ON INTEGRABILITY IN THE LIL FOR DEGENERATE U–STATISTICS

BY EVARIST GINÉ*

University of Connecticut

Runninghead: LIL for U-statistics.

Abstract

It is shown by example that square integrability of the kernel h of a completely degenerate U–statistic is not a necessary condition for the law of the iterated logarithm to hold. h must however enjoy some integrability. Sufficient conditions for a degenerate U–statistic of order 2 to satisfy the LIL, weaker than square integrability of the kernel, are also given.

August 1994

Key words and phrases: law of the iterated logarithm, completely degenerate (or canonical) U-statistics

* Research partially supported by NSF Grant No. DMS-93–00725. This work was completed at the University of Bielefeld, supported by the Deutsche Forshungsgemeinshaft, Sonderforschungbereich 343, 'Diskrete Structuren in der Mathematik'.

1. Introduction. Recently Delhing (1989) and Arcones and Giné (1993b) proved that, for *completely degenerate* kernels, if $Eh^2 < \infty$ then the law of the iterated logarithm

$$\limsup_{n\to\infty} \frac{1}{(2n\log\log n)^{\frac{m}{2}}} \Big| \sum_{(i_1,\dots,i_m)\in I_n^m} h(X_{i_1},\dots,X_{i_m})\Big| < \infty \quad \text{a.s.} \tag{1.1}$$

holds and that, moreover, the limit set of the sequence without the absolute values is, almost surely,

$$K_m(h) = \{Eh(X_1,\dots,X_m)g(X_1)\cdots g(X_m) : Eg(X_1)^2 \le 1\}. \tag{1.2}$$

Here the notation is standard: X, X_i are independent identically distributed $(S,\mathcal{S})$ valued random variables; $h : S^m \to \mathbf{R}$ is a measurable function invariant under permutations of the coordinates (a *symmetric* kernel) and satisfies

$$Eh(X, x_2, \dots, x_m) = 0$$

almost surely for $\mathcal{L}(X)^{m-1}$ (a *completely degenerate* or canonical kernel for the law of X); and I_n^m is the set of all m–dimensional vectors whose coordinates form a set of m *different* natural numbers not larger than n. In all that follows we use the notation $\log x$ to really mean $1 \vee \log x$.

If $m \ge 2$ the set K_m may be a bounded set and yet $Eh^2 = \infty$. This indicates that, for $m \ge 2$, the condition $Eh^2 < \infty$ may not be necessary for the LIL (although it is, of course, for $m = 1$). Another indication that this may be the case is the Borell type exponential bound for Rademacher chaos in the line, Theorem 4.11 in Ledoux and Talagrand (1991) with $B = I\!R$ (see (2.10) below). This inequality implies that the bounded LIL for

$$\sum_{I_n^2} \varepsilon_i\varepsilon_j h(X_i, X_j)$$

reduces to the following two limits

$$\limsup_n \frac{1}{n^2\log\log n} \sum_{I_n^2} h^2(X_i, X_j) < \infty \text{ a.s.}$$

and

$$\limsup_n \frac{1}{n} \sup_{|u|\le 1} \Big|\sum_{I_n^2} u_i u_j h(X_i, X_j)\Big| < \infty \text{ a.s.}$$

(where $u = (u_1, \dots, u_n)$ and $|u|^2 = \sum_{i=1}^n u_i^2$), none of which requires $Eh^2 < \infty$. In fact the first holds if $E(h^2/\log\log|h|) < \infty$, even with a.s. limit zero. But we do not know how to handle the second in general.

The main object of this note is to present an example of a degenerate kernel h of two variables such that $Eh^2 = \infty$, actually, $E(h^2/\log\log|h|) = \infty$, but such that the corresponding U–statistic satisfies the LIL. In order to avoid complications the case $m > 2$

is not considered but we are confident that the same type of examples can be given in general. This is presented in Section 2.

We show in Section 3 that $Eh^2/(\log\log|h|)^{m+\delta} < \infty$ for some $\delta > 0$ is necessary for the limit (1.1) to hold in probability; hence some integrability is necessary for the LIL.

In the last section, based on the Ledoux–Talagrand exponential inequality for Rademacher chaos variables, we prove a LIL result that contains the LIL for square integrable kernels. Although it is the strongest LIL for degenerate U–statistics to date, this result is not satisfactory as the conditions are not explicitly in terms of the law of $h(X_1, X_2)$ alone.

I thank M. Arcones, W. Madych and J. Zinn for discussions and comments concerning the content of this paper.

2. The Example. In this section logarithms are supposed to have base 2 although we do not change notation (this is not important, but some constants become nicer). Let X, X_i be i.i.d. uniform on $[0,1]$ and let ε, ε_i be i.i.d. Rademacher variables (i.e., $P\{\varepsilon = 1\} = P\{\varepsilon = -1\} = \frac{1}{2}$), independent from the X's. Set $\tilde{X}_i = (X_i, \varepsilon_i)$, a random varible with values in $S := [0,1] \times \{-1,1\}$. Let

$$I_r = (1 - 2^{1-r}, 1 - 2^{-r}],\ r \in \mathbb{N},$$

and let I_r denote as well its own indicator function. We define, on S^2, the function

$$h((x,\alpha),(y,\beta)) = \sum_{r=1}^{\infty} \frac{2^r}{r^{\frac{1}{2}}} I_r(x) I_r(y) \alpha\beta,\ x, y \in [0,1],\ \alpha, \beta = -1, 1. \tag{2.1}$$

Then h is finite,

$$E\frac{h^2(\tilde{X}_1, \tilde{X}_2)}{\log\log|h(\tilde{X}_1, \tilde{X}_2)|} = \infty \tag{2.2}$$

and

$$E\frac{h^2(\tilde{X}_1, \tilde{X}_2)}{(LL|h(\tilde{X}_1, \tilde{X}_2)|)^{1+\delta}} < \infty \text{ for all } \delta > 0. \tag{2.3}$$

Moreover, if f and g are square integrable on $[0,1]$, we have, by Hölder's inequality,

$$\begin{aligned}
\int_0^1 \int_0^1 |f(x)g(y) \sum_{r=1}^{\infty} \frac{2^r}{r^{\frac{1}{2}}} I_r(x) I_r(y)|\, dx\, dy &\le \sum_r \frac{2^r}{r^{\frac{1}{2}}} \Big(\int_{I_r} |f|\Big)\Big(\int_{I_r} |g|\Big) \\
&\le \sum_r \frac{1}{r^{\frac{1}{2}}} \Big(\int_{I_r} f^2\Big)^{\frac{1}{2}} \Big(\int_{I_r} g^2\Big)^{\frac{1}{2}} \\
&\le \sum_r \Big(\int_{I_r} f^2\Big)^{\frac{1}{2}} \Big(\int_{I_r} g^2\Big)^{\frac{1}{2}} \\
&= \|f\|_2 \|g\|_2.
\end{aligned}$$

Hence,

$$\text{l.u.b.}\big\{E|h(\tilde{X}_1, \tilde{X}_2) g(\tilde{X}_1) g(\tilde{X}_2)| : Eg^2 \le 1\big\} \le 1. \tag{2.4}$$

So, h is a prime candidate for our example.

2.1. THEOREM. *The function h defined by (2.1), which satisfies (2.2)–(2.4), also satisfies*

$$\limsup_n \frac{1}{n\log\log n}\Big|\sum_{1\le i\ne i\le j\le n} h(\tilde X_i,\tilde X_j)\Big| < \infty \quad \text{a.s.} \tag{2.5}$$

PROOF. In this proof we do not keep unnecessary accounting of constants, so c may not be the same at each appearance, and δ will denote a positive small number, also possibly different at different occurrences.

