

THE EDGEWORTH EXPANSION FOR U -STATISTICS OF DEGREE TWO

BY P. J. BICKEL,¹ F. GÖTZE² AND W. R. VAN ZWET²

*University of California, Berkeley, University of Bielefeld and
University of Leiden*

An Edgeworth expansion with remainder $o(N^{-1})$ is established for a U -statistic with a kernel h of degree 2. The assumptions involved appear to be very mild; in particular, the common distribution of the summands $h(X_i, X_j)$ is not assumed to be smooth.

1. Introduction. Let X_1, X_2, \dots, X_N be independent and identically distributed (i.i.d.) random variables assuming values in a measurable space $(\mathcal{X}, \mathcal{B})$ with a common distribution P_X . Let $h: \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ be measurable and symmetric in its two arguments, i.e., $h(x, y) = h(y, x)$. For $N \geq 2$, a U -statistic of degree 2 is defined as

$$(1.1) \quad U_N = \sum_{i=1}^{N-1} \sum_{j=i+1}^N h(X_i, X_j).$$

Note that we do not follow the usual convention of dividing the sum in (1.1) by the number $\binom{N}{2}$ of its terms. Since our results concern the standardized version of U , this does not make any difference.

We assume throughout that

$$(1.2) \quad Eh(X_1, X_2) = 0, \quad Eh^2(X_1, X_2) < \infty,$$

and define

$$(1.3) \quad g(x) = E(h(X_1, X_2)|X_1 = x), \quad \psi(x, y) = h(x, y) - g(x) - g(y),$$

$$(1.4) \quad \hat{U}_N = (N-1) \sum_{i=1}^N g(X_i), \quad \Delta_N = \sum_{i=1}^{N-1} \sum_{j=i+1}^N \psi(X_i, X_j),$$

so that

$$(1.5) \quad U_N = \hat{U}_N + \Delta_N.$$

Since $E(\psi(X_1, X_2)|X_1) = 0$ almost surely (a.s.), the random variables $g(X_i)$ and

Received January 1985; revised November 1985.

¹Research supported by the U.S. Office of Naval Research, Contract N0014-80-C-1063 and by the Netherlands Organization for Pure Scientific Research.

²Research supported by the U.S. Office of Naval Research, Contract N00014-80-C-1063.

AMS 1980 subject classifications. Primary 62E20; secondary 60F05.

Key words and phrases. Edgeworth expansion, second order asymptotics, U -statistics.

$\psi(X_i, X_j)$, $1 \leq i < j \leq N$, are pairwise uncorrelated and hence

$$(1.6) \quad \begin{aligned} \sigma_N^2 &= \sigma^2(U_N) = \sigma^2(\hat{U}_N) + \sigma^2(\Delta_N) \\ &= N(N-1)^2 E g^2(X_1) + \frac{1}{2} N(N-1) E \psi^2(X_1, X_2). \end{aligned}$$

If it is assumed that

$$(1.7) \quad \sigma_g^2 = E g^2(X_1) > 0,$$

then $\sigma^2(\hat{U}_N)$ dominates the right-hand side of (1.6) and $\sigma_N^{-1}U_N$ is asymptotically normal as $N \rightarrow \infty$ [cf. Hoeffding (1948), where U -statistics were introduced].

The speed of convergence to normality was investigated by Bickel (1974), Chan and Wierman (1977), Callaert and Janssen (1978) and Helmers and van Zwet (1982) who showed in increasing generality that

$$(1.8) \quad \sup_x |P(\sigma_N^{-1}U_N \leq x) - \Phi(x)| = O(N^{-1/2}),$$

where Φ denotes the standard normal distribution function (d.f.). If (1.2) and (1.7) are satisfied, so that asymptotic normality is ensured, then $E|g(X_1)|^3 < \infty$ suffices to establish (1.8). Moreover, the assumption $Eh^2(X_1, X_2) < \infty$ may be relaxed, provided σ_N is replaced by $\sigma(\hat{U}_N)$ in (1.8).

The next step in the asymptotic analysis of $\sigma_N^{-1}U_N$, is to obtain an Edgeworth expansion for its d.f., and for statistical purposes one typically needs such an expansion up to a remainder term which is $o(N^{-1})$. To be specific, let

$$(1.9) \quad \kappa_3 = \sigma_g^{-3} \{ E g^3(X_1) + 3 E g(X_1) g(X_2) \psi(X_1, X_2) \},$$

$$(1.10) \quad \begin{aligned} \kappa_4 &= \sigma_g^{-4} \{ E g^4(X_1) - 3 \sigma_g^4 + 12 E g^2(X_1) g(X_2) \psi(X_1, X_2) \\ &\quad + 12 E g(X_1) g(X_2) \psi(X_1, X_3) \psi(X_2, X_3) \}. \end{aligned}$$

Straightforward calculation shows that if $Eh^4(X_1, X_2) < \infty$ —which we shall *not* generally require in this paper—then $\kappa_3 N^{-1/2}$ and $\kappa_4 N^{-1}$ are asymptotic expressions with error $o(N^{-1})$ for the third and fourth cumulants of $\sigma_N^{-1}U_N$, respectively. Define

$$(1.11) \quad \begin{aligned} F_N(x) &= \Phi(x) - \phi(x) \left\{ \frac{\kappa_3}{6} N^{-1/2} (x^2 - 1) + \frac{\kappa_4}{24} N^{-1} (x^3 - 3x) \right. \\ &\quad \left. + \frac{\kappa_3^2}{72} N^{-1} (x^5 - 10x^3 + 15x) \right\}, \end{aligned}$$

where ϕ denotes the standard normal density. We wish to show that

$$(1.12) \quad \sup_x |P(\sigma_N^{-1}U_N \leq x) - F_N(x)| = o(N^{-1})$$

as $N \rightarrow \infty$.

The validity of the Edgeworth expansion (1.11)–(1.12) was established by Janssen (1978) and by Callaert, Janssen and Veraverbeke (1980) under a complicated condition which these authors were able to verify only for certain cases where the distribution of $h(X_1, X_2)$ possesses an absolutely continuous part. An

inspection of special cases, however, quickly reveals that the expansion may be valid even when h assumes only two values. In this respect the situation appears to be more favorable than it is for sums of i.i.d. random variables, where the lattice case has to be excluded. The explanation of this phenomenon is simple: the left-hand side of (1.12) cannot be smaller than the largest jump of the d.f. of U_N and in the lattice case the jumps are of the order $N^{-1/2}$ for sums, but $N^{-3/2}$ for most U -statistics. An exception is, of course, the U -statistic $\{\sum 1_B(X_i)\}^2$ which is distributed like the square of a binomial random variable, so that the jumps are of the order $N^{-1/2}$.

The aim of the present paper is to establish the Edgeworth expansion under very mild assumptions that are easy to verify and do not involve smoothness of the distribution of $h(X_1, X_2)$. Suppose that there exist positive numbers $\delta, \delta_1, \delta_2, \delta_3, C$ and positive and continuous functions $\chi_j: (0, \infty) \rightarrow (0, \infty), j = 1, 2$, satisfying

$$(1.13) \quad \lim_{t \rightarrow \infty} \chi_1(t) = 0,$$

$$(1.14) \quad \lim_{t \rightarrow \infty} \chi_2(t) \geq \delta_1 > 0,$$

as well as a real number r such that

$$(1.15) \quad r \geq 2 + \delta > 2,$$

$$(1.16) \quad E|\psi(X_1, X_2)|^r \leq C,$$

$$(1.17) \quad Eg^4(X_1)1_{[t, \infty)}(|g(X_1)|) \leq \chi_1(t) \quad \text{for all } t > 0,$$

$$(1.18) \quad |Ee^{ig(X_1)}| \leq 1 - \chi_2(t) < 1 \quad \text{for all } t > 0.$$

Let $\lambda_1, \lambda_2, \dots$ denote the eigenvalues of the kernel ψ with respect to P_X , ranked according to descending absolute values and with multiple eigenvalues repeated. Thus, for some orthonormal sequence of eigenfunctions $\omega_1, \omega_2, \dots$,

$$(1.19) \quad \int \psi(x, y) \omega_j(x) dP_X(x) = \lambda_j \omega_j(y), \quad |\lambda_1| \geq |\lambda_2| \geq \dots$$

Assume, in addition to (1.13)–(1.18), that there exists a natural number k such that

$$(1.20) \quad |\lambda_k| \geq \delta_2 > 0.$$

Finally, assumptions (1.15), (1.16) and (1.20) are linked by requiring that

$$(1.21) \quad (r - 2)(k - 4) \geq 8 + \delta_3 > 8.$$

We note that (1.18) implies the existence of a positive number δ_4 depending only on χ_2 and such that

$$(1.22) \quad Eg^2(X_1) \geq \delta_4 > 0,$$

so that the conditions for asymptotic normality of $\sigma_N^{-1}U_N$ are satisfied. We shall

prove

THEOREM 1.1. *Suppose that positive numbers $\delta, \delta_1, \delta_2, \delta_3, C$ and positive continuous functions χ_1 and χ_2 exist such that (1.13)–(1.21) are satisfied. Then there exists a sequence $\varepsilon_N \downarrow 0$ depending only on $\delta, \delta_1, \delta_2, \delta_3, C, \chi_1$ and χ_2 such that for $N = 2, 3, \dots$,*

$$(1.23) \quad \sup_x |P(\sigma_N^{-1}U_N \leq x) - F_N(x)| \leq \varepsilon_N N^{-1},$$

where σ_N^2 and F_N are given by (1.6) and (1.9)–(1.11).

