

ON HOFFMANN-JØRGENSEN'S INEQUALITY FOR U-PROCESSES

Evarist Giné* and Joel Zinn*

Summary

The object of this note is to prove an analogue for U-processes of Hoffmann-Jørgensen's (1974) tail inequality for sums of independent symmetric random vectors. The result obtained is best possible in a certain sense but is less useful than the original inequality.

1. Hoffmann-Jørgensen type inequalities for certain non-convex functionals.

Let (V_n, \mathcal{V}_n) , $n \in \mathbf{N}$, be measurable linear spaces, let $E_n = V_n \times \mathbf{R}$, $\mathcal{E}_n = \mathcal{V}_n \otimes \mathcal{B}$ and let $(E^\infty, \mathcal{E}^\infty) = (\prod_{n=1}^\infty E_n, \otimes_{n=1}^\infty \mathcal{E}_n)$. We will denote points $((x_1, \varepsilon_1), \dots, (x_n, \varepsilon_n), \dots) \in E^\infty$ by (x, ε) . We say that $x, x' \in \prod_{n=1}^\infty V_n$ are disjoint if for each $i \in \mathbf{N}$ either $x_i = 0$ or $x'_i = 0$. Let $q_t : E^\infty \rightarrow \mathbf{R}_+$, $t \geq 0$, be a collection of maps satisfying:

- (1) q_t is *measurable* for all $t \geq 0$;
- (2) $0 \leq q_t(x, \varepsilon) \leq 1$ for all $t \geq 0$ and $(x, \varepsilon) \in E^\infty$;
- (3) $q_t(x, \varepsilon)$ is *even* for all $t \geq 0$ and $x \in V_n$ that is, $q_t(x, \varepsilon) = q_t(x, -\varepsilon)$, $\varepsilon \in \mathbf{R}^\infty$;
- (4) q is *non-decreasing* in the sense that $q_t(x, \varepsilon)$ is non-decreasing in t for each (x, ε) , and $q_t(x, \alpha\varepsilon)$ is non decreasing in $\alpha \geq 0$ for all $t \geq 0$ and $(x, \varepsilon) \in E^\infty$;
- (5) for all $t \geq 0$, $(x, \varepsilon), (x', \varepsilon') \in E^\infty$ with x and x' disjoint, and $\alpha, \beta \geq 0$ with $\alpha + \beta = 1$,
 $q_t((x, \alpha\varepsilon) + (x', \beta\varepsilon')) \leq q_t(x, \varepsilon) + q_t(x', \varepsilon')$ and $q_t((x, \varepsilon) + (x', 0)) = q_t(x, \varepsilon)$.

For lack of a better word, if q satisfies properties (1)-(5) we will say that q is a

* Partially supported by NSF Grants No. DMS-9000132 and DMS-8601250. Part of this research was carried out at the Mathematical Sciences Research Institute, Berkeley, under NSF Grant No. DMS-8505550.

*normlike** class of functions.

Here is an example of a class of functions with the above properties: Let Z be a S -valued random variable and let \mathcal{H}_i , $i \in \mathbf{N}$, be families of measurable functions $S^2 \rightarrow \mathbf{R}$ such that the variables

$$q_t(x, \varepsilon) = \Pr\left\{ \sup_{h_i \in \mathcal{H}_i} \left| \sum_{i=1}^{\infty} \varepsilon_i h_i(x_i, Z) \right| > t \right\}$$

exist and are measurable functions of $x_i \in S$ and $\varepsilon_i \in \mathbf{R}$, $i \in \mathbf{N}$. Then q is a family of *normlike** functions.

*The content of the next two propositions is that *normlike** classes of functions verify Lévy's inequality and Hoffmann-Jørgensen's inequalities only with different constants.* First we give some notation: For $(x, \varepsilon) \in E^\infty$ with coordinates (x_i, ε_i) , and for $k \in \mathbf{N}$, we let

$$\pi_k](x, \varepsilon) = ((x_1, \varepsilon_1), \dots, (x_k, \varepsilon_k), 0, \dots, 0, \dots).$$

We also let $\pi_k = \pi_k] - \pi_{k-1}]$ and $\pi_{k,n} = \pi_n] - \pi_{k-1}]$ ($k < n$).