In order to prove that (2.5) holds, it suffices to show that for some $c, n_0 < \infty$,

$$\sum_{n\ge n_0} \Pr\Big\{\Big|\sum_{I_{2^n}^2} h(\tilde X_i,\tilde X_j)\Big| > c2^n\log n\Big\} < \infty. \tag{2.6}$$

To see this, we notice that (2.5) is equivalent to its decoupled version by the result of Montgomery–Smith and de la Peña (1993), then we proceed in a way similar to Lemma 2.4 in Arcones and Giné (1993b) (basically, using Lévy's inequality twice) but using probabilities instead of expectations and, finally, undecouple back. Details on decoupling expressions such as $\sup_{n\ge k}|a_k\sum_{I_n^k} h(\tilde X_i,\tilde X_j)|$ can also be found in the last mentioned paper.

The first step in the proof of (2.6) is truncation, which in turn we do in three steps. Since

$$\sum_n \Pr\{\#[i\le 2^n : X_i\in\cup_{r=r_n+1}^\infty I_r]\ge 2\} \le \sum_n \binom{2^n}{2}\frac{1}{2^{2r_n}} < \sum_n 2^{2(n-r_n)},$$

it follows that, for

$$r_n = n + \frac{1}{2}\log n + \log(\log n)^{\frac12+\delta},$$

(any $\delta>0$), we have

$$\sum_n \Pr\Big\{\Big|\sum_{I_{2^n}^2}\sum_{r=r_n+1}^\infty r^{-\frac12}2^r I_r(X_i)I_r(X_j)\varepsilon_i\varepsilon_j\Big| > c2^n\log n\Big\} < \infty. \tag{2.7}$$

In fact we want to bring r_n in (2.7) below the level n. Let us denote s_n the next intermediate level. Since

$$\sum_{r=s_n+1}^{r_n} 2^r r^{-\frac12}\sum_{I_{2^n}^2} I_r(X_i)I_r(X_j) \le 2^{r_n} r_n^{-\frac12}\big(\#[i\le 2^n : X_i\in\cup_{r=s_n+1}^{r_n} I_r]\big)^2,$$

we have, using the inequality $P\{\mathrm{Bin}(n,p)\geq k\}\leq\binom{n}{k}p^k\leq(enp/k)^k$,

$$
\begin{aligned}
\Pr\Big\{\big|\sum_{I_{2^n}^2}\sum_{r=s_n+1}^{r_n} r^{-\frac{1}{2}}2^r I_r(X_i)I_r(X_j)\varepsilon_i\varepsilon_j\big| > c2^n\log n\Big\}\\
&\leq \Pr\Big\{\#[i\leq 2^n : X_i\in\cup_{r=s_n+1}^{r_n} I_r] > c\big(r_n^{\frac{1}{2}}2^{n-r_n}\log n\big)^{\frac{1}{2}}\Big\}\\
&\leq \Big[\frac{e2^{n-s_n}}{c\big(r_n^{\frac{1}{2}}2^{n-r_n}\log n\big)^{\frac{1}{2}}}\Big]^{c\big(r_n^{\frac{1}{2}}2^{n-r_n}\log n\big)^{\frac{1}{2}}}\\
&\lesssim \Big[\frac{2^{n-s_n}}{(\log n)^{\frac{1}{4}-\frac{\delta}{2}}}\Big]^{(\log n)^{\frac{1}{4}-\frac{\delta}{2}}}
\end{aligned}
$$

as long as $2^{n-s_n}<(\log n)^{\frac{1}{4}-\frac{\delta}{2}}$. The values

$$s_n = n+(\log n)^{\frac{3}{4}+\frac{\delta}{2}}$$

make the series of the above probabilities converge. For the next level t_n we have, as above,

$$
\begin{aligned}
\Pr\Big\{\big|\sum_{I_{2^n}^2}\sum_{r=t_n+1}^{s_n} r^{-\frac{1}{2}}2^r I_r(X_i)I_r(X_j)\varepsilon_i\varepsilon_j\big| > c2^n\log n\Big\}\\
&\leq \Big[\frac{e2^{n-t_n}}{c\big(s_n^{\frac{1}{2}}2^{n-s_n}\log n\big)^{\frac{1}{2}}}\Big]^{c\big(s_n^{\frac{1}{2}}2^{n-s_n}\log n\big)^{\frac{1}{2}}}\\
&\lesssim \Big[\frac{2^{n-t_n}}{n^{\frac{1}{4}-\delta'}}\Big]^{n^{\frac{1}{4}-\delta'}}
\end{aligned}
$$

for any $\delta', t_n>0$ such that $2^{n-t_n}<n^{\frac{1}{4}-\delta'}$. Then, $t_n=n-\alpha\log n$, for $0<\alpha<\frac{1}{4}-\delta'$, makes the series of these probabilities converge. We thus conclude that the series in (2.7) converges for

$$r_n = n-\alpha\log n,\quad \text{for some } \alpha>0. \tag{2.8}$$

(By repeating the argument we can make α arbitrarily close to, but smaller than, 1/2, however $\alpha>0$ is all we need.)

The final step consists in proving

$$\sum_n \Pr\Big\{\big|\sum_{I_{2^n}^2}\sum_{r=r_0}^{r_n} r^{-\frac{1}{2}}2^r I_r(X_i)I_r(X_j)\varepsilon_i\varepsilon_j\big| > c2^n\log n\Big\}<\infty, \tag{2.9}$$

with r_n given by (2.8) and r_0 fixed (the first r_0 terms offer no problem, obviously). Then, (2.7) and (2.9) give (2.6), thus the proposition. In order to show that (2.9) holds we will apply the Borell type exponential inequality for Rademacher chaos in Theorem 4.11

of Ledoux and Talagrand (1991). On account of the equivalence of the L_p norms for Rademacher chaos, this inequality, for the Banach space $B = I\!R$, is as follows: there exist $0 < c_i < \infty$, $1 \le i \le 3$, such that for any double sequence of numbers $x_{i,j}$ with $x_{i,i} = 0$ for all i, and for all $t > 0$,

$$\Pr\Big\{\big|\sum_{i,j} \varepsilon_i\varepsilon_j x_{i,j}\big| > c_1(M(t+1) + \sigma t^2)\Big\} \le c_2 e^{-c_3 t^2}, \tag{2.10}$$

where

$$M^2 = \sum_{i,j} x_{i,j}^2 \quad \text{and} \quad \sigma = \sup_{|u|\le 1}\big|\sum_{i,j} u_i u_j x_{i,j}\big|,$$

with $|u| = \big(\sum_i u_i^2\big)^{\frac{1}{2}}$. Applied to the terms of the series in (2.9) this gives that for some C, C_i, $i \le 3$,

$$\begin{aligned}
&\Pr\Big\{\big|\sum_{I_{2^n}^2}\sum_{r=r_0}^{r_n} r^{-\frac{1}{2}} 2^r I_r(X_i)I_r(X_j)\varepsilon_i\varepsilon_j\big| > C2^n \log n\Big\} \\
&\le C_1 e^{-2\log n} + \Pr\big\{M_n > C_2 2^n (\log n)^{\frac{1}{2}}\big\} + \Pr\big\{\sigma_n > C_3 2^n\big\}
\end{aligned} \tag{2.11}$$

