The Annals of Probability
1992, Vol. 20, No. 4, 1877–1892

DECOUPLING AND KHINTCHINE'S INEQUALITIES FOR U-STATISTICS

BY VICTOR H. DE LA PEÑA

Columbia University

In this paper we introduce a fairly general decoupling inequality for U-statistics. Let $\{X_i\}$ be a sequence of independent random variables in a measurable space $(S, \mathscr{S})$, and let $\{\tilde{X}_i\}$ be an independent copy of $\{X_i\}$. Let $\Phi(x)$ be any convex increasing function for $x \geq 0$. Let Π_{ij} be families of functions of two variables taking $(S \times S)$ into a Banach space $(D, \|\cdot\|)$. If the $f_{ij} \in \Pi_{ij}$ are Bochner integrable and

$$\max_{1 \leq i \neq j \leq n} E\Phi\left(\sup_{f_{ij} \in \Pi_{ij}} \|f_{ij}(X_i, X_j)\|\right) < \infty,$$

then, under measurability conditions,

$$E\Phi\left(\sup_{\mathbf{f} \in \mathbf{\Pi}} \left\| \sum_{1 \leq i \neq j \leq n} f_{ij}(X_i, X_j) \right\|\right) \leq E\Phi\left(8 \sup_{\mathbf{f} \in \mathbf{\Pi}} \left\| \sum_{1 \leq i \neq j \leq n} f_{ij}(X_i, \tilde{X}_j) \right\|\right),$$

where $\mathbf{f} = (f_{ij},\ 1 \leq i \neq j \leq n)$ and $\mathbf{\Pi} = (\Pi_{ij},\ 1 \leq i \neq j \leq n)$. In the case where $\mathbf{\Pi}$ is a family of functions of two variables satisfying $f_{ij} = f_{ji}$ and $f_{ij}(X_i, X_j) = f_{ij}(X_j, X_i)$, the reverse inequality holds (with a different constant). As a corollary, we extend Khintchine's inequality for quadratic forms to the case of degenerate U-statistics. A new maximal inequality for degenerate U-statistics is also obtained. The multivariate extension is provided.

1. Introduction. Let $\{X_i\}$ be a sequence of independent random variables in a measurable space $(S, \mathscr{S})$. Let Π_{ij} be families of Bochner integrable functions of two variables taking $(S \times S)$ into D, where $(D, \|\cdot\|)$ is again a Banach space.

Let $\mathbf{f} = (f_{ij},\ 1 \leq i \neq j \leq n)$ and $\mathbf{\Pi} = (\Pi_{ij},\ 1 \leq i \neq j \leq n)$. Let Φ be an arbitrary convex increasing function for $x \geq 0$. Set

$$U_n = \sum_{1 \leq i \neq j \leq n} f_{ij}(X_i, X_j)$$

and

$$U_n(\Pi) = \sup_{\mathbf{f} \in \Pi} \left\| \sum_{1 \leq i \neq j \leq n} f_{ij}(X_i, X_j) \right\|$$

[the usual U-statistics can be obtained by letting $f_{ij} = f/\binom{n}{2}$]. In this paper we introduce new decoupling inequalities (to be defined later) for $E\Phi(U_n(\mathbf{\Pi}))$ and hence for $E\Phi(\|U_n\|)$. These inequalities make the problem of approximating

Received May 1990; revised March 1991.
AMS 1980 *subject classifications*. Primary 60E15; secondary 10C10.
Key words and phrases. Khintchine's inequalities, U-statistics, decoupling.

the preceding expectations easier. Beginning with McConnell and Taqqu (1986), a new and rich area of probability has been evolving. This is the area of so-called decoupling inequalities. In the context of our study, a decoupled version of U_n is

$$U_n^D = \sum_{1 \le i \ne j \le n} f_{ij}\big(X_i, \tilde{X}_j\big),$$

where $\{\tilde{X}_i\}$ is an independent copy of $\{X_i\}$. A decoupling inequality consists in comparing $E\Phi(\|U_n\|)$ to $E\Phi(\|U_n^D\|)$, with the idea that $E\Phi(\|U_n^D\|)$ should be easier to compute than $E\Phi(\|U_n\|)$. The special case where $\{X_i\}$ is a sequence of independent symmetric random variables and convex increasing functions Φ, and $U_n = \sum_{1 \le i \ne j \le n} a_{ij} X_i X_j$, with $\{a_{ij}\}$ a sequence of constants, is the one treated in McConnell and Taqqu (1986). Their interest in decoupling inequalities was motivated by the study of double stochastic integrals. In an unpublished work, de la Peña and Klass (1990) extend their results to include all mean-zero random variables and convex increasing functions Φ, based in part on Kwapień (1987), Bourgain and Tzafriri (1987) and unpublished work of Zinn (1989).

