sum_{i \in [j,k]} X_i.$$

While proving (2.1) and (2.2) we may assume that $\|X_j\| \le N^2$ with probability 1. Otherwise we may replace X_j by $X_j \mathbb{I}\{\|X_j\| \le N^2\} - \mathbf{E}\, X_j \mathbb{I}\{\|X_j\| \le N^2\}$. For example, in the case of (2.1), by such a replacement the error is bounded from above by

$$AN^{-1/2} \sum_{j=1}^{N} \mathbf{E}\,\|X_j\| \mathbb{I}\{\|X_j\| \ge N^2\} = A\sqrt{N}\mathbf{E}\,\|X_1\| \mathbb{I}\{\|X_1\| \ge N^2\} \le AN^{-7/2}.$$

Indeed, we may split $X_j = X_j \mathbb{I}\{\|X_j\| \le N^2\} + X_j \mathbb{I}\{\|X_j\| > N^2\}$, use the representation $R = \sqrt{N}(H(\overline{X}) - H(0) - H'(0)\overline{X})$, apply $|H(u) - H(v)| \le M_1\|u - v\|$ and use the triangle inequality.

Let us prove (2.1). Note that $|R| \le M_2\sqrt{N}\|\overline{X}\|^2$. Thus in the case of the Banach space $\mathcal{B}$ of type 2, the bound (2.1) is a consequence of Lemma 3.1. The proof of (2.1) without the type 2 assumption is slightly more complicated, see Section 3.

Let us prove (2.2). Notice that $\overline{X}_{[j,k-1]} = \overline{X}_{[j,k]} + N^{-1}X_k$ and that

$$|R_{j,k} - R_{j,k-1}| = \sqrt{N}|H(\overline{X}_{[j,k-1]}) - H(\overline{X}_{[j,k]}) - H'(0)N^{-1}X_k|.$$

Expanding in powers of $N^{-1}X_k$ we see that instead of (2.2) it is sufficient to show that

$$\mathbf{E}\,|H'(\overline{X}_{[j,k]})X_k - H'(0)X_k|^{3/2} \le Am^{3/2}N^{-3/4}. \tag{2.3}$$

Let us write $\overline{X}_{[j,k]} = U + \Delta$, where U is the sum of terms $N^{-1}X_l$ in the sum $\overline{X}_{[j,k]}$ such that $|l - k| > m$, and where Δ denotes the remaining part of the sum $\overline{X}_{[j,k]}$. Expanding in powers of Δ and using $(a_1 + \cdots + a_m)^{3/2} \le \sqrt{m}(a_1^{3/2} + \cdots + a_m^{3/2})$, we reduce (2.3) to

$$\mathbf{E}\,|H(U)X_k - H'(0)X_k|^{3/2} \le Am^{3/2}N^{-3/4} \tag{2.4}$$

Using (1.5) we may replace X_k in (2.4) by its independent copy $\widehat{X}_k$. An expansion in powers of U and an application of the Hölder inequality show that (2.4) follows from

$$\mathbf{E}\,(H''(\tau U)U\widehat{X}_k)^2 \le Am^2N^{-1}. \tag{2.5}$$

If the Banach space is of type 2 then we may apply Lemma 2.3 in order to estimate $\mathbf{E}\,\|U\|^2$, and (2.5) implies the result. Let us continue the proof of (2.2) for arbitrary Banach spaces. Expanding the square in (2.5) we see that (2.5) is a corollary of

$$N^{-1}\sum_i\sum_l |\mathbf{E}\,H''(\tau U)X_i\hat{X}_k H''(\tau U)X_l\hat{X}_k| \le Am^2, \tag{2.6}$$

where the sums $\sum_i$ and $\sum_l$ are taken over the indices i and l in the sum U. By the triangle inequality we may remove in (2.6) summands with $|i-l|\le m$. Next we may remove from the sum U summands with indices near to i or l. An application of (1.5) will complete the proof.

2.2 Functionals of empirical distribution functions. The results of this subsection follow straightforwardly from those for functions of sample means.

Throughout this subsection we shall use the following notation. By $\eta_1,\eta_2,\dots$ we denote a sequence of real random variables stationary in the strong sense. Let F denote the distribution function of η_1, and let F_N be the empirical distribution function corresponding to the sample $\eta_1,\dots,\eta_N$. Define the random processes

$$X_i(t),\quad t\in\mathbf{R},\quad 1\le i\le N,\quad \text{by}\quad X_i(t)\overset{\text{def}}{=}\mathbb{I}\{\eta_i<t\}-F(t).$$

Finally let

$$E_N\overset{\text{def}}{=}\sqrt{N}(F_N-F)=(X_1+\dots+X_N)/\sqrt{N}$$

denote the empirical process.

Assume that a functional H takes real values and that $H(F)$ and $H(F_N)$ are well defined. Define the statistic

$$T\overset{\text{def}}{=}\sqrt{N}(H(F_N)-H(F)).$$

We may write $F_N-F=E_N/\sqrt{N}=\overline{X}$. Introducing the functional

$$G_F(h)\overset{\text{def}}{=}H(F+h)\qquad\text{we have}\qquad T=\sqrt{N}(G_F(\overline{X})-G_F(0)).$$

Let us define derivatives of H via derivatives of G_F as $H^{(s)}(F+x)\overset{\text{def}}{=}G_F^{(s)}(x)$. In order to define derivatives of G_F introduce a Banach space $\mathcal{B}$, which may depend on F and should be chosen in dependence on H and the particular problem. We shall assume that $G_F:\mathcal{B}\to\mathbf{R}$ admits Frechet derivatives, and we shall require that $\mathcal{B}$ contains the sample functions $X_i(t)$ a.s. Furthermore we assume that X_i, $\overline{X}$ r.v. are well-defined and take values in $\mathcal{B}$.

Denote

$$M_s\overset{\text{def}}{=}\sum_{j=1}^{s}\sup_{x\in B}\big\|H^{(j)}(x)\big\|.$$

Theorem 2.1 implies the following

Theorem 2.2. *Assume that the sequence $\eta_1, \eta_2, \dots$ satisfies the mixing condition (1.4). Let*

$$\mathbf{E}\, X_1 = 0, \qquad \rho \stackrel{\text{def}}{=} \mathbf{E}\, \|X_1\|^3 < \infty.$$

Define $g(x) \stackrel{\text{def}}{=} H'(F)x$ and assume that $\sigma^2 > 0$. If $M_3 < \infty$ then there exists a constant A depending only on $\mathcal{B}$, K, β, ρ, M_3 and σ such that

$$\sup_x \Big| \mathbf{P}\big\{ T < \sigma x \big\} - \Phi(x) \Big| \leqslant C\, N^{-1/2} \log^4 N.$$

If the Banach space $\mathcal{B}$ is of type 2 then the condition $M_3 < \infty$ may be relaxed to $M_2 < \infty$.