where

$$M_n^2 = \sum_{I_{2^n}^2}\sum_{r=r_0}^{r_n} r^{-1} 2^{2r} I_r(X_i)I_r(X_j) \text{ and } \sigma_n = \sup_{|u|\le 1}\big|\sum_{I_{2^n}^2}\sum_{r=r_0}^{r_n} u_i u_j r^{-\frac{1}{2}} 2^r I_r(X_i)I_r(X_j)\big|,$$

for $n \in I\!N$. We remark that the constants C_2 and C_3 can be taken to be arbitrarily large, just by increasing C. First we estimate σ_n. For $|u| \le 1$ we have

$$\begin{aligned}
\sum_{I_{2^n}^2}\sum_{r=r_0}^{r_n} \big|u_i u_j r^{-\frac{1}{2}} 2^r I_r(X_i)I_r(X_j)\big| &\le \sum_{r=r_0}^{r_n} 2^r r^{-\frac{1}{2}} \big(\sum_{i\le 2^n : X_i \in I_r} |u_i|\big)^2 \\
&\le \sum_{r=r_0}^{r_n} 2^r r^{-\frac{1}{2}} \big(\#[i \le 2^n : X_i \in I_r]\big)\big(\sum_{i\le 2^n : X_i\in I_r} u_i^2\big) \\
&\le \max_{r_0 \le r \le r_n} 2^r r^{-\frac{1}{2}} \big(\#[i \le 2^n : X_i \in I_r]\big).
\end{aligned}$$

Hence,

$$\begin{aligned}
\Pr\big\{\sigma_n > C_3 2^n\big\} &\le \sum_{r=r_0}^{r_n} \Pr\big\{\#[i \le 2^n : X_i \in I_r] > C_3 2^{n-r} r^{\frac{1}{2}}\big\} \\
&\le \sum_{r=r_0}^{r_n} \big(\frac{e}{C_3 r^{\frac{1}{2}}}\big)^{C_3 2^{n-r} r^{\frac{1}{2}}}.
\end{aligned}$$

For r_0 large enough (independently of n) and $r_0 \le r \le r_n$ we have $\frac{e}{C_3 r^{\frac{1}{2}}} < 1$ and $C_3 2^{n-r} r^{\frac{1}{2}} > 2^{n-r_n} + 1 = n^\alpha + 1$. Therefore,

$$\Pr\{\sigma_n > C_3 2^n\} \le \frac{1}{(r_0-1)^{n^\alpha}},$$

that is,

$$\sum_n \Pr\{\sigma_n > C_3 2^n\} < \infty. \tag{2.12}$$

Finally, we estimate M_n i.e., we prove that

$$\sum_n \Pr\{M_n^2 > C_2^2 2^{2n} \log n\} < \infty. \tag{2.13}$$

We show first that we can degenerate the U–statistic corresponding to M_n^2. The centering term is small enough so that we can center:

$$EM_n^2 = \sum_{I_{2^n}^2} \sum_{r=r_0}^{r_n} r^{-1} \le 2^{2n} \log n < C_2^2 2^{2n} \log n.$$

Next we see that we can ignore the linear terms in the Hoeffding decomposition of M^2: since $\mathrm{Var}\big(\sum_{r=r_0}^{r_n} r^{-1} 2^r I_r\big) \le \sum_{r=r_0}^{r_n} 2^r r^{-2} \lesssim 2^{r_n} r_n^{-2}$, we have

$$\begin{aligned}\Pr\Big\{2^n\Big|\sum_{i=1}^{2^n}\sum_{r=r_0}^{r_n} r^{-1}2^r I_r(X_i) - 2^n \sum_{r=r_0}^{r_n} r^{-1}\Big| > 2^{2n}\log n\Big\} &\le \frac{2^{n+r_n} r_n^{-2}}{2^{2n}(\log n)^2} \\ &\lesssim \frac{2}{n^{2+\alpha}(\log n)^2},\end{aligned}$$

which is the general term of a convergent series. Therefore, letting $\pi_{2,2} f(x,y) = f(x,y) - Ef(X,y) - Ef(x,Y) + Ef$ be the operator projecting over degenerate functions in Hoeffding's decomposition, the previous two estimates show that the proof of (2.13) reduces to the verification of

$$\sum_n \Pr\Big\{\Big|\sum_{I_{2^n}^2} \pi_{2,2}\Big(\sum_{r=r_0}^{r_n} r^{-1} 2^{2r} I_r(X_i) I_r(X_j)\Big)\Big| > 2^{2n}\log n\Big\} < \infty. \tag{2.14}$$

Since degenerate U–statistics consist of orthogonal terms, Chebyshev's inequality gives that the general term of this series is dominated by

$$\begin{aligned}\frac{2^{2n} E\big(\sum_{r=r_0}^{r_n} r^{-1} 2^{2r} I_r(X_1) I_r(X_2)\big)^2}{2^{4n}(\log n)^2} &= \frac{2^{2n}\sum_{r=r_0}^{r_n} r^{-2} 2^{4r} E\big(I_r(X_1)I_r(X_2)\big)}{2^{4n}(\log n)^2} \\ &= \frac{2^{2n}\sum_{r=r_0}^{r_n} 2^{2r} r^{-2}}{2^{4n}(\log n)^2} \\ &\lesssim \frac{2^{2(n+r_n)}}{2^{4n}(\log n)^2 r_n^2} \\ &\lesssim \frac{2}{(\log n)^2 n^{2+2\alpha}}.\end{aligned}$$

Therefore, the series in (2.14) converges. Thus, (2.13) holds and this, together with (2.12), gives, by (2.11), convergence of the series in (2.9). $\square$

It is easy to see that c in (2.7) can be made arbitrarily small. But C_2 in (2.11) must at least be 1 because of the size of EM_n^2. So, this proof does not seem to show that we have

$$\sum_n \Pr\Big\{\big|\sum_{I_{2^n}^2}\sum_{r=r_0}^{\infty} r^{-\frac{1}{2}} 2^r I_r(X_i)I_r(X_j)\varepsilon_i\varepsilon_j\big| > C(r_0)2^n\log n\Big\} < \infty$$

for $C(r_0) \to 0$ as $r_0 \to \infty$. In particular we do not know if this example satisfies the compact LIL with limit set $K_2(h)$. On the other hand, if we take the coefficients of h in (2.1) to be $2^r/(r^{1/2}\alpha_r)$ with $\alpha_r \to \infty$ slowly, e.g. $\alpha_r = \log\log\log r$, then carefully going over the last proof shows that indeed the above series is finite for all r_0 and some $C(r_0) \to 0$. Using the LIL for square integrable kernels and the computation above (2.4) (starting at $r = r_0$ and taking $r_0^{-1/2}$ as common factor in the third and fourth line) in some of the arguments in the proof of Theorem 3.1 in Arcones and Giné (1993b), we conclude that, in this case,

$$\lim\mathrm{set}\Big\{\frac{1}{2n\log\log n}\sum_{I_{2^n}^2} h(\tilde{X}_1,\tilde{X}_2)\Big\} = K_2(h)$$

almost surely. In other words, *there exist degenerate kernels h that satisfy (2.2), and even (2.3), as well as the compact LIL with limit set $K_2(h)$.*

3. Necessity of some integrability. Giné and Zinn's (1994) proof of necessity of $Eh^2 < \infty$ for the CLT in the degenerate case can be adapted to show that some integrability is necessary for the stochastic boundedness of the LIL sequence, hence also for the LIL. The only difference is: at a certain point in that proof we use the LLN; we cannot use it here and instead we must rely on an extension of Hoffmann–Jørgensen's inequality to U–statistics with positive kernels (Giné and Zinn, 1992a, for degenerate kernels). The proof for positive kernels is formally the same as the proof in the middle of page 88 in the paper just mentioned: one just invokes Hoffmann-Jørgensen's inequality for sums of non–negative independent real random variables and for sums of non–negative integrable functions in L_1 [Hoffmann–Jørgensen's inequality holds for these two types of variables because for them, the maximum of the norms (absolute value, L_1 norm) of the partial sums is the norm of the total sum, thus rendering Lévy's inequality unnecessary in this case: e.g. making this replacement in inequality (6.8), page 156, Ledoux and Talagrand (1991), gives it]. We state the inequality separately, but only for decoupled U–statistics (all we need), although, by decoupling, it holds also for regular (undecoupled) U–statistics.

