A REMARK ON CONVERGENCE IN DISTRIBUTION OF U-STATISTICS

by Evarist Giné* and Joel Zinn†
University of Connecticut and Texas A&M University

Runninghead: U-statistics.

Abstract

It is proved that, for h measurable and symmetric in its arguments and X_i i.i.d., if the sequence $\{n^{-\frac{m}{2}}\sum_{\substack{i_1,\dots,i_m\le n\\ i_j\ne i_k \text{ if } j\ne k}} h(X_{i_1},\dots,X_{i_m})\}_{n=1}^{\infty}$, is stochastically bounded, then $Eh^2<\infty$ and $Eh(X_1,x_2,\dots,x_m)=0$ a.s.

1. Introduction. Whereas the limit theory for sums of i.i.d. random variables is well understood in the sense that there are necessary and sufficient analytic conditions for each of the main limit theorems to hold, the limit theory for U-statistics is far from complete. There are very sharp sufficient conditions for e.g. the law of large numbers and the central limit theorem for U-statistics, but either they are not necessary (e.g. for the law of large numbers: Giné and Zinn, 1992) or it is not known whether they are (e.g. for the CLT). In this note we show that the usual sufficient condition for weak convergence of completely degenerate U-statistics, namely finiteness of the second moment of the defining function, is also necessary (in fact we prove a stronger statement). The same problem for U-statistics which are not completely degenerate is not considered here and seems to require techniques different from those used in this note.

Let $(S,\mathcal{S},P)$ be a measure space, let X, X_i be i.i.d.(P) S-valued random variables, let $m\in I\!N$ and let $h:\ S^m\to I\!R$ be a measurable function *symmetric* in its arguments, that is $h(x_1,\dots,x_m)=h(x_{\sigma_1},\dots,x_{\sigma_m})$ for any permutation σ of $\{1,\dots,m\}$. We let, as usual,

$$U_n(h)=\frac{1}{\binom{n}{m}}\sum_{i_1<\dots<i_m\le n} h(X_{i_1},\dots,X_{i_m})=\frac{(n-m)!}{n!}\sum_{I_n} h(X_{i_1},\dots,X_{i_m}),$$

AMS 1991 *subject classifications*. Primary: 60F05. Secondary: 60E15.

Key words and phrases: U-statistics, necessary conditions for convergence in distribution, decoupling.

* Research partially supported by NSF Grants No. DMS-9000132 and DMS-8505550. Part of this research was carried out at the MSRI, Berkeley.

† Research partially supported by NSF Grants No. DMS-8902418 and DMS-8505550. Part of this research was carried out at the MSRI, Berkeley.

where $I_n = \{(i_1, \ldots, i_m) : \ i_r \leq n, \ i_r \neq i_s \text{ if } r \neq s\}$. The object of this note is to prove the following:

THEOREM 1. *If the sequence* $\{n^{\frac{m}{2}} U_n(h)\}_{n=1}^{\infty}$ *is stochastically bounded then* $Eh^2(X_1, \ldots, X_m) < \infty$ *and* $Eh(X_1, x_2 \ldots, x_m) = 0$ *for almost every* $(x_2, \ldots, x_m) \in S^{m-1}$ *(and therefore,* $\{n^{\frac{m}{2}} U_n(h)\}_{n=1}^{\infty}$ *converges in distribution).*

The case $m = 1$ of Theorem 1 is just the necessity of $EX^2 < \infty$ for the CLT (Feller, 1935; Khinchin, 1935; Lévy, 1935). There are several proofs of this classical result, the most elementary being perhaps one based on symmetrization, Lévy's inequality and the converse Kolmogorov inequality or Hoffmann-Jørgensen's inequality. This proof does not seem to extend beyond sums of independent random variables. Our proof of Theorem 1 is based on randomization and Khinchin's inequality and, specialized to the case $m = 1$, it provides a new, very simple proof of the classical result.