2.3 Linear combinations of order statistics. We are going to apply the result for functionals of empirical distribution functions using notation of Subsection 2.2.

Consider the statistic

$$l_N \stackrel{\text{def}}{=} N^{-1} \sum_{i=1}^{N} c_{iN} \eta_{i:N},$$

where $\eta_{1:N} \leq \cdots \leq \eta_{N:N}$ denote the order statistics of $\eta_1, \dots, \eta_N$, and coefficients $c_{1N}, \dots, c_{NN}$ are generated by a weight function $J : [0,1] \to \mathbf{R}$,

$$c_{iN} \stackrel{\text{def}}{=} N \int_{(i-1)/N}^{i/N} J(u)du.$$

Define

$$T \stackrel{\text{def}}{=} \sqrt{N}(l_N - \mu), \qquad \text{where} \quad \mu = \int_{-\infty}^{\infty} x J(F(x)) dF(x).$$

If $\mathbf{E}\, |\eta_1| < \infty$ the boundedness of J is sufficient for the following representation (see Govindarajulu and Mason (1983))

$$l_N - \mu = \int_{-\infty}^{\infty} [\Psi(F_N(t)) - \Psi(F(t))] dt, \quad \text{where} \quad \Psi(x) = \int_x^1 J(u)du.$$

Therefore we may write

$$l_N - \mu = H(F_N) - H(F),$$

where

$$H(h) = \int_{-\infty}^{0} [\Psi(h(t)) - \Psi(0)] dt + \int_0^{\infty} \Psi(h(t)) dt.$$

Let $|| \cdot ||_p$ denote the norm of the space $\mathbf{L}^p(\mathbf{R})$. Let $\mathbf{B}$ be the Banach space of functions with norm

$$||x|| = ||x||_1 + ||x||_2 + ||x||_3.$$

If the function J is smooth then the functional $h \mapsto T(P+h) : \mathbf{B} \to \mathbf{R}$ is Frechét differentiable and

$$M_s \leq L_{s-1} \overset{\text{def}}{=} \sum_{j=0}^{s-1} \sup_x |J^{(j)}(x)|, \qquad s \leq 3.$$

It is easy to verify that the random process X_1 (see previous subsection for the definition of X_1) satisfies

$$\mathbf{E}\,||X_1||_p^s \leq c(p,s)(\mathbf{E}\,|\eta_1|^{s/p} + (\mathbf{E}\,|\eta_1|)^{s/p}), \quad \text{for all} \quad p \geq 1, \ s > 0.$$

Theorem 2.3. *Assume that a stationary sequence $\eta_1, \eta_2, \dots$ satisfies the mixing condition (1.4). Let $\rho \overset{\text{def}}{=} \mathbf{E}\,|\eta_1|^3 < \infty$. Define $a_0 = 1$, $a_j \overset{\text{def}}{=} 2$, for $j > 0$,*

$$\sigma^2 \overset{\text{def}}{=} \int_{\mathbf{R}^2} J(F(t))J(F(s)) \sum_{j=0}^{\infty} a_j [\mathbf{P}\,\{\eta_1 < t, \eta_{1+j} < s\} - F(t)F(s)] ds dt$$

and assume that $\sigma^2 > 0$. If $L_2 < \infty$ then there exists a constant A depending only on K, β, ρ, L_2 and σ such that

$$\sup_x \Big| \mathbf{P}\,\big\{ T < \sigma x \big\} - \Phi(x) \Big| \leqslant C\, N^{-1/2} \log^4 N.$$

The condition $L_2 < \infty$ may be relaxed to $L_1 < \infty$.

Under the condition $L_2 < \infty$ the proof of Theorem 2.3 requires only a straightforward application of Theorem 2.2. The weaker condition $L_1 < \infty$ implies an inequality for the empirical processes E_N like that characterizing the type 2 spaces, and the result still remains true (cf. Bentkus, Götze and van Zwet (1994)).

2.4 Studentized sample means. In this subsection we shall assume that random variables in the stationary sequence $X_1, X_2, \dots$ take real values and satisfy the mixing condition (1.4). Assume as well that $\mathbf{E}\,X_1 = 0$ and $\mathbf{E}\,X_1^4 \leq \rho_4 < \infty$. Denote

$$\overline{X} = N^{-1}(X_1 + \dots + X_N), \qquad S \overset{\text{def}}{=} \sqrt{N}\,\overline{X}, \qquad \sigma_N^2 = \mathbf{E}\,S^2,$$

and assume that $\lim_{N\to\infty} \sigma_N^2 \overset{\text{def}}{=} \sigma^2 > 0$. Consider the estimator

$$s^2 = \frac{1}{N} \sum_{j=1}^{N} \sum_{l:|j-l|\leq m} X_j X_l$$

of σ. It is consistent and asymptotically unbiased provided $m \approx A \log N$ with a sufficiently large constant A. Put $s \overset{\text{def}}{=} \sqrt{s^2}$ if $s^2 \geq 0$, and $s \overset{\text{def}}{=} 0$ if $s^2 < 0$. Introduce the Studentized statistic t,

$$t \overset{\text{def}}{=} \frac{S}{s}, \qquad \text{for } s > 0, \qquad \text{and} \quad t \overset{\text{def}}{=} 0, \qquad \text{for } s = 0.$$

Theorem 2.4. *There exists a constant $A = A(K, \beta, \sigma, \rho_4)$ such that*

$$\sup_x \Big| \mathbf{P}\left\{ t < x \right\} - \Phi(x) \Big| \leqslant A\, N^{-1/2} \log^{9/2} N.$$

Proof. We shall derive the result from Theorem 1.1. Without loss of generality we shall assume that $\sigma^2 = 1$.