3.1. LEMMA. *Let $X_i, X_i^{(j)}$ be i.i.d. S–valued random variables. Let $h : S^m \to \mathbb{R}_+ \cup \{0\}$ be measurable non–negative functions symmetric in their entries. Then, for all $p > 1$, there exist universal constants $c_{1,p}$, $c_{2,p}$ and $c_{3,p}$ finite, positive, such that*

$$E\Big[\sum_{i_1,\dots,i_m\le n} h(X_{i_1}^{(1)},\dots,X_{i_m}^{(m)})\Big]^p \le c_{1,p}E\Big[\max_{i_1\le n}\sum_{i_2,\dots,i_m\le n} h(X_{i_1}^{(1)},\dots,X_{i_m}^{(m)})\Big]^p + c_{2,p}t_0^p,$$

for any t_0 *satisfying* $\Pr\{\sum_{i_1,\dots,i_m\le n} h(X_{i_1}^{(1)},\dots,X_{i_m}^{(m)}) \ge t_0\} \le c_{3,p}$.

The integrability condition is as follows:

3.2. Proposition. *If the sequence*

$$\Big\{\frac{1}{(n\log\log n)^{\frac{m}{2}}}\Big|\sum_{(i_1,\dots,i_m)\in I_n^m} h(X_{i_1},\dots,X_{i_m})\Big|\Big\} \tag{3.1}$$

is stochastically bounded then

$$E\frac{h^2}{(\log\log|h|)^{m+\delta}}<\infty$$

for all $\delta>0$, *and* h *is degenerate.*

Proof. The proof for general m is not different from the proof for $m=2$, so we consider only this case. The first part of the proof of Theorem 1 in Giné and Zinn (1994) shows that the stochastic boundedness of the sequence (3.1) ($m=2$) implies that the sequence

$$\Big\{\frac{1}{(n\log\log n)^2}\sum_{i,j\le n} h^2(X_i^{(1)},X_j^{(2)})\Big\}$$

is stochastically bounded. By positivity we can truncate at any levels and still have tightness, uniformly over the truncations. Then, the previous lemma (for any $p>1$) and Hölder give

$$\sup_n E\frac{1}{(n\log\log n)^2}\sum_{i,j\le n} h^2 I_{h^2\le n(\log\log n)^2}(X_i^{(1)},X_j^{(2)})<\infty,$$

that is,

$$Eh^2 I_{h^2\le n(\log\log n)^2}\le c(\log\log n)^2$$

for some $c<\infty$ and all n. Therefore,

$$\begin{aligned} E\frac{h^2}{(\log\log|h|)^{2+\delta}} &= \sum_k E\Big[\frac{h^2}{(\log\log|h|)^{2+\delta}} I_{2^{2^k}2^{2k}<h^2\le 2^{2^{k+1}}2^{2(k+1)}}\Big] \\ &\lesssim \sum_k \frac{1}{2^{(2+\delta)k}} Eh^2 I_{h^2\le 2^{2^{k+1}}2^{2(k+1)}}<\infty, \end{aligned}$$

proving necessity of the integrability condition. Consider now Hoeffding's decomposition of the U–statistic with kernel h, multiplied by $n/\log\log n$; the constant term is then of the order $n/\log\log n$ unless it is zero; the linear term is a.s. $o(n/\log\log Ln)$ by the law of large numbers and, by the CLT, it is at least $O_p(n^{1/2}/\log\log n)$ unless it vanishes; the degenerate term is at most $o(n^\delta/\log\log n)$ a.s. (for any $\delta>0$) by the integrability condition and the Marzinkiewicz law of large numbers for degenerate U–statistics (Giné and Zinn, 1992b). Hence the constant term and the linear term are each either zero or a.s. unbounded and no cancellations occurs among the three different terms since they are each of different order. Therefore, if the bounded LIL holds for h, the constant term and the linear term must vanish, proving that h is degenerate (for more details, see e.g. a similar argument in Giné and Zinn, 1994, used to prove degeneracy in the CLT setting). $\square$

Note that if the normings in (3.1) were $n^{m/2}$ instead of $(n \log\log n)^{m/2}$, then this proof would give $Eh^2 < \infty$: in this case, tightness of the normalized sums of squares and Lemma 3.1 give

$$\sup_n E\frac{1}{n^2}\sum_{i,j\leq n} h^2 I_{h^2\leq n}(X_i^{(1)}X_i^{(2)}) < \infty,$$

that is $\sup_n Eh^2 I_{h^2\leq n} < \infty$ i.e., $Eh^2 < \infty$ (and similarly for general m). This is therefore another ending to the proof of the main theorem in Giné and Zinn (1994).

If instead of the stochastic boundedness of the sequence (3.1), we asume the bounded LIL, then, with some extra care (e.g., as in Cuzick et al., 1993), the first inequality in the above proof can be strengthened to

$$E\sup_n \frac{1}{(n\log\log n)^2}\sum_{I_n^2} h^2 I_{h^2\leq n(\log\log n)^2}(X_i^{(1)}, X_j^{(2)}) < \infty,$$

but it is not clear that this implies more integrability. We have the impression that, at the very least, the proposition is true with $\delta = 0$. But, even if δ is removed, there is a gap between the present result and the integrability enjoyed by the example in Section 2.

4. Towards a better LIL. The Arcones and Giné proof of the LIL for square integrable kernels rests on the hypercontractivity inequality for Rademacher chaos. A better estimate of the size of a Rademacher chaos variable is given by the Ledoux–Talagrand inequality (2.10) and this makes the proof of the LIL for the example in Section 2 possible. We indicate here a general proof of the LIL for not necessarily square integrable degenerate kernels based on inequality (2.10). In what follows X, X_i, X', X_i' denote i.i.d. S–valued random variables, $\varepsilon_i, \varepsilon_i'$ are i.i.d. Rademacher variables independent of the X variables, E denotes total expectation, and E_1 and E_2, expectation with respect to the X, ε variables and the X', ε' variables respectively, $u = (u_1, \ldots, u_n)$, $u' = (u_1, \ldots, u_n)$ are vectors in $\mathbb{R}^n$, n varying, and $|u|, |u'|$ denote their Euclidean norm.