In Section 2 we show that the tails of the original U-statistic dominate the tails of a decoupled, randomized version of it. This is an elementary but useful fact. In Section 3 we prove Theorem 1 as follows: once established that the decoupled, randomized U-statistics are tight, Khinchin's inequality together with the Paley-Zygmund inequality allow us to conclude that the U-statistics based on h^2 form also a tight sequence and therefore, by positivity, the U-statistics based on the truncations $h^2 I(h^2 \leq c)$ are tight uniformly in n and c; this yields $Eh^2 < \infty$ by the law of large numbers for U-statistics with integrable defining functions (in fact, with the bounded defining functions $h^2 I(h^2 \leq c)$) .

A version of Theorem 1 also holds for Banach space valued functions h, if the Banach space is of cotype 2 (see e.g. Araujo and Giné, 1980, for the definition). This remark is made in Section 4.

The reader who is only interested in Theorem 1 for $m = 2$, may skip Section 2 and read instead Remark 1 in Section 4.

Let us now briefly consider the general case. By Hoeffding's decomposition, as soon as h is integrable, the U-statistic with kernel h decomposes into a sum of completely degenerate U-statistics with kernels $\binom{m}{k}(\pi_k h)(x_1, \ldots, x_k) := \binom{m}{k}(\delta_{x_1} - P) \ldots (\delta_{x_k} - P) P^{m-k} h$, $0 \leq k \leq m$, and our result applies to each of these terms. However the general problem should be formulated along the following lines: Is it true that if, for some $1 \leq r \leq m$, the sequence $\{n^{\frac{r}{2}} U_n(h)\}_{n=1}^{\infty}$ is stochastically bounded then $E|h| < \infty$, $\pi_k h \equiv 0$ for $k < r$, $E(\pi_r h)^2 < \infty$ and $n^{\frac{r}{2}} U_n(\pi_k h) \to 0$ in probability for $r < k \leq m$? The present work answers this question in the affirmative for $r = m$ but our methods alone do not seem to be adequate to answer it for $r < m$.

2. A (one-sided) decoupling inequality. Let $(S, \mathcal{S}, P)$ and X, X_i be as above, let B be a measurable linear space, and let $h : \ S^m \to B$ be a measurable function, *symmetric* in its arguments. Denote multiindices $(i_1, \ldots, i_m) \in I\!N^m$ by $\mathbf{i}$ and vectors $(x_{i_1}, \ldots, x_{i_m})$

by $\mathbf{x_i}$. Let $I = \cup_{n=1}^{\infty} I_n = \{(i_1, \ldots, i_m) : \ i_r \neq i_s \text{ if } r \neq s\}$. Given a *finite* set $A \subset I\!N$ we let

$$S_A := S_A(m) := \sum_{I \cap A^m} h(\mathbf{x_i}) \ \Big(= \sum_{i_j \in A, \ i_j \neq i_k \ \text{if} \ j \neq k} h(x_{i_1}, \ldots, x_{i_m})\Big).$$

Given $A_1, \ldots, A_r \subset I\!N$, *disjoint* and *finite*, and $(m_1, \ldots, m_r)$, $m_i \in I\!N \cup \{0\}$, $\sum_{i=1}^r m_i = m$, we will let $S_{A_1,\ldots,A_r}(m_1, \ldots, m_r)$ be $\sum h(\mathbf{x_i})$, $\mathbf{i}$ in the intersection with I of any of the $\frac{m!}{m_1! \ldots m_r!}$ cartesian products of m_1 factors equal to A_1, ... , m_r factors equal to A_r. Formally, if $\mathcal{P}(m; m_1, \ldots, m_r)$ is the set of partitions of $\{1, \ldots, m\}$ into r sets $P_1, \ldots, P_r$ with $|P_j| = m_j$, $P_j = \emptyset$ if $m_j = 0$, then

$$S_{A_1,\ldots,A_r}(m_1, \ldots, m_r) = \sum_{(P_1,\ldots,P_r) \in \mathcal{P}(m; m_1,\ldots,m_r)} \ \sum_{\mathbf{i} \in I, \ i_j \in A_k \ \text{if} \ j \in P_k} h(\mathbf{x_i})$$