Zinn (1985) studied the modified version

$$\sum_{1 \le i < j \le n} f_{ij}(X_i, X_j)\varepsilon_i \varepsilon_j,$$

where the f_{ij} have values in R^1, $\{\varepsilon_i\}$ is an independent sequence of Bernoulli random variables, $P(\varepsilon_i = 1) = P(\varepsilon_i = -1) = \frac{1}{2}$ and $\{\tilde{\varepsilon}_i\}$ is an independent copy of $\{\varepsilon_i\}$. In that paper he proved

$$E\left|\sum_{1 \le i < j \le n} f_{ij}(X_i, X_j)\varepsilon_i \varepsilon_j\right|^\alpha \cong_\alpha E\left|\sum_{1 \le i < j \le n} f_{ij}\big(X_i, \tilde{X}_j\big)\varepsilon_i \tilde{\varepsilon}_j\right|^\alpha, \qquad 0 < \alpha < 2.$$

McConnell and Taqqu (1987) extended Zinn's result to the case where f_{ij}: $[0,1] \times [0,1] \Rightarrow E$ are Bochner integrable and the X_i's are i.i.d. uniform on $[0,1]$. They deal with convex functions instead of powers. The preceding may also be generalized to nonnegative or degenerate f_{ij} for $\alpha > 1$ (without the need for randomizing with $\{\varepsilon_i\}$) by using the results of Hitczenko (1988).

Nolan and Pollard (1987, 1988) introduced the U-process $U_n(f) = \sum_{1 \le i \ne j \le n} f(X_i, X_j)$ for $\{X_i\}$ i.i.d. and $f \in \mathscr{F}$, where $\mathscr{F}$ is a class of real-valued, symmetric functions on $S \times S$ with a nonnegative envelope F:

$$F(\cdot,\cdot) \ge |f(\cdot,\cdot)| \quad \text{if } f \in \mathscr{F}.$$

Define

$$U_n'(\mathscr{F}) = \sup_{f \in \mathscr{F}} \left\|\sum_{1 \le i \ne j \le n} f(X_i, X_j)\right\|.$$

In Lemma 1 of their first paper, they present an upper bound on $E\Phi(U_n'(\mathscr{F}))$ and used it to obtain uniform (for $f \in \mathscr{F}$) almost sure convergence results for the original U-process. Their bound is related to our decoupling inequality, which is more general.

Our decoupling inequality includes as special cases a generalization of the result in McConnell and Taqqu (1986, 1987), the application of Hitczenko's work to the problem at hand and, in some sense, the related result of Nolan and Pollard (1987). The main results are proved by using conditioning arguments, in particular, conditional Jensen's inequality. Throughout, we will assume all mathematical objects of interest are measurable. The Appendix contains a set of measurability conditions that would be sufficient for the proofs to be valid. Following Nolan and Pollard (1987), we refer the reader to Chapter 10 of Dudley (1984) for further technical points on measurability.

This paper is divided into five sections. In Section 2, we present the main result (in more generality than in the introduction) and a proof of the upper bound. In Section 3 we present a multivariate extension. The upper bound in the general case is proved by induction and a direct proof for the lower bound is given. In Section 4 we present several applications. The first shows that the results of McConnell and Taqqu (1987) are a special case of ours. In the special case of degenerate U-statistics, we introduce a generalization of Khintchine's inequality for quadratic forms and a new maximal inequality. The fifth section is an appendix.

Throughout this paper we will use the following notation:

$$E_\sigma Y = E[Y \mid \sigma],$$

where Y is a random variable and σ denotes a sigma field.

2. Main result.

THEOREM 1. *Let $X_1, X_2, \dots, X_n$ be a sequence of independent random variables in a measurable space $(S, \mathscr{S})$ and suppose that $\{\tilde{X}_i\}_{i=1}^n$ is an independent copy of $\{X_i\}_{i=1}^n$. Let Π_{ij}^m be families of Bochner integrable functions f_{ij}^m such that $f_{ij}^m \in \Pi_{ij}^m$ maps $S \times S \Rightarrow D$ with $(D, \|\cdot\|)$ a Banach space. Let N_n be an arbitrary subset of $\{1, 2, \dots, n\}$. For $x \geq 0$, let $\Phi(x)$ be any convex increasing function such that*

$$\max_{1 \leq i \neq j \leq n} E\Phi\left(\max_{m \in N_n} \sup_{f_{ij}^m \in \Pi_{ij}^m} \left\| f_{ij}^m(X_i, X_j) \right\| \right) < \infty.$$

Then, for $\mathbf{f}^m = (f_{ij}^m, 1 \leq i \neq j \leq n)$, $\mathbf{\Pi}^m = (\Pi_{ij}^m, 1 \leq i \neq j \leq m)$,

$$E\Phi\left(\max_{m \in N_n} \sup_{\mathbf{f}^m \in \mathbf{\Pi}^m} \left\| \sum_{1 \leq i \neq j \leq m} f_{ij}^m(X_i, X_j) \right\| \right)$$
$$\leq E\Phi\left(8 \max_{m \in N_n} \sup_{\mathbf{f}^m \in \mathbf{\Pi}^m} \left\| \sum_{1 \leq i \neq j \leq m} f_{ij}^m\left(X_i, \tilde{X}_j\right) \right\| \right).$$