4.1. Theorem. *Let $h : S^2 \to \mathbb{R}$ be a symmetric measurable kernel, completely degenerate for the law of X. If*

$$E\frac{h^2(X,X')}{\log\log|h(X,X')|} < \infty \tag{4.1}$$

and

$$\limsup_{n\to\infty}\frac{1}{n}\sup_{|u|\leq 1}\Big|\sum_{I_n^2} u_i u_j h I_{h^2\leq (i\vee j)^2\log\log i\vee j}(X_i,X_j)\Big| < \infty \text{ a.s.}, \tag{4.2}$$

then

$$\limsup_{n\to\infty} \frac{1}{n\log\log n} |\sum_{I_n^2} h(X_i, X_j)| < \infty \text{ a.s.} \tag{4.3}$$

By Cauchy–Schwartz, for each n, the n–th random variable in (4.2) is dominated by the square root of $\frac{1}{n^2}\sum_{I_n^2} h^2(X_i, X_j)$, which, by the law of large numbers for U–statistics, tends to Eh^2 a.s. if h has finite second moment. So, this theorem implies the bounded LIL under finite second moment. This theorem does not apply to the example of Section 2, which has less integrability than required here, but it is easy to check (given the computations in Section 2) that it applies, for instance, to the kernel obtained from (2.1) by multiplying each denominator $r^{\frac{1}{2}}$ by e.g. $(\log r)^{\frac{1}{2}}$ (the resulting kernel still misses to be square integrable). However, in general, condition (4.2) is not easy to verify, and we do not know how to reduce it to a condition on simple characteristics of the law of $h(X_1, X_2)$. For instance, it would be interesting to know whether continuity or, more conservatively, compactness of the operator $g(\cdot) \to E_X h(X, \cdot)g(X)$ from $L_2(\mathcal{L}(X))$ into itself, implies (4.2), at least in the presence of (4.1).

Proof of Theorem 4.1. By the decoupling inequalities of Montgomery–Smith and de la Peña (1993), the LIL (4.3) is equivalent to its decoupled analogue

$$\limsup_{n\to\infty} \frac{1}{n\log\log n} |\sum_{I_n^2} h(X_i, X_j')| < \infty \text{ a.s.} \tag{4.4}$$

[The limit being a.s. a constant by the Hewitt–Savage zero–one law, (4.3) is equivalent to

$$\lim_{k\to\infty} \Pr\{\sup_{n\geq k} \frac{1}{n\log\log n} |\sum_{I_n^2} h(X_i, X_j)| > c\} = 0,$$

for some $c < \infty$, and likewise for (4.4); for each k, the random variable in this expression is the norm of a generalized U–statistic with kernels in $c_0(\mathbb{R})$, as indicated in the proof of Theorem 2.5 in Arcones and Giné, 1993b, and these statistics satisfy the decoupling inequalities].

Condition (4.1) implies

$$\sum_n n \Pr\{h^2 > n^2 \log\log n\} < \infty,$$

hence

$$\Pr\{\max_{i\leq j} h^2 I_{h^2 > j^2 \log\log j}(X_i, X_j') > 0 \text{ i.o.}\} = 0.$$

This implies

$$\frac{1}{n\log\log n} \sum_{I_n^2} |h|^p I_{h^2 > (i\vee j)^2 \log\log i\vee j}(X_i, X_j') \to 0 \text{ a.s.} \tag{4.5}$$

for all $p > 0$. In particular, the proof of the theorem reduces to showing

$$\limsup_{n\to\infty} \frac{1}{n \log\log n} \Big| \sum_{I_n^2} h I_{h^2 \le (i\vee j)^2 \log\log i\vee j}(X_i, X_j') \Big| < \infty \text{ a.s.} \tag{4.6}$$

To simplify notation we write h_r for $hI_{h^2 \le r^2 \log\log r}$. Next we show how, using the symmetrization Lemma 7.1 in Ledoux and Talagrand (1991), we can randomize with Rademacher variables in (4.6) (this argument was first made in an earlier version of Arcones and Giné (1993b) for degenerate kernels and I thank M. Arcones for allowing me to indicate it here; in addition, in the present situation we must take care of the centerings). Lemma 7.1 in the Ledoux–Talagrand book, can be modified to read: If $\{Y_n\}$, $\{Y_n'\}$ are independent sequences of random vectors, $\limsup \|Y_n - Y_n'\| \le M$ a.s. and $\limsup E\|Y_n\| \le A$ then $\limsup \|Y_n\| \le M + A$ a.s. Using this lemma conditionally and then randomizing the conditionally symmetric random variables (possible since the limsup's involved are a.s. constant), we obtain:

$$\begin{aligned}\limsup_n \frac{1}{n \log\log n} \Big| \sum_{I_n^2} h_{i\vee j}(X_i, X_j') \Big| \le & 2 \limsup_n \frac{1}{n \log\log n} \Big| \sum_{I_n^2} \varepsilon_j' h_{i\vee j}(X_i, X_j') \Big| \\ & + \limsup_n \frac{1}{n \log\log n} E_2 \Big| \sum_{I_n^2} h_{i\vee j}(X_i, X_j') \Big|. \quad (4.7)\end{aligned}$$

Now, by well known symmetrization inequalities for sums of independent random variables, we have

$$E_2 \Big| \sum h_{i\vee j} \Big| \le E_2 \Big| \sum h_{i\vee j} - E_2 \sum h_{i\vee j} \Big| + \Big| \sum E_2 h_{i\vee j} \Big| \le 2E_2 \Big| \sum \varepsilon_j' h_{i\vee j} \Big| + \Big| \sum E_2 h_{i\vee j} \Big|.$$

Let us consider the last summand in this inequality: using degeneracy of h, the integrability condition (4.1) and the law of large numbers, we have

$$\begin{aligned}\frac{1}{n \log\log n} \Big| \sum_{I_n^2} E_2 h_{i\vee j}(X_i, X') \Big| &= \frac{1}{n \log\log n} \Big| \sum_{I_n^2} E_2 (h - h_{i\vee j})(X_i, X') \Big| \\ &\lesssim \frac{1}{n \log\log n} \sum_{I_n^2} \frac{(\log\log i\vee j)^{\frac12}}{i \vee j} E_2 \frac{h^2}{\log\log h}(X_i, X') \\ &\le \frac{1}{n(\log\log n)^{\frac12}} \Big(\sum_{i<j\le n} \frac{1}{j} E_2 \frac{h^2}{\log\log h}(X_i, X') + \sum_{j<i\le n} \frac{1}{i} E_2 \frac{h^2}{\log\log h}(X_i, X') \Big) \\ &\le \frac{1}{n(\log\log n)^{\frac12}} \Big(\sum_{j=2}^{n} \Big(\frac{1}{j} \sum_{i=1}^{j+1} E_2 \frac{h^2}{\log\log h}(X_i, X') \Big) + \sum_{i=1}^{n-1} E_2 \frac{h^2}{\log\log h}(X_i, X') \Big) \\ &= O\Big(\frac{1}{(\log\log n)^{\frac12}} \Big) \text{ a.s.}\end{aligned}$$

Hence, (4.7) becomes

$$\limsup_n \frac{1}{n\log\log n}\Big|\sum_{I_n^2} h_{i\vee j}(X_i,X_j')\Big| \le 2\limsup_n \frac{1}{n\log\log n}\Big|\sum_{I_n^2}\varepsilon_j' h_{i\vee j}(X_i,X_j')\Big|$$
$$+\,2\limsup_n \frac{1}{n\log\log n}E_2\Big|\sum_{I_n^2}\varepsilon_j' h_{i\vee j}(X_i,X_j')\Big|.$$