Notice that

$$\mathbf{P}\left\{|s^2 - 1| \geq \varepsilon\right\} \leq A\varepsilon^{-1} m^{3/2} N^{-1/2}, \qquad \text{for any} \quad \varepsilon > 0. \tag{2.7}$$

Indeed, by Chebyshev's inequality it is sufficient to estimate $\mathbf{E}\,|s^2 - \sigma^2| \leq \mathbf{E}\,|s^2 - \mathbf{E}\,s^2| + |\mathbf{E}\,s^2 - 1|$, and to show that

$$|\mathbf{E}\,s^2 - 1| \leq Am^2 N^{-1}, \qquad \mathbf{E}\,(s^2 - \mathbf{E}\,s^2)^2 \leq m^3 N^{-1}. \tag{2.8}$$

While proving (2.8) we may assume that $|X_j| \leq N$. Otherwise we may replace X_j by $X_j \mathbb{I}\{|X_j| \leq N\} - \mathbf{E}\,X_j \mathbb{I}\{|X_j| \leq N\}$. By the triangle inequality the error is bounded by Am/N. Let us prove, for example, the second inequality in (2.8). By the triangle inequality

$$\mathbf{E}\,(s^2 - \mathbf{E}\,s^2)^2 \leq mN^{-2} \sum_{s=0}^{m} \mathbf{E}\,(\sum_{j=1}^{N-s} Z_{j,s})^2, \qquad Z_{j,s} \overset{\text{def}}{=} X_j X_{j+s} - \mathbf{E}\,X_j X_{j+s}.$$

The sequence $X_j X_{j+s}$, $j = 1, 2, \dots$ is stationary and satisfies the mixing condition with coefficients $\beta(m - s)$, for $m \geq s$. Thus we may apply Lemma 3.1 (with $2m$ instead of m) and get $\mathbf{E}\,(\sum_{j=1}^{N-s} Z_{j,s})^2 \leq AmN$. Collecting the estimates we obtain (2.8).

Define a $\mathbf{C}^\infty$ function by

$$\theta(x) \overset{\text{def}}{=} 2, \ \text{for} \ x \leq 0, \quad \theta(x) = \frac{1}{\sqrt{x}}, \ \text{for} \ \frac{1}{2} \leq x \leq \frac{3}{2}, \quad \theta(x) = 0, \ \text{for} \ x \geq 2.$$

Due to (2.7)

$$\sup_x \Big| \mathbf{P}\left\{ t < x \right\} - \mathbf{P}\left\{ S\theta(s^2) < x \right\} \Big| \leqslant Am^{3/2} N^{-1/2},$$

and it is sufficient to prove that

$$\sup_x \Big| \mathbf{P}\left\{ S\theta(s^2) < x \right\} - \Phi(x) \Big| \leqslant AmN^{-1/2}. \tag{2.9}$$

While proving (2.9) we may assume that $|X_j| \le N^2$, for $1 \le j \le N$. Otherwise we may replace X_j by their truncated and centered versions because θ is a bounded function such that $|\theta(u) - \theta(v)| \le A|u - v|$.

Write

$$T \stackrel{\text{def}}{=} S\theta(s^2) = S + R, \qquad \text{with} \quad R \stackrel{\text{def}}{=} S(\theta(s^2) - 1).$$

In order to apply Theorem 1.1 we have to verify that

$$\mathbf{E}\,|R| \le A m^2 N^{-1/2}, \qquad \mathbf{E}\,|R_{j,k} - R_{j,k-1}|^{3/2} \le A m^{9/4} N^{-3/2}, \tag{2.10}$$

for $|j - k| \le \log^3 N$, with

$$R_{j,k} \stackrel{\text{def}}{=} S_{j,k}(\theta(s^2_{j,k}) - 1),$$

where $S_{j,k}$ denotes the sum S without summands $N^{-1/2}X_i$, $j \le i \le k$, and $s^2_{j,k}$ denotes the sum s^2 but without summands $N^{-1}X_iX_l$ such that at least one of $i, l \in \{j, \dots, k\}$.

Using $\theta(1) = 1$ and $|\theta(u) - \theta(v)| \le A|u - v|$ we have

$$\mathbf{E}\,|R| \le A\mathbf{E}\,|S|\,|s^2 - 1| \le A m^2 N^{-1/2},$$

and the bound for $\mathbf{E}\,|R|$ follows provided we apply Hölder's inequality, Lemma 3.1 to bound $\mathbf{E}\,S^2 \le A m^{1/2}$, and use (2.8).

Let us prove the second inequality in (2.10). It is sufficient to show that

$$\mathbf{E}\,|(S_{j,k} - S_{j,k-1})(\theta(s^2_{j,k}) - 1)|^{3/2} \quad \text{and} \quad \mathbf{E}\,|S_{j,k-1})(\theta(s^2_{j,k}) - \theta(s^2_{j,k-1}))| \tag{2.11}$$

are bounded by $A m^{9/4} N^{-3/2}$. Let us estimate the first expectation in (2.11). Notice that $|S_{j,k} - S_{j,k-1}| = N^{-1/2}|X_k|$ and that we may represent $s^2_{j,k} = q_{j,k} + \Delta_{j,k}$, where $\Delta_{j,k}$ denotes the sum of terms $N^{-1}X_iX_l$ of the sum $s^2_{j,k}$ such that i or l is near to k. Using $|\theta(u) - \theta(v)| \le A|u - v|^{5/6}$ and the triangle inequality, we may neglect $\Delta_{j,k}$, and it is sufficient to show that

$$\mathbf{E}\,|X_k(\theta(q^2_{j,k}) - 1)|^{3/2} \le A m^{9/4} N^{-3/4}.$$

Using (1.5) we may replace X_k by its independent copy, and it remains to show that

$$\mathbf{E}\,|\theta(q^2_{j,k}) - 1|^{3/2} \le A\mathbf{E}\,|q^2_{j,k} - 1|^{3/2} \le A m^{9/4} N^{-3/4}.$$

But this bound may be proved like (2.8).

In order to estimate the second expectation in (2.11) notice that the difference $s^2_{j,k} - s^2_{j,k-1}$ may contain at most $A \log N$ summands $N^{-1}X_iX_l$ and proceed similarly.

2.5 U-statistics. For notational simplicity we shall consider U-statistics of second order only, that is

$$U = S + R, \quad \text{with} \quad S = \frac{1}{\sqrt{N}} \sum_{1}^{N} g(X_j) \quad \text{and} \quad R = N^{-3/2} \sum_{1 \leqslant i < j \leqslant N} \psi(X_i, X_j),$$

where the function $\psi : \mathcal{X}^2 \to \mathbb{R}^1$ is symmetric, $\psi(x, y) = \psi(y, x)$, and $\mathbf{E}\,\psi(x, X_j) = 0$, for all $x \in \mathcal{X}$ (functions g and ψ may depend on N).