If $f_{ij}^m \in \Pi_{ij}^m$ satisfy the symmetry conditions

(1) $$f_{ij}^m = f_{ji}^m \quad \text{and} \quad f_{ij}^m(X_i, X_j) = f_{ij}^m(X_j, X_i),$$

then the reverse bound holds:

$$E\Phi\left(\tfrac{1}{4}\max_{m\in N_n}\sup_{\mathbf{f}^m\in\Pi^m}\left\|\sum_{1\le i\ne j\le m} f_{ij}^m\left(X_i,\tilde{X}_j\right)\right\|\right)$$
$$\le E\Phi\left(\max_{m\in N_n}\sup_{\mathbf{f}^m\in\Pi^m}\left\|\sum_{1\le i\ne j\le m} f_{ij}^m(X_i,X_j)\right\|\right).$$

PROOF. To make the proof easier to follow, we will restrict attention to the special case $f_{ij}^m = f_{ij}$, for all m, and $N_n = \{n\}$. The proof requires the following result.

LEMMA 1. *Let*

$$\mathscr{G} = \sigma(Z_i, i = 1, 2, \dots, n), \tag{2}$$

where $\{Z_i\}$ *is a sequence of independent random vectors with* $Z_i = (X_i, \tilde{X}_i)$ *with probability* $\frac{1}{2}$ *and* $Z_i = (\tilde{X}_i, X_i)$ *with probability* $\frac{1}{2}$. *More formally,*

$$Z_i = \frac{(1+b_i)}{2}\left(X_i,\tilde{X}_i\right) + \frac{(1-b_i)}{2}\left(\tilde{X}_i, X_i\right),$$

where the b_i *are symmetric Bernoulli random variables independent of each other and of* $\{X_i\}$, $\{\tilde{X}_i\}$. *Then*,

$$\begin{aligned} E_{\mathscr{G}} f_{ij}(X_i,X_j) &= E_{\mathscr{G}} f_{ij}\left(X_i,\tilde{X}_j\right) \\ &= \frac{1}{4}\Big\{f_{ij}(X_i,X_j) + f_{ij}\left(X_i,\tilde{X}_j\right) \\ &\qquad + f_{ij}\left(\tilde{X}_i,X_j\right) + f_{ij}\left(\tilde{X}_i,\tilde{X}_j\right)\Big\}. \end{aligned} \tag{3}$$

PROOF. The first line in (3) holds since the conditional distributions of $f_{ij}(X_i, X_j)$ and $f_{ij}(X_i, \tilde{X}_j)$ given $\mathscr{G}$ are the same. The second line follows by noticing that the sum of those four terms is measurable with respect to $\mathscr{G}$. □

The preceding constitutes a variation on a result of Kwapień (1987). Now, we give the proof of the upper bound. The proof of the lower bound will be given in Section 3.

PROOF OF THE UPPER BOUND OF THEOREM 1. We use the following identity:

$$\begin{aligned} &\sum_{1\le i\ne j\le n} f_{ij}(X_i,X_j) \\ &\quad = \sum_{1\le i\ne j\le n}\Big\{E_{\mathscr{X}} f_{ij}(X_i,X_j) + E_{\mathscr{X}} f_{ij}\left(X_i,\tilde{X}_j\right) \\ &\qquad\qquad + E_{\mathscr{X}} f_{ij}\left(\tilde{X}_i,X_j\right) + E_{\mathscr{X}} f_{ij}\left(\tilde{X}_i,\tilde{X}_j\right)\Big\} \\ &\qquad - \sum_{1\le i\ne j\le n}\Big\{E_{\mathscr{X}} f_{ij}\left(X_i,\tilde{X}_j\right) + E_{\mathscr{X}} f_{ij}\left(\tilde{X}_i,X_j\right) + E_{\mathscr{X}} f_{ij}\left(\tilde{X}_i,\tilde{X}_j\right)\Big\}, \end{aligned}$$

where $\mathscr{X} = \sigma(X_1, \dots, X_n)$. In what follows, we let sup denote $\sup_{\mathbf{f}\in\Pi}$.

From the preceding and the convexity of Φ, it follows that

$$E\Phi\left(\sup\left\|\sum_{1\le i\ne j\le n} f_{ij}(X_i, X_j)\right\|\right)$$

$$\le \frac{1}{2}E\Phi\left(2\sup\left\|\sum_{1\le i\ne j\le n} E_{\mathscr{X}}\left\{f_{ij}(X_i, X_j) + f_{ij}\left(X_i, \tilde{X}_j\right)\right.\right.\right.$$

$$\left.\left.\left.+f_{ij}\left(\tilde{X}_i, X_j\right) + f_{ij}\left(\tilde{X}_i, \tilde{X}_j\right)\right\}\right\|\right)$$

$$+\frac{1}{2}E\Phi\left(2\sup\left\|\sum_{1\le i\ne j\le n} E_{\mathscr{X}}\left\{f_{ij}\left(X_i, \tilde{X}_j\right) + f_{ij}\left(\tilde{X}_i, X_j\right) + f_{ij}\left(\tilde{X}_i, \tilde{X}_j\right)\right\}\right\|\right)$$