Proposition 1. *Let X_k be independent V_k -valued random variables, $k \leq n < \infty$, and let $\{\xi_k\}$ be a sequence of independent symmetric real random variables independent of $\{X_k\}$. Let q be a *normlike class of functions.* Then, letting $(X, \xi) = ((X_1, \xi_1), \dots, (X_n, \xi_n), 0, \dots, 0, \dots)$, we have that for all t and $r > 0$,*

$$\Pr\left\{ \max_{k \leq n} q_t \circ \pi_k](X, \xi) > r \right\} \leq 2 \Pr\left\{ q_t(X, \xi) > \frac{r}{2} \right\}$$

and

$$\Pr\left\{ \max_{k \leq n} q_t \circ \pi_k(X, \xi) > r \right\} \leq 2 \Pr\left\{ q_t(X, \xi) > \frac{r}{2} \right\}.$$

Proof. Let $\tau = \inf\{k : q_t \circ \pi_k](X, \xi) > r\}$ and $A_k = \{\tau = k\}$. Then, a) $\mathcal{L}(q_t(X, \xi), A_k) = \mathcal{L}(q_t(\pi_k](X, \xi) + \pi_{k+1,n}(X, -\xi), A_k)$ since these vectors are each the same measurable function respectively of (X, ξ) and $\pi_k](X, \xi)$

$+\pi_{k+1,n}(X, -\xi)$, and these two E^∞ -valued random variables have the same distribution (by symmetry and independence); and b) on A_k , $q_t \circ \pi_k](X, \xi) > r$ by definition and therefore, on A_k , either $q_t(X, \xi) > \frac{r}{2}$ or $q_t(\pi_k](X, \xi) + \pi_{k+1,n}(X, -\xi)) > \frac{r}{2}$ since

$$\begin{aligned} r &< q_t \circ \pi_k](X, \xi) \\ &= q_t(((X, \xi) + \pi_k](X, \xi) + \pi_{k+1,n}(X, -\xi))/2) \\ &\leq q_t(X, \xi) + q_t(\pi_k](X, \xi) + \pi_{k+1,n}(X, -\xi)). \end{aligned}$$

Then, by a) and b),

$$\Pr(A_k) \leq 2 \Pr[A_k \cap \{q_t(X, \xi) > \frac{r}{2}\}]$$

and therefore,

$$\begin{aligned} \Pr\{\max_{k \leq n} q_t \circ \pi_k](X, \xi) > r\} &= \sum_{k=1}^n \Pr(A_k) \\ &\leq 2 \sum_{k=1}^n \Pr[A_k \cap \{q_t(X, \xi) > \frac{r}{2}\}] \\ &\leq 2 \Pr\{q_t(X, \xi) > \frac{r}{2}\} \end{aligned}$$

by disjointness of the A_k 's. This proves the first inequality; the second one is proved similarly, using $q_t \circ \pi_k$ to define τ . ■

Obviously, Proposition 1 applies also to $q_t \circ \pi_{k,n}(X, \xi)$. Moreover, n can be taken to be ∞ .

Theorem 2. (Hoffmann-Jorgensen's inequalities.) *Under the hypotheses of Proposition 1, the following inequalities hold:*

a) for all $r, t > 0$ and $n \in \mathbf{N}$,

$$\Pr\{q_t(X, \xi/3) > 3r\} \leq 4[\Pr\{q_t(X, \xi) > \frac{r}{2}\}]^2 + \Pr\{\max_{i \leq n} q_t \circ \pi_i(X, \xi) > r\}.$$

b) For all $\lambda, t > 0$ and $n \in \mathbf{N}$,

$$E[q_t(X, \xi/3)]^\lambda \leq 2 \cdot 4^{\frac{\lambda}{1+\lambda}} \cdot 6^\lambda [E q_t(X, \xi)]^{\frac{2\lambda}{1+\lambda}} + 3^\lambda E[\max_{i \leq n} q_t \circ \pi_i(X, \xi)]^\lambda.$$

Proof. Using the facts that

$$*i)^* \quad q_t(X, \xi/3) \leq q_t \circ \pi_{k-1}(X, \xi) + q_t \circ \pi_k(X, \xi) + q_t \circ \pi_{k+1,n}(X, \xi),$$

$$*ii)^* \quad q_t \circ \pi_{k-1}(X, \xi) \leq r \text{ on } \{\tau = k\},$$

$$*iii)^* \quad \{\tau = k\} \text{ and } \pi_{k+1,n}(X, \xi) \text{ are independent,}$$