The laborious way in which we have phrased the assumptions as well as the conclusion of the theorem is caused by our insistence to define uniformity classes: for any class of pairs (h, P_X) for which the assumptions are satisfied for fixed δ, δ_j, C and χ_j , (1.12) holds uniformly. It will therefore continue to hold if we let h and P_X vary with N , provided $(h_N, P_{X,N}), N = 1, 2, \dots$, are all in such a class. If we do not insist on uniformity and simply consider a fixed pair (h, P_X) , then the result is much easier to state:

COROLLARY 1.1. *Suppose that there exist a number $r > 2$ and an integer k such that $(r - 2)(k - 4) > 8$ and that the following assumptions are satisfied*

$$(1.24) \quad E|\psi(X_1, X_2)|^r < \infty,$$

$$(1.25) \quad E|g(X_1)|^4 < \infty,$$

$$(1.26) \quad \limsup_{|t| \rightarrow \infty} |Ee^{itg(X_1)}| < 1,$$

$$(1.27) \quad \psi \text{ possesses } k \text{ nonzero eigenvalues with respect to } P_X.$$

Then (1.12) holds.

In the theorem as well as in the corollary, the role of all but one of the conditions is immediately clear. Since $Eg^2(X_1) > 0$ [cf. (1.22)] and $E\psi^2(X_1, X_2) < \infty$, \hat{U}_N is the dominating term on the right in (1.5) and the conditions on $g(X_1)$ establish an Edgeworth expansion for \hat{U}_N . The moment assumption $E|\psi(X_1, X_2)|^r < \infty$ for some $r > 2$ allows us to correct the expansion for the remainder term Δ_N in (1.5). The existence of k nonzero eigenvalues of ψ , however, plays a much more subtle part which we shall discuss after the proof of the theorem has been given. We note that this kind of assumption first occurs in this context in Götze (1979).

If we are content to have an Edgeworth expansion with remainder $o(N^{-1/2})$ instead of $o(N^{-1})$, then we can do without the eigenvalue assumption. At the same time we may, of course, replace 4 by 3 in (1.17) and delete (1.14) so that

(1.18) becomes a nonlattice condition. Define

$$(1.28) \quad \tilde{F}_N(x) = \Phi(x) - \frac{1}{6}\kappa_3 N^{-1/2}\phi(x)(x^2 - 1),$$

where κ_3 is given by (1.9).

THEOREM 1.2. *Suppose that positive numbers δ, C and positive continuous functions χ_1 and χ_2 exist such that (1.13) and (1.15)–(1.18) are satisfied, with g^4 replaced by $|g|^3$ in (1.17). Then there exists a sequence $\varepsilon_N \downarrow 0$ depending only on δ, C, χ_1 and χ_2 such that for $N = 2, 3, \dots$,*

$$(1.29) \quad \sup_x |P(\sigma_N^{-1}U_N \leq x) - \tilde{F}_N(x)| \leq \varepsilon_N N^{-1/2}.$$

To prove Theorem 1.1 we shall have to study the characteristic function (c.f.) of $\sigma_N^{-1}U_N$. This is done separately for small (and intermediate) and for large values of the argument in Sections 2 and 3, respectively. After the extensive previous work on the asymptotics of U -statistics, the arguments in the first part are almost standard; the essential difficulties arise in the second part. Combination of the results of Sections 2 and 3 immediately yields Theorem 1.1. Theorem 1.2 follows from an analysis closely resembling that of Section 2, the only difference being that the use of the fourth moment of $g(X_1)$ should now be avoided. The proof that this can be done is easy and we omit it.

In Section 4 we discuss various aspects of assumption (1.20) and in Section 5 we give an application of Theorem 1.1. Two technical results—a moment inequality and a concentration inequality—which are needed in Section 3 but which may be of wider interest, are dealt with in the Appendix.

2. The c.f. for small values of the argument. Let ϕ_N denote the c.f. of $\sigma_N^{-1}U_N$,

$$(2.1) \quad \phi_N(t) = E \exp\{it\sigma_N^{-1}U_N\}$$

and, for κ_3 and κ_4 as in (1.9)–(1.10), let

$$(2.2) \quad \phi_N^*(t) = e^{-t^2/2} \left[1 - \frac{i\kappa_3}{6} N^{-1/2} t^3 + \frac{\kappa_4}{24} N^{-1} t^4 - \frac{\kappa_3^2}{72} N^{-1} t^6 \right]$$

be the Fourier–Stieltjes transform $\int \exp(itx) dF_N(x)$ of F_N in (1.11). By Esseen’s smoothing lemma [cf. Feller (1971), page 538] we have proved (1.23) if we construct sequences $\{T_N\}$ and $\{\varepsilon'_N\}$ depending only on $\delta, \delta_1, \delta_2, \delta_3, C, \chi_1$ and χ_2 such that $N^{-1}T_N \rightarrow \infty, \varepsilon'_N \rightarrow 0$, and

$$(2.3) \quad \int_{-T_N}^{T_N} \left| \frac{\phi_N(t) - \phi_N^*(t)}{t} \right| dt \leq \varepsilon'_N N^{-1}.$$

We begin by studying $\phi_N(t)$ for small $|t|$ and prove

LEMMA 2.1. *Suppose that (1.13)–(1.18) are satisfied. Then there exists a sequence $\varepsilon'_N \downarrow 0$ depending only on $\delta, \delta_1, C, \chi_1$ and χ_2 such that for*

$$(2.4) \quad t_N = N^{(r-1)/r}(\log N)^{-1},$$

$$(2.5) \quad \int_{-t_N}^{t_N} \left| \frac{\phi_N(t) - \phi_N^*(t)}{t} \right| dt \leq \varepsilon'_N N^{-1}.$$

PROOF. To prevent the laborious formulation of our results from occurring throughout the proofs also, we shall make extensive use of o and O symbols rather than explicit error bounds. It will be tacitly understood that every statement involving o and O holds uniformly for all h and P_X satisfying the assumptions of the lemma to be proved for a fixed choice of the δ, δ_j, C and χ_j involved, and also uniformly for the values of t being considered.

Assume without loss of generality that $\delta \in (0, 1]$ and define

$$(2.6) \quad \varepsilon = \frac{\delta}{3(2 + \delta)} \in (0, 1/9].$$

Combining (2.1), (1.5) and

$$(2.7) \quad \left| e^{ix} - \sum_{\nu=0}^m \frac{(ix)^\nu}{\nu!} \right| \leq \frac{2}{m!} |x|^{m+\theta} \quad \text{for every } \theta \in [0, 1],$$

we can write

$$(2.8) \quad \phi_N(t) = E \exp\{it\sigma_N^{-1}\hat{U}_N\} (1 + it\sigma_N^{-1}\Delta_N - \frac{1}{2}t^2\sigma_N^{-2}\Delta_N^2) + O(E|t\sigma_N^{-1}\Delta_N|^{2+\delta}).$$

Let

$$\gamma_N(t) = E \exp\{it\sigma_N^{-1}(N-1)g(X_1)\}$$

denote the c.f. of $\sigma_N^{-1}(N-1)g(X_1)$. In view of (1.6), (1.22) and the fact that $E|\Delta_N|^{2+\delta} = O(N^{2+\delta})$ [cf. Callaert and Janssen (1978)] we may rewrite (2.8) as

$$(2.9) \quad \begin{aligned} \phi_N(t) &= \gamma_N^N(t) + \gamma_N^{N-2}(t)it\sigma_N^{-1}\binom{N}{2}E \exp\left\{it\sigma_N^{-1}(N-1)\sum_{j=1}^2 g(X_j)\right\}\psi(X_1, X_2) \\ &\quad - \frac{1}{2}\gamma_N^{N-2}(t)t^2\sigma_N^{-2}\binom{N}{2}E \exp\left\{it\sigma_N^{-1}(N-1)\sum_{j=1}^2 g(X_j)\right\}\psi^2(X_1, X_2) \\ &\quad - 3\gamma_N^{N-3}(t)t^2\sigma_N^{-2}\binom{N}{3}E \exp\left\{it\sigma_N^{-1}(N-1)\sum_{j=1}^3 g(X_j)\right\} \\ &\quad \quad \times \psi(X_1, X_3)\psi(X_2, X_3) \\ &\quad - 3\gamma_N^{N-4}(t)t^2\sigma_N^{-2}\binom{N}{4}\left[E \exp\left\{it\sigma_N^{-1}(N-1)\sum_{j=1}^2 g(X_j)\right\}\psi(X_1, X_2)\right]^2 \\ &\quad + O(|N^{-1/2}t|^{2+\delta}). \end{aligned}$$