[by using conditional Jensen's inequality on the first term]

$$\le \frac{1}{2}E\Phi\left(2\sup\left\|\sum_{1\le i\ne j\le n} f_{ij}(X_i, X_j) + f_{ij}\left(X_i, \tilde{X}_j\right)\right.\right.$$

$$\left.\left.+f_{ij}\left(\tilde{X}_i, X_j\right) + f_{ij}\left(\tilde{X}_i, \tilde{X}_j\right)\right\|\right)$$

$$+\frac{1}{2}E\Phi\left(2\sup\left\|\sum_{1\le i\ne j\le n} E_{\mathscr{X}}\left\{f_{ij}\left(X_i, \tilde{X}_j\right)\right.\right.\right.$$

$$\left.\left.\left.+f_{ij}\left(\tilde{X}_i, X_j\right) + f_{ij}\left(\tilde{X}_i, \tilde{X}_j\right)\right\}\right\|\right)$$

[by (3) and the convexity of Φ]

$$\le \frac{1}{2}E\Phi\left(8\sup\left\|\sum_{1\le i\ne j\le n} E_{\mathscr{Y}} f_{ij}\left(X_i, \tilde{X}_j\right)\right\|\right)$$

$$+\frac{1}{6}\left\{E\Phi\left(6\sup\left\|\sum_{1\le i\ne j\le n} E_{\mathscr{X}} f_{ij}\left(X_i, \tilde{X}_j\right)\right\|\right)\right.$$

$$+E\Phi\left(6\sup\left\|\sum_{1\le i\ne j\le n} E_{\mathscr{X}} f_{ij}\left(\tilde{X}_i, X_j\right)\right\|\right)$$

$$\left.+\Phi\left(6\sup\left\|\sum_{1\le i\ne j\le n} E f_{ij}\left(\tilde{X}_i, \tilde{X}_j\right)\right\|\right)\right\}$$

by conditional Jensen's inequality and the identity $Ef_{ij}(\tilde{X}_i, \tilde{X}_j) = Ef_{ij}(X_i, \tilde{X}_j)$]

$$\leq \frac{1}{2} E\Phi\left(8 \sup \left\| \sum_{1 \leq i \neq j \leq n} f_{ij}(X_i, \tilde{X}_j) \right\|\right)$$

$$+ \frac{2}{6} E\Phi\left(6 \sup \left\| \sum_{1 \leq i \neq j \leq n} f_{ij}(X_i, \tilde{X}_j) \right\|\right)$$

$$+ \frac{1}{6} E\Phi\left(6 \sup \left\| \sum_{1 \leq i \neq j \leq n} Ef_{ij}(X_i, \tilde{X}_j) \right\|\right)$$

[by regular Jensen's inequality on the last term and regrouping]

$$\leq \frac{1}{2} E\Phi\left(8 \sup \left\| \sum_{1 \leq i \neq j \leq n} f_{ij}(X_i, \tilde{X}_j) \right\|\right) + \frac{1}{2} E\Phi\left(6 \sup \left\| \sum_{1 \leq i \neq j \leq n} f_{ij}(X_i, \tilde{X}_j) \right\|\right)$$

[since Φ is convex increasing]

$$\leq E\Phi\left(8 \sup \left\| \sum_{1 \leq i \neq j \leq n} f_{ij}(X_i, \tilde{X}_j) \right\|\right).$$

The proof of the lower bound is presented in Section 3 in its multivariate form. □

REMARKS. The fact that the lower bound does not hold for general f_{ij} follows trivially by using

$$f_{ij}(X_i, X_j) = f(X_i, X_j) = X_j - X_i,$$

because then, $\sum_{1 \leq i \neq j \leq n} f(X_i, X_j) = 0$. If the kernels are not symmetric, one may still obtain a lower bound by using the symmetrized kernels $\hat{f}_{ij}(X_i, X_j) = (f_{ij}(X_i, X_j) + f_{ij}(X_j, X_i))/2$, for $i < j$, and letting $\hat{f}_{ji} = \hat{f}_{ij}$.

Regarding the upper bound, the range of summation can be replaced by any subset of $\{1, 2, \dots, n\}^2$ as long as $i \neq j$. The set $\{1 \leq i < j \leq n\}$ is such an example.

3. Multivariate extension. The following result is a generalization of Theorem 1. Here and in the sequel, the expression $i_1 \neq i_2 \neq \cdots \neq i_m$ is used to mean that all of $i_1, i_2, \dots, i_m$ are different from each other.

