Resampling U-statistics using p-stable laws

Dehling, Herold; Denker, Manfred; Woyczynski, Wojbor A.

Published in:
Journal of Multivariate Analysis

DOI:
10.1016/0047-259X(90)90057-O

JOURNAL OF MULTIVARIATE ANALYSIS 34, 1–13 (1990)

Resampling U-Statistics Using p-Stable Laws

HEROLD DEHLING

University of Groningen, Groningen, The Netherlands, and Boston University

MANFRED DENKER

Universität Göttingen, Göttingen, West Germany, and Indiana University

AND

WOJBOR A. WOYCZYNSKI*

Case Western Reserve University

Communicated by the Editors

It is well known that symmetric statistics based on a kernel with finite second moment have a limit law which can be described by a multiple Wiener–Ito integral. However, if the kernel has less than second moments, no weak limit law holds in general. In the present paper we show that by a suitable change of the empirical process this process has a p-stable multiple integral as its limit.

1. INTRODUCTION

In this paper we shall study the asymptotic distribution of symmetric statistics of the form

$$U_n(h) = \binom{n}{m}^{-1} \sum_{1 \leqslant i_1 < \cdots < i_m \leqslant n} h(X_{i_1}, ..., X_{i_m}) Y_{i_1} \cdots Y_{i_m}, \tag{1.1}$$

Received July 27, 1987; revised October 31, 1988.

AMS 1980 subject classifications: 62G99, 60H05, 60F17.

Key words and phrases: U-statistic, weak convergence, normal domain of attraction of the p-stable law, p-stable multiple stochastic integral.

* This author thanks the faculty and staff of the Institut für Mathematische Stochastik of the University of Göttingen for their hospitality while this paper was being written. Partly supported by a grant from ONR.

0047-259X/90 $3.00

where $X_1, X_2, \dots$ are independent, identically distributed random variables, and where $Y_1, Y_2, \dots$ is another sequence of i.i.d. variables, independent of the process (X_n) and such that Y_i, $i = 1, 2, \dots$, belongs to the domain of normal attraction of some p-stable law ($Y_i \in \mathrm{NDA}(p)$), $p \in (1, 2)$. Such a process, also called U-statistic in nonparametric inference, may be alternatively expressed in the form

$$I_n^m(h) = n^{-m/p} \int \cdots \int h(x_1, \dots, x_m)\, dF_n(x_1) \cdots dF_n(x_m), \tag{1.2}$$

where

$$F_n(x) := \sum_{1 \leqslant i \leqslant n} 1\{X_i \leqslant x\} \cdot Y_i. \tag{1.3}$$

$n^{-1/p} F_n$ will be called the resampled empirical process, since it arises from the classical empirical measure given by the random points $X_1, \dots, X_n$, "resampled" by independent random weights defined by $Y_1, \dots, Y_n$. Similarly, $U_n(h)$ is called the resampled symmetric statistic, "resampled" from the usual statistic based on $X_1, X_2, \dots$:

$$\binom{n}{m}^{-1} \sum_{1 \leqslant i_1 < \cdots < i_m \leqslant n} h(X_{i_1}, \dots, X_{i_m}). \tag{1.4}$$

Integration in (1.2) is not to be extended over the diagonals. Alternatively, we may assume that h vanishes on all diagonals, i.e., $h(x_1, \dots, x_m) = 0$ if $x_i = x_j$ for some $i \neq j$. Without loss of generality we may always assume that the X_i's are uniformly distributed on $[0, 1]$ (see Denker *et al.* [6]).

It turns out that the limit distribution of the resampled symmetric statistics (1.1), properly normalized, can be represented as a multiple p-stable stochastic integral of the form

$$I^m(h) = \int \cdots \int h(x_1, \dots, x_m)\, dM(x_1) \cdots dM(x_m), \tag{1.5}$$

where $\{M(x), x \geqslant 0\}$ is a p-stable motion, i.e., a process with independent stationary increments such that $E[\exp(itM(1))] = \exp(-|t|^p)$. (For the notion of multiple stochastic integrals we refer to Rosinski and Woyczynski [15], cf. also Kwapien and Woyczynski [10].) This development, roughly speaking, parallels the results for Gaussian multiple integrals, but permits handling U-statistics without finite variance. U-statistics—and resampled symmetric statistics as a special case—form a backwards martingale [1] when $h \in L^1$. This implies the a.s. and L^1-convergence of $U_n(h)$ towards $Eh(X_1, \dots, X_m) Y_1 \cdots Y_m$. On the other hand, for degenerate h, $\binom{n}{m} U_n(h)$ has a forward martingale structure. This property, together with a variance estimate, gives the weak convergence to a multiple Wiener–Ito integral in the case of h having finite second moments. (This goes back at least to the work of Filippova [8].) In the present case, without the second moment assumption, to the best of our knowledge no weak convergence results exist in the literature.