$$*iv)^* \quad \{q_t(X, \xi/3) > 3r\} \subset \{q_t(X, \xi) > r\}, \text{ and}$$

$$*v)^* \quad \text{Proposition 1 twice,}$$

we obtain

$$\begin{aligned} \Pr\{q_t(X, \xi/3) > 3r\} &\leq \sum_{k \leq n} \Pr\{q_t(X, \xi/3) > 3r, \tau = k, \max_{i \leq n} q_t \circ \pi_i(X, \xi) \leq r\} \\ &\quad + \Pr\{\max_{i \leq n} q_t \circ \pi_i(X, \xi) > r\} \end{aligned}$$

and

$$\begin{aligned} &\sum_{k \leq n} \Pr\{q_t(X, \xi/3) > 3r, \tau = k, \max_{i \leq n} q_t \circ \pi_i(X, \xi) \leq r\} \\ &\leq \sum_{k \leq n} \Pr\{\tau = k, q_t \circ \pi_{k+1,n}(X, \xi) > r\} \\ &= \sum_{k \leq n} \Pr\{\tau = k\} \Pr\{q_t \circ \pi_{k+1,n}(X, \xi) > r\} \\ &\leq 2 \Pr[\cup_{k \leq n} \{\tau = k\}] \Pr\{q_t(X, \xi) > \frac{r}{2}\} \\ &\leq 4[\Pr\{q_t(X, \xi) > \frac{r}{2}\}]^2, \end{aligned}$$

proving a).

To prove b) we observe that, by part a), for all $a \geq 0$,

$$\begin{aligned} E[q_t(X, \xi/3)]^\lambda &= \lambda \int_0^1 r^{\lambda-1} \Pr\{q_t(X, \xi/3) > r\} dr \\ &\leq a^\lambda + 4\lambda \int_a^1 r^{\lambda-1} \Pr\{q_t(X, \xi) > \frac{a}{6}\} \Pr\{q_t(X, \xi) > \frac{r}{6}\} dr \\ &\quad + \lambda \int_0^1 r^{\lambda-1} \Pr\{\max_{i \leq n} q_t \circ \pi_i(X, \xi) > \frac{r}{3}\} dr \\ &\leq a^\lambda + 4 \cdot 6^\lambda \Pr\{q_t(X, \xi) > \frac{a}{6}\} E[q_t(X, \xi)]^\lambda \\ &\quad + 3^\lambda E[\max_{i \leq n} q_t \circ \pi_i(X, \xi)]^\lambda. \end{aligned}$$

Defining a by the equation $a^\lambda = \frac{4 \cdot 6^{\lambda+1} E[q_t(X, \xi)]^\lambda E q_t(X, \xi)}{a}$ the above gives

(by Markov's inequality)

$$E[q_t(X, \xi/3)]^\lambda \leq 2[4 \cdot 6^{\lambda+1} E[q_t(X, \xi)]^\lambda E q_t(X, \xi)]^{\frac{\lambda}{\lambda+1}} + 3^\lambda E[\max_{i \leq n} q_t \circ \pi_i(X, \xi)]^\lambda.$$

Now part b) follows from the observation that, since $0 \leq q_t(X, \xi) \leq 1$ and $\lambda > 1$, $E[q_t(X, \xi)]^\lambda \leq E q_t(X, \xi)$. ■

The above results hold for outer probabilities and expectations without any measurability on $\{q_t\}$ since they are based on inequalities satisfied by outer probabilities.

2. Inequalities for U-processes.* We will restrict the exposition, for the sake of simplicity, to independent, identically distributed random variables and symmetric functions. If needed, inequalities for U-processes based on non-symmetric functions and independent variables with arbitrary distributions can be obtained by adapting the proofs below.* Let (S, \mathcal{S}) be a measurable space, let $m \in \mathbf{N}$ and let $\{X_i, X_i^{(1)}, \dots, X_i^{(m)} : i \in \mathbf{N}\}$ be independent identically distributed S -valued random variables. A function $h : S^m \rightarrow B$ is *symmetric* if for any permutation σ of $1, \dots, m$, $h(x_1, \dots, x_m) = h(x_{\sigma(1)}, \dots, x_{\sigma(m)})$, and a collection of classes of functions on S^m indexed by $\mathbf{i} := (i_1, \dots, i_m) \in \{1, \dots, n\}^m$, $\mathcal{H}_{\mathbf{i}}$, is *symmetric* if 1) the functions in each class $\mathcal{H}_{\mathbf{i}}$ are symmetric and 2) for each permutation σ of $1, \dots, m$, $\mathcal{H}_{\mathbf{i}} \ni h_{\mathbf{i}} = h_{\sigma(\mathbf{i})} \in \mathcal{H}_{\sigma(\mathbf{i})}$ where $\sigma(i_1, \dots, i_m) := (i_{\sigma(1)}, \dots, i_{\sigma(m)})$. We assume that the functions $h_{\mathbf{i}} \in \mathcal{H}_{\mathbf{i}}$ take values in a Banach space B . We define, for $h_{\mathbf{i}} \in \mathcal{H}_{\mathbf{i}}$,