THEOREM 2. *Let $X_1, X_2, \dots, X_n$ be a sequence of independent random variables in a measurable space $(S, \mathscr{S})$, and suppose that $\{X_1^j, \dots, X_n^j\}_{j=1}^k$ are independent copies of $\{X_i\}_{i=1}^n$. Let $\Pi_{i_1 \dots i_k}^m$ be families of Bochner integrable functions $f_{i_1 \dots i_k}^m$ such that $f_{i_1 \dots i_k}^m \in \Pi_{i_1, \dots, i_k}^m$ maps $S \times S \times \cdots \times S \Rightarrow D$ with $(D, \|\cdot\|)$ a Banach space. Let N_n be an arbitrary subset of $\{1, 2, \dots, n\}$. For*

$x \geq 0$, *let* $\Phi(x)$ *be any convex increasing function such that*

$$\max_{1 \leq i_1 \neq i_2 \neq \cdots \neq i_k \leq n} E\Phi\left(\max_{m \in N_n} \sup_{f_{i_1 \cdots i_k}^m \in \Pi_{i_1 \cdots i_k}^m} \left\| f_{i_1 i_2 \cdots i_k}^m (X_{i_1}, \ldots, X_{i_k}) \right\| \right) < \infty.$$

Then for $\mathbf{f}^m = (f_{i_1 \cdots i_k}^m,\ 1 \leq i_1 \neq i_2 \neq \cdots \neq i_k \leq n)$, $\mathbf{\Pi}^m = (\Pi_{i_1 \cdots i_k}^m,\ 1 \leq i_1 \neq \cdots \neq i_k \leq n)$, $m \in N_n$,

$$E\Phi\left(\max_{m \in N_n} \sup_{\mathbf{f}^m \in \mathbf{\Pi}^m} \left\| \sum_{1 \leq i_1 \neq i_2 \neq \cdots \neq i_k \leq m} f_{i_1 \cdots i_k}^m (X_{i_1}, \ldots, X_{i_k}) \right\| \right)$$

$$\leq E\Phi\left(C_k \max_{m \in N_n} \sup_{\mathbf{f}^m \in \mathbf{\Pi}^m} \left\| \sum_{1 \leq i_1 \neq i_2 \neq \cdots \neq i_k \leq m} f_{i_1 \cdots i_k}^m (X_{i_1}^1, \ldots, X_{i_k}^k) \right\| \right),$$

with $C_k = 2^k(k^k - 1)((k-1)^{k-1} - 1) \times \cdots \times 3$. *In case the kernels satisfy the degeneracy property*

$$E_{\mathscr{X}^1} f_{i_1 \cdots i_k}^m (X_{i_1}^1, \ldots, X_{i_k}^k) = 0,$$

where $\mathscr{X}^i = \sigma(X_1^i, \ldots, X_n^i)$, *then the constant* C_k *can be taken to be* k^k.

If $f_{i_1 \cdots i_k}^m \in \Pi_{i_1 \cdots i_k}^m$ *satisfy the symmetry conditions*

(4)

$$f_{i_1 \cdots i_k}^m = f_{i_{s_1} \cdots i_{s_k}}^m \quad \text{and} \quad f_{i_1 \cdots i_k}^m (X_{i_1}, \ldots, X_{i_k}) = f_{i_1 \cdots i_k}^m (X_{i_{s_1}}, \ldots, X_{i_{s_k}}),$$

for all permutations $(i_{s_1}, \ldots, i_{s_k})$ *of* $(i_1, \ldots, i_k)$, *then the reverse bound for* $k \geq 2$ *holds*:

$$E\Phi\left(\frac{1}{2^{(2k-2)}(k-1)!} \max_{m \in N_n} \sup_{\mathbf{f}^m \in \Pi^m} \left\| \sum_{1 \leq i_1 \neq i_2 \neq \cdots \neq i_k \leq m} f_{i_1 \cdots i_k}^m (X_{i_1}^1, \ldots, X_{i_k}^k) \right\| \right)$$

$$\leq E\Phi\left(\max_{m \in N_n} \sup_{\mathbf{f}^m \in \Pi^m} \left\| \sum_{1 \leq i_1 \neq i_2 \neq \cdots \neq i_k \leq m} f_{i_1 \cdots i_k}^m (X_{i_1}^1, \ldots, X_{i_k}^1) \right\| \right).$$

PROOF. We will treat the upper and lower bounds separately. The upper bound is proved by induction on k. For the lower bound, a direct proof is given. The proof of the lower bound presents new interesting technical difficulties.

PROOF OF THE UPPER BOUND IN THE MULTIVARIATE CASE. Instead of $\sum_{1 \leq i_1 \neq i_2 \neq \cdots \neq i_k \leq n}$ we will use $\sum$. Instead of $f_{i_1, \ldots, i_k}^m$ we will use just f. Instead of $\max_{m \in N_n} \sup_{\mathbf{f}^m \in \mathbf{\Pi}^m}$ we will use just sup. Recall that $X_1^1, \ldots, X_n^1, X_1^2, \ldots, X_n^2, \ldots, X_1^k, \ldots, X_n^k$ are independent copies of $X_1, \ldots, X_n$. We will use induction over k. For $k = 2$ see Theorem 1. Assume that the upper bound is valid for $2,\dots,k-1$. Let $\mu = \mu_{i_1\cdots i_k} = Ef(X_{i_1}^{j_1},\dots,X_{i_k}^{j_k})$. We have