Theorem 2.5. *Assume that conditions (1.2)–(1.4) are fulfilled and that*

$$\sup_N \sup_{1 \leqslant l < j \leqslant N} \mathbf{E}\,\psi^2(X_l, X_j) + \sup_N \mathbf{E}\,\psi^2(\widehat{X}_1, X_2) \leqslant L < \infty.$$

Then

$$\delta_N \overset{\text{def}}{=} \sup_x \Big|\, \mathbf{P}\big\{U < x\sigma_N\big\} - \Phi(x) \,\Big| \leqslant A N^{-1/2} \log^{7/2} N,$$

where the constant A may depend on K, β, ρ, Σ and L only.

Corollary 2.6. *Let the conditions of Theorem 2.5 hold and let g be independent of N. Assume that the variance $\sigma^2 > 0$. Then*

$$\delta_N' \overset{\text{def}}{=} \sup_x \Big|\, \mathbf{P}\big\{U < \sigma x\big\} - \Phi(x) \,\Big| \leqslant A N^{-1/2} \log^{7/2} N,$$

where a constant A may depend on K, β, ρ, σ and L only.

Yoshihara (1984) proved the bound $\delta_N' = O(N^{-1/2} \log^2 N)$ assuming that g and ψ are independent of N and that $\sup_{l<j} \mathbf{E}\,|\psi(X_l, X_j)|^3 + \mathbf{E}\,|\psi(\widehat{X}_1, X_2)|^3 < \infty$.

Proof of Theorem 2.5. We shall apply Theorem 1.1. It is sufficient to show that

$$\mathbf{E}\,|R| \leq A m N^{-1/2} \qquad \text{and} \qquad \gamma \leq A\sqrt{m} N^{-1}, \tag{2.12}$$

with

$$R_{j,k} \overset{\text{def}}{=} N^{-3/2} \sum_{B \cap [j,k] = \emptyset} \psi(X_i, X_l),$$

where the sum is taken over all two point subsets $B = \{i, l\} \subset \{1, \ldots, N\}$ such that $B \cap [j, k] = \emptyset$. Notice that

$$R_{j,k} - R_{j,k-1} = N^{-3/2} \sum_{l \in [1,N] \setminus [j,k]} \psi(X_k, X_l).$$

Therefore the inequalities (2.12) follow from

$$\mathbf{E}\Big|\sum_{1\leqslant i<j\leqslant N}\psi(X_i,X_j)\Big|\leqslant A\,mN \tag{2.13}$$

and

$$\mathbf{E}\Big|\sum_{l\in[1,N]\setminus[j,k]}\psi(X_k,X_l)\Big|^{3/2}\leqslant A\,(mN)^{3/4}. \tag{2.14}$$

While proving (2.13) and (2.14) we may and shall assume that $|\psi(x,y)|\le 4N^3$. Indeed, otherwise we may replace $\psi(x,y)$ by

$$\psi(x,y)\mathbb{I}\{|\psi(x,y)|\le N^3\}-\lambda(x)-\lambda(y)+\mathbf{E}\,\lambda(\hat X_1)$$

with $\lambda(x)\overset{\mathrm{def}}{=}\mathbf{E}\,\psi(x,X_1)\mathbb{I}\{|\psi(x,X_1)|\le N^3\}$.

Lemma 3.2 together with the Hölder inequality yields (2.13).

Let us prove (2.14). By the triangle inequality we may remove in the sum in (2.14) all summands with indices l such that $|l-k|\le m$. Using (1.5) we may replace X_k by its independent copy $\hat X_k$. Thus by Hölders inequality (2.14) follows from

$$\mathbf{E}\,\Big(\Big(\sum\psi(\hat X_k,X_l)\Big)^2\,|\hat X_k\Big)\leqslant A\,mN,$$

where the sum is taken over all l such that $|l-k|>m$ and $l\in[1,N]\setminus[j,k]$. An application of Lemma 3.1 concludes the proof of (2.14) and of the Lemma.

3. Auxiliary results

Lemma 3.1. *Assume (1.4) and let r.v. X_i take values in a Banach space $\mathcal{B}$ of type 2. Let $\mathbf{E}\,X_i=0$ and $\mathbf{P}\,\{\|X_i\|\le D\}=1$, for all $1\le i\le N$, with some $D<\infty$. Then*

$$\mathbf{E}\,\|\sum_{i=1}^N X_i\|^2\le 2\,Cm\sum_{i=1}^N\mathbf{E}\,\|X_i\|^2+2\beta(m)\,D^2N^3,$$

for any natural m.

Proof. Let us split $\{1,\dots,N\}=\Delta_1\cup\Delta_2\cup\dots\cup\Delta_s$ into the union of disjoint subintervals Δ_j of length m with $s\approx N/m$, for $1\le j\le s-1$. Denote

$$Y_j=\sum_{i\in\Delta_j}X_i,\qquad Z_1=\sum_{j:\,1\le 2j\le s}Y_{2j},\qquad Z_2=\sum_{j:\,1\le 2j+1\le s}Y_{2j+1}.$$

Then $\sum_{j=1}^{N} U_j = Z_1 + Z_2$, and it is sufficient to estimate $\mathbf{E}\,\|Z_1\|^2$ and $\mathbf{E}\,\|Z_2\|^2$. We shall estimate $\mathbf{E}\,\|Z_1\|^2$ only. Let $\hat{Y}_1, \hat{Y}_2, \dots$ denote a sequence of independent copies of $Y_1, Y_2, \dots$. Applying (1.5) and estimating $\|X_j\| \le D$ we have

$$\mathbf{E}\,\|Z_1\|^2 = \mathbf{E}\,\|Y_2 + Z_1 - Y_2\|^2 \le \mathbf{E}\,\|\widehat{Y}_2 + Z_1 - Y_2\|^2 + D^2N^2\beta(m).$$

Repeating this procedure (with $Y_4, Y_6, \dots$ instead of Y_2) we get

$$\mathbf{E}\,\|Z_1\|^2 \le \mathbf{E}\,\|\sum_{j:\,1\le 2j\le s} \widehat{Y}_{2j}\|^2 + D^2N^3\beta(m).$$

Now the definition of the type 2 of Banach spaces and the triangle inequality yield

$$\mathbf{E}\,\|\sum_{j:\,1\le 2j\le s} \widehat{Y}_{2j}\|^2 \le Cm\sum_{i=1}^{N}\mathbf{E}\,\|X_i\|^2.$$

Collecting the estimates we conclude the proof of the Lemma.