The extension of our results to von Mises functionals, i.e., integrals as in (1.2) but including the diagonals, can be carried out by standard arguments and will not be discussed here.

Probabilistic aspects of symmetric statistics have been studied recently by several authors, e.g., the a.s. invariance principle is derived in Dehling *et al.* [5], Nolan and Pollard [13] proved uniform a.s. convergence over certain classes of kernels h and McConnell [12] provides a necessary and sufficient condition for two-parameter convergence in the strong law of large numbers for U-statistics.

In Section 2 we prove a continuity theorem, i.e., an inequality for integrals of the form (1.2) which shows that I_n^m is a bounded map from $L^r([0,1]^m) \to L^q(\Omega)$ $(r>p>q)$. This is augmented by some auxiliary technical results on these integrals. Theorem 2.1 has an immediate application to the bootstrap method showing that a.s. the bootstrap distributions converge to the theoretical limit distribution. This proof also carries over to the case of square integrable kernels, providing an easy argument for a result in Bickel and Freedman [2].

In Section 3 we shall show that the resampled empirical process F_n in (1.3) converges weakly in the Skorohod topology to a stable motion M, and, as a result, $U_n(h)$ converges weakly to $I^m(h)$.

We do not know how to obtain almost sure approximations for $U_n(h)$ as in Dehling *et al.* [5]. However, an invariance principle in probability can be obtained (Section 4), extending the results of Section 3.

2. The Continuity Theorem

With the same notation as in Section 1 we obtain:

THEOREM 2.1. *For any $q<p<r$ there exists a constant $C=C(p,q,r,m)$ such that*

$$\|I_n^m(h)\|_{L^q(\Omega)} \leqslant C\,\|h\|_{L^r([0,1]^m)}. \tag{2.1}$$

Proof. Since the general case is similar, we consider only the case $q=1$. Since h vanishes on the diagonals, it suffices to look at tetragonal sets in order to estimate $I_n^m(h)$.

From the decoupling inequality (see Kwapien and Woyczynski [10], especially Theorem 3.1 in Krakowiak and Szulga [9]) we obtain that

$$I \equiv \|I_n^m(h)\|_{L^1(\Omega)} = E\left|\sum_{1\leqslant i_1<\cdots<i_m\leqslant n} h(X_{i_1},...,X_{i_m})\frac{Y_{i_1}}{n^{1/p}}\cdot\,\cdots\,\cdot\frac{Y_{i_m}}{n^{1/p}}\right|$$

$$\leqslant KE_XE_1\cdots E_m\left|\sum_{1\leqslant i_1<\cdots<i_m\leqslant n} h(X_{i_1},...,X_{i_m})\frac{Y_{i_1}^{(1)}}{n^{1/p}}\cdot\,\cdots\,\cdot\frac{Y_{i_m}^{(m)}}{n^{1/p}}\right|,$$

where K denotes some constant, and where $(Y_i^{(1)}),...,(Y_i^{(m)})$ denote m independent copies of (Y_i). Here E_X denotes the expectation with respect to (X_i) and E_j with respect to $(Y_i^{(j)})$.

Integrating with respect to E_m and using the properties of the p-stable law we obtain the upper bound

$$I\leqslant KE_XE_1\cdots E_{m-1}\left(\sum_{1\leqslant i_m\leqslant n}\left|\sum_{1\leqslant i_1<\cdots<i_m} h(X_{i_1},...,X_{i_m})\right.\right.$$
$$\left.\left.\times\frac{Y_{i_1}^{(1)}}{n^{1/p}}\cdot\,\cdots\,\cdot\frac{Y_{i_{m-1}}^{(m-1)}}{n^{1/p}}\right|^p\frac{1}{n}\right)^{1/p}$$

so that, since $r>p$,

$$I\leqslant KE_XE_1\cdots E_{m-1}\left(\sum_{1\leqslant i_m\leqslant n}\left|\sum_{1\leqslant i_1<\cdots<i_m} h(X_{i_1},...,X_{i_m})\right.\right.$$
$$\left.\left.\times\frac{Y_{i_1}^{(1)}}{n^{1/p}}\cdot\,\cdots\,\cdot\frac{Y_{i_{m-1}}^{(m-1)}}{n^{1/p}}\right|^r\frac{1}{n}\right)^{1/r}$$
$$=KE_XE_1\cdots E_{m-1}\left\|\sum_{1\leqslant i_m\leqslant n}\sum_{1\leqslant i_1<\cdots<i_{m-1}<i_m} h(X_{i_1},...,X_{i_m})\right.$$
$$\left.\times\frac{Y_{i_1}^{(1)}}{n^{1/p}}\cdot\,\cdots\,\cdot\frac{Y_{i_{m-1}}^{(m-1)}}{n^{1/p}}1_{[i_m-1/n,\,i_m/n]}(x_m)\right\|_{L^r(dx_m)}.$$