$$E\Phi\Big(\sup\Big\|\sum f\big(X_{i_1}^1, X_{i_2}^1,\dots,X_{i_k}^1\big)\Big\|\Big)$$

$$\le \frac{1}{2}E\Phi\left(2\sup\left\|\sum\left\{f\big(X_{i_1}^1, X_{i_2}^1,\dots,X_{i_k}^1\big) + \sum_{\substack{\text{not all}\\ j\text{'s equal}}} E_{\mathscr{X}^1} f\big(X_{i_1}^{j_1}, X_{i_2}^{j_2},\dots,X_{i_k}^{j_k}\big)\right\} + (k-1)\mu\right\|\right)$$

$$+\frac{1}{2}E\Phi\left(2\sup\left\|\sum\left\{\sum_{\substack{\text{not all}\\ j\text{'s equal}}} E_{\mathscr{X}^1} f\big(X_{i_1}^{j_1}, X_{i_2}^{j_2},\dots,X_{i_k}^{j_k}\big)\right\} + (k-1)\mu\right\|\right)$$

[because of Jensen's inequality and convexity]

$$\le \frac{1}{2}E\Phi\left(2\sup\left\|\sum\sum_{1\le j_1,\dots,j_k\le k} f\big(X_{i_1}^{j_1},\dots,X_{i_k}^{j_k}\big)\right\|\right)$$

$$+\frac{1}{2(k^k-1)}\sum_{\substack{\text{not all}\\ j\text{'s equal}}} E\Phi\Big(2(k^k-1)\sup\Big\|\sum f\big(X_{i_1}^{j_1},\dots,X_{i_k}^{j_k}\big)\Big\|\Big)$$

$$+\frac{k-1}{2(k^k-1)}E\Phi\Big(2(k^k-1)\sup\Big\|\sum f\big(X_{i_1}^1,\dots,X_{i_k}^k\big)\Big\|\Big).$$

Conditioning on $\mathscr{X}^j = \sigma(X_1^j,\dots,X_n^j)$ for appropriate j's, we can bound all the summands in the second term by the summands involving $f(X_{i_1}^1,\dots,X_{i_k}^k)$, using the induction hypothesis. For example, in case $k=3$ we have

$$E\Phi\left(52\sup\left\|\sum_{1\le i_1\ne i_2\ne i_3\le n} f\big(X_{i_1}^1, X_{i_2}^1, X_{i_3}^2\big)\right\|\right)$$

$$= EE\left[\Phi\left(52\sup\left\|\sum_{1\le i_1\ne i_2\ne i_3\le n} f\big(X_{i_1}^1, X_{i_2}^1, X_{i_3}^2\big)\right\|\right)\middle| X_1^2,\dots,X_n^2\right]$$

$$= EE\left[\Phi\left(52\sup\left\|\sum_{1\le i_1\ne i_2\le n} g\big(X_{i_1}^1, X_{i_2}^1\big)\right\|\right)\middle| X_1^2,\dots,X_n^2\right],$$

where

$$g\big(X_{i_1}^1, X_{i_2}^1\big) = \sum_{\substack{i\le i\le n\\ i\ne i_1,i_2}} f\big(X_{i_1}^1, X_{i_2}^1, X_i^2\big).$$

Hence, using the result for $k = 2$,

$$E\Phi\Big(52\sup\Big\|\sum f\big(X^1_{i_1}, X^1_{i_2}, X^2_{i_3}\big)\Big\|\Big) \le E\Phi\Big(52K\sup\Big\|\sum f\big(X^1_{i_1}, X^3_{i_2}, X^2_{i_3}\big)\Big\|\Big) = E\Phi\Big(52K\sup\Big\|\sum f\big(X^1_{i_1}, X^2_{i_2}, X^3_{i_3}\big)\Big\|\Big),$$

where K is the upper bound constant in case $k = 2$. Because of the exponential nature of the bounding constants, it's easy to see that the largest constant is used to bound the terms involving $f(X^1_{i_1}, \dots, X^1_{i_{k-1}}, X^k_{i_k})$ (i.e., all but one observation are from the copy of the same kind) and that constant is $2^{k-1}((k-1)^{k-1} - 1)((k-2)^{k-2} - 1) \times \cdots \times 3$ (obtained by the conditioning argument in the preceding example, using the induction hypothesis). Hence, from the multivariate analogue of relation (3) in the proof of Theorem 1 we obtain the upper bound by using conditional Jensen's inequality applied to the expression

$$\tfrac{1}{2}E\Phi\Big(2k^k\sup\Big\|\sum E_{\mathscr{G}_k} f\big(X^1_{i_1}, \dots, X^k_{i_k}\big)\Big\|\Big) + \tfrac{1}{2}E\Phi\Big(C_k\sup\Big\|\sum f\big(X^1_{i_1}, \dots, X^k_{i_k}\big)\Big\|\Big) \le E\Phi\Big(C_k\sup\Big\|\sum f\big(X^1_{i_1}, \dots, X^k_{i_k}\big)\Big\|\Big),$$

where $\mathscr{G}_k$ is the analogue of the σ-algebra in (2),

$$\mathscr{G}_k = \sigma(Z_i, i = 1, \dots, n), \tag{5}$$

where $\{Z_i\}$ are independent random vectors with $Z_i = (X^{j_1}_i, X^{j_2}_i, \dots, X^{j_k}_i)$ with probability $1/k!$ for each permutation $(j_1, j_2, \dots, j_k)$ of $(1, 2, \dots, k)$ and where $C_k = 2^k(k^k - 1)((k-1)^{k-1} - 1) \times \cdots \times 3$.