Proof of (2.1) *for arbitrary Banach spaces.* Expanding into the Taylor series we may write

$$R = N^{-1/2}\mathbf{E}_\tau \sum_{i=1}^{N}(H'(\tau\overline{X}) - H'(0))X_i,$$

and (2.1) follows from

$$\mathbf{E}\,|\sum_{i=1}^{N}(H'(\tau\overline{X}) - H'(0))X_i| \le Am. \tag{3.1}$$

Write $\overline{X}_i = \overline{X} - \Delta$, where $\Delta \overset{\text{def}}{=} N^{-1}\sum_{j\in[i-m,i+m]} X_j$. Expanding in powers of Δ, using the triangle inequality and applying the Hölder inequality, we see that instead of (3.1) it is sufficient to prove

$$\mathbf{E}\,\big(\sum_{i=1}^{N}\mathbf{E}_\tau(H'(\tau\overline{X}_i) - H'(0))X_i\big)^2 \le Am^2. \tag{3.2}$$

Expanding the square in (3.2) we arrive at

$$\sum_{i=1}^{N}\sum_{j=1}^{N}|\mathbf{E}\,(H'(\tau\overline{X}_i) - H'(0))X_i(H'(\tau\overline{X}_j) - H'(0))X_j| \le Am^2,$$

and it is sufficient to prove that

$$|\mathbf{E}\,(H'(\tau\overline{X}_i)-H'(0))X_i(H'(\tau\overline{X}_j)-H'(0))X_j| \le Am\,N^{-1}, \quad \text{for} \quad |i-j| \le 2m, \tag{3.3}$$

and

$$|\mathbf{E}\,(H'(\tau\overline{X}_i)-H'(0))X_i(H'(\tau\overline{X}_j)-H'(0))X_j| \le Am^2\,N^{-2}, \quad \text{for} \quad |i-j| > 2m. \tag{3.4}$$

Let us prove (3.3). Split the sum $\overline{X}_i = \overline{X}_{ij} + \Delta_j$ with Δ_j denoting the sum of $N^{-1}X_l$ in $\overline{X}_i$ such that $|l-j| \le m$. A similar representation we may write for $\overline{X}_j$ with Δ_i. Expanding in powers of Δ_j and of Δ_i we see that (3.3) follows from

$$|\mathbf{E}\,(H'(\tau\overline{X}_{ij})-H'(0))X_i(H'(\tau\overline{X}_{ij})-H'(0))X_j| \le Am\,N^{-1}, \qquad \text{for} \quad |i-j| \le 2m.$$

By (1.5) we may replace the pair $\eta \overset{\text{def}}{=} (X_i, X_j)$ by its independent copy, say $(\hat{X}_i, \hat{X}_j)$, use the Taylor expansion and reduce ((3.3) to

$$\sum_{l,k} |\mathbf{E}\,H''(\tau\tau_1\overline{X}_{ij})X_l\hat{X}_i H''(\tau\tau_2\overline{X}_{ij})X_k\hat{X}_j| \le AmN, \qquad |i-j| \le 2m,$$

where the sum is taken over all l and k present as indices in the sum $\overline{X}_{ij}$. A repetition of the previous arguments conclude the proof of (3.3).

The proof of (3.4) is similar (although more technical) to that of (3.3), and we omit it.

Lemma 3.2. *Assume the mixing condition (1.4). Let* $\sup_{x,y} |\psi(x,y)| \le D$ *and* $\mathbf{E}\,\psi(x, X_1) = 0$*, for all* x*. Then*

$$J \overset{\text{def}}{=} \mathbf{E}\,\Big(\sum_{1\le i<j\le N} \psi(X_i,X_j)\Big)^2 \le Am^2N^2L + AD^2N^4\beta(m),$$

where $L \overset{\text{def}}{=} \max_{1\le i<j\le N} \mathbf{E}\,\psi^2(X_i,X_j)$.

Proof. We may write $J \overset{\text{def}}{=} \mathbf{E}\,(\sum_B \psi(X_i,X_j))^2$, where the sum is taken over all two point subsets $B = \{i,j\}$ such that $B \subset \{1,\dots,N\}$. Let B_m denote the m-neighborhood of $B \subset \{1,\dots,N\}$, that is $B_m \overset{\text{def}}{=} \{s \in \{1,\dots,N\} : \ |i-s| \le m \text{ or } |j-s| \le m\}$. Let $d(B) \overset{\text{def}}{=} |i-j|$ be the diameter of B. Then $J \le 2J_1 + 2J_2$, where

$$J_1 \overset{\text{def}}{=} \mathbf{E}\left(\sum_{d(B)\le m} \psi(X_i,X_j)\right)^2, \quad \text{and} \quad J_2 \overset{\text{def}}{=} \mathbf{E}\left(\sum_{d(B)>m} \psi(X_i,X_j)\right)^2.$$

The number of two point subsets $B \subset \{1, \dots, N\}$ such that $d(B) \le m$ does not exceed mN. Thus $(a_1 + \dots + a_s)^2 \le s(a_1^2 + \dots + a_s^2)$ implies $J_1 \le Am^2N^2L$. Furthermore,

$$J_2 = \sum_{d(B)>m,\ d(D)>m} \mathbf{E}\,\psi(X_i, X_j)\psi(X_k, X_l),$$

where $B = \{i, j\}$ and $D = \{k, l\}$ denote two point subsets of $\{1, \dots, N\}$. The number of pairs of sets B and D such that $D \subseteq B_m$ is bounded from above by Am^2N^2. Therefore it is sufficient to show that

$$\sum_{d(B)>m,\ d(D)>m,\ D\not\subseteq B_m} \mathbf{E}\,\psi(X_i, X_j)\psi(X_k, X_l) \le AD^2N^4\beta(m). \tag{3.5}$$

But relations $d(B) > m$, $d(D) > m$ and $D \not\subseteq B_m$ imply that at least one of k or l, say k, satisfies $|k - i| > m$ and $|k - j| > m$. Thus by (1.5) we may replace X_k by its independent copy $\hat{X}_k$, and (3.5) follows since $\mathbf{E}\,\psi(\hat{X}_k, x) = 0$, for all x.