Now we can repeat this procedure using, in the last step, that

$$\begin{aligned}
\frac{1}{n\log\log n}E_2\Big|E_1\big(\sum_{I_n^2}\varepsilon_j' h_{i\vee j}(X_i,X_j')\big)\Big| &= \frac{1}{n\log\log n}E_2\Big|E_1\big(\sum_{I_n^2}\varepsilon_j'(h-h_{i\vee j})(X_i,X_j')\big)\Big| \\
&\le \frac{1}{n\log\log n}E_2\Big[\sum_{j=2}^{n}(j-1)|E_1(h-h_j)(X,X_j')| + \sum_{j=1}^{n-1}\Big|\sum_{i=j+1}^{n}E_1(h-h_i)(X,X_j')\Big|\Big] \\
&\le \frac{1}{n\log\log n}\Big[\sum_{j=2}^{n}(j-1)\frac{(\log\log j)^{\frac12}}{j} + \sum_{j=1}^{n-1}\sum_{i=j+1}^{n}\frac{(\log\log i)^{\frac12}}{i}\Big]E\frac{h^2}{\log\log|h|} \\
&< \frac{2}{(\log\log n)^{\frac12}}E\frac{h^2}{\log\log|h|} = O\big(\frac{1}{(\log\log n)^{\frac12}}\big),
\end{aligned}$$

and obtain, as the final result, the following randomization inequality:

$$\begin{aligned}
\limsup_n \frac{1}{n\log\log n}\Big|\sum_{I_n^2} h_{i\vee j}(X_i,X_j')\Big| \le\, & 4\limsup_n \frac{1}{n\log\log n}\Big|\sum_{I_n^2}\varepsilon_i\varepsilon_j' h_{i\vee j}(X_i,X_j')\Big| \\
&+ 8\limsup_n \frac{1}{n\log\log n}E_2\Big|\sum_{I_n^2}\varepsilon_i\varepsilon_j' h_{i\vee j}(X_i,X_j')\Big| \\
&+ 4\limsup_n \frac{1}{n\log\log n}E\Big|\sum_{I_n^2}\varepsilon_i\varepsilon_j' h_{i\vee j}(X_i,X_j')\Big| \\
&=: 4(I)+8(II)+4(III) \qquad (4.8)
\end{aligned}$$

Once we have randomized we are ready to conditionally apply the properties of Rademacher sums and Rademacher chaos variables. We have, e.g. by iterating Khinchin's inequality,

$$\begin{aligned}
E\Big|\sum_{I_n^2}\varepsilon_i\varepsilon_j' h_{i\vee j}(X_i,X_j')\Big| &\lesssim E\big(\sum_{I_n^2}h_n^2(X_i,X_j')\big)^{\frac12} \le n(Eh_n^2)^{\frac12} \\
&\le n(\log\log n)^{\frac12}\Big(\frac{t^2}{\log\log n} + E\big[\frac{h^2}{\log\log|h|}I_{|h|>t}\big]\Big)^{\frac12},
\end{aligned}$$

and therefore, by letting $n\to\infty$ and then $t\to\infty$,

$$E\Big|\sum_{I_n^2}\varepsilon_i\varepsilon_j' h_{i\vee j}(X_i,X_j')\Big| = o\big(n(\log\log n)^{\frac12}\big). \qquad (4.9)$$

Thus, $(III) = 0$ in (4.8). (I) and (II) in (4.8) require some more work.

Proof of $(II) < \infty$ *a.s.* We can write

$$
\begin{aligned}
E_2\Big|\sum_{I_n^2} \varepsilon_i \varepsilon_j' h_{i\vee j}(X_i, X_j')\Big| &\le E_2\Big|\sum_{i=2}^{n} \varepsilon_i \Big(\sum_{j=1}^{i-1} \varepsilon_j' h_i(X_i, X_j')\Big)\Big| \\
&\quad + E_2\Big|\sum_{i=1}^{n-1} \varepsilon_i \Big(\sum_{j=i+1}^{n} \varepsilon_j' h_i(X_i, X_j')\Big)\Big| \\
&:= \Big\|\sum_{i=2}^{n} \varepsilon_i Y_i\Big\| + \Big\|\sum_{i=1}^{n-1} \varepsilon_i Z_{n,i}\Big\|, \qquad (4.10)
\end{aligned}
$$

where $\{Y_i\}$ are independent $L_1(E_2)$–valued random vectors, and likewise for $\{Z_{n,i}\}$ for each n. To show that

$$\limsup_n \frac{1}{n \log\log n} \Big\|\sum_{i=2}^{n} \varepsilon_i Y_i\Big\| < \infty \text{ a.s.} \qquad (4.11)$$

we apply Lemma 2.1 in Marcus and Zinn (1984), taking into account that for symmetric variables and regular norming sequences their condition (2.3) is superfluous. This lemma shows that (4.11) is equivalent to

$$\sum_n \Pr\Big\{\Big\|\sum_{i=2}^{2^n} \varepsilon_i Y_i\Big\| > c 2^n \log n\Big\} < \infty \text{ a.s.} \qquad (4.12)$$

for some $c < \infty$.To prove (4.12) we use the exponential inequality for B–valued Rademacher sums in Ledoux and Talagrand (1991), Theorem 4.7, which we now recall in the less precise form that we need:

$$\Pr\Big\{\Big\|\sum \varepsilon_i x_i\Big\| \ge c_1\big(E\Big\|\sum \varepsilon_i x_i\Big\| + \tilde{\sigma} t\big)\Big\} < c_2 e^{-c_3 t^2} \qquad (4.13)$$

for some finite positive c_i, $i \le 3$, and where $\tilde{\sigma}^2 = \sup_{f \in B', \|f\| \le 1} \sum f^2(x_i)$. In our case, for each n,

$$
\begin{aligned}
E\Big\|\sum \varepsilon_i x_i\Big\| &= E_\varepsilon E_2\Big|\sum_{i=2}^{2^n} \varepsilon_i \Big(\sum_{j=1}^{i-1} \varepsilon_j' h_i(X_i, X_j')\Big)\Big| \\
&\simeq E_2\Big[\sum_{i=2}^{2^n}\Big(\sum_{j=1}^{i-1} \varepsilon_j' h_i(X_i, X_j')\Big)^2\Big]^{\frac{1}{2}} \\
&\le \Big[\sum_{i=2}^{2^n} E_2\Big(\sum_{j=1}^{i-1} \varepsilon_j' h_i(X_i, X_j')\Big)^2\Big]^{\frac{1}{2}} \\
&= \Big(\sum_{i=2}^{2^n} (i-1) E_2 h_i^2(X_i, X')\Big)^{\frac{1}{2}} \le \Big(2^n \sum_{i=2}^{2^n} E_2 h_{2^n}^2(X_i, X')\Big)^{\frac{1}{2}} \\
&\le \Big[2^n \sum_{i=2}^{2^n} E_2\big(h^2 I_{h^2 \le t}(X_i, X')\big) + 2^n \log n \sum_{i=2}^{2^n} E_2\Big(\frac{h^2}{\log\log h} I_{h^2 > t}(X_i, X')\Big)\Big]^{\frac{1}{2}},
\end{aligned}
$$

and

$$\tilde\sigma^2 = \sup_{\|f\|_\infty\le 1}\sum_{i=2}^{2^n}\Big(E_2\big[f\cdot\sum_{j=1}^{i-1}\varepsilon'_j h_i(X_i,X'_j)\big]\Big)^2$$
$$\le \sum_{i=2}^{2^n}\Big(E_2\big|\sum_{j=1}^{i-1}\varepsilon'_j h_i(X_i,X'_j)\big|\Big)^2 \lesssim 2^n\sum_{i=2}^{2^n}E_2h_{2^n}^2,$$

where the last inequality follows from the previous estimate as the L_p norms of Rademacher sums are all comparable. Taking $t=(2c_3^{-1}\log n)^{1/2}$ in (4.13), as in (2.11) above, the last two estimates show that (4.12) (hence (4.11)) will be proved if

$$\sum_n \Pr\big\{2^n\sum_{i=2}^{2^n}E_2\big(h^2 I_{h^2\le t}(X_i,X')\big)>c2^{2n}\log n\big\}<\infty$$

and

$$\sum_n \Pr\big\{2^n\log n\sum_{i=2}^{2^n}E_2\big(\frac{h^2}{\log\log h}I_{h^2>t}(X_i,X')\big)>c2^{2n}\log n\big\}<\infty.$$