Since $L^r(dx_m)$ $(r>p)$ is of stable type p (cf., e.g., [17, p. 369]), we get the bound

$$I\leqslant K_1E_XE_1\cdots E_{m-2}\left(\sum_{1\leqslant i_{m-1}\leqslant n}\left\|\sum_{1\leqslant i_1<\cdots<i_{m-2}<i_{m-1}<i_m} h(X_1,...,X_{i_m})\right.\right.$$
$$\left.\left.\times\frac{Y_{i_1}^{(1)}}{n^{N_p}}\cdot\,\cdots\,\cdot\frac{Y_{i_{m-1}}^{(m-1)}}{n^{1/p}}1_{[i_m-1/n,\,i_m/n]}\right\|_{L^r(dx_m)}^p\frac{1}{n}\right)^{1/p}$$

$$\leqslant K_1 E_X E_1 \cdots E_{m-2}\left(\left\| \sum_{1 \leqslant i_{m-1} \leqslant n} \sum_{1 \leqslant i_1 < \cdots < i_{m-2} < i_{m-1} < i_m} h(X_{i_1}, ..., X_{i_m}) \right.\right.$$

$$\times \frac{Y^{(1)}_{i_1}}{n^{1/p}} \cdot \cdots \cdot \frac{Y^{(m-2)}_{i_{m-2}}}{n^{1/p}} 1_{[i_{m-1} - 1/n, i_{m-1}/n]}(x_{m-1})$$

$$\left.\left. \cdot 1_{[i_m - 1/n, i_m/n]}(x_m) \right\|_{L^r(dx_{m-1}; L^r(dx_m))}\right),$$

where $L^r(dx; E)$ denotes the space of all functions f with values in a Banach space E such that $\int \|f(x)\|^r \, dx < \infty$. Since $L^r(dx_{m-1}; L^r(dx_n))$ is also of stable type p [17, p. 373], we can repeat the above procedure to obtain finally,

$$\|I^m_n(h)\|_{L^1(\Omega)} \leqslant C E_X \left\| \sum_{1 \leqslant i_1 < \cdots < i_m \leqslant n} h(X_{i_1}, ..., X_{i_m}) \right.$$

$$\left. \cdot 1_{[i_1 - 1/n, i_1/n]}(x_1) \cdot \cdots \cdot 1_{[i_m - 1/n, i_m/n]}(x_m) \right\|_{L^r(dx_1; L^r(dx_2; ...; L^r(dx_m) \cdots))}$$

$$= C E_X \left(\sum_{1 \leqslant i_1 < \cdots < i_m \leqslant n} |h(X_{i_1}, ..., X_{i_m})|^r \frac{1}{n^m} \right)^{1/r}$$

$$= C \left(\int \cdots \int |h(x_1, ..., x_m)|^r \, dx_1 \cdots dx_m \right)^{1/r}$$

which is the required estimate. Q.E.D.

Below, in the proof of the invariance principle we will also need the following maximal inequality which is an immediate corollary to Theorem 2.1 and to the fact that U-statistics have the martingale structure.

COROLLARY 2.2. *Let* $\mathbf{n} = (n_1, ..., n_m)$, *and define*

$$I^m_{\mathbf{n}}(h) = (n_1 \cdot \cdots \cdot n_m)^{-1/p} \int \cdots \int h(x_1, ..., x_m) \, dF_{n_1}(x_1) \cdots dF_{n_m}(x_m). \tag{2.2}$$

Then, for all $t > 0$ *we have*

$$P\{\max_{\mathbf{k} \leqslant \mathbf{n}} |I^m_{\mathbf{k}}(h)| > t\} \leqslant C \frac{\|h\|^q_{L^r([0,1]^m)}}{t^q}. \tag{2.3}$$

In the remaining part of this section we shall show how Theorem 2.1 can be applied to obtain a.s. convergence of the bootstrap distributions. In the case of square integrable kernels the analogue to Theorem 2.1 is well known (e.g., [8]); consequently the short argument below together with the Hoeffding decomposition for multiple integrals give an alternative proof of Theorem 3.1 in Bickel and Freedman [2].