In the case where the kernels are degenerate, the constant $C_k = k^k$ is obtained by the following argument.

By the degeneracy property,

$$E\Phi\Big(\sup\Big\|\sum f\big(X^1_{i_1}, X^1_{i_2}, \dots, X^1_{i_k}\big)\Big\|\Big)$$

$$= E\Phi\left(\sup\left\|\sum\left\{f\big(X^1_{i_1}, X^1_{i_2}, \dots, X^1_{i_k}\big) + \sum_{\substack{\text{not all}\\ j\text{'s equal}}} E_{\mathscr{X}^1} f\big(X^1_{i_1}, X^2_{i_2}, \dots, X^k_{i_k}\big) + (k-1)\mu\right\}\right\|\right)$$

[because of conditional Jensen's inequality and convexity]

$$\le E\Phi\left(\sup\left\|\sum\sum_{1\le j_1, \dots, j_k \le k} f\big(X^{j_1}_{i_1}, \dots, X^{j_k}_{i_k}\big)\right\|\right).$$

[by the multivariate version of (3)]

$$= E\Phi\Big(k^k\sup\Big\|\sum E_{\mathscr{G}_k} f\big(X^1_{i_1}, \dots, X^k_{i_k}\big)\Big\|\Big)$$

[by conditional Jensen's inequality]

$$\le E\Phi\Big(k^k\sup\Big\|\sum f\big(X^1_{i_1}, \dots, X^k_{i_k}\big)\Big\|\Big).$$

PROOF OF THE LOWER BOUND: MULTIVARIATE CASE. Suppose the symmetry assumptions (4) on $\{f_{i_1}\cdots_{i_k}\}$ hold. Under (4),

$$(6)\quad \sum_{1\le i_1\ne i_2\ne\cdots\ne i_k\le n} f_{i_1\cdots i_k}\left(X_{i_1}^1,\dots,X_{i_k}^k\right)=\frac{1}{k!}\sum f_{i_1,\dots,i_k}\left(X_{i_1}^{j_1},\dots,X_{i_k}^{j_k}\right),$$

where the right-hand side summation is over $(j_1,\dots,j_k)\in$ {all permutations of $(1,\dots,k)$} and $1\le i_1\ne i_2\ne\cdots\ne i_k\le n$. Instead of

$$\sum_{\substack{\{j_1,\dots,j_k\}\in J\\ 1\le i_1\ne i_2\ne\cdots\ne i_k\le n}} f_{i_1,\dots,i_k}\left(X_{i_1}^{j_1},\dots,X_{i_k}^{j_k}\right)$$

we use $\Sigma\{j_1,\dots,j_k\in J\}$. Then the lower bound part of the theorem can be written as

$$E\Phi\left(\sup\left\|\sum\{j_1=1,\dots,j_k=k\}\right\|\right)$$
$$\le E\Phi\left(2^{2k-2}(k-1)!\sup\left\|\sum\{j_1=\cdots=j_k=1\}\right\|\right).$$

Now, by (6) and convexity,

$$E\Phi\left(\sup\left\|\sum\{j_1=1,\dots,j_k=k\}\right\|\right)$$
$$\le\frac{1}{2}E\Phi\left(\frac{2}{k!}\sup\left\|\sum\{j_1,\dots,j_k\in\{1,\dots,k\}\}\right\|\right)$$
$$+\frac{1}{2}E\Phi\left(\frac{2}{k!}\sup\left\|\sum\{j_1,\dots,j_k\in\{1,\dots,k\},\ <k\ j\text{'s different}\}\right\|\right).$$