$$(1) \quad U_n^{d,s} = U_n^{d,s}(\{h_{\mathbf{i}}\}) := \sum_{\mathbf{i}} h_{\mathbf{i}}(X_{i_1}^{(1)}, \dots, X_{i_m}^{(m)}) \varepsilon_{i_1}^{(1)} \dots \varepsilon_{i_m}^{(m)}$$

where $\{\varepsilon_i, \varepsilon_i^{(1)}, \dots, \varepsilon_i^{(m)} : i \in \mathbf{N}\}$ are independent Rademacher random variables independent of the X_i 's and the $X_i^{(j)}$'s. $U_n^{d,s}$ is a general decoupled, unnormalized U-statistic (or V-statistic: in the decoupled case, diagonals are not important in general). The norm symbol, $\|\cdot\|$, will mean the following:

$$\|U_n^{d,s}\| := \sup_{h_{\mathbf{i}} \in \mathcal{H}_{\mathbf{i}}} \left\| \sum_{\mathbf{i}} h_{\mathbf{i}}(X_{i_1}^{(1)}, \dots, X_{i_m}^{(m)}) \varepsilon_{i_1}^{(1)} \dots \varepsilon_{i_m}^{(m)} \right\|_B.$$

We say that the collection of classes $\{\mathcal{H}_{\mathbf{i}}\}$ is *measurable* if the variables $\|U_n^{d,s}\|$, as well as the variables

$$M_n^{d,s} := \max_{i_m \leq n} \left\| \sum_{i_1, \dots, i_{m-1} \leq n} h_{\mathbf{i}}(X_{i_1}^{(1)}, \dots, X_{i_m}^{(m)}) \varepsilon_{i_1}^{(1)} \dots \varepsilon_{i_{m-1}}^{(m-1)} \right\|$$

are measurable functions of $\{(X_i^{(j)}, \varepsilon_i^{(j)})\}$. Hoffmann-Jørgensen's inequality (Hoffmann-Jørgensen, 1974) together with Theorem 1 above, give:

Theorem 3. *If $\{\mathcal{H}_i\}$ is symmetric and measurable, then there exist finite positive universal constants $c_1(m)$ and $c_2(m)$ such that, for all $t > 0$,*

$$(2) \quad Pr\{\|U_n^{d,s}\| > 3^m t\} \leq c_1 [Pr\{\|U_n^{d,s}\| > t\}]^{\frac{2^m}{2^m-1}} + c_2 Pr\{M_n^{d,s} > t\}.$$

The constants can be taken to be $c_1(m) = 2^{m-1} \cdot 4^{1+\frac{2}{3}+\dots+\frac{2^{m-1}}{2^m-1}} \cdot 6^{2+\frac{2}{3}+\dots+\frac{2^{m-1}}{2^m-1}}$, and $c_2(m) = 1 + 3^2 \cdot 4 + \dots + 3^{\frac{2^m-1}{2^m-1}-1} c_1(m-1)$.

Proof. For each $k \leq m$ we denote $P^{(k)}$ integration with respect to only the variables $X^{(k)}, \varepsilon^{(k)}$. Hoffmann-Jørgensen's inequality applied conditionally (for $P^{(m)}$) gives:

$$P^{(m)}\{\|U_n^{d,s}\| > 3^m t\} \leq 4[P^{(m)}\{\|U_n^{d,s}\| > 3^{m-1} t\}]^2 + P^{(m)}\{M_n^{d,s} > t\}.$$

For $X_i^{(1)}, \dots, X_i^{(m-2)}, \varepsilon_i^{(1)}, \dots, \varepsilon_i^{(m-2)}$ fixed we consider the following functions of the $X_i^{(m-1)}, \varepsilon_i^{(m-1)}$ variables:

$$q_t(X^{(m-1)}, \varepsilon^{(m-1)}) = P^{(m)}\{\|U_n^{d,s}\| > t\}, \quad t \geq 0.$$