Recall that $X_1, ..., X_n$ are i.i.d. random variables with uniform distribution function F. Denote by $H_n(t) = (1/n) \sum 1_{\{X_i \leq t\}}$ its empirical distribution function. Now, let $X_1^*, ..., X_n^*$ be an i.i.d. sequence of r.v.'s with common pdf H_n. If h is a symmetric kernel, then

$$n^{-m/p} \sum h(X_{i_1}, ..., X_{i_m}) Y_{i_1} \cdots Y_{i_m} \overset{\mathscr{L}}{=} n^{-m/p} \int \cdots \int h(x_1, ..., x_m) \prod_{i=1}^{m} dF_n(x_i) \tag{5.1}$$

and

$$\begin{aligned} & n^{-m/p} \sum h(X_{i_1}^*, ..., X_{i_m}^*) Y_{i_1} \cdot \cdots \cdot Y_{i_m} \\ & \overset{\mathscr{D}}{=} n^{-m/p} \int \cdots \int h(H_n^{-1}(x_1), ..., H_n^{-1}(x_m)) \prod_{i=1}^{m} dF_n(x_i), \end{aligned} \tag{5.2}$$

where $F_n(t) = \sum 1_{\{X_i \leq t\}} Y_i$ denotes the resampled empirical process as in (1.3).

THEOREM 2.3. *If $h \in L^r([0, 1]^m)$ for some $r > p$, then, with probability one*

$$d(\mathscr{L}(I_n^m(h)), \mathscr{L}(I_n^m(h(H_n^{-1}(\cdot), ..., H_n^{-1}(\cdot))))) \to 0,$$

where d denotes some metric for the topology of weak convergence of measures, and where $\mathscr{L}(Z)$ denotes the distribution of the r.v. Z.

Proof. By Theorem 2.1 we have that

$$\begin{aligned} & E\,|I_n^m(h) - I_n^m(h(H_n^{-1}(\cdot), ..., H_n^{-1}(\cdot)))|^q \\ & \leq C\, \|h - h(H_n^{-1}(\cdot), ..., H_n^{-1}(\cdot))\|_{L^r([0,1]^m)}^q. \end{aligned}$$

If h is bounded the upper bound tends to zero by the Lebesgue dominated convergence theorem since $H_n^{-1} \to F^{-1}$ a.s. Now, any h can be approximated by bounded functions so that another application of Theorem 2.1 gives the result. Q.E.D.

3. WEAK CONVERGENCE OF THE RESAMPLED EMPIRICAL PROCESS AND SYMMETRIC STATISTIC

In this section we first prove weak convergence of the resampled empirical process to a p-stable motion. Applying the continuity theorem, we get as a corollary the weak convergence of the resampled symmetric statistic.

THEOREM 3.1. *There exists a p-stable motion $M(t)$ such that*

$$n^{-1/p}F_n \to M \tag{3.1}$$

weakly in $D([0, 1])$ with respect to the Skorohod topology.

Proof. We first show convergence of the finite dimensional distributions. For simplicity, we write out the proof only for $d=2$.

Let $0 \leqslant s < t \leqslant 1$ be fixed and define

$$J_1(n) = \{i : 1 \leqslant i \leqslant n;\ X_i \leqslant s\}$$
$$J_2(n) = \{i : 1 \leqslant i \leqslant n;\ s < X_i \leqslant t\}.$$

Then

$$\begin{aligned} &n^{-1/p}(F_n(s), F_n(t) - F_n(s)) \\ &= n^{-1/p}\left(\sum_{i \leqslant n} 1_{\{X_i \leqslant s\}} Y_i,\ \sum_{i \leqslant n} 1_{\{s < X_i \leqslant t\}} Y_i\right) \\ &= n^{-1/p}\left(\sum_{i \in J_1(n)} Y_i,\ \sum_{i \in J_2(n)} Y_i\right) \\ &= n^{-1/p}\Bigg(|J_1(n)|^{1/p}\, |J_1(n)|^{-1/p} \sum_{i \in J_1(n)} Y_i, \\ &\qquad |J_2(n)|^{1/p}\, |J_2(n)|^{-1/p} \sum_{i \in J_2(n)} Y_i\Bigg). \end{aligned}$$

The two coordinates of this vector are conditionally independent given (X_i). Consequently, in view of Fubini's theorem, it suffices to check that each of them, conditioned on (X_i), converges weakly to the required limit, (X_i)—a.s.