(for h general, not necessarily symmetric; if h is symmetric the rightmost sums are all equal). The following identity is obvious: For $A \subset I\!N$ finite and A_i, $i = 1, \ldots, r$, disjoint, with $\cup_{i=1}^r A_i = A$,

$$S_A = \sum_{(m_1,\ldots,m_r): \ \sum m_i = m} S_{A_1,\ldots,A_r}(m_1, \ldots, m_r). \tag{1}$$

We can now prove the following elementary lemma:

LEMMA 1. *Let A_i, $i = 1, \ldots, m$, be finite disjoint subsets of $I\!N$ and let $A = \cup_{i=1}^m A_i$. Then,*

$$m! \sum_{\mathbf{i} \in A_1 \times \ldots \times A_m} h(\mathbf{x_i}) = S_{A_1,\ldots,A_m}(1, \ldots, 1)$$

$$= S_A - \sum_{r=1}^m S_{A \setminus A_r} + \sum_{1 \le r_1 < r_2 \le m} S_{A \setminus (A_{r_1} \cup A_{r_2})} - \ldots \pm \sum_{r=1}^m S_{A_r}. \tag{2}$$

PROOF. The first identity is a direct consequence of the symmetry of h. The second, whose proof follows, does not require symmetry. Let us extend, for convenience, the definition of $S_{A_1,\ldots,A_m}(m_1, \ldots, m_m)$ to subsets C of $M := \{(m_1, \ldots, m_m) : \ \sum_{i=1}^m m_i = m$, $m_i \in I\!N \cup \{0\}\}$, as $S_{A_1,\ldots,A_m}(C) = \sum_{(m_1,\ldots,m_m) \in C} S_{A_1,\ldots,A_m}(m_1, \ldots, m_m)$. Let $C_r = \{(m_1, \ldots, m_m) \in M : \ m_r = 0\}$, $r = 1, \ldots, m$. Then $M = \{(1, \ldots, 1)\} \cup C_1 \cup \ldots \cup C_m$, and therefore, equation (1), together with the inclusion-exclusion formula, gives

$$\begin{aligned} S_A &= S_{A_1,\ldots,A_m}(M) \\ &= S_{A_1,\ldots,A_m}(1, \ldots, 1) + S_{A_1,\ldots,A_m}(\cup_{r=1}^m C_r) \\ &= S_{A_1,\ldots,A_m}(1, \ldots, 1) + \sum_{r=1}^m S_{A_1,\ldots,A_m}(C_r) \\ &\quad - \sum_{1 \le r_1 < r_2 \le m} S_{A_1,\ldots,A_m}(C_{r_1} \cap C_{r_2}) + \ldots \pm \sum_{r=1}^m S_{A_1,\ldots,A_m}(\cap_{j \neq r} C_j) \end{aligned}$$

(note $\cap_{j=1}^m C_j = \emptyset$). But, again by (1), $S_{A_1,\ldots,A_m}(C_{r_1} \cap \ldots \cap C_{r_k}) = S_{A \setminus (A_{r_1} \cup \ldots \cup A_{r_k})}$, and the lemma follows. $\square$

Let now $\{X_i^{(j)},\ i \in I\!N\},\ j \le m$, be m independent copies of the sequence $\{X_i,\ i \in I\!N\}$ (i.e., these random vectors are all i.i.d.(P)). If $A_j,\ j \le m$, are disjoint and $|A_j| = n_j$, we obviously have

$$\mathcal{L}\Big(\sum_{\mathbf{i} \in A_1 \times \ldots \times A_m} h(\mathbf{X_i})\Big) = \mathcal{L}\Big(\sum_{\mathbf{i} \in [1,n_1] \times \ldots \times [1,n_m]} h(X_{i_1}^{(1)}, \ldots, X_{i_m}^{(m)})\Big). \tag{3}$$

Because of the simple observation (3), (2) gives a relationship between the original and the decoupled U-statistics. We will also need to randomize the decoupled U-statistics; to this end, we let $\{\varepsilon_i^{(j)},\ i \in I\!N,\ j \le m\}$ be an independent array of Rademacher variables, independent of the variables $\{X_i^{(j)}\}$.