Next we expand the exponentials and find, e.g.,

$$\begin{aligned}
 & E \exp \left\{ it\sigma_N^{-1}(N-1) \sum_{j=1}^2 g(X_j) \right\} \psi(X_1, X_2) \\
 &= E \left[\prod_{j=1}^2 \left(\exp \{ it\sigma_N^{-1}(N-1)g(X_j) \} - 1 - it\sigma_N^{-1}(N-1)g(X_j) \right. \right. \\
 &\quad \left. \left. + 2it\sigma_N^{-1}(N-1) \left(\exp \{ it\sigma_N^{-1}(N-1)g(X_1) \} \right. \right. \right. \\
 &\quad \quad \left. \left. \left. - \sum_{\nu=0}^2 \{ it\sigma_N^{-1}(N-1)g(X_1) \}^\nu / \nu! \right) \right) g(X_2) \right. \\
 &\quad \left. - t^2\sigma_N^{-2}(N-1)^2 g(X_1)g(X_2) \right. \\
 &\quad \left. - it^3\sigma_N^{-3}(N-1)^3 g^2(X_1)g(X_2) \right] \psi(X_1, X_2) \\
 &= -t^2\sigma_N^{-2}(N-1)^2 E g(X_1)g(X_2) \psi(X_1, X_2) \\
 &\quad - it^3\sigma_N^{-3}(N-1)^3 E g^2(X_1)g(X_2) \psi(X_1, X_2) \\
 &\quad + O(N^{-2}t^4 + |N^{-1/2}t|^{3(1+2\epsilon)}),
 \end{aligned}$$

with ϵ as in (2.6). To see this, use (2.7), (1.15)–(1.17), (1.22), (1.6) and

$$\begin{aligned}
 & E g^2(X_1)g^2(X_2) |\psi(X_1, X_2)| \leq E g^4(X_1) \{ E \psi^2(X_1, X_2) \}^{1/2}, \\
 & E |g(X_1)|^{2+6\epsilon} |g(X_2) \psi(X_1, X_2)| \leq \left[E g^4(X_1) E |g(X_1)|^{(2+\delta)/(1+\delta)} \right]^{(1+\delta)/(2+\delta)} \\
 &\quad \times \left[E |\psi(X_1, X_2)|^{2+\delta} \right]^{1/(2+\delta)}.
 \end{aligned}$$

The other exponentials in (2.9) may be expanded in a similar fashion and after some further simplification (2.9) reduces to

$$\begin{aligned}
 \phi_N(t) &= \gamma_N^N(t) + \gamma_N^{N-2}(t) \left(-\frac{1}{2} it^3 \sigma_N^{-3} N^4 E g(X_1)g(X_2) \psi(X_1, X_2) \right. \\
 &\quad \left. + \frac{1}{2} t^4 \sigma_N^{-4} N^5 E g^2(X_1)g(X_2) \psi(X_1, X_2) - \frac{1}{4} t^2 \sigma_N^{-2} N^2 E \psi^2(X_1, X_2) \right) \\
 (2.10) \quad &+ \frac{1}{2} \gamma_N^{N-3}(t) t^4 \sigma_N^{-4} N^5 E g(X_1)g(X_2) \psi(X_1, X_3) \psi(X_2, X_3) \\
 &- \frac{1}{8} \gamma_N^{N-4}(t) t^6 \sigma_N^{-6} N^8 [E g(X_1)g(X_2) \psi(X_1, X_2)]^2 \\
 &+ O \left(|\gamma_N(t)|^{N-4} |t| P(|t|) N^{-1-3\epsilon} + |N^{-1/2}t|^{2+\delta} \right),
 \end{aligned}$$

where P is a fixed polynomial.

For $\sigma_g^2 = \mathbf{E}g^2(X_1)$ as in (1.7), let

$$\gamma(t) = \mathbf{E} \exp\{it\sigma_g^{-1}g(X_1)\}$$

denote the c.f. of $\sigma_g^{-1}g(X_1)$. From the classical theory of Edgeworth expansions for sums of i.i.d. random variables we know that (1.17) and (1.22) imply that for sufficiently small $\epsilon' > 0$ and for $|t| \leq \epsilon'N^{1/2}$,

$$(2.11) \quad \gamma^N(N^{-1/2}t) = e^{-t^2/2} \left[1 - \frac{i\tilde{\kappa}_3}{6}N^{-1/2}t^3 + \frac{\tilde{\kappa}_4}{24}N^{-1}t^4 - \frac{\tilde{\kappa}_3^2}{72}N^{-1}t^6 \right] + o(N^{-1}|t|e^{-t^2/4}),$$

where

$$\tilde{\kappa}_3 = \sigma_g^{-3}\mathbf{E}g^3(X_1), \quad \tilde{\kappa}_4 = \sigma_g^{-4}\mathbf{E}g^4(X_1) - 3$$

are the third and fourth cumulants of $\sigma_g^{-1}g(X_1)$. Since $\gamma_N(t) = \gamma(\sigma_g\sigma_N^{-1}(N-1)t)$, an easy calculation shows that for $m = 0, 2, 3, 4$,

$$(2.12) \quad \gamma_N^{N-m}(t) = \gamma^N(N^{-1/2}t) + e^{-t^2/2} \left[\frac{1}{4}\sigma_g^{-2}\mathbf{E}\psi^2(X_1, X_2) + \frac{m}{2} \right] N^{-1}t^2 + o(N^{-1}|t|e^{-t^2/4})$$

for $|t| \leq \epsilon'N^{1/2}$. Substitution of (2.11), (2.12) and (1.6) in (2.10) shows after some rearrangement that for $|t| \leq \epsilon'N^{1/2}$,

$$(2.13) \quad \phi_N(t) = \phi_N^*(t) + o(N^{-1}|t|P(|t|)e^{-t^2/4}) + O(N^{-1-\delta/2}|t|^{2+\delta}),$$

where ϕ_N^* is given by (2.2), (1.9) and (1.10) and P is a fixed polynomial. It follows that for ϵ as given by (2.6),

$$(2.14) \quad \int_{-N^\epsilon}^{N^\epsilon} \left| \frac{\phi_N(t) - \phi_N^*(t)}{t} \right| dt = o(N^{-1}).$$

Obviously,

$$\int_{|t| \geq N^\epsilon} \left| \frac{\phi_N^*(t)}{t} \right| dt = o(N^{-1})$$

and it therefore remains to be shown that for t_N as in (2.4),

$$(2.15) \quad \int_{N^\epsilon \leq |t| \leq t_N} \left| \frac{\phi_N(t)}{t} \right| dt = o(N^{-1}).$$

Define, for $m = 1, \dots, N-1$,

$$(2.16) \quad \Delta_N(m) = \sum_{i=1}^m \sum_{j=i+1}^N \psi(X_i, X_j).$$

As $E|\Delta_N(m)|^r = O((mN)^{r/2})$ [cf. Callaert and Janssen (1978) for $r = 3$], we obtain

$$(2.17) \quad |\phi_N(t)| \leq \left| E \exp\{it\sigma_N^{-1}(U_N - \Delta_N(m))\} \sum_{\nu=0}^{[r]} \frac{(it\sigma_N^{-1}\Delta_N(m))^\nu}{\nu!} \right| + O\left(\left|\frac{m^{1/2}t}{N}\right|^r\right),$$

where $[r]$ denotes the integer part of r . Since $(N - 1)\sum_{i=1}^m g(X_i)$ are the only terms in $(U_N - \Delta_N(m))$ involving X_1, \dots, X_m , we find that for $m \geq 2\nu$,

$$(2.18) \quad |E \exp\{it\sigma_N^{-1}(U_N - \Delta_N(m))\} \Delta_N^\nu(m)| \leq |\gamma_N^{m-2\nu}(t)| (mN)^\nu E|\psi(X_1, X_2)|^\nu.$$

Also, for sufficiently small $\tilde{\epsilon} > 0$ and $|t| \leq \tilde{\epsilon}N^{1/2}$, we have

$$(2.19) \quad |\gamma_N(t)| \leq 1 - \frac{t^2}{3N} \leq \exp\left\{-\frac{t^2}{3N}\right\}.$$

First take $N^\epsilon \leq |t| \leq \tilde{\epsilon}N^{1/2}$ and $m = m(t) = [3rN \log N/t^2] + 1$. For sufficiently large N , we see that indeed $1 \leq m \leq N - 1$ and (2.17)–(2.19) yield

$$(2.20) \quad |\phi_N(t)| = O\left(\left(\frac{\log N}{N}\right)^{r/2}\right)$$

for $N^\epsilon \leq |t| \leq \tilde{\epsilon}N^{1/2}$.