By the strong law of large numbers

$$n^{-1}\, |J_1(n)| \to s \qquad \text{a.s.}$$
$$n^{-1}\, |J_2(n)| \to t - s \qquad \text{a.s.}$$

Therefore, by the assumption on (Y_i),

$$n^{-1/p} \sum_{i \in J_1(n)} Y_i \xrightarrow{\mathscr{D}} M(1) s^{1/p} \stackrel{\mathscr{D}}{=} M(s),$$

and

$$n^{-1/p} \sum_{i \in J_2(n)} Y_i \xrightarrow{\mathscr{D}} M(1)(t-s)^{1/p} \stackrel{\mathscr{D}}{=} M(t) - M(s),$$

and thus

$$n^{-1/p}(F_n(s), F_n(s)-F_n(t)) \to (M(s), M(t)-M(s)).$$

It remains to show uniform tightness of the process $(n^{-1/p}F_n)$. For this we will use the following lemma of Skorohod [16]:

LEMMA 3.2. *Let $\{\xi(t): 0 \leqslant t \leqslant 1\}$ be a process with sample paths in $D[0, 1]$ and with independent increments. Denote*

$$\Delta^p(c, \delta) = \sup \min\{P(|\xi(t)-\xi(t_1)| > \delta), P(|\xi(t_2)-\xi(t)| > \delta)\}$$

and

$$\Delta(c) = \sup \min\{|\xi(t)-\xi(t_1)|, |\xi(t_2)-\xi(t)|\},$$

where both suprema extend over all (t, t_1, t_2) with $0 \leqslant t \leqslant 1$ and $t-c \leqslant t_1 < t < t_2 \leqslant t+c$.

If $0 < c \leqslant 1$ satisfies $\Delta^p(c, \delta/20) \leqslant 1/4$, then for any positive integer $l \geqslant 3/c$ we have

$$P(\Delta(1/l) > \delta) \leqslant 10^3 \Delta^p(3/l, \delta/12)/c. \tag{3.2}$$

Using this lemma we can argue as follows: Denote by $P_{\mathbf{x}}$ the conditional probability given $(X_i) = (x_i) = \mathbf{x}$. Under $P_{\mathbf{x}}$, the process F_n has independent increments and we have that

$$\begin{aligned}
&P_{\mathbf{x}}\Big(\sup_{\substack{t_1 \leqslant t \leqslant t_2 \\ |t_2-t_1| \leqslant 2/l}} \min\{|F_n(t)-F_n(t_1)|, |F_n(t_2)-F_n(t)|\} > \delta\Big) \\
&\quad \leqslant 10^3 c^{-1} \sup_{\substack{t_1 \leqslant t \leqslant t_2 \\ |t_2-t_1| \leqslant 6/l}} \min\{P_{\mathbf{x}}(|F_n(t)-F_n(t_1)| > \delta/12), \\
&\qquad\qquad P_{\mathbf{x}}(|F_n(t_2)-F_n(t)| > \delta/12)\} \\
&\quad \leqslant K\left(\frac{12}{\delta}\right)^p \sup_{|s-t| \leqslant 6/l} |J_n(s, t)|,
\end{aligned}$$

where K denotes a constant and where $J_n(s, t) = \{i \leqslant n : x_i \in (s, t]\}$. Choosing $\delta = \delta' n^{1/p}$, we obtain that

$$\begin{aligned}
&P\Big(\sup_{\substack{t_1 \leqslant t \leqslant t_2 \\ |t_2-t_1| \leqslant 2/l}} \min\{|F_n(t)-F_n(t_1)|, |F_n(t_2)-F_n(t)|\} > \delta' n^{1/p}\Big) \\
&\quad \leqslant K \cdot E \sup_{|s-t| \leqslant 6/l} \frac{1}{n} |J_n(s, t)| \to 0 \quad \text{as} \quad l \to \infty.
\end{aligned}$$

Q.E.D.

Using an approximation of h by simple kernels we obtain as a corollary to Theorems 2.1 and 3.1 the weak convergence of the multiple stochastic integrals $I_n^m(h)$. The details of this argument are the same as in the Gaussian case [8, 6].

THEOREM 3.3. *If* $\|h\|_{L^r([0,1]^m)} < \infty$ *for some* $r > p$ *then*

$$I_n^m(h) \to I^m(h) \qquad \text{weakly as } n \to \infty. \tag{3.3}$$

4. An Invariance Principle for the Resampled Symmetric Statistics

In this section we extend Theorem 3.3 approximating the integrals $n^{m/p} I_n^m(h)$ in probability by multiple stable integral

$$I^m(h, Z_n) := \int \cdots \int h(x_1, ..., x_m)\, dZ_n(x_1) \cdots dZ_n(x_m),$$

where Z_n is a certain sequence of p-stable motions. Here Z_n plays the same role as the Kiefer process does in the Gaussian case (see Dehling *et al.* [5]). Z_n can be written as a sum of i.i.d. stable processes M_j, i.e., $Z_n = \sum_{j \leqslant n} M_j$, where each M_j has the same distribution as M in Theorem 3.1. We have the following properties:

$$\begin{aligned} &\text{(i)} \quad n^{-1/p} Z_n \overset{\mathscr{D}}{=} M \\ &\text{(ii)} \quad \{Z_n - Z_{n-1} : n \geqslant 2\} \text{ is an i.i.d. sequence.} \end{aligned} \tag{4.1}$$