THEOREM 2. *Let K be a convex symmetric subset of B.*
(a) If $D_j,\ j = 1, \ldots, m$, are subsets of $\{1, \ldots, n\}$ then

$$\Pr\Big\{\sum_{\mathbf{i} \in D_1 \times \ldots \times D_m} h(X_{i_1}^{(1)}, \ldots, X_{i_m}^{(m)}) \in \frac{2^m - 1}{m!} K^c\Big\} \le (2^m - 1) \max_{k \le mn} \Pr\Big\{\sum_{\mathbf{i} \in I_k} h(\mathbf{X_i}) \in K^c\Big\}. \tag{4}$$

(b)

$$\Pr\Big\{\sum_{i_1,\ldots,i_m \le n} \varepsilon_{i_1}^{(1)} \ldots \varepsilon_{i_m}^{(m)} h(X_{i_1}^{(1)}, \ldots, X_{i_m}^{(m)}) \in \frac{2^m(2^m - 1)}{m!} K^c\Big\} \le 2^m(2^m - 1) \max_{k \le nm} \Pr\Big\{\sum_{\mathbf{i} \in I_k} h(\mathbf{X_i}) \in K^c\Big\}. \tag{5}$$

PROOF. (a) follows immediately from Lemma 1 taking $A_1 = D_1,\ A_2 = n + D_2, \ldots, A_m = n(m-1) + D_m$ (see (3)). (b) follows from (a) and Fubini's theorem because $\sum_{i_1,\cdots,i_m \le n} \varepsilon_{i_1}^{(1)} \ldots \varepsilon_{i_m}^{(m)} h(X_{i_1}^{(1)}, \ldots, X_{i_m}^{(m)})$ is a linear combination with coefficients ± 1 of 2^m terms of the form $\sum_{\mathbf{i} \in D_1 \times \ldots \times D_m} h(X_{i_1}^{(1)}, \ldots, X_{i_m}^{(m)})$, with $D_j = \{i \le n :\ \varepsilon_i^{(j)} = 1\}$ or $D_j = \{i \le n :\ \varepsilon_i^{(j)} = -1\}$. $\square$

Lemma 1 and Theorem 2 could be stated in more generality; for instance, it is clear that analogous results can be stated for multiple stochastic integrals.

It would be interesting to have inequalities analogous to those in Theorem 2, but in the opposite direction.

3. Proof of Theorem 1. The stochastic boundedness of the sequence $\big\{S_n := n^{-\frac{m}{2}} \sum_{\mathbf{i} \in I_n} h(\mathbf{X_i}) :\ n \in I\!N\big\}$ implies, by Theorem 2, that the sequence

$$\big\{\tilde{S}_n := n^{-\frac{m}{2}} \sum_{i_1,\ldots,i_m \le n} \varepsilon_{i_1}^{(1)} \ldots \varepsilon_{i_m}^{(m)} h(X_{i_1}^{(1)}, \ldots, X_{i_m}^{(m)}) :\ n \in I\!N\big\}$$

is also stochastically bounded. Let

$$[S_n]^2 := n^{-m} \sum_{i_1,\ldots,i_m \le n} h^2(X_{i_1}^{(1)}, \ldots, X_{i_m}^{(m)}), \quad n \in I\!N.$$

The next step consists in showing that the sequence $\{[S_n]^2\}$ is also stochastically bounded. To prove this we use two well known inequalities.