$\{q_t : t \geq 0\}$ is normlike. Applying Theorem 2b) to q_t with $\lambda = \lambda_1 = 2$ we obtain

$$\begin{aligned} P^{(m-1)} [P^{(m)}\{\|U_n^{d,s}\| > 3^{m-1} t\}]^2 &\leq 2 \cdot 4^{\frac{2}{3}} \cdot 6^2 [P^{(m-1)} P^{(m)}\{\|U_n^{d,s}\| > 3^{m-2} t\}]^{\frac{4}{3}} \\ &+ 3^2 P^{(m-1)} \left[\max_{i_{m-1} \leq n} P^{(m)} \left\| \sum_{i_1, \dots, i_{m-2}, i_m \leq n} h_{\mathbf{i}}(X_{i_1}^{(1)}, \dots, X_{i_m}^{(m)}) \varepsilon_{i_1}^{(1)} \dots \varepsilon_{i_{m-2}}^{(m-2)} \varepsilon_{i_m}^{(m)} \right\| \right. \\ &\quad \left. > 3^{m-2} t \right]^2. \end{aligned}$$

Plugging this inequality into the previous one (after estimating its max term using the symmetry of the functions h) yields:

$$\begin{aligned} P^{(m-1)} P^{(m)}\{\|U_n^{d,s}\| > 3^m t\} &\leq 2 \cdot 4^{1+\frac{2}{3}} \cdot 6^2 [P^{(m-1)} P^{(m)}\{\|U_n^{d,s}\| > 3^{m-2} t\}]^{\frac{4}{3}} \\ &+ (1 + 4 \cdot 3^2) P^{(m-1)} P^{(m)}\{M_n^{d,s} > t\}. \end{aligned}$$

Next, we apply Theorem 2 to $P^{(m-1)}P^{(m)}\{\|U_n^{d,s}\| > 3^{m-2}t\}$, conditionally on $X^{(1)}, \dots, X^{(m-3)}, \varepsilon^{(1)}, \dots, \varepsilon^{(m-3)}$, and with $\lambda = \lambda_2 = \frac{2\lambda_1}{2\lambda_1-1} = \frac{4}{3}$, and plug the result into the previous inequality, as above. Repeating this procedure $m-1$ times gives the theorem. ■

Theorem 3 gives, with computations similar to those of Theorem 2b), which we omit,

$$(3) \quad E\|U_n^{d,s}\|^p \leq 2 \cdot 3^{mp} c_2(m) \cdot E(M_n^{d,s})^p + 2 \cdot 3^{mp} c_1(m) t_0^p$$

for all $p > 0$ and for any t_0 such that $Pr\{\|U_n^{d,s}\| > t_0\} \leq (2 \cdot 3^{mp} c_1(m))^{1-2^m}$. By Paley-Zygmund's inequality, inequality (2) is equivalent, up to constants, to

$$(4) \quad E\|U_n^{d,s}\|^p \leq K \left[E(M_n^{d,s})^p + (E\|U_n^{d,s}\|^r)^{\frac{p}{r}} \right]$$

for all $0 \leq r < p < \infty$ (where $K = K(r, p)$ is a finite constant; it is easy to obtain an upper bound for it, but we will not do so). ((4) follows from (3) by Chebyshev's inequality; to see that (4) implies (3) for possibly different constants, we just repeat the easy argument for the Paley-Zygmund inequality applied to this situation: Just note

$$E\|U_n^{d,s}\|^r \leq t^r + E\|U_n^{d,s}\|^r I(\|U_n^{d,s}\| > t) \leq t^r + (E\|U_n^{d,s}\|^p)^{\frac{r}{p}} \Pr(\|U_n^{d,s}\| > t)^{1-\frac{r}{p}};$$

taking $t = t_0$ such that $Pr(\|U_n^{d,s}\| > t_0)^{\frac{p}{r}-1} \leq (2^{\frac{p}{r}} K)^{-1}$ and replacing the resulting bound for $E\|U_n^{d,s}\|^r$ in (4) gives an inequality of the form (3)).

We now let \mathcal{H} be a class of symmetric functions and $\mathcal{H}_i = \{hI[\mathbf{i} \in I_m^n] : h \in \mathcal{H}\}$ where $I_m^n = \{(i_1, \dots, i_m) \in \{1, \dots, n\}^m : i_r \neq i_s \text{ for } r \neq s\}$. Then, inequalities (2) to (4) hold for

$$U_n^{d,s}(h) = \sum_{\mathbf{i} \in I_m^n} h(X_{i_1}^{(1)}, \dots, X_{i_m}^{(m)}) \varepsilon_{i_1}^{(1)} \cdots \varepsilon_{i_m}^{(m)}, \quad h \in \mathcal{H},$$

the usual decoupled, symmetrized U-process based on P and \mathcal{H} . We just write (4) down for $1 = r < p$:

(4')

$$E\|U_n^{d,s}(h)\|^p \leq K \left[E \max_{i_m \leq n} \left\| \sum_{i_1, \dots, i_{m-1}: (i_1, \dots, i_m) \in I_m^n} h(X_{i_1}^{(1)}, \dots, X_{i_m}^{(m)}) \varepsilon_{i_1}^{(1)} \dots \varepsilon_{i_{m-1}}^{(m-1)} \right\|^p + (E\|U_n^{d,s}(h)\|)^p \right].$$

Next, let

$$U_n(h) := \sum_{\mathbf{i} \in I_m^n} h(X_{i_1}, \dots, X_{i_m}), \quad h \in \mathcal{H},$$

that is, $\{U_n(h) : h \in \mathcal{H}\}$ is the regular B -valued U -process based on P and "indexed" by \mathcal{H} . If \mathcal{H} consists of P -canonical functions (i.e., functions h such that $Ph(x_1, \dots, x_{m-1}, \cdot) = 0$ P^{m-1} -a.s.), then we can drop the ε 's from the three terms in (3'), undecouple the first and third terms there by the result in de la Peña (1990) and, by the same result, replace the max term by

$$E \max_{i_m \leq n} \left\| \sum_{i_1, \dots, i_{m-1}: (i_1, \dots, i_m) \in I_m^n} h(X_{i_1}, \dots, X_{i_{m-1}}, X_{i_m}^{(m)}) \right\|^p.$$

Undecoupling of the m -th argument follows from Lemma 1 in Hitczenko (1988) upon observing that the sequences

$$\left\{ \left\| \sum_{i_1, \dots, i_{m-1}: (i_1, \dots, i_{m-1}, j) \in I_m^n} h(X_{i_1}, \dots, X_{i_{m-1}}, X_j^{(m)}) \right\|^p \right\}_{j=1}^n$$

and

$$\left\{ \left\| \sum_{i_1, \dots, i_{m-1}: (i_1, \dots, i_{m-1}, j) \in I_m^n} h(X_{i_1}, \dots, X_{i_{m-1}}, X_j) \right\|^p \right\}_{j=1}^n$$

are tangent with respect to the σ -algebras $\sigma(X_1^{(m)}, \dots, X_j^{(m)}, X_1, \dots, X_j)$, $j = 1, \dots, n$. Decoupling requires some extra measurability (e.g. de la Peña, loc. cit.), which we assume for the result that follows. The preceding observations give:

Corollary 4. *If the measurable class \mathcal{H} of symmetric B -valued functions on S^m consists of P -canonical functions then, for $p > 1$,*

(5)

$$E\|U_n(h)\|^p \leq K \left[E \max_{i_m \leq n} \left\| \sum_{i_1, \dots, i_{m-1}: (i_1, \dots, i_m) \in I_m^n} h(X_{i_1}, \dots, X_{i_m}) \right\|^p + (E\|U_n(h)\|)^p \right]$$

for some $K < \infty$. The last term can be replaced by a quantile of $\|U_n\|$ (as in (3)).

M. Arcones showed us a shorter proof of Corollary 4 (which however does not yield Theorem 3). Here is his proof for $m = 2$ (the proof for general m is analogous): Applying Hoffmann-Jørgensen's inequality with expected values, conditionally on the X' 's in the first step and then for the L_1 -norm in the second step, and Paley-Zygmund inequality in the last step, as above, we have (with X , X' instead of $X^{(1)}$ and $X^{(2)}$),

$$\begin{aligned}
E \left\| \sum_{(i,j) \in I_2^n} h(X_i, X'_j) \right\|^p &= P' P \left\| \sum_{i=1}^n \left(\sum_{j \neq i, j=1}^n h(X_i, X'_j) \right) \right\|^p \\
&\lesssim P' \left[\left(P \left\| \sum_{i=1}^n \left(\sum_{j \neq i, j=1}^n h(X_i, X'_j) \right) \right\| \right)^p + P \max_{i \leq n} \left\| \sum_{j \neq i, j=1}^n h(X_i, X'_j) \right\|^p \right] \\
&\lesssim \left[P' \left(P \left\| \sum_{i=1}^n \left(\sum_{j \neq i, j=1}^n h(X_i, X'_j) \right) \right\| \right)^p + P' \max_{i \leq n} \left(P \left\| \sum_{j \neq i, j=1}^n h(X_i, X'_j) \right\| \right)^p \right. \\
&\quad \left. + E \max_{i \leq n} \left\| \sum_{j \neq i, j=1}^n h(X_i, X'_j) \right\|^p \right] \\
&\lesssim \left(E \left\| \sum_{(i,j) \in I_2^n} h(X_i, X'_j) \right\| \right)^p + E \max_{i \leq n} \left\| \sum_{j \neq i, j=1}^n h(X_i, X'_j) \right\|^p \\
&\lesssim t_0^p + E \max_{i \leq n} \left\| \sum_{j \neq i, j=1}^n h(X_i, X'_j) \right\|^p.
\end{aligned}$$

The following reformulation of Corollary 4 may be useful in uniform integrability proofs.