THEOREM 4.1. *Let* $h \in L^r([0,1]^m)$, *where* $r > p$. *Then there exists a sequence* $\{Z_n(t) : 0 \leqslant t \leqslant 1\}$, $n \geqslant 1$, *of stable processes satisfying* (4.1) *such that*

$$n^{-m/p} \sup_{k \leqslant n} |k^{m/p} I_k^m(h) - I^m(h, Z_k)| \to 0 \tag{4.2}$$

in probability, as $n \to \infty$.

Proof. Let $\pi_k = \{0 = s_0(k) < s_1(k) < \cdots < s_k(k) = 1\}$ be a nested sequence of partitions of $[0, 1]$. We first approximate h by simple functions h^k with sets of constancy being rectangles given by the partition π_k satisfying

$$\begin{aligned} &\|h - h^k\|_r \leqslant \eta_k \downarrow 0 \\ &\max |h^k| \leqslant a_k \uparrow. \end{aligned} \tag{4.3}$$

For any process $\{X(s), 0 \leqslant s \leqslant 1\}$ we denote by $\Delta^k X = (\Delta_1^k X, ..., \Delta_k^k X)$ the k-dimensional increment vector $\{X(s_1(k)), X(s_2(k)) - X(s_1(k)), ..., X(s_k(k)) - X(s_{k-1}(k))\}$ defined by the points of π_k.

For the moment let k be fixed. Denote by $\{X_i^{(j)}, Y_i^{(j)}, i \geqslant 1\}_{j \geqslant 1}$ independent copies of the basic observations $\{X_i, Y_i : i \geqslant 1\}$, and let $F_n^{(j)}$ be the corresponding resampled empirical processes. It has been shown in the proof of Theorem 3.1 that $\Delta^k F_n^{(k)}$ converges weakly to $\Delta^k M$. By Theorem 1 of Philipp [14, Corrigendum 1986] and by Lemma 2 of Dudley and Philipp [7] there exist i.i.d stable processes $M_j^{(k)}(t), j \geqslant 1$, having the same distribution as M and satisfying $n^{-1/p} \max_{j \leqslant n} \|\Delta^k F_j^{(k)} - \Delta^k \sum_{l \leqslant j} M_l^{(k)}\|_1 \to 0$ in probability. Here $\|\cdot\|_1$ denotes the l^1-norm on $\mathbb{R}^k$, i.e., $\|x\|_1 = \sum_{i \leqslant k} |x_i|$. Hence, there exists an integer $n = n_0(k)$ such that for all $n \geqslant n_0$,

$$P\left\{\max_{j \leqslant n} \left\| \Delta^k F_j^{(k)} - \Delta^k \sum_{l \leqslant j} M_l^{(k)} \right\| > \gamma_k n^{1/p}\right\} \leqslant \gamma_k, \tag{4.4}$$

where

$$\gamma_k = a_k^{-1} k^{-m} 2^{-k}. \tag{4.5}$$

Note at this point that (4.4) continues to hold if Δ^k is replaced by Δ^κ for any $\kappa \leqslant k$. Now, we put $t_k = \sum_{l \leqslant k} n_0(l)$ and we define the basic observations X_i and $Y_i (i \geqslant 1)$ and the stable process Z in the following way:

$$\begin{aligned} &\text{For} \quad i \in (t_k, t_{k+1}] \cap \mathbb{Z} : X_i = X_{i-t_k}^{(k)},\ Y_i = Y_{i-t_k}^{(k)},\ M_i = M_{i-t_k}^{(k)}, \\ &\text{and} \quad Z_n = \sum_{j \leqslant n} M_j. \end{aligned} \tag{4.6}$$

Next, we choose a sequence $\kappa(k) \uparrow \infty$ such that

$$k_\kappa(k)/t_k \to 0 \qquad \text{as} \quad k \to \infty. \tag{4.7}$$

For brevity, we introduce the following notation: If $K = (k_1, ..., k_m)$, $N = (n_1, ..., n_m)$, let

$$I_{K,N}(h) = \int \cdots \int h(x_1, ..., x_m) \prod_{i=1}^{m} (F_{n_i} - F_{k_i})(dx_i) \tag{4.8}$$

and

$$I_{k,n}(h) = I_{(k,...,k),(n,...,n)}(h).$$

(Note that $I_{i,n}(h)$ differs from $I^m(h)$ by a norming constant.) Let $t_k < n \leqslant t_{k+1}$. Then, with $t_\kappa = t_{\kappa(k)}$ we get