Next we take $\tilde{\epsilon}N^{1/2} \leq |t| \leq t_N$. In view of (1.14), (1.18) and the continuity of χ_2 , there exists $\eta > 0$ such that for sufficiently large N ,

$$(2.21) \quad |\gamma_N(t)| \leq 1 - \eta.$$

Choose $m = -r \log N / \log(1 - \eta)$. For sufficiently large N , (2.17), (2.18) and (2.21) imply that

$$(2.22) \quad |\phi_N(t)| = O((\log N)^{r/2} N^{-r} |t|^r)$$

for $\tilde{\epsilon}N^{1/2} \leq |t| \leq t_N$. Since (2.20) and (2.22) hold uniformly not only for fixed $\delta, \delta_1, C, \chi_1$ and χ_2 but also for the values of t being considered, (2.15) follows and the proof of Lemma 2.1 is complete. \square

3. The c.f. for large values of the argument. In this section we prove

LEMMA 3.1. *Suppose that (1.13), (1.15)–(1.17) and (1.19)–(1.22) are satisfied. Then there exists a sequence $\epsilon_N''' \downarrow 0$ depending only on $\delta, \delta_2, \delta_3, \delta_4, C$ and χ_1 such that for t_N as in (2.4) and*

$$(3.1) \quad T_N = N \log N,$$

$$(3.2) \quad \int_{t_N \leq |t| \leq T_N} \left| \frac{\phi_N(t)}{t} \right| dt \leq \epsilon_N''' N^{-1}.$$

PROOF. We begin by noting that (1.21) implies that $k \geq 5$ and in view of (1.15) we may assume without loss of generality that

$$(3.3) \quad 2 + \delta \leq r \leq 10 + \delta_3, \quad 5 \leq k \leq 5 + \frac{8 + \delta_3}{\delta}.$$

Though of course not essential, these restrictions make it easier to go from error bounds in terms of r and k to bounds in terms of δ and δ_3 as required in the statement of the lemma.

In Section 2, the proof that $|\phi_N(t)|$ is sufficiently small for $N^\epsilon \leq |t| \leq t_N$ was based on the fact that for these values of t the behaviour of $|\phi_N(t)|$ is still determined to some extent by that of the c.f. of \hat{U}_N , and hence by the c.f. of $g(X_1)$. For larger values of $|t|$, however, the influence of the remainder term Δ_N may completely destroy that of \hat{U}_N . It seems that we have no more use for the $g(X_i)$ and we shall remove them by a conditioning argument.

Define random variables Y_1, \dots, Y_N such that $X_1, \dots, X_N, Y_1, \dots, Y_N$ are i.i.d. and let $V_i = (X_i, Y_i)$, $i = 1, \dots, N$. Let n be an integer with $1 \leq n \leq (N - 1)/4$. Then

$$\begin{aligned}
 |\phi_N(t)|^2 &\leq E \left| E(\exp\{it\sigma_N^{-1}U_N\} | X_1, \dots, X_{4n}) \right|^2 \\
 &\leq E \left| E \left(\exp \left(it\sigma_N^{-1} \sum_{j=1}^{l-1} \sum_{l=4n+1}^N h(X_j, X_l) \right) \middle| X_1, \dots, X_{4n} \right) \right|^2 \\
 &= E \exp \left\{ it\sigma_N^{-1} \left[\sum_{j=1}^{4n} \sum_{l=4n+1}^N (h(X_j, X_l) - h(X_j, Y_l)) \right. \right. \\
 &\quad \left. \left. + \sum_{j=4n+1}^{l-1} \sum_{l=4n+1}^N (h(X_j, X_l) - h(Y_j, Y_l)) \right] \right\} \\
 (3.4) \quad &\leq E \left| E \left(\exp \left(it\sigma_N^{-1} \sum_{j=1}^{4n} \sum_{l=4n+1}^N (h(X_j, X_l) - h(X_j, Y_l)) \right) \middle| V_{4n+1}, \dots, V_N \right) \right|^2 \\
 &= E \exp \left\{ it\sigma_N^{-1} \sum_{j=1}^{2n} \sum_{l=4n+1}^N (h(X_j, X_l) \right. \\
 &\quad \left. - h(Y_j, X_l) - h(X_j, Y_l) + h(Y_j, Y_l)) \right\} \\
 &= E \exp \left\{ it\sigma_N^{-1} \sum_{j=1}^{2n} \sum_{l=4n+1}^N (\psi(X_j, X_l) \right. \\
 &\quad \left. - \psi(Y_j, X_l) - \psi(X_j, Y_l) + \psi(Y_j, Y_l)) \right\} \\
 &= E \exp \left\{ it\sigma_N^{-1} \sum_{j=1}^{2n} \sum_{l=4n+1}^N \Psi(V_j, V_l) \right\},
 \end{aligned}$$

where for $v_j = (x_j, y_j)$, $j = 1, \dots, N$, we have defined

$$(3.5) \quad \Psi(v_j, v_l) = \psi(x_j, x_l) - \psi(y_j, x_l) - \psi(x_j, y_l) + \psi(y_j, y_l).$$

Our next step is to truncate the random variables $X_1, \dots, X_{2n}, Y_1, \dots, Y_{2n}$, while losing half of them in the process. Consider a measurable set $B \in \mathcal{B}$ with

$$(3.6) \quad \alpha = P(V_1 \in B \times B) = P_X^2(B).$$

For every $\alpha \in [0, 1]$, $x \in [0, 1]$ and $\rho \in (0, 1)$ we have

$$(3.7) \quad \alpha x + (1 - \alpha) \leq x^\rho \vee \left(\frac{\rho}{\alpha}\right)^{\rho/(1-\rho)},$$

where $(x \vee y)$ denotes the larger of x and y . It follows from (3.4)–(3.7) that

$$\begin{aligned} |\phi_N(t)|^2 &\leq E \left[E \left(\exp \left(it\sigma_N^{-1} \sum_{l=4n+1}^N \Psi(V_1, V_l) \right) \middle| V_{4n+1}, \dots, V_N \right) \right]^{2n} \\ &\leq E \left[\alpha E \left(\exp \left(it\sigma_N^{-1} \sum_{l=4n+1}^N \Psi(V_1, V_l) \right) \middle| V_{4n+1}, \dots, V_N; V_1 \in B \times B \right) \right. \\ &\qquad \qquad \qquad \left. + (1 - \alpha) \right]^{2n} \\ &\leq E \left[E \left(\exp \left(it\sigma_N^{-1} \sum_{l=4n+1}^N \Psi(V_1, V_l) \right) \middle| V_{4n+1}, \dots, V_N; V_1 \in B \times B \right) \right]^{2\rho n} \\ &\qquad \qquad \qquad + \left(\frac{\rho}{\alpha}\right)^{2\rho n/(1-\rho)} \end{aligned} \tag{3.8}$$

for every $\rho \in (0, 1)$. Take $\rho = \frac{1}{2}$ and define $\tilde{V}_j = (\tilde{X}_j, \tilde{Y}_j)$, $j = 1, \dots, n$, in such a way that $\tilde{X}_1, \dots, \tilde{X}_n, \tilde{Y}_1, \dots, \tilde{Y}_n$ are i.i.d. with common distribution

$$(3.9) \quad P(\tilde{X}_j \in A) = P(\tilde{Y}_j \in A) = \frac{P_X(A \cap B)}{P_X(B)}$$

and independent of V_{4n+1}, \dots, V_N . Then (3.8) may be rewritten as

$$\begin{aligned} (3.10) \quad |\phi_N(t)|^2 &\leq E \exp \left\{ it\sigma_N^{-1} \sum_{j=1}^n \sum_{l=4n+1}^N \Psi(\tilde{V}_j, V_l) \right\} + (2\alpha)^{-2n} \\ &= E \left[E \left(\exp \{ it\sigma_N^{-1} Z_n \} \middle| \tilde{V}_1, \dots, \tilde{V}_n \right) \right]^{N-4n} + (2\alpha)^{-2n}, \end{aligned}$$

where

$$\begin{aligned}
 (3.11) \quad Z_n &= \sum_{j=1}^n \Psi(\tilde{V}_j, V_N) \\
 &= \sum_{j=1}^n [\psi(\tilde{X}_j, X_N) - \psi(\tilde{Y}_j, X_N) - \psi(\tilde{X}_j, Y_N) + \psi(\tilde{Y}_j, Y_N)].
 \end{aligned}$$

It remains to choose the set B and we take

$$(3.12) \quad B = \left\{ x: \int |\psi(x, y)|^r dP_X(y) \leq C_1 \tau \right\},$$

for a large but fixed $\tau > 0$ to be specified later.