Using Khinchin's inequality (e.g. Kahane, 1968) first in the Banach space L_1 spanned by the $\varepsilon_j^{(2)}$ variables, and then twice in $I\!R$, we obtain that for any $\{a_{i,j}\} \subset I\!R$ and any $n \in I\!N$,

$$\begin{aligned} E\Big|\sum_{i,j\le n} a_{i,j}\varepsilon_i^{(1)}\varepsilon_j^{(2)}\Big| &= E_1\Big[E_2\Big|\sum_{i=1}^n\Big(\sum_{j=1}^n a_{i,j}\varepsilon_j^{(2)}\Big)\varepsilon_i^{(1)}\Big|\Big] \\ &\ge c\Big[E_1\Big(E_2\Big|\sum_{i=1}^n\Big(\sum_{j=1}^n a_{i,j}\varepsilon_j^{(2)}\Big)\varepsilon_i^{(1)}\Big|\Big)^2\Big]^{\frac12} \\ &\ge c\Big[2E_1\sum_{j=1}^n\Big(\sum_{i=1}^n a_{i,j}\varepsilon_i^{(1)}\Big)^2\Big]^{\frac12} \\ &\ge \sqrt{2}c\Big(\sum_{i,j\le n} a_{i,j}^2\Big)^{\frac12} \end{aligned}$$

where c is the constant in Khinchin's inequality for L_1. By iteration, it follows that there exists a universal constant c_m such that, for any $a_{i_1,\ldots,i_m} \in I\!R$ and $n \in I\!N$,

$$E\Big|\sum_{i_1,\ldots,i_m\le n} a_{i_1,\ldots,i_m}\varepsilon_{i_1}^{(1)}\cdots\varepsilon_{i_m}^{(m)}\Big| \ge c_m\Big(\sum_{i_1,\ldots,i_m\le n} a_{i_1,\ldots,i_m}^2\Big)^{\frac12}. \tag{6}$$

(This inequality also follows from Bonami's (1970) hypercontractivity inequality for Rademacher polynomials, but the above derivation is more elementary.) This moment inequality, by an easy argument of Paley and Zygmund (e.g. Kahane, 1968) yields an inequality for tails, which is what we need. The Paley-Zygmund argument is as follows: let ξ be a real random variable; then we obviously have by Jensen's inequality that, for any $t > 0$, $E|\xi| \le t + \big(E\xi^2\big)^{\frac12}\big(\mathrm{P}\{|\xi| > t\}\big)^{\frac12}$, and this yields

$$\Pr\{|\xi| > t\} \ge \left(\frac{(E|\xi| - t)^+}{(E\xi^2)^{\frac12}}\right)^2. \tag{7}$$

Then, by inequality (6),

$$E_\varepsilon|\tilde{S}_n| \ge c[S_n],$$

so that, by inequality (7), for all $t > 0$,

$$\mathrm{P}_\varepsilon\{|\tilde{S}_n| > t\} \ge \Big[\frac{(c[S_n] - t)^+}{[S_n]}\Big]^2 \ge \frac{c^2}{4} I\big([S_n] > \frac{2t}{c}\big).$$

Integrating, we obtain

$$\Pr\{|\tilde{S}_n| \geq t\} \geq \frac{c^2}{4} \Pr\{[S_n] > \frac{2t}{c}\},$$

showing that the sequence $\{[S_n]^2 : \ n \in I\!N\}$ is stochastically bounded (since $\{\tilde{S}_n\}$ is).

The law of large numbers for U-statistics (e.g., Serfling, 1980) gives that for every $c < \infty$,

$$n^{-m} \sum_{i_1,\ldots,i_m \leq n} [h^2 I(h^2 \leq c)](X_{i_1}^{(1)}, \ldots, X_{i_m}^{(m)}) \to Eh^2 I(h^2 \leq c) \ \text{ a.s.}$$