$$\max_{j \leqslant n} n^{-m/p} |j^{m/p} I_j^m(h) - I^m(h, Z_j)| \leqslant \text{I} + \text{II} + \text{III} + \text{IV}, \tag{4.9}$$

where

$$\mathrm{I} = n^{-m/p} \max_{t_\kappa \leqslant j \leqslant n} |j^{m/p} I_j^m(h) - I_{t_\kappa, j}(h)|$$

$$+ n^{-m/p} \max_{t_\kappa \leqslant j \leqslant n} |I^m(h, Z_j) - I^m(h, Z_j - Z_{t_\kappa})|,$$

$$\mathrm{II} = n^{-m/p} \max_{j \leqslant t_\kappa} j^{m/p} |I_j^m(h)| + n^{-m/p} \max_{j \leqslant t_\kappa} I^m(h, Z_j)|$$

$$\mathrm{III} = n^{-m/p} \max_{t_\kappa \leqslant j \leqslant n} |I_{t_\kappa, j}(h - h^\kappa)| + \max_{t_\kappa \leqslant j \leqslant n} |I^m(h - h^\kappa, Z_j - Z_{t_\kappa})|$$

$$\mathrm{IV} = n^{-m/p} \max_{t_\kappa \leqslant j \leqslant n} |I^m_{t_\kappa, j}(h^\kappa) - I^m(h^\kappa, Z_j - Z_{t_\kappa})|.$$

We shall first treat I, II, and III using the inequalities of Section 2. A simple computation shows that

$$|j^{m/p} I_j^m(h) - I_{t_\kappa, j}(h)| \leqslant \sum_{i=1}^m |I_{0, j(\mathbf{1} - e_i) + t_\kappa e_i}(h)| + m t_\kappa^{m/p} I_{t_\kappa}^m(h),$$

where e_i is the ith unit vector in $\mathbb{R}^m$ and $\mathbf{1} = (1, ..., 1) \in \mathbb{R}^m$. Now Corollary 2.2, together with (4.7), proves that

$$\max_{t_\kappa \leqslant j \leqslant n} n^{-m/p} |j^{m/p} I_j^m(h) - I_{t_\kappa, j}(h)| \to 0$$

in probability as $n \to \infty$. The same reasoning applies to the second term in I and to II. Using (4.3) and Corollary 2.2 we obtain III $\to 0$ in probability. Hence, it only remains to deal with IV. We write

$$h^\kappa(x_1, ..., x_m) = \sum_{i_1, ..., i_m} h_{i_1, ..., i_m} 1_{[s_{i_1}(\kappa), s_{i_1+1}(\kappa))}(x_1) \cdot \cdots \times 1_{[s_{i_m}(\kappa), s_{i_m+1}(\kappa))}(x_m), \tag{4.10}$$

where, by (4.3), we have that $|h_{i_1, ..., i_m}| \leqslant a_\kappa$. Writing $F_{k,n} = F_n - F_k$ for the sake of brevity, we obtain for any $(i_1, ..., i_m)$,

$$\max_{t_\kappa \leqslant j \leqslant n} n^{-m/p} \left| \prod_{l=1}^m \Delta_l^\kappa F_{t_\kappa, j} - \prod_{l=1}^m \Delta_l^\kappa (Z_j - Z_{t_\kappa}) \right|$$

$$\leqslant \sum_{\substack{L \subset \{1, ..., m\} \\ \mathrm{card}\, L \geqslant 1}} \max_{t_\kappa \leqslant j \leqslant n} \left| \prod_{l \in L} n^{-1/p} (\Delta_l^\kappa F_{t_\kappa, j} - \Delta_l^\kappa (Z_j - Z_{t_\kappa})) \right|$$

$$\times \prod_{l \notin L} n^{-1/p} \max_{t_k \leqslant j \leqslant n} |\Delta_l^\kappa (Z_j - Z_{t_\kappa})|. \tag{4.11}$$

Each term in the last product can be estimated as follows: Let $t_\mu < j \leqslant t_{\mu+1}$. Then

$$n^{-1/p}\,|\Delta_l^\kappa F_{t_\kappa, j} - \Delta_l^\kappa Z_{t_\kappa, j}| \leqslant n^{-1/p}\,|\Delta_l^\kappa F^{(\kappa)}_{n_0(\kappa+1)} - \Delta_l^\kappa Z^{(\kappa)}_{n_0(\kappa+1)}|$$
$$+ n^{-1/p}\,|\Delta_l^\kappa F^{(\kappa+1)}_{n_0(\kappa+2)} - \Delta_l^\kappa Z^{(\kappa)}_{n_0(\kappa+2)}|$$
$$+ n^{-1/p}\,|\Delta_l^\kappa F^{(\mu)}_{j-t_\mu} - \Delta_l^\kappa Z^{(\mu)}_{j-t_\mu}|.$$