Corollary 5. 1) Under the hypotheses and with the same notation of Theorem 3, for $p > 0$, $u > 0$,

$$\begin{aligned}
&\left(1 - c_1 3^{mp} \Pr \left\{ \|U_n^{d,s}\| > \frac{u}{3^m} \right\}^{\frac{1}{2^m-1}} \right) E \|U_n^{d,s}\|^p I(\|U_n^{d,s}\| > u) \\
&\leq c_1 3^{mp} \Pr \left\{ \|U_n^{d,s}\| > \frac{u}{3^m} \right\}^{\frac{1}{2^m-1}} \left(\frac{u}{3^m} \right)^p \Pr \left\{ \|U_n^{d,s}\| > \frac{u}{3^m} \right\} \\
&\quad + c_2 3^{mp} E M_n^{d,s} I(M_n^{d,s} > \frac{u}{3^m}).
\end{aligned}$$

2) If a collection of processes $\{U_n^{d,s}\}_{n=1}^\infty$, each associated to classes of functions $\{\mathcal{H}_{\mathbf{i},n} : \mathbf{i} \in \{1, \dots, n\}^m\}$, $n \in \mathbf{N}$, possibly different for each n and satisfying the hypotheses of Theorem 3, is stochastically bounded and the sequence $\{(M_n^{d,s})^p\}_{n=1}^\infty$ is uniformly integrable, then the sequence $\{\|U_n^{d,s}\|^p\}_{n=1}^\infty$ is also uniformly integrable.

3) 2) holds for $\{U_n\}$ with $p \geq 1$ and with $M_n^{d,s}$ replaced by $M_n := \max_{i_m \leq n} \|\sum_{i_1, \dots, i_{m-1} : (i_1, \dots, i_m) \in I_m^n} h_n(X_{i_1}, \dots, X_{i_m})\|$ if the classes \mathcal{H}_n satisfy the hypotheses of Corollary 4.

Proof. 1) follows from Theorem 3 by integrating as in the proof of Theorem 2 b).

2) follows directly from 1) by noting that Theorem 3 also implies that if $\{U_n^{d,s}\}$ is stochastically bounded then

$$\limsup_{n \rightarrow \infty} u^p \Pr\{\|U_n^{d,s}\| > u\} \leq c_2 3^{mp} \limsup_{n \rightarrow \infty} u^p \Pr\{M_n^{d,s} > u\}.$$

3) follows from 2) and decoupling since for each $p \geq 1$ there is a non-negative convex function $\Phi_{p,u}$ such that $|x|^p I(|x| > u) \leq \Phi_{p,u}(x) \leq 2^p |x|^p I(|x| > \frac{u}{2})$. ■

In inequalities (1)-(4), one of the right hand side terms is obtained from the term at the left by replacing just one of the m sums by a max. We may ask if it is possible to actually replace more than one sum by the corresponding max in these inequalities. The following example shows that the answer is negative. Let $\{X_i\}_{i=1}^\infty$ be an i.i.d. sequence of symmetric p -stable random variables, $0 < p < 2$, and let $\{X_i^{(r)}\}$, $r = 1, \dots, m$, be, as usual, independent i.i.d. copies of $\{X_i\}$. We then have

$$\begin{aligned} & E\left(* * \sum_{i_1, \dots, i_m \leq n} \varepsilon_{i_1}^{(1)} \cdots \varepsilon_{i_m}^{(m)} I_{\{|X_{i_1}^{(1)} + \cdots + X_{i_m}^{(m)}| \geq a\}}\right)^4 \\ & \gtrsim n^{1+2(m-1)} P\{|X_1 + \cdots + X_m| \geq a, |X_1 + X_{m+2*} + \cdots + X_{2m}| \geq a\} \\ & \gtrsim \frac{n^{2m-1}}{4a^p}, \end{aligned}$$

as is seen using symmetry and stability in the inclusion

$$\begin{aligned}
& \{x_1 \geq a, \sum_{i=2}^m x_i \geq 0, \sum_{i=m+1}^* 2m * x_i \geq 0\} \\
& \cup \{x_1 < -a, \sum_{i=2}^m x_i \leq 0, \sum_{i=m+*2*}^* 2m * x_i \leq 0\} \\
& \subset \{|x_1 + \dots + x_m| \geq a, |x_1 + x_{m+1} + \dots + x_{2m}| \geq a\}.
\end{aligned}$$