(note that a decoupled U-statistic based on h is just a regular U-statistic based on the function H on $(S^m)^m$ defined as $H(\mathbf{x}_1, \ldots, \mathbf{x}_m) = h(x_1^{(1)}, \ldots, x_m^{(m)})$.) This limit (actually in probability), the stochastic boundedness of $\{[S_n]^2\}$, and positivity give

$$\begin{aligned}\sup_{c>0} I\big[Eh^2 I(h^2 \leq c) > t\big] &\leq \sup_{c>0} \sup_n \Pr\big\{n^{-m} \sum_{\mathbf{i} \in I_n} [h^2 I(h^2 \leq c)](X_{i_1}^{(1)}, \ldots, X_{i_m}^{(m)}) > t\big\} \\ &\leq \sup_n \Pr\big\{[S_n]^2 > t\big\} \to 0 \ \text{ as } t \to \infty.\end{aligned}$$

Hence, there is $t_0 < \infty$ such that $\sup_{c>0} Eh^2 I(h^2 \leq c) \leq t_0$, i.e. $Eh^2 < \infty$.

Let us recall Hoeffding's decomposition:

$$U_n(h) = \sum_{k=0}^{m} \binom{m}{k} U_n(\pi_k h)$$

where $(\pi_k h)(x_1, \ldots, x_k) = (\delta_{x_1} - P) \times \ldots \times (\delta_{x_k} - P) \times P^{m-k}(h)$. $\pi_0 h$ is simply $P^m h = Eh$, and for $k > 0$, $\pi_k h$ is $P-$*canonical*, that is, $E(\pi_k h)(X_1, x_2, \ldots, x_k) = 0$ a.s.; note also $E(\pi_k h)^2 \leq Eh^2 < \infty$. So, the central limit theorem for degenerate U-statistics (Rubin and Vitale, 1980; e.g. Bretagnolle, 1983, or Dynkin and Mandelbaum, 1983) gives convergence in distribution of $\{n^{\frac{k}{2}} U_n(\pi_k h)\}$, with a non-zero limit if and only if $E(\pi_k h)^2 \neq 0$. Therefore, for each $k \geq 0$, the k-th term in the Hoeffding decomposition above is either exactly $0_P(n^{-\frac{k}{2}})$, or $\pi_k h = 0$ a.s. Since, by hypothesis, $U_n(h)$ is $0_P(n^{-\frac{m}{2}})$, it follows that $\pi_k h = 0$ a.s. for $k = 0, 1, \ldots, m-1$. For $k = 0$ this gives $P^m h = Eh = 0$; for $k = 1$, this gives $(\delta_x - P) \times P^{m-1}(h) = 0$ a.s. or, since $P^m h = 0$, $(P^{m-1}h)(x) = 0$ a.s., etc. That is, $(Ph)(x_1, \ldots, x_{m-1}) = 0$ for P^{m-1} almost all $(x_1, \ldots, x_{m-1})$, thus proving that h is P-canonical. Then, the above mentioned CLT for U-statistics gives the convergence in distribution of $\{n^{\frac{m}{2}} U_n(h)\}$, and this completes the proof of Theorem 1.

4. Remarks. (1) In the case $m = 1$ the proof of Theorem 1 is easier in the sense that Section 2 is not needed, inequality (6) is just Khinchin's inequality in $I\!R$, and the last part of the proof uses the law of large numbers and the central limit theorem for sums of i.i.d. random variables (instead of the limit theorems for U-statistics). The argument replacing

Section 2 is as follows: for ξ_i i.i.d. and ε_i independent Rademacher, independent of $\{\xi_i\}$,

$$\Pr\Big\{|n^{-\frac{1}{2}}\sum_{i=1}^{n}\varepsilon_i\xi_i|>2t\Big\}\le E_\varepsilon\Big[\mathrm{P}_\xi\Big\{|n^{-\frac{1}{2}}\sum_{i\le n:\ \varepsilon_i=1}\xi_i|>t\Big\}+\mathrm{P}_\xi\Big\{|n^{-\frac{1}{2}}\sum_{i\le n:\ \varepsilon_i=-1}\xi_i|>t\Big\}\Big]$$
$$\le 2\sup_n \Pr\Big\{|n^{-\frac{1}{2}}\sum_{i=1}^{n}\xi_i|>t\Big\}.$$