Let us now consider the conditional expectation in (3.10). Since $|\exp\{ix\} - 1 - ix + \frac{1}{2}x^2| \leq x^2/6 + |x|^r$ for $r > 2$, we have

$$\begin{aligned}
 (3.13) \quad 0 \leq E(\exp\{it\sigma_N^{-1}Z_n\}|\tilde{V}_1, \dots, \tilde{V}_n) &\leq 1 - \frac{1}{3}t^2\sigma_N^{-2}E(Z_n^2|\tilde{V}_1, \dots, \tilde{V}_n) \\
 &\quad + |t|^r\sigma_N^{-r}E(|Z_n|^r|\tilde{V}_1, \dots, \tilde{V}_n).
 \end{aligned}$$

By (3.11) and (3.12)

$$(3.14) \quad E(\Psi(\tilde{V}_1, V_N)|\tilde{V}_1) = E(\Psi(\tilde{V}_1, V_N)|V_N) = 0 \quad \text{a.s.},$$

$$(3.15) \quad E(|\Psi(\tilde{V}_1, V_N)|^r|\tilde{V}_1) \leq 4^r C_1 \tau \quad \text{a.s.}$$

It follows from Lemma A.1 in the Appendix together with (3.3) that for every integer $m \geq 1$

$$E [E(|Z_n|^r|\tilde{V}_1, \dots, \tilde{V}_n)]^m = O(n^{rm/2}).$$

Taking $m = 10k/\delta_3$, we find by (3.3) and Markov's inequality that

$$(3.16) \quad P(E(|Z_n|^r|\tilde{V}_1, \dots, \tilde{V}_n) \geq n^{r/2}N^{\delta_3/(4k)}) = O(N^{-5/2}).$$

Next we turn to the quadratic term in (3.13). Let $\lambda_1, \lambda_2, \dots$ be the eigenvalues of ψ with respect to P_X with $|\lambda_1| \geq |\lambda_2| \geq \dots$ and let $\omega_1, \omega_2, \dots$ be an orthonormal sequence of eigenfunctions corresponding to $\lambda_1, \lambda_2, \dots$, i.e., (1.19) holds and for all ν and ν' ,

$$(3.17) \quad \int \omega_\nu(x) dP_X(x) = 0, \quad \int \omega_\nu(x)\omega_{\nu'}(x) dP_X(x) = \delta_{\nu, \nu'},$$

where $\delta_{\nu, \nu'} = 0$ or 1 according as $\nu \neq \nu'$ or $\nu = \nu'$. Assume (1.20) is satisfied. We have

$$(3.18) \quad \psi(x, y) = \sum_{\nu=1}^k \lambda_\nu \omega_\nu(x)\omega_\nu(y) + R(x, y),$$

where R is a symmetric function of its two variables satisfying

$$(3.19) \quad \int R(x, y)\omega_\nu(y) dP_X(y) = 0 \quad \text{for } \nu = 1, \dots, k.$$

As a consequence we find

$$\begin{aligned}
 E(Z_n^2 | \tilde{V}_1, \dots, \tilde{V}_n) &= 2 \int \left[\sum_{j=1}^n \{ \psi(\tilde{X}_j, y) - \psi(\tilde{Y}_j, y) \} \right]^2 dP_X(y) \\
 &= 2 \int \left[\sum_{j=1}^n \left\{ \sum_{\nu=1}^k \lambda_\nu \omega_\nu(y) (\omega_\nu(\tilde{X}_j) - \omega_\nu(\tilde{Y}_j)) \right. \right. \\
 &\qquad \qquad \qquad \left. \left. + R(\tilde{X}_j, y) - R(\tilde{Y}_j, y) \right\} \right]^2 dP_X(y) \\
 &= 2 \sum_{j=1}^n \sum_{j'=1}^n \sum_{\nu=1}^k \sum_{\nu'=1}^k \int [\lambda_\nu \omega_\nu(y) (\omega_\nu(\tilde{X}_j) - \omega_\nu(\tilde{Y}_j)) \\
 &\qquad \qquad \qquad + R(\tilde{X}_j, y) - R(\tilde{Y}_j, y)] \\
 &\qquad \qquad \qquad \times [\lambda_{\nu'} \omega_{\nu'}(y) (\omega_{\nu'}(\tilde{X}_{j'}) - \omega_{\nu'}(\tilde{Y}_{j'})) \\
 &\qquad \qquad \qquad + R(\tilde{X}_{j'}, y) - R(\tilde{Y}_{j'}, y)] dP_X(y) \\
 &= 2 \sum_{\nu=1}^k \lambda_\nu^2 \left[\sum_{j=1}^n \{ \omega_\nu(\tilde{X}_j) - \omega_\nu(\tilde{Y}_j) \} \right]^2 \\
 &\qquad \qquad \qquad + 2k^2 \int \left[\sum_{j=1}^n \{ R(\tilde{X}_j, y) - R(\tilde{Y}_j, y) \} \right]^2 dP_X(y) \\
 &\geq 2\delta_2^2 \sum_{\nu=1}^k \left[\sum_{j=1}^n \{ \omega_\nu(\tilde{X}_j) - \omega_\nu(\tilde{Y}_j) \} \right]^2.
 \end{aligned}
 \tag{3.20}$$

We shall have to investigate the covariance matrix Σ of the random vector $(\omega_1(\tilde{X}_1) - \omega_1(\tilde{Y}_1), \dots, \omega_k(\tilde{X}_1) - \omega_k(\tilde{Y}_1))$. First note that (1.16) and (1.20) imply that for $\nu = 1, \dots, k$

$$\begin{aligned}
 E|\omega_\nu(X_1)|^r &\leq \lambda_\nu^{-r} E \left| \int \psi(X_1, y) \omega_\nu(y) dP_X(y) \right|^r \\
 &\leq \delta_2^{-r} E \left\{ \int \psi^2(X_1, y) dP_X(y) \right\}^{r/2} \leq \delta_2^{-r} C.
 \end{aligned}
 \tag{3.21}$$

Let $\sigma_{\nu, \nu'} = E(\omega_\nu(\tilde{X}_1) - \omega_\nu(\tilde{Y}_1))(\omega_{\nu'}(\tilde{X}_1) - \omega_{\nu'}(\tilde{Y}_1))$, $\nu, \nu' = 1, \dots, k$, denote the elements of Σ . For $\nu \neq \nu'$, (3.21) and Hölder's inequality ensure that

$$\begin{aligned}
 |\sigma_{\nu, \nu'}| &\leq \left| \frac{1}{\alpha} E(\omega_\nu(X_1) - \omega_\nu(Y_1))(\omega_{\nu'}(X_1) - \omega_{\nu'}(Y_1)) 1_{(B \times B)^c}(X_1, Y_1) \right| \\
 &\leq 4\delta_2^{-2} C^{2/r} \alpha^{-1} (1 - \alpha)^{(r-2)/r},
 \end{aligned}
 \tag{3.22}$$

whereas for $\nu = \nu'$ we find similarly

$$(3.23) \quad -4\delta_2^{-2}C^2/r\alpha^{-1}(1-\alpha)^{(r-2)/r} \leq \sigma_{\nu,\nu} - 2 \leq 2\alpha^{-1}(1-\alpha).$$

Now we may still choose τ in (3.12) and since

$$(3.24) \quad 1 - \alpha^{1/2} = P_X(B^c) \leq \tau^{-1},$$

by (3.12) and Markov's inequality, we can force α to be arbitrarily close to 1 by taking τ large. In view of (3.22)–(3.24) and (3.3), we can choose $\tau = \tau(\delta, \delta_2, \delta_3, C)$ in such a way that

$$(3.25) \quad 2\alpha \geq e^{1/2},$$

$$(3.26) \quad |\sigma_{\nu,\nu'} - 2\delta_{\nu,\nu'}| \leq k^{-1} \quad \text{for all } \nu, \nu' = 1, \dots, k.$$

If ρ_k denotes the smallest eigenvalue of Σ , then (3.26) yields

$$(3.27) \quad \rho_k \geq 2 - \left\{ \sum_{\nu=1}^k \sum_{\nu'=1}^k (\sigma_{\nu,\nu'} - 2\delta_{\nu,\nu'})^2 \right\}^{1/2} \geq 1.$$

Also (3.21), (3.25) and (3.3) imply that $E|\omega_\nu(\tilde{X}_1) - \omega_\nu(\tilde{Y}_1)|^{2+\delta}$, $\nu = 1, \dots, k$, as well as k are bounded. It follows that we may apply Lemma A.2 in the Appendix to the right-hand side of (3.20) to obtain

$$(3.28) \quad \begin{aligned} P(E(Z_n^2 | \tilde{V}_1, \dots, \tilde{V}_n) \leq nN^{-4/k}(\log N)^{-6/k}) \\ = O(N^{-2}(\log N)^{-3} + n^{-k/2}). \end{aligned}$$