We also have

$$\begin{aligned}
& \left[E \left(\sum_{i_1, \dots, i_m \leq n} \varepsilon_{i_1}^{(1)} \dots \varepsilon_{i_m}^{(m)} I_{[|X_{i_1}^{(1)} + \dots + X_{i_m}^{(m)}| \geq a]} \right)^2 \right]^2 \\
& = (n^m P\{|X_1 + \dots + X_m| \geq a\})^2 \\
& \approx \frac{n^{2m}}{a^{2p}}.
\end{aligned}$$

And, using e.g. hypercontractivity of the real Rademacher chaos (Bonami, 1970),

$$\begin{aligned}
& E \max_{i_1, i_2 \leq n} \left(\sum_{i_3, \dots, i_m \leq n} \varepsilon_{i_3}^{(3)} \dots \varepsilon_{i_m}^{(m)} I_{[|X_{i_1}^{(1)} + \dots + X_{i_m}^{(m)}| \geq a]} \right)^4 \\
& \leq n^2 E \left(\sum_{i_3, \dots, i_m \leq n} \varepsilon_{i_3}^{(3)} \dots \varepsilon_{i_m}^{(m)} I_{[|X_{i_1}^{(1)} + \dots + X_{i_m}^{(m)}| \geq a]} \right)^4 \\
& \leq 3^{2m} n^2 E_X \left[E_\varepsilon \left(\sum_{i_3, \dots, i_m \leq n} \varepsilon_{i_3}^{(3)} \dots \varepsilon_{i_m}^{(m)} I_{[|X_{i_3}^{(3)} + \dots + X_{i_m}^{(m)}| \geq a]} \right)^2 \right]^2 \\
& = 3^{2m} n^2 E \left(\sum_{i_3, \dots, i_m \leq n} I_{[|X_{i_3}^{(3)} + \dots + X_{i_m}^{(m)}| \geq a]} \right)^2 \\
& = 3^{2m} n^2 \sum_{\substack{i_3, \dots, i_m \leq n \\ j_3, \dots, j_m \leq n}} E I_{[|X_{i_1}^{(1)} + \dots + X_{i_m}^{(m)}| \geq a]} I_{[|X_{j_1}^{(1)} + \dots + X_{j_m}^{(m)}| \geq a]} \\
& \leq 3^{2m} n^{2+2(m-2)} P\{|X_{i_3}^{(3)} + \dots + X_{i_m}^{(m)}| \geq a\} \\
& \lesssim 3^{2m} m \frac{n^{2(m-1)}}{a^p}.
\end{aligned}$$

If we take, for each $n < \infty$, $a^p = a_n^p = n^{1+\delta}$ for $0 < \delta < 1$, the previous three bounds become respectively of the order of $n^{2m-2-\delta}$, $n^{2m-2-2\delta}$ and $n^{2m-3-\delta}$, i.e., as $n \rightarrow \infty$, the first term is of larger order than the other two. We have thus shown that it is not possible to replace more than one sum by a max in inequalities (2)-(5). This considerably hinders the applicability of the results above because it is not easy, in general, to prove uniform integrability of $\{(M_n^{d,s})^p\}$. See, however,

Arcones and Giné (1991) for some applications of these results to limit theorems for U-processes.

Acknowledgement. We thank M. Arcones for useful conversations regarding Corollary 4, and in particular for providing the above mentioned direct proof of it.

References

- Arcones, M.A. and Giné, E. (1991). Limit theorems for U-processes. Preprint.
Bonami, A. (1970). Etude des coefficients de Fourier des fonctions de $L_p(G)$. *Ann. Inst. Fourier* **20**, 335-402.
Hitczenko, P. (1988). Comparison of moments for tangent sequences of random variables. *Probab. The. Rel. Fields* **78**, 223-230.
Hoffmann-Jørgensen, J. (1974). Sums of independent Banach space valued random variables. *Studia Math.* **52**, 159-189.
de la Peña, V. (1990). Decoupling and Khintchine's inequalities for U-statistics. *Ann. Probability*, to appear.

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

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