4. Proof of Theorem 1.1

We shall denote by a a generic sufficiently small positive constant which may depend only on K, β, Σ, and ρ. Write as well

$$f(t) \overset{\text{def}}{=} \mathbf{E}\exp\{itT\}, \qquad g(t) \overset{\text{def}}{=} \exp\{-t^2/2\}.$$

The proof of the theorem combines the techniques of Tikhomirov (1980) for sums of weakly dependent r.v., and those used by Götze (1991), Bentkus, Götze and van Zwet (1994)) in the i.i.d. case for symmetric statistics.

While proving the Theorem we may assume that $|g(X_j)| \le \sqrt{N}$. Otherwise we may replace $g(X_j)$ by

$$g(X_j)\mathbb{I}\{|g(X_j)| \le \sqrt{N}\} - \mathbf{E}\,g(X_j)\mathbb{I}\{|g(X_j)| \le \sqrt{N}\}.$$

Without loss of generality we shall assume as well that $\sigma_N^2 = \mathbf{E}\,S^2 = 1$.

We shall prove that the characteristic function f for $|t| \le a\sqrt{N}/\log^2 N = a\sqrt{N}/m^2$ satisfies the ordinary differential equation,

$$f'(t) = -tf(t) + \varepsilon(t)f(t) + \varepsilon_0(t), \qquad f(0) = 1 \tag{4.1}$$

with some functions ε and ε_0 such that

$$|\varepsilon(t)| \le \frac{Am^2t^2}{\sqrt{N}}, \tag{4.2}$$

and

$$|\varepsilon_0(t)| \le AN^{1/2}|t|m^2\gamma + A\mathbf{E}\,|R| + AN^{-1/2} + Am|t|N^{-1/2}. \tag{4.3}$$

The equation has the unique solution

$$f(t) = g(t)\exp\left\{\int_0^t \varepsilon(u)\,du\right\} + g(t)\int_0^t \exp\left\{\frac{u^2}{2} + \int_u^t \varepsilon(z)\,dz\right\}\varepsilon_0(u)\,du. \tag{4.4}$$

Let us derive Theorem 1.1 from (4.4). It follows from (4.4) that

$$|f(t) - g(t)| \le I_1 + I_2,$$

where

$$I_1 \overset{\text{def}}{=} g(t)\left|\exp\left\{\int_0^t \varepsilon(u)\,du\right\} - 1\right| \le Am^2N^{-1/2}t^2\exp\{-t^2/4\}, \tag{4.5}$$

and

$$\begin{aligned} I_2 &\overset{\text{def}}{=} g(t)\int_0^t \exp\left\{\frac{u^2}{2} + \int_u^t \varepsilon(z)\,dz\right\}\,|\varepsilon_0(u)|\,du \\ &\le A(N^{1/2}m^2 + mN^{-1/2})\,\gamma\,\min\{1;|t|\} + A\big(\mathbf{E}\,|R| + N^{-1/2}\big)\min\{|t|^{-1};|t|\}. \end{aligned} \tag{4.6}$$

Estimates (4.5) and (4.6) together with the Esseen inequality for characteristic functions imply the result of the Theorem.

In order to prove the inequality in (4.5) apply $|\exp\{z\} - 1| \le |z|\exp\{|z|\}$ and (4.2) on the interval $|t| \le a\sqrt{N}/m^2$ with a sufficiently small a. Similarly, the estimate for I_2 we derive using $|\int_u^t \varepsilon(z)\,dz| \le aA(t^2 - u^2)$ and $aA < 1/4$.

It remains to prove (4.1)–(4.3). We shall write $B \sim D$ if $B = D + \varepsilon(t)f(t) + \varepsilon_0(t)$, for $|t| \le a\sqrt{N}/m^2$, with some functions ε and ε_0 bounded as in (4.2) and (4.3). Thus we have to prove that $f'(t) \sim -tf(t)$. Differentiating we have

$$f'(t) = i\,\mathbf{E}\,S\exp\{itT\} + i\,\mathbf{E}\,R\,\exp\{itT\} \sim I_3 \overset{\text{def}}{=} \frac{i}{\sqrt{N}}\sum_{j=1}^N \mathbf{E}\,g(X_j)\,\exp\{itT\}$$

with the error bounded by $\mathbf{E}\,|R|$.

Define $S_{j,0} \overset{\text{def}}{=} S$ and

$$\Delta_{j,1} \overset{\text{def}}{=} N^{-1/2}\sum_{l\in\Omega_1} X_l \qquad \text{with} \quad \Omega_1 \overset{\text{def}}{=} \{l : 1 \le l \le N,\ \ |l - j| \le m\}.$$

Put $S_{j,1} \overset{\text{def}}{=} S - \Delta_{j,1}$. By induction we may define

$$\Delta_{j,s} \overset{\text{def}}{=} N^{-1/2} \sum_{l \in \Omega_s} X_l \quad \text{with} \quad \Omega_s \overset{\text{def}}{=} \{l : 1 \le l \le N,\ (s-1)m < |l-j| \le sm\},$$

and $S_{j,s} \overset{\text{def}}{=} S_{j,s-1} - \Delta_{j,s}$. Furthermore, for a natural $r = A_1 \log N$ (to be chosen later),

$$Q_j \overset{\text{def}}{=} R_{j-rm,j+rm}, \qquad \delta_j \overset{\text{def}}{=} R - Q_j, \qquad T_{j,s} \overset{\text{def}}{=} S_{j,s} + Q_j.$$

An application of $(a_1 + \cdots + a_p)^{3/2} \le \sqrt{p}(a_1^{3/2} + \cdots + a_p^{3/2})$ and of the triangle inequality implies

$$\mathbf{E}\,|\delta_j|^{3/2} \le A m^3 \gamma^{3/2}. \tag{4.7}$$

An expansion in powers of δ_j and an application of (4.7) together with Hölder's inequality show that

$$I_3 \sim I_4 \overset{\text{def}}{=} \frac{i}{\sqrt{N}} \sum_{j=1}^{N} \mathbf{E}\, g(X_j)\, \exp\{itT_{j,0}\}$$

with the error bounded by $A|t|m^2 N^{1/2}\gamma$.