Let us now combine the results obtained in (3.10), (3.13), (3.16) and (3.28). First we note that (3.25) ensures that the term $(2\alpha)^{-2n}$ in (3.10) is $O(e^{-n})$ and that σ_N^2 is of exact order N^3 by (1.16), (1.17) and (1.22). Take t_N and T_N as in (2.4) and (3.1), choose any t such that $t_N \leq |t| \leq T_N$ and then define

$$(3.29) \quad n = n(t) = \left\lfloor \frac{\sigma_N^2 N^{(4/k)-1} (\log N)^{2+(6/k)}}{t^2} \right\rfloor,$$

where $[x]$ denotes the integer part of x . As $t_N \leq |t| \leq T_N$, it follows from (2.4), (3.1) and (1.21) that

$$(3.30) \quad \begin{aligned} \sigma_N^2 N^{(4/k)-3} (\log N)^{6/k} - 1 \leq n = O(N^{(4/k)+(2/r)} (\log N)^{4+(6/k)}) \\ = O(N^{1-\delta_3/(kr)} (\log N)^6), \end{aligned}$$

and in view of (3.3) this means that $1 \leq n \leq (N-1)/4$ for sufficiently large N , so that (3.29) is indeed a possible choice of n . Similarly, one easily checks that (3.29), (1.21) and (3.3) imply that for sufficiently large N ,

$$(3.31) \quad |t|^r \sigma_N^{-r} n^{r/2} N^{\delta_3/(4k)} \leq \frac{1}{6} t^2 \sigma_N^{-2} n N^{-4/k} (\log N)^{-6/k}.$$

Together (3.10), (3.13), (3.16), (3.28), (3.31), (3.29) and (3.30) show that for

sufficiently large N

$$\begin{aligned} |\phi_N(t)|^2 &\leq \left[1 - \frac{1}{6}t^2 n \sigma_N^{-2} N^{-4/k} (\log N)^{-6/k}\right]^{N-4n} + O(N^{-2}(\log N)^{-3} + n^{-k/2}) \\ &\leq \exp\left\{-\frac{1}{6}N^{-1}(N-4n)(\log N)^2\right\} + O(N^{-2}(\log N)^{-3} + n^{-k/2}) \\ &= O(N^{-2}(\log N)^{-3}), \end{aligned}$$

so that

$$(3.32) \quad |\phi_N(t)| = O(N^{-1}(\log N)^{-3/2})$$

uniformly for $t_N \leq |t| \leq T_N$. This proves Lemma 3.1 and Theorem 1.1 at the same time. \square

4. The eigenvalue assumption. In Section 1 we noted that the meaning of assumption (1.20) concerning the eigenvalues of ψ , is not intuitively clear. From the analysis in Sections 2 and 3, however, we can at least see the part that it plays in the proof of Theorem 1.1. As we pointed out at the beginning of the proof of Lemma 3.1, the analysis of $|\phi_N(t)|$ for $|t| \leq N^{(r-1)/r}(\log N)^{-1}$ proceeds by showing that up to that point, the properties of \hat{U}_N determine the behaviour of $|\phi_N(t)|$, because the influence of $|t|\sigma_N^{-1}\Delta_N$ is still small. For larger values of $|t|$, \hat{U}_N does not play a role any longer and we have to show that $|t|\sigma_N^{-1}\Delta_N$ is large enough to take over the task of making $|\phi_N(t)|$ small. Since, in general, sums of independent random variables can be unpleasantly close to zero with probabilities that are nonnegligible for our purposes, assumption (1.20) is there to prevent this.

Still, we are unable to show that without assumption (1.20), the theorem would indeed fail. Our search for a counterexample, however, has convinced us that such an example would have to be extremely pathological.

To compute the eigenvalues $\lambda_1, \dots, \lambda_k$ of ψ can of course be laborious, but fortunately this is not necessary in order to verify assumption (1.20). Consider functions f_1, \dots, f_k with

$$(4.1) \quad \int f_j^2(x) dP_X(x) \leq 1, \quad j = 1, \dots, k,$$

and define random variables

$$(4.2) \quad W_j = \int \psi(X_1, y) f_j(y) dP_X(y).$$

Let Σ_W denote the covariance matrix of the random vector $W = (W_1, \dots, W_k)$ and suppose that it has a smallest eigenvalue $\tilde{\lambda}_k$ satisfying

$$(4.3) \quad \tilde{\lambda}_k \geq \delta_2 > 0.$$

LEMMA 4.1. *Suppose that in the set of conditions of Theorem 1.1 we replace (1.19)–(1.20) by the assumption that f_1, \dots, f_k exist such that (4.1)–(4.3) are satisfied. Then the set of conditions obtained is equivalent to the original set.*

PROOF. If (1.19)–(1.20) hold, we may choose $f_j = \omega_j$, $W_j = \lambda_j \omega_j(X_1)$ and hence $\tilde{\lambda}_k = \lambda_k^2 \geq \delta_2^2$. Replacing δ_2 by $\delta_2^{1/2}$ yields (4.3).

Conversely, suppose that (4.1) and (4.3) hold for certain f_1, \dots, f_k . Let \mathcal{F} denote the linear space spanned by f_1, \dots, f_k and define $\|f\|$ and Tf by

$$\|f\|^2 = \int f^2 dP_X, \quad (Tf)(x) = \int \psi(x, y) f(y) dP_X(y).$$

For $f = \sum_{j=1}^k c_j f_j$, we have

$$\begin{aligned} \|Tf\|^2 &= E \left(\sum_{j=1}^k c_j W_j \right)^2 = c' \Sigma_W c \geq \tilde{\delta}_2 \sum_{j=1}^k c_j^2, \\ \|f\|^2 &\leq \sum_{j=1}^k c_j^2 \sum_{j=1}^k \|f_j\|^2 \leq k \sum_{j=1}^k c_j^2 \end{aligned}$$

in view of (4.2), (4.3) and (4.1). Together this yields

$$(4.4) \quad \|Tf\|^2 \geq \frac{\tilde{\delta}_2}{k} \|f\|^2 \quad \text{for every } f \in \mathcal{F}.$$

On the other hand, (4.3) ensures that f_1, \dots, f_k are linearly independent in $L_2(P_X)$ and hence \mathcal{F} must contain functions orthogonal to $\omega_1, \omega_2, \dots, \omega_{k-1}$ defined in (1.19). But this implies that

$$(4.5) \quad \inf_{f \in \mathcal{F}} \frac{\|Tf\|^2}{\|f\|^2} \leq \lambda_k^2,$$

where $\lambda_1, \lambda_2, \dots$, are given by (1.19). Combining (4.4) and (4.5) we find

$$(4.6) \quad |\lambda_k| \geq \left(\frac{\tilde{\delta}_2}{k} \right)^{1/2}.$$

Because of (1.15) and (1.21) we may assume k to be bounded [cf. (3.3)] and the proof is complete. \square

Of course (4.3) will usually be easier to verify than (1.20). The situation is even simpler in Corollary 1.1 or, more generally, in all cases where ψ is fixed. Assumption (1.27) may then be replaced by the nonsingularity of Σ_W , i.e., by the fact that W_1, \dots, W_k are not almost surely linearly dependent. A simple sufficient condition for the existence of such W_1, \dots, W_k is that there exist points y_1, \dots, y_k in the support of F such that the functions $\psi(\cdot, y_1), \dots, \psi(\cdot, y_k)$ are linearly independent.

5. An example. Let X_1, \dots, X_N be i.i.d. random variables with a common continuous d.f. F on \mathbb{R} . Let

$$(5.1) \quad R_i^+ = \sum_{j=1}^N 1_{\{|X_j| \leq |X_i\}}$$

and let W_N^+ denote Wilcoxon's one-sample signed rank statistic for testing the hypothesis that the distribution of X_1 is symmetric about zero, thus

$$(5.2) \quad W_N^+ = \sum_{i=1}^N 1_{\{X_i \geq 0\}} R_i^+.$$

If we define

$$(5.3) \quad \begin{aligned} U_N &= W_N^+ - EW_N^+ \\ &= W_N^+ - N(1 - F(0)) \\ &\quad - N(N - 1) \int_0^\infty (F(x) - F(-x)) dF(x), \end{aligned}$$

then U_N is clearly a U -statistic. An easy computation yields

$$(5.4) \quad U_N = (N - 1) \sum_{i=1}^N g_N(X_i) + \sum_{i=1}^N \sum_{j=i+1}^N \psi(X_i, X_j),$$

where

$$(5.5) \quad \begin{aligned} g_N(x) &= 1 - F(-x) - \int (1 - F(-x)) dF(x) \\ &\quad + \frac{1}{N - 1} \{1_{[0, \infty)}(x) - 1 + F(0)\}, \end{aligned}$$

$$(5.6) \quad \begin{aligned} \psi(x, y) &= 1_{[0, \infty)}(x + y) - (1 - F(-x)) - (1 - F(-y)) \\ &\quad + \int (1 - F(-x)) dF(x). \end{aligned}$$

Note that $Eg_N(X_1) = 0$ and $E(\psi(X_1, X_2)|X_1) = 0$ a.s.