But the law of large numbers implies that these series converge for each $0<t<\infty$, again by Lemma 3.2 in Marcus and Zinn (1984). (4.11) is proved. The second sum in (4.10) reduces, in the end, to similar computations, but it is formally different. We have

$$\limsup_n \frac{1}{n\log\log n}\|\sum_{i=1}^{n-1}\varepsilon_i Z_{n,i}\|<\infty \text{ a.s.} \tag{4.14}$$

if

$$\lim_{k\to\infty}\Pr\big\{\sup_{r\ge k}\max_{n\in I(r)}\frac{1}{2^r\log r}\|\sum_{i=1}^{n-1}\varepsilon_i Z_{n,i}\|>c\big\}=0$$

for some $c<\infty$, where $I(r)=\{2^r+1,\dots,2^{r+1}\}$. Hence, by Lévy's inequality, (4.14) holds if

$$\sum_r \Pr\big\{\max_{m\in I(r)}\|\sum_{i=1}^{2^{r+1}-1}\varepsilon_i Z_{m,i}\|>c2^r\log r\big\}<\infty, \tag{4.15}$$

with $Z_{m,i}$ defined as zero if $m<i+1$. Now letting $\mathbf{Z}_i=(Z_{2^r+1,i},\dots,Z_{2^{r+1},i})\in\ell_\infty^{2^r}(L_1(E_2))$ the probabilities in (4.15) become $\Pr\big\{\|\sum_{i=1}^{2^{r+1}-1}\varepsilon_i\mathbf{Z}_i\|>c2^r\log r\big\}$ and we can apply to them inequality (4.13) as before, but now in the Banach space $\ell_\infty^{2^r}(L_1)$. We have, using Lévy's inequality for integrals (and interpreting sums from a larger to a smaller index as zero),

$$
\begin{aligned}
E_\varepsilon \Big\| \sum_{i=1}^{2^{r+1}-1} \varepsilon_i \mathbf{Z}_i \Big\| &= E_\varepsilon \max_{m\in I(r)} E_2 \Big| \sum_{i=1}^{2^{r+1}-1} \varepsilon_i \Big(\sum_{j=i+1}^{m} \varepsilon_j' h_j(X_i, X_j') \Big) \Big| \\
&\le E_\varepsilon E_2 \max_{m\in I(r)} \Big| \sum_{j=2}^{m} \varepsilon_j' \sum_{i=1}^{j-1} \varepsilon_i h_j(X_i, X_j') \Big| \\
&\le 2 E_\varepsilon E_2 \Big| \sum_{j=2}^{2^{r+1}} \varepsilon_j' \sum_{i=1}^{j-1} \varepsilon_i h_j(X_i, X_j') \Big| \\
&\simeq 2 E_\varepsilon E_2 \Big(\sum_{j=2}^{2^{r+1}} \Big(\sum_{i=1}^{j-1} \varepsilon_i h_j(X_i, X_j') \Big)^2 \Big)^{\frac{1}{2}} \\
&\lesssim \Big[2^{r+1} \sum_{i=1}^{2^{r+1}} E_2 h_{2^r}^2 (X_i, X') \Big]^{\frac{1}{2}}.
\end{aligned}
$$

Likewise, the new $\tilde\sigma$ admits the same bound:

$$
\begin{aligned}
\tilde\sigma^2 &\le \sum_{i=1}^{2^{r+1}-1} \max_{m\in I(r)} \Big[E_2 \Big| \sum_{j=i+1}^{m} \varepsilon_j' h_j(X_i, X_j') \Big| \Big]^2 \\
&\le \sum_{i=1}^{2^{r+1}-1} \Big[E_2 \Big| \sum_{j=i+1}^{2^{r+1}} \varepsilon_j' h_j(X_i, X_j') \Big| \Big]^2 \\
&\le 2^{r+1} \sum_{i=1}^{2^{r+1}} E_2 h_{2^r}^2 (X_i, X').
\end{aligned}
$$

Therefore, the series (4.15) converges and (4.14) is proved. (4.12) and (4.14) give $(II) < \infty$ a.s. by (4.11).

Proof of $(I) < \infty$ *a.s.* By (un–)decoupling it suffices to prove

$$
\limsup_n \frac{1}{\log\log n} \Big| \sum_{I_n^2} \varepsilon_i \varepsilon_j h_{i\vee j}(X_i, X_j) \Big| < \infty \text{ a.s.} \tag{4.16}
$$

and, as observed above (e.g., below (2.6)), the a.s. finiteness of this limit (which equals a constant a.s.) is implied by the condition

$$
\sum_n \Pr\Big\{ \Big| \sum_{I_{2^n}^2} \varepsilon_i \varepsilon_j h_{i\vee j}(X_i, X_j) \Big| > c 2^n \log n \Big\} < \infty. \tag{4.17}
$$

Proceeding as in (2.10) and (2.11), it follows that this series converges if

$$
\sum_n \Pr\Big\{ \sum_{I_{2^n}^2} h_{i\vee j}^2 (X_i, X_j) > c 2^{2n} \log n \Big\} < \infty \tag{4.18}
$$

and

$$\sum_n \Pr\{\sup_{|u|\le 1}\big|\sum_{I^2_{2^n}} u_i u_j h_{i\vee j}(X_i,X_j)\big| > c2^n\} < \infty \tag{4.19}$$

for some c finite. Observing that

$$\sup_{|u|\le 1}\big|\sum_{I^2_{2^n}} u_i u_j h_{i\vee j}(X_i,X_j)\big| \ge \sup_{|u|\le 1}\big|\sum_{i,j\in I(n-1), i\ne j} u_i u_j h_{i\vee j}(X_i,X_j)\big|$$

(as the latter is a sup over unit vectors $(u_1,\ldots,u_{2^n})$'s with the first half of their coordinates zero), (4.19) follows from hypothesis (4.2) by Glivenko–Cantelli. It is also clear that, by positivity and Glivenko–Cantelli, (4.18) is implied by

$$\limsup_n \frac{1}{n^2\log\log n}\sum_{I^2_n} h^2_{i\vee j}(X_i,X_j) < \infty \text{ a.s.} \tag{4.20}$$