Splitting $T_{j,0} = T_{j,1} + \Delta_{j,1}$ we may write

$$I_4 = \frac{i}{\sqrt{N}} \sum_{j=1}^{N} \mathbf{E}\, g(X_j) J_0\, \exp\{itT_{j,1}\} + \frac{i}{\sqrt{N}} \sum_{j=1}^{N} \mathbf{E}\, g(X_j)\, J_1 \exp\{itT_{j,1}\},$$

where $J_0 \overset{\text{def}}{=} 1$ and $J_s \overset{\text{def}}{=} \exp\{it\Delta_{j,s}\} - 1$, for $s \ge 1$. Repeating the procedure we get (we shall choose $r = A_1 \log N$)

$$f'(t) \sim I_4 = \sum_{s=0}^{r-1} \sum_{j=1}^{N} I(j,s) + \sum_{j=1}^{N} I_1(j,r), \qquad \text{for any natural} \quad r, \tag{4.8}$$

with

$$I(j,s) \overset{\text{def}}{=} \frac{i}{\sqrt{N}}\, \mathbf{E}\, g(X_j) J_1 \ldots J_s\, \exp\{itT_{j,s+1}\},$$
$$I_1(j,r) \overset{\text{def}}{=} \frac{i}{\sqrt{N}}\, \mathbf{E}\, g(X_j) J_1 \ldots J_r\, \exp\{itT_{j,r}\}.$$

The expectation in $I(j,s)$ is taken over a function of a set of random variables which does not contain X_l with indices l such that $sm < |l-j| \le (s+1)m$. Thus by (1.5) we may replace $\exp\{itT_{j,s+1}\}$ by its independent copy, and obtain

$$\sum_{s=0}^{r-1}\sum_{j=1}^{N} I(j,s) \sim \sum_{s=0}^{r-1}\sum_{j=1}^{N} I_2(j,s) \tag{4.9}$$

with

$$I_2(j,s) \overset{\text{def}}{=} \frac{i}{\sqrt{N}}\,\mathbf{E}\,g(X_j)J_1\dots J_s\,\mathbf{E}\,\exp\{itT_{j,s+1}\}.$$

Indeed, $|J_s| \le 2$ and $|g(X_j)| \le \sqrt{N}$, and the error in (4.9) is bounded from above by $A\sum_{s=0}^{r-1} N2^s\beta(m) \le N2^r\beta(m) \le AN^{-1/2}$ since $r = A_1 \log N$ and since we may choose the constant $A = A(A_1)$ in $m = A\log N$ sufficiently large.

We have

$$I_2(j,0) = 0 \qquad\qquad \text{since} \quad \mathbf{E}\,g(X_j) = 0. \tag{4.10}$$

Collecting J_l with odd or even l into separate groups, applying the Hölder inequality and then by (1.5) replacing the multipliers $|J_l| \le 2$ by their independent copies, we obtain

$$(\mathbf{E}\,|J_1\dots J_s|^{3/2})^{2/3} \le \prod_{l=1}^{s}(\mathbf{E}\,|J_l|^3)^{1/3} + A3^s\beta^{1/3}(m), \qquad s \ge 1. \tag{4.11}$$

The choices of m and r and $|J_s| \le |t\Delta_{j,s}|$ together with (4.11) imply

$$(\mathbf{E}\,|J_1\dots J_s|^{3/2})^{2/3} \le A\min\Big\{(a/m)^s;\ \big(m|t|N^{-1/2}\big)^s\Big\} + AN^{-3}. \tag{4.12}$$

An application of (4.12) to $I_1(j,r)$ together with (4.8) and (4.10) yield

$$f'(t) \sim \sum_{s=1}^{r-1}\sum_{j=1}^{N} I_2(j,s), \qquad \text{since} \qquad \sum_{j=1}^{N} I_1(j,r) \sim 0, \tag{4.13}$$

with the error bounded by $AN^{-1/2}$.

Denote $T'_{j,s} \overset{\text{def}}{=} T_{j,s} + \delta_j = S_{j,s} + R$. An expansion in powers of δ_j, an application of (4.7) and (4.12) together with (4.13) imply

$$f'(t) \sim \sum_{s=1}^{r-1}\sum_{j=1}^{N} I_3(j,s), \qquad I_3(j,s) \overset{\text{def}}{=} \frac{i}{\sqrt{N}}\,\mathbf{E}\,g(X_j)J_1\dots J_s\mathbf{E}\,\exp\{itT'_{j,s+1}\} \tag{4.14}$$

with the error bounded by $A|t|m^2\sqrt{N}\gamma$.

Define

$$I_4(s) \stackrel{\text{def}}{=} i\,\mathbf{E}\,g(X_j)J_1\dots J_s$$

and notice that, for $j \in \Gamma_s \stackrel{\text{def}}{=} \{j : sm < j < N - sm\}$, the expectation $I_4(s)$ is independent of j since the sequence $X_1, X_2, \dots$ is stationary. Thus

$$\sum_{j\in\Gamma_s} I_3(j,s) = I_4(s) \sum_{j\in\Gamma_s} N^{-1/2}\mathbf{E}\,\exp\{itT'_{j,s+1}\}. \tag{4.15}$$

Using (4.12) we have

$$\sum_{s=1}^{r-1} \sum_{1\le j\le N,\, j\notin\Gamma_s} N^{-1/2}\,\mathbf{E}\,|g(X_j)J_1\dots J_s| \le AN^{-1/2}\sum_{s=1}^{r-1} ms(a/m)^s \le AN^{-1/2}. \tag{4.16}$$

Relations (4.14)–(4.16) imply

$$f'(t) \sim \sum_{s=1}^{r-1} \sqrt{N} I_4(s) f_s(t), \qquad f_s(t) \stackrel{\text{def}}{=} \frac{1}{N}\sum_{j=1}^{N} \mathbf{E}\,\exp\{itT'_{j,s+1}\} \tag{4.17}$$

with the error bounded by $AN^{-1/2}$.

Let us show that

$$|f_s(t) - f(t)| \le A\,m\,s\,N^{-1/2}|tf(t)| + AsN^{-1/2}. \tag{4.18}$$

Define $\mu_{j,s} \stackrel{\text{def}}{=} S - S_{j,s+1}$. Then $T'_{j,s+1} = T - \mu_{j,s}$, and

$$f_s(t) - f(t) = D_1 f(t) + D_2, \qquad \text{with} \quad \xi_j \stackrel{\text{def}}{=} \exp\{-it\mu_{j,s}\},$$

where

$$D_1 \stackrel{\text{def}}{=} \frac{1}{N}\sum_{j=1}^{N}\mathbf{E}\,\exp\{-it\mu_{j,s}\} - 1 \qquad D_2 \stackrel{\text{def}}{=} \frac{1}{N}\mathbf{E}\sum_{j=1}^{N}(\xi_j - \mathbf{E}\,\xi_j)\exp\{itT\}.$$