(The case $m=1$ is only included here for comparison purposes and we do not claim that this is the best proof of necessity of finite variances for the CLT.) The proof of Theorem 2 in the case $m=2$ is somewhat less involved than the general case, and can be easily read off from Section 2. However, for $m=2$, there is an even simpler argument to control the tails of the distribution of the randomized (but not decoupled) U-statistic in terms of those of the original one, as follows: If $A,B\subset I\!N$ are disjoint, let

$$S_A=\sum_{i,j\in I\cap A}h(X_i,X_j)\text{ and }S_{A,B}=\sum_{(i,j)\in A\times B\cup B\times A}h(X_i,X_j).$$

Now, if $\{\varepsilon_i\}$ is a Rademacher sequence independent of $\{X_i\}$, define $A_n(\varepsilon)=\{i\le n:\ \varepsilon_i=1\}$ and $B_n(\varepsilon)=\{i\le n:\ \varepsilon_i=-1\}$, and observe

$$\sum_{i,j\le n}\varepsilon_i\varepsilon_j h(X_i,X_j)=S_{A_n(\varepsilon)}+S_{B_n(\varepsilon)}-S_{A_n(\varepsilon),B_n(\varepsilon)}$$
$$=2S_{A_n(\varepsilon)}+2S_{B_n(\varepsilon)}-\sum_{i\ne j\le n}h(X_i,X_j).$$

This gives

$$\Pr\Big\{|\sum_{i,j\le n}\varepsilon_i\varepsilon_j h(X_i,X_j)|>5t\Big\}\le \mathrm{P}_X\{|\sum_{i\ne j\le n}h(X_i,X_j)|>t\}+E_\varepsilon\mathrm{P}_X\{|S_{A_n(\varepsilon)}|>t\}$$
$$+E_\varepsilon\mathrm{P}_X\{|S_{B_n(\varepsilon)}|>t\}$$
$$\le 3\max_{k\le n}\Pr\{|\sum_{i\ne j\le k}h(X_i,X_j)|>t\}.$$

This inequality can be used instead of Theorem 2 in the proof of Theorem 1 with only one change: now the analogue of inequality (6) does not follow from recursive use of Khinchin's inequality as above, but from Bonami's (1970) work.

(2) The *symmetry* condition on h cannot be completely dropped in Theorem 1: if $h(x,y)$ is antisymmetric i.e., $h(x,y)=-h(y,x)$, then $\sum_{i\ne j\le n}h(X_i,X_j)=0$. If h is not symmetric it can be symmetrized, for instance,

$$n^{-1}\sum_{i\ne j\le n}h(X_i,X_j)=(2n)^{-1}\sum_{i\ne j\le n}\big(h(X_i,X_j)+h(X_j,X_i)\big)$$

ant tightness of this sequence does imply, by Theorem 1,

$$E\big(h(X_1,X_2)+h(X_2,X_1)\big)^2<\infty,$$

but, as seen in the extreme antysimmetric case, this does not generally imply $Eh^2<\infty$.

(3) If B is a cotype 2 Banach space, then there is an analogue to Theorem 1. The result of Section 2 is in fact stated for B-valued h. Inequality (6) is also valid in cotype 2 spaces, in the following form: There exist positive constants $c_m=c_m(B)$, depending on m and the space B, such that

$$E\Big\|\sum_{i_1,\dots,i_m\le n} a_{i_1,\dots,i_m}\varepsilon^{(1)}_{i_1}\dots\varepsilon^{(m)}_{i_m}\Big\| \ge c_m\Big(\sum_{i_1,\dots,i_m\le n}\|a_{i_1,\dots,i_m}\|^2\Big)^{\frac12}, \tag{6'}$$