So, we have reduced the proof of $(I)<\infty$ a.s., to showing that (4.20) is satisfied. Actually we show, with a method of proof borrowed from Giné and Zinn (1992b), that this sequence converges a.s. to zero. By symmetry, the summation set I^2_n can be replaced by $\{(i,j): i < j \le n\}$. Next, we see that $h^2_{i\vee j}$ in (4.20) can be degenerated. The centerings tend to zero a.s.:

$$\begin{aligned}\frac{2}{n^2\log n}\sum_{i<j\le n} Eh^2_j(X_i,X_j) &\le \frac{2}{\log n}Eh^2_n \\ &\le \frac{2t}{\log\log n} + 2E\big(\frac{h^2}{\log\log h} I_{h^2>t}\big) \to 0\end{aligned}$$

as $n\to\infty$ and then $t\to 0$. And so do the first order terms:

$$\begin{aligned}\frac{1}{n^2\log n}\Big(\sum_{i<j\le n} E_2h^2_j(X_i,X') + \sum_{j<i\le n} E_2h^2_i(X_i,X')\Big) &\le \frac{2(n-1)}{n^2\log\log n}\sum_{i=1}^n E_2h^2_n(X_i,X') \\ &\le \frac{2}{n}\Big(\sum_{i=1}^n \frac{t}{\log\log n} + \sum_{i=1}^n E_2\big(\frac{h^2}{\log\log h} I_{h^2>t}(X_i,X')\big)\Big) \to 0 \text{ a.s.}\end{aligned}$$

as $n\to\infty$ and then $t\to 0$ by the integrability condition (4.1) and the law of large numbers. Hence, (4.20), with limit zero, is implied by

$$\lim_n \frac{1}{n^2\log\log n}\sum_{i<j\le n}\pi_{2,2}h^2_j(X_i,X_j) = 0 \text{ a.s.} \tag{4.21}$$

where, as in Section 2, $\pi_{2,2}f(x,y) = f(x,y) - Ef(X,y) - Ef(y,X) + Ef$ is a degenerate kernel. By Kronecker's lemma, it is enough to see

$$\sum_{j=1}^\infty \frac{1}{j^2\log\log j}\sum_{i=1}^j \pi_{2,2}\big(h^2 I_{h^2\le j^2\log\log j}\big)(X_i,X_j) < \infty \text{ a.s.}$$

The partial sums of this series are a martingale relative to $\sigma(X_1, \cdots, X_n)$ and therefore it suffices to show L_1–boundedness (or L_2–boundedness). Since $E\big[\pi_{2,2}(h^2 I_{h^2\le j^2\log\log j})(X_i, X_j)\cdot\pi_{2,2}(h^2 I_{h^2\le \ell^2\log\log \ell})(X_k, X_\ell)\big] = 0$ unless $i = k$ and $j = \ell$, the L_2 boundedness of the above series follows from:

$$\begin{aligned}\sum_j \frac{1}{j^3(\log\log j)^2} Eh^4 I_{h^2\le j^2\log\log j} &= \sum_j \frac{1}{j^3(\log\log j)^2}\sum_{i\le j} Eh^4 I_{(i-1)^2\log\log(i-1)<h^2\le i^2\log\log i} \\ &\approx \sum_{i=1}^{\infty} \frac{1}{i^2(\log\log i)^2} Eh^4 I_{(i-1)^2\log\log(i-1)<h^2\le i^2\log\log i} \\ &\lesssim \sum_{i=1}^{\infty} E\frac{h^2}{\log\log|h|} I_{(i-1)^2\log\log(i-1)<h^2\le i^2\log\log i} \\ &= E\frac{h^2}{\log\log|h|} < \infty.\end{aligned}$$

This proves (4.21) and completes the proof of $(I) < \infty$ a.s.

We have thus proved that the limsup in (4.8) is a.s. finite; hence, by randomization, that (4.6) holds; by truncation, that (4.4) holds; and, by decoupling, that (4.3) is satisfied. □

Previous arguments show that condition (4.2) is equivalent to the existence of $c < \infty$ such that

$$\sum_n \Pr\Big\{\sup_{|u|\le 1}\Big|\sum_{I^2_{2^n}} u_i u_j h_{i\vee j}(X_i, X_j)\Big| > c2^n\Big\} < \infty, \tag{4.19}$$

and *it is this condition that one must verify in practice* (e.g. to check that Theorem 4.1 applies to the modification of the example in Section 2 mentioned immediately below the statement of the theorem). The proof above, or simple modifications of it, show:

4.2. Corollary. *In Theorem 4.1, condition (4.2) can be replaced by any one of the following:*

$$\limsup_{n\to\infty}\frac{1}{n}\sup_{|u|,|u'|\le 1}\sum_{I_n^2} u_i u'_j h_{i\vee j}(X_i, X'_j) < \infty \quad a.s., \tag{4.2$'$}$$

$$\limsup_{n\to\infty}\frac{1}{n}\sup_{|u|\le 1}\Big|\sum_{I_n^2} u_i u_j h(X_i, X_j)\Big| < \infty \text{ a.s.}, \tag{4.2$''$}$$

or

$$\limsup_{n\to\infty}\frac{1}{n}\sup_{|u|,|u'|\le 1}\sum_{I_n^2} u_i u'_j h(X_i, X'_j) < \infty \text{ a.s.} \tag{4.2$'''$}$$

Several problems remain, particularly, 1) we do not know how to handle (4.2) (or the other conditions in Corollary 4.2) in general, and 2) we do not know either whether (4.1) is the weakest integrability condition that, together with a condition of the type of (4.2), implies the LIL.

The chaos inequality for $m > 2$, which we have not seen formally stated, should contain several norms of mixed type (some, but not necessarily all, of the $\varepsilon^{(j)}$ replaced by $u^{(j)}$), just as the corresponding Gaussian chaos tail bound does (see e.g. in Arcones and Giné, 1993a). So, in principle one can prove results like Theorem 4.1 for $m > 2$, but these will involve somewhat more complicated hypotheses.

References

Arcones, M. and giné, E. (1993a). On decoupling, series expansions, and tail behavior of chaos processes. *J. Theoret. Probab.* **6** 101–122.

Arcones, M. and Giné, E. (1993b). On the law of the iterated logarithm for canonical U–statistics and processes. Preprint.

Cuzick, J.; Giné, E.; Zinn, J. (1993). Laws of large numbers for quadratic forms, maxima of products and truncated sums of i.i.d. random variables. *Ann. Probab.*, to appear.

Dehling, H. (1989). Complete convergence of triangular arrays and the law of the iterated logarithm for degenerate U–statistics. *Stat. Probab. Letters* **7** 319–321.

de la Peña, V. (1992). Decoupling and Khintchine inequalities for U–statistics. *Ann. Probab.* **20** 1877–1892.

de la Peña, V. and Montgomery–Smith, S. (1993). Decoupling inequalities for the tail probabilities of multivariate U–statistics. *Ann. Probab.*, to appear.

Giné, E. and Zinn, J. (1992a). On Hoffmann-Jørgensen's inequality for U–processes. *Probability in Banach Spaces 8* 80–91. Birkhäuser, Boston.

Giné, E. and Zinn, J. (1992b). Marcinkiewicz type laws of large numbers and convergence of moments for U–statistics. *Probability in Banach Spaces 8* 273–291. Birkhäuser, Boston.

Giné, E. and Zinn, J. (1994). A remark on convergence in distribution of U–statistics. *Ann. Probab.* **22** 117–125.

Ledoux, M. and Talagrand, M. (1991). *Probability in Banach Spaces.* Springer, Berlin.

Marcus, M. B. and Zinn, J. (1984). The bounded law of the iterated logarithm for the weighted empirical distribution process in the non i.i.d. case. *Ann. Probab.* **12** 335–360.

Department of Mathematics and
Department of Statistics
University of Connecticut
Storrs, CT 06269