Expanding in powers of $t\mu_{j,s}$ we get

$$|D_1| \le \frac{1}{N}\sum_{j=1}^{N}\mathbf{E}\,|t\mu_{j,s}| \le Ams|t|N^{-1/2}.$$

By Hölder's inequality

$$|D_2|^2 \le N^{-2}\mathbf{E}\,|\sum_{j=1}^{N}(\xi_j - \mathbf{E}\,\xi_j)|^2 = N^{-2}\sum_{j=1}^{N}\sum_{k=1}^{N}\mathbf{E}\,(\xi_j - \mathbf{E}\,\xi_j)(\bar{\xi}_k - \mathbf{E}\,\bar{\xi}_k) \le As^2N^{-1}$$

since in the last sum summands, say $P_{j,k}$, with indices j and k such that $|j-k| \le 4ms$ satisfy

$$|P_{j,k}| \le (\mathbf{E}\,|\xi_j - \mathbf{E}\,\xi_j|^2\mathbf{E}\,|\xi_k - \mathbf{E}\,\xi_k|^2)^{1/2} \le At^2(\mathbf{E}\,|\mu_{j,s}|^2\mathbf{E}\,|\mu_{k,s}|^2)^{1/2} \le As^2m^{-2},$$

and since $|P_{j,k}| \le A/N$ in the case $|j-k| > 4ms$ (to see this apply (1.5)). Collecting these estimates we get (4.18).

Relations (4.17), (4.18) and (4.12) imply

$$f'(t) \sim \sum_{s=1}^{r-1}\sqrt{N}I_4(s)f(t)$$

with the error bounded by $Am^2t^2N^{-1/2}|f(t)| + Am|t|N^{-1/2}$. By (4.12)

$$\sum_{s=2}^{r-1}|I_4(s)| \le At^2m^2N^{-1}$$

and

$$f'(t) \sim \sqrt{N}I_4(1)f(t) = i\sqrt{N}\mathbf{E}\,g(X_j)J_1f(t).$$

A Taylor expansion applied to J_1 and a comparison of the coefficient of t with $-1 = -\mathbf{E}\,S^2$ show that $f'(t) \sim -tf(t)$ with the error bounded by $AN^{-1/2}(1 + m^2t^2)|f(t)|)$, which concludes the proof of the Theorem.

References

Bentkus V. and Götze F., *The Berry–Esseen bound for Student's statistics*, Universität Bielefeld, SFB 343, Preprint 94–033, To appear in Ann. Prob. (1994).

Bentkus V., Götze F. and Zitikis R., *Lower estimates of the convergence rate for U-statistics*, Ann. Prob. (1994), no. 4 vol **22**.

Bentkus V., Götze F. and van Zwet W., *An Edgeworth expansion for symmetric statistics*, Universität Bielefeld, SFB 343, Preprint 94–020 (1994).

Bolthausen E. and Götze F., *The rate of convergence for multivariate sampling statistics*, Ann. Statist. vol **21** (1993), 1692–1710.

Cartan H., *Calcul Differentiel. Formes Differentielles.*, Hermann, Paris, 1971.

Denker M., *Statistical decision procedures and ergodic theory*, Ergodic theory and related topics. Mathem. results. Vitte 1981 vol **12**, 1982, pp. 35-47.

Denker M., and Keller G., *On U-statistics and von Mises statistics for weakly dependent processes*, Z. Wahrsch. Verw. Geb. vol **64** (1983), 505–522.

Eberlein E., *Weak convergence of partial sums of absolutely regular sequences*, Stat. & Probab. Letters vol **2** (1984), 291–293.

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

Götze F., *On the rate of convergence in the multivariate CLT*, Ann. Probab. vol **19** (1991), 724–739.

Götze F. and Hipp C., *Asymptotic expansions for sums of weakly dependent random vectors vol* **64**, Zeit. Wahrsch. verw. Geb. (1983), 211–239.

Götze F. and Hipp C., *Asymptotic expansions for sums of weakly dependent random vectors*, Universität Bielefeld, SFB 343, Preprint 92–045 (1992).

Heinrich L., *Bounds for the absolute regularity coefficient of a stationary renewal process*, Yokohama Mathem. J. vol **40** (1992), 25–33.

Ibragimov I.A. and Linnik Yu.V., *Independent and stationary sequences of random variables*, Wolters–Nordhoff, Groningen, 1971.

Ibragimov I.A., *A note on the CLT for dependent random variables*, Probab. Theory Appl. vol **20** (1975), 135–141.

Stein Ch., *A bound on the error in the normal approximation to the distribution of a sum of dependent random variables*, Proc. Sixth Berkeley Symp. Math. Statist. and Prob. vol **2**, 1972, pp. 583-602.

Sunklodas J., *Aproximation of distributions of sums of weakly dependent random variables by the normal distribution*, Limit theorems of probability theory, vol **81**, VINITI, Moscow, 1991, pp. 140–199.

Tikhomirov A.N., *On the rate of convergence in the central limit theorem for weakly dependent variables*, Theory Probab. Appl. vol **25** (1980), 790–809.

Veretennikov A. Yu., *Bounds for the mixing rate in the theory of stochastic equations*, Theory Probab. Appl. vol **32** (1987), 273–281.

Yoshihara K.-i., *Limit behavior for stationary absolutely regular processes*, Zeit. Wahrsch. verw. Geb. vol 35 (1976), 237-252.

Yoshihara K.-i, *The Berry–Esseen theorems for U-statistics generated by absolutely regular processes*, Yokohama Mathematical Journal vol **32** (1984), 89–111.

van Zwet W.R., *A Berry-Esseen bound for symmetric statistics*, Z. Wahrsch. verw. Gebiete, vol **66** (1984), 425–440.

VIDMANTAS BENTKUS
FAKULTÄT FÜR MATHEMATIK
UNIVERSITÄT BIELEFELD
POSTFACH 100131
33501 BIELEFELD 1
GERMANY
E-mail address: bentkus@mathematik.uni-bielefeld.de

FRIEDRICH GÖTZE
FAKULTÄT FÜR MATHEMATIK
UNIVERSITÄT BIELEFELD
POSTFACH 100131
33501 BIELEFELD 1
GERMANY
E-mail address: goetze@mathematik.uni-bielefeld.de

ALEKSANDER TIKHOMIROV
FACULTY OF MATHEMATICS
UNIVERSITY OF SYKTYVKAR
OKTJABRSKYI PROSPEKT 55
167001 SYKTYVKAR
RUSSIA
E-mail address: tichomir@sucnit.komi.su