Having decomposed U_N in the manner of Section 1, we check the conditions of Theorem 1.1. Since both g_N and ψ are bounded, (1.13) and (1.15)–(1.17) are satisfied for every r . Next, (1.14) and (1.18) will hold if the distribution of $F(-X_1)$ has an absolutely continuous component. It remains to verify (1.20) for some $k \geq 5$. In view of Lemma 4.1 and the fact that ψ does not depend on N it suffices to find functions f_1, \dots, f_k with $\int f_j^2 dF \leq 1$ such that the random variables,

$$(5.7) \quad W_j = \int \psi(X_1, y) f_j(y) dF(y), \quad j = 1, \dots, k,$$

are not almost surely linearly dependent. Take

$$(5.8) \quad f_j(x) = F^j(x), \quad j = 1, \dots, k,$$

so that

$$(5.9) \quad W_j = \frac{1}{j + 1} \left\{ F(-X_1) - F^{j+1}(-X_1) - \int (F(-x) - F^{j+1}(-x)) dF(x) \right\}.$$

Then

$$(5.10) \quad \sum_{j=1}^k c_j W_j = \sum_{i=0}^{k+1} a_i F^i(-X_1),$$

with

$$a_0 = - \sum_{j=1}^k c_j \int (F(-x) - F^{j+1}(-x)) dF(x),$$

$$a_1 = \sum_{j=1}^k \frac{c_j}{j+1}, \quad a_i = - \frac{c_{j-1}}{j} \quad \text{for } i = 2, \dots, k+1.$$

Since the distribution of $F(-X_1)$ is supposed to have an absolutely continuous part, (5.10) can vanish almost surely only if $a_0 = a_1 = \dots = a_{k+1} = 0$ which implies $c_1 = \dots = c_k = 0$. It follows that assumption (1.20) holds every k .

Thus we have established the validity of the Edgeworth expansion with remainder $o(N^{-1})$ for Wilcoxon's one-sample rank statistic under the assumptions that F is continuous and that the distribution of $F(-X_1)$ has an absolutely continuous component. We stress the fact that previous results on Edgeworth expansions for U -statistics would fail in this case because U_N has a pure lattice distribution. Edgeworth expansions for one-sample rank statistics were obtained in Albers, Bickel and van Zwet (1976) by a completely different method.

APPENDIX

In this appendix we prove a moment inequality and a concentration inequality which are needed in Section 3 of the present paper, but which may also be of independent interest.

LEMMA A.1. *Let P and Q be probability measures on arbitrary sample spaces \mathcal{X} and \mathcal{Y} and let X_1, \dots, X_n be i.i.d. with common distribution P . Let $\psi: \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}$ satisfy $\int \psi(x, y) dP(x) = 0$ for Q -almost all $y \in \mathcal{Y}$, and $\int \psi(x, y) dQ(y) = 0$ for P -almost all $x \in \mathcal{X}$. Then, for every real $p \geq 2$ and integer $k \geq 1$, there exists a positive number $A = A(p, k)$ which is bounded for bounded p and k and such that*

$$E \left\{ \int \left| \sum_{i=1}^n \psi(X_i, y) \right|^p dQ(y) \right\}^k \leq A n^{kp/2} E \left\{ \int |\psi(X_1, y)|^p dQ(y) \right\}^k.$$

PROOF. If the expectation on the right equals $+\infty$, then there is nothing to prove. Assume therefore that

$$C = E \left\{ \int |\psi(X_1, y)|^p dQ(y) \right\}^k < \infty.$$

Let Y_1, \dots, Y_k be i.i.d. with common distribution Q and independent of

X_1, \dots, X_n . Then

$$\begin{aligned}
 B &= E \left\{ \int \left| \sum_{i=1}^n \psi(X_i, y) \right|^p dQ(y) \right\}^k = E \left| \prod_{j=1}^k \sum_{i=1}^n \psi(X_i, Y_j) \right|^p \\
 &= E \left| \sum_{i_1=1}^n \cdots \sum_{i_k=1}^n \prod_{j=1}^k \psi(X_{i_j}, Y_j) \right|^p.
 \end{aligned}$$

Let m_1, \dots, m_r be integers ≥ 2 with $\sum_{v=1}^r m_v = k - l$, $l \geq 0$, and let $I(m_1, \dots, m_r)$ denote the collection of sequences $i_1, \dots, i_k \in \{1, 2, \dots, n\}$ which contain $(l + r)$ distinct values, out of which l occur with multiplicity 1 and r with multiplicities m_1, \dots, m_r , respectively. Define

$$(A.1) \quad Z(m_1, \dots, m_r) = \sum \cdots \sum_{(i_1, \dots, i_k) \in I(m_1, \dots, m_r)} \prod_{j=1}^k \psi(X_{i_j}, Y_j),$$

and note that each term in this sum has the same distribution. There are at most n^r different ways of choosing the indices with multiplicities m_1, \dots, m_r and at most $k!$ different ways of permuting i_1, \dots, i_k . It follows that

$$\begin{aligned}
 (A.2) \quad E|Z(m_1, \dots, m_r)|^p &\leq (k!n^r)^p E \left| \sum_{1 \leq i_1 < \dots < i_l \leq n-r} \prod_{j=1}^l \psi(X_{i_j}, Y_j) \right. \\
 &\quad \left. \times \prod_{j=1}^{m_1} \psi(X_{n-r+1}, Y_{l+j}) \times \cdots \times \prod_{j=1}^{m_r} \psi(X_n, Y_{l+m_1+\dots+m_{r-1}+j}) \right|^p \\
 &\leq (k!n^r)^p \prod_{v=1}^r E \left\{ \int |\psi(X_1, y)|^p dQ(y) \right\}^{m_v} \\
 &\quad \times E \left| \sum_{1 \leq i_1 < \dots < i_l \leq n-r} \prod_{j=1}^l \psi(X_{i_j}, Y_j) \right|^p \\
 &\leq (k!n^r)^p C^{(k-l)/k} E|W_l(n-r)|^p,
 \end{aligned}$$

where, for $l = 1, 2, \dots, k$ and $t = l, l + 1, \dots, n$,

$$W_l(t) = \sum \cdots \sum_{1 \leq i_1 < i_2 < \dots < i_l \leq t} \prod_{j=1}^l \psi(X_{i_j}, Y_j)$$

and we define $W_l(l - 1) = 0$ for $l = 1, 2, \dots, k$ and $W_0(t) = 1$ for $t = 0, 1, \dots, n$.

For fixed $l \geq 1$, $W_l(t)$, $t = l - 1, l, \dots, n$, is a martingale with $W_l(l - 1) = 0$. It follows from an inequality of Dharmadhikari, Fabian and Jogdeo (1968) that

for $l \geq 1$ and $t = l, l + 1, \dots, n$

$$\begin{aligned}
 E|W_l(t)|^p &\leq a(p)(t-l+1)^{p/2-1} \sum_{s=l}^t E|W_l(s) - W_l(s-1)|^p \\
 &= a(p)(t-l+1)^{p/2-1} E|\psi(X_1, Y_1)|^p \sum_{s=l}^t E|W_{l-1}(s-1)|^p \\
 (A.3) \quad &\leq \left\{ a(p)t^{p/2-1} E|\psi(X_1, Y_1)|^p \right\}^l \sum_{s_1=l}^t \sum_{s_2=l-1}^{s_1-1} \cdots \sum_{s_l=1}^{s_{l-1}-1} 1 \\
 &\leq \left\{ a(p) E|\psi(X_1, Y_1)|^p \right\}^l t^{lp/2}
 \end{aligned}$$

for $a(p) = 2^{2p^2}$. Clearly (A.3) will continue to hold for $l = 0$ and $t = 0, \dots, n$ provided we define $0^0 = 1$. Combining this with (A.2) we find

$$E|Z(m_1, \dots, m_r)|^p \leq (k!n^r)^p C^{(k-l)/k} a^l(p) C^{l/k} n^{lp/2}$$

and as $2r + l \leq k$,

$$B \leq (k!)^{p+1} 2^{2kp^2} n^{kp/2} C.$$

The lemma is proved. \square

LEMMA A.2. *Let X_1, \dots, X_n be i.i.d. k -dimensional random vectors with common distribution P with a positive definite covariance matrix Σ with smallest eigenvalue ρ_k . Define $S_n = n^{-1/2} \sum_{i=1}^n X_i$. Then there exists a positive number B depending only on k and P such that for every $\epsilon > 0$ and $n = 1, 2, \dots$,*

$$P(\|S_n\| \leq \epsilon) \leq B(\epsilon^k + n^{-k/2}).$$

B is constant over any class of distributions with k bounded, ρ_k bounded away from zero and $E\|X\|^{2+\delta}$ bounded for a fixed $\delta > 0$.