Hence, by (4.4) and the remark following (4.5) we have

$$\max_{j \leqslant n} n^{1/p}\,|\Delta_l^\kappa F_{t_\kappa, j} - \Delta_l^\kappa Z_{t_\kappa, j}| \leqslant \sum_{j=\kappa}^{k} \gamma_j \leqslant \frac{1}{a_\kappa \kappa^m} \cdot 2^{-\kappa+1},$$

except on a set of probability less than $2^{-\kappa+1}/(a_\kappa \kappa^m)$. Hence, the left-hand side of (4.11) is bounded by

$$2^m (a_\kappa \kappa^m)^{-1} \cdot 2^{-\kappa/2}$$

except on a set of probability less than $\varepsilon_\kappa = 2^m m((a_\kappa \kappa^m)^{-1}\, 2^{-\kappa} + C2^{-p\kappa/m})$. Using this last bound, together with (4.10) and (4.3) we finally get that

$$\mathrm{IV} \leqslant 2^m 2^{-\kappa/2},$$

except on a set of probability less than ε_κ. Hence $\mathrm{IV} \to 0$ in probability.

Q.E.D.

References

[1] Berk, R. H. (1966). Limiting behavior of posterior distributions when the model is incorrect. *Ann. Math. Statist.* **37** 51–58.

[2] Bickel, P. J., and Freedman, D. A. (1981). Some asymptotic theory for the bootstrap. *Ann Statist.* **9** 1196–1217.

[3] Billingsley, P. (1968). *Convergence of Probability Measures*, Wiley, New York.

[4] Dehling, H. (1986). Almost sure approximation for *U*-statistics. In *Dependence in Probability and Statistics* (E. Eberlein and M. S. Taqqu, Eds.), pp. 119–135, Birkhäuser, Boston.

[5] Dehling, H., Denker, M., and Philipp, W. (1984). Invariance principles for von Mises and *U*-statistics, *Z. Wahrsch. Verw. Gebiete* **67** 139–167.

[6] Denker, M., Grillenberger, C., and Keller, G. (1985). A note on invariance principles for von Mises statistics, *Metrika* **32** 197–214.

[7] Dudley, R M., and Philipp, W. (1983). Invariance principles for sums of Banach space valued random elements and empirical processes, *Z. Wahrsch. Verw. Gebiete* **62** 509–552.

[8] Filippova, A. A. (1962). Mises' theorem on the asymptotic behavior of functionals of empirical distribution functions and its statistical applications. *Theory Probab. Appl.* **7** 24–57.

[9] KRAKOWIAK, W., AND SZULGA, J. (1986). A multiple stochastic integral with respect to a strictly p-stable random measure. *Ann. Probab.* **16** 764–777.

[10] KWAPIEN, S., AND WOYCZYNSKI, W. A. (1986). Decoupling of martingale transforms and stochastic integrals for processes with independent increments. In *Probability Theory and Harmonic Analysis* (J. A. Chao and W. A. Woyczynski, Ed.), pp. 139–148, Dekker, New York.

[11] KWAPIEN, S., AND WOYCZYNSKI, W. A. (1987). Double stochastic integrals, random quadratic forms, and random series in Orlicz spaces. *Ann. Probab.* **15** 1072–1096.

[12] MCCONNELL, T. R. (1987). Two-parameter strong laws and maximal inequalities for U-statistics. *Proc. Roy. Soc. Edinburgh, Sect. A* **107** 133–151.

[13] NOLAN, D., AND POLLARD, D. (1987). U-processes: Rates of convergence. *Ann. Statist.* **15** 780–799.

[14] PHILIPP, W. (1980). Weak and L^p-invariance principles for sums of B-valued random variables, *Ann. Probab.* **8** 68–82; Corrigendum, *ibid.* (1986). **14**.

[15] ROSINSKI, J., AND WOYCZYNSKI, W. A. (1986). On Ito stochastic integrals with respect to p-stable motion: Inner clock, integrability of sample paths, double and multiple integrals, *Ann. Probab.* **14** 271–286.

[16] SKOROHOD, A. V. (1957), Limit theorems for stochastic processes with independent increments, *Theory Probab. Appl.* **2** 138–171.

[17] WOYCZYNSKI, W. A. (1978), Geometry and Martingales in Banach spaces. Part II. Independent increments, *Adv. in Probab.*, Vol. 4, pp. 267–518. Dekker, New York.