because Khinchin's inequality holds in any Banach space and by the defining cotype 2 inequality (these two facts allow for the arguments above (6) in the proof of Theorem 1). The law of large numbers for B-valued U-statistics $U_n(H)$ holds as long as $E\|H\|<\infty$ (Arcones and Giné, 1991)). So, the proof of Theorem 1, with only formal changes, yields that if $\{\|n^{\frac m2}U_n(h)\|\}$ is stochastically bounded then $E\|h\|^2<\infty$. The final part of the proof of Theorem 1 applied to $f(h)$, $f\in B'$ shows $Eh(X_1,x_2\dots,x_m)=0$ for almost every $(x_2,\dots,x_m)\in S^{m-1}$ (B can be assumed to be separable, so that the unit ball of B' is separable for the weak-star topology, and this is all that is needed to take care of the sets of P^{m-1}-measure zero on which $Ef\big(h(X_1,x_2\dots,x_m)\big)=0$.) We have thus proved:

Theorem 3. *Let B be a cotype 2 Banach space and let h be a B-valued measurable, symmetric function on S^m. If the sequence $\{\|n^{\frac m2}U_n(h)\|\}_{n=1}^\infty$ is stochastically bounded, then $E\|h(X_1,\dots,X_m)\|^2<\infty$ and $Eh(X_1,x_2\dots,x_m)=0$ for almost every $(x_2,\dots,x_m)\in S^{m-1}$.*

If B is not of cotype 2 then Theorem 3 is not even true for $m=1$. The proof of Theorem 1, only with formal changes that we skip, shows that in a general Banach space B, if the sequence $\{\|n^{\frac m2}U_n(h)\|\}_{n=1}^\infty$ is stochastically bounded, then

$$\sup_{f\in B',\|f\|\le1} E\big[f\big(h(X_1,\dots,X_m)\big)\big]^2<\infty.$$

Acknowledgement. We would like to thank Peter Stiller for several useful conversations regarding Lemma 1 above.

References

Araujo, A. and Giné, E. (1980). *The central limit theorem for real and Banach valued random variables.* Wiley, New York.

Arcones, M. A. and Giné, E. (1991). Limit theorems for U-processes. *Ann. Probability,* to appear.

BONAMI, A. (1970). Étude des coefficients de Fourier des fonctions de $L^p(G)$. *Ann. Inst. Fourier* **20**, 335-402.

BRETAGNOLLE, J. (1983). Lois limites du bootstrap de certaines fonctionelles. *Ann. Inst. H. Poincaré* Sect. B. **3** 256-261.

DYNKIN, E. B. AND MANDELBAUM, A. (1983). Symmetric statistics, Poisson point processes and multiple Wiener integrals. *Ann. Statist.* **11** 739-745.

FELLER, W. (1935). Ueber den zentralen Grenswertsatz der Wahrscheinlichkeitsrechnung. *Math. Zeits.* **40** 521-559.

GINÉ, E. AND ZINN, J. (1992). Marcinkiewicz type laws of large numbers and convergence of moments for U-statistics, in *Probability in Banach Spaces, 8*, pp. 273-291. Progress in Probability Series, No. 30, Birkhäuser, Boston.

KAHANE, J. P. (1968). *Some random series of functions.* D. C. Heath, Lexington, Massachusetts.

KHINCHIN, A. YA. (1935). Sul dominio di attrazione della legge di Gauss. *Giorn. Ist. Ital. Attuari* **6** 371-393.

LÉVY, P. (1935). Proprietés asymptotiques des sommes de variables aléatoires indépen–dantes ou enchaînés. *J. Math. Pures Appl.* **14** 347-402.

RUBIN, M. AND VITALE, R. A. (1980). Asymptotic distribution of symmetric statistics. *Ann. Statist.* **8**, 165-170.

SERFLING, R. J. (1980). *Approximation Theorems of Mathematical Statistics.* Wiley, New York.

Evarist Giné
Department of Mathematics
University of Connecticut
Storrs, CT 06269

Joel Zinn
Department of Mathematics
Texas A&M University
College Station, TX 77843