PROOF. Let \tilde{P} be the distribution of $(X_1 - X_2)$ and for $t \in \mathbb{R}^k$ let

$$\psi(t) = E \exp\{it'(X_1 - EX_1)\}.$$

It follows that

$$\begin{aligned}
 |\psi(t)|^2 &= \int_{\mathbb{R}^k} e^{it'x} d\tilde{P}(x) = \int_{\mathbb{R}^k} \cos(t'x) d\tilde{P}(x), \\
 \frac{1 - |\psi(t)|^2}{\|t\|^2} &\geq \int_{\|x\| \leq \theta^{-1}} \frac{1 - \cos(t'x)}{\|t\|^2} d\tilde{P}(x)
 \end{aligned}$$

for every $\theta > 0$. For $\|t\| \leq \theta$, we have $|t'x| \leq 1$ for $\|x\| \leq \theta^{-1}$ and $1 - \cos(t'x) \geq \frac{1}{2} \cos(1)(t'x)^2$. Hence for $\|t\| \leq \theta$ we see that

$$\begin{aligned}
 (A.4) \quad \frac{1 - |\psi(t)|^2}{\|t\|^2} &\geq \frac{1}{2} \cos(1) \int_{\|x\| \leq \theta^{-1}} \frac{(t'x)^2}{\|t\|^2} d\tilde{P}(x) \\
 &= \frac{1}{2} \cos(1) \int_{\|x\| \leq \theta^{-1}} (\tau'x)^2 d\tilde{P}(x),
 \end{aligned}$$

where $\tau = t/\|t\|$, so that $\|\tau\| = 1$. By the dominated convergence theorem we obtain

$$\lim_{\theta \downarrow 0} \int_{\|x\| \leq \theta^{-1}} (\tau'x)^2 d\tilde{P}(x) = E\{\tau'(X_1 - X_2)\}^2 = 2\sigma^2(\tau'X_1) \geq 2\rho_k,$$

and hence for sufficiently small $\theta_0 > 0$ and $\|t\| \leq \theta_0$, we find

$$(A.5) \quad |\psi(t)|^2 \leq 1 - \frac{1}{2}\rho_k \cos(1)\|t\|^2 \leq \exp\left\{-\frac{1}{2}\rho_k \cos(1)\|t\|^2\right\}.$$

Let $U = (U_1, \dots, U_k)$ be a random vector which is independent of X_1, \dots, X_N and which has i.i.d. components U_1, \dots, U_k with a common density $g(u) = (1 - \cos u)/(\pi u^2)$ and corresponding c.f.

$$\gamma(t) = Ee^{itU_1} = (1 - |t|)1_{[0,1]}(|t|).$$

Choose a_k such that $P(|U_1| \leq a_k) = 2^{-1/k}$ and $\varepsilon \geq a_k k^{1/2}/(\theta_0 n^{1/2})$. It is clear that

$$\begin{aligned} P(\|S_n + n^{-1/2}\theta_0^{-1}U\| \leq 2\varepsilon) &\geq P(\|U\| \leq \varepsilon\theta_0 n^{1/2})P(\|S_n\| \leq \varepsilon) \\ &\geq P(\|U\| \leq a_k k^{1/2})P(\|S_n\| \leq \varepsilon) \geq \frac{1}{2}P(\|S_n\| \leq \varepsilon), \end{aligned}$$

and using (A.5) we arrive at

$$\begin{aligned} (A.6) \quad P(\|S_n\| \leq \varepsilon) &\leq 2P(\|S_n + n^{-1/2}\theta_0^{-1}U\| \leq 2\varepsilon) \\ &\leq \frac{2}{\pi^k} \int_{\mathbb{R}^k} \prod_{j=1}^k \left| \frac{\sin(2\varepsilon t_j)}{t_j} \right| |\psi(n^{-1/2}t)|^n \prod_{j=1}^k \gamma\left(\frac{t_j}{\theta_0 n^{1/2}}\right) dt \\ &\leq \frac{2}{\pi^k} \int_{-\theta_0 n^{1/2}}^{\theta_0 n^{1/2}} \dots \int_{-\theta_0 n^{1/2}}^{\theta_0 n^{1/2}} (2\varepsilon)^k \exp\left\{-\frac{1}{4}\rho_k \cos(1)\|t\|^2\right\} dt \\ &\leq 2(2\varepsilon/\pi)^k \int_{\mathbb{R}^k} \exp\left\{-\frac{1}{4}\rho_k \cos(1)\|t\|^2\right\} dt = 2^{2k+1}(\pi \cos(1)\rho_k)^{-k/2} \varepsilon^k \end{aligned}$$

for all $|\varepsilon| \geq a_k k^{1/2}/(\theta_0 n^{1/2})$. For $|\varepsilon| < a_k k^{1/2}/(\theta_0 n^{1/2})$ (A.6) yields the trivial bound

$$\begin{aligned} (A.7) \quad P(\|S_n\| \leq \varepsilon) &\leq P(\|S_n\| \leq a_k k^{1/2}/(\theta_0 n^{1/2})) \\ &\leq 2^{2k+1} \left[\frac{ka_k^2}{\pi \cos(1)\rho_k \theta_0^2} \right]^{k/2} n^{-k/2}. \end{aligned}$$

Addition of (A.6) and (A.7) proves the lemma for fixed P .

If we assume that $E\|X_1\|^{2+\delta} \leq C$, then this implies that for every $\theta > 0$ and $\|\tau\| = 1$,

$$\int_{\|x\| > \theta^{-1}} (\tau'x)^2 d\tilde{P}(x) \leq \int_{\|x\| > \theta^{-1}} \|x\|^2 d\tilde{P}(x) \leq 2^{(2+\delta)}C\theta^\delta.$$

Returning to (A.4)–(A.5) we now see that we can choose

$$(A.8) \quad \theta_0 = 2^{-(2+\delta)/\delta}(\rho_k/C)^{1/\delta}$$

and ensure the validity of (A.5) for $\|t\| \leq \theta_0$. Substituting (A.8) in (A.7) and assuming in addition that k and ρ_k^{-1} are bounded, we conclude that B in Lemma A.2 is also bounded and the proof of the lemma is complete. \square

Acknowledgments. It is a pleasure to acknowledge the help of J. Bretagnolle which very much improved our proof of the concentration inequality of Lemma A.2. The authors are indebted to the Associate Editor and the referees for their constructive criticism.

REFERENCES

- ALBERS, W., BICKEL, P. J. and VAN ZWET, W. R. (1976). Asymptotic expansions for the power of distribution-free tests in the one-sample problem. *Ann. Statist.* **4** 108–156.
- BICKEL, P. J. (1974). Edgeworth expansions in nonparametric statistics. *Ann. Statist.* **2** 1–20.
- CALLAERT, H. and JANSSEN, P. (1978). The Berry–Esseen theorem for U -statistics. *Ann. Statist.* **6** 417–421.
- CALLAERT, H., JANSSEN, P. and VERAVERBEKE, N. (1980). An Edgeworth expansion for U -statistics. *Ann. Statist.* **8** 299–312.
- CHAN, Y.-K. and WIERMAN, J. (1977). On the Berry–Esseen theorem for U -statistics. *Ann. Probab.* **5** 136–139.
- DHARMADHIKARI, S. W., FABIAN, V. and JOGDEO, K. (1968). Bounds on the moments of martingales. *Ann. Math. Statist.* **39** 1719–1723.
- FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications* **2**, 2nd ed. Wiley, New York.
- GÖTZE, F. (1979). Asymptotic expansions for bivariate von Mises functionals. *Z. Wahrsch. verw. Gebiete* **50** 333–355.
- HELMERS, R. and VAN ZWET, W. R. (1982). The Berry–Esseen bound for U -statistics. *Statistical Decision Theory and Related Topics, III* (S. S. Gupta and J. O. Berger, eds.) **1** 497–512. Academic, New York.
- HOEFFDING, W. (1948). A class of statistics with asymptotically normal distributions. *Ann. Math. Statist.* **19** 293–325.
- JANSSEN, P. L. (1978). De Berry–Esseen Stelling en een Asymptotische Ontwikkeling voor U -statistieken. Ph.D. thesis, Dept. of Mathematics, Catholic Univ. of Leuven.

P. J. BICKEL
DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720

F. GÖTZE
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF BIELEFELD
POSTFACH 8640
4800 BIELEFELD 12
FEDERAL REPUBLIC OF GERMANY

W. R. VAN ZWET
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF LEIDEN
POSTBUS 9512
2300 RA LEIDEN
THE NETHERLANDS