Since for $\mathcal{G}_l$ as in (5) the following analog of (3) is valid:

$$(7)\quad \sum_{i_1,\dots,i_k\ \text{fixed}}\{j_1,\dots,j_k\in\{1,\dots,l\}\}=l^kE_{\mathcal{G}_l}f_{i_1,\dots,i_k}\left(X_{i_1}^1,\dots,X_{i_k}^1\right),$$

for $l=2,\dots,k$, we can use Jensen's inequality to bound the first term by

$$\frac{1}{2}E\Phi\left(\frac{2}{k!}k^k\sup\left\|\sum\{j_1=\cdots=j_k=1\}\right\|\right),$$

which is what we want. The second term is less than or equal to

$$\frac{1}{2(k-1)}E\Phi\left(\frac{2(k-1)}{k!}\sup\left\|\sum\{j_1,\dots,j_k\in\{1,\dots,k\},\ \text{all } j\text{'s equal}\}\right\|\right)$$
$$+\frac{1}{2(k-1)}E\Phi\left(\frac{2(k-1)}{k!}\sup\left\|\sum\{j_1,\dots,j_k\in\{1,\dots,k\},\right.\right.$$
$$\left.\left.\text{exactly 2 } j\text{'s different}\}\right\|\right)+\cdots$$
$$+\frac{1}{2(k-1)}E\Phi\left(\frac{2(k-1)}{k!}\sup\left\|\sum\{j_1,\dots,j_k\in\{1,\dots,k\},\right.\right.$$
$$\left.\left.\text{exactly } k-1\ j\text{'s different}\}\right\|\right).$$

The last term of this is also less than or equal to

$$\frac{1}{2(k-1)}E\Phi\left(\frac{2(k-1)}{k!}\binom{k}{k-1}\sup\left\|\sum\{j_1,\dots,j_k\in\{1,\dots,k-1\},\ \text{exactly } k-1\ j\text{'s different}\}\right\|\right)$$

$$\leq \frac{1}{4(k-1)}E\Phi\left(\frac{4}{k!}(k-1)k\sup\left\|\sum\{j_1,\dots,j_k\in\{1,\dots,k-1\}\}\right\|\right)$$
$$+\frac{1}{4(k-1)}E\Phi\left(\frac{4}{k!}(k-1)k\sup\left\|\sum\{j_1,\dots,j_k\in\{1,\dots,k-1\},\ < k-1\ j\text{'s different}\}\right\|\right).$$

The first term can again be bounded using (7) with $l = k-1$. We deal with the second term exactly as before, when we had "less than k j's different." Proceeding inductively in the same way, we finally get that the "last" last term is bounded by

$$\frac{1}{C_k}E\Phi\left(2^{k-3}(k-1)!\sup\left\|\sum\{j_1,\dots,j_k\in\{1,2\},\ \text{exactly two } j\text{'s different}\}\right\|\right)$$
$$\leq \frac{1}{2C_k}E\Phi\left(2^{k-2}(k-1)!\sup\left\|\sum\{j_1,\dots,j_k\in\{1,2\}\}\right\|\right)$$
$$+\frac{1}{2C_k}E\Phi\left(2^{k-2}(k-1)!\sup\left\|\sum\{j_1,\dots,j_k\in\{1,2\},\ \text{all } j\text{'s equal}\}\right\|\right)$$
$$\leq \frac{1}{2C_k}E\Phi\left(2^{2k-2}(k-1)!\sup\left\|\sum\{j_1=\cdots=j_k=1\}\right\|\right)$$
$$+\frac{1}{2C_k}E\Phi\left(2^{k-1}(k-1)!\sup\left\|\sum\{j_1=\cdots=j_k=1\}\right\|\right),$$

where C_k is a constant we do not have to worry about since it came from the convexity argument, so that when we add up all of the terms the outside constants $1/2C_k$ will add up to 1. Applying the preceding procedure to all other terms we missed, it is easy to see that it really is always the last term that carries the largest (inside) constant, so that all of them will be bounded by

$$\frac{1}{2C_k}E\Phi\left(2^{2k-2}(k-1)!\sup\left\|\sum\{j_1=\cdots=j_k=1\}\right\|\right).$$

This completes the proof. □

4. Applications. The first application generalizes the results of Zinn (1985) and McConnell and Taqqu (1987). Moreover, the constants we get are better.

COROLLARY 1. *In the notation of Theorem 2, assume that the g's satisfy the symmetry conditions (4), and suppose $\{\varepsilon_i^1\}_{i=1}^n, \{\varepsilon_i^2\}_{i=1}^n, \dots, \{\varepsilon_i^k\}_{i=1}^n$ are independent copies of a sequence of independent Bernoulli random variables, $\{\varepsilon_i\}$ with $P(\varepsilon_i = 1) = P(\varepsilon_i = -1) = \frac{1}{2}$, where the ε's are taken to be independent of the X's. Then,*

$$E\Phi\left(\frac{1}{2^{(2k-2)}(k-1)!}\left\|\sum_{1\le i_1\ne i_2\ne\cdots\ne i_k\le m} g_{i_1\cdots i_k}\left(X_{i_1}^1,\dots,X_{i_k}^k\right)\varepsilon_{i_1}^1\cdots\varepsilon_{i_k}^k\right\|\right)$$
$$\le E\Phi\left(\left\|\sum_{1\le i_1\ne i_2\ne\cdots\ne i_k\le m} g_{i_1\cdots i_k}\left(X_{i_1}^1,\dots,X_{i_k}^1\right)\varepsilon_{i_1}^1\cdots\varepsilon_{i_k}^1\right\|\right)$$
$$\le E\Phi\left(k^k\left\|\sum_{1\le i_1\ne i_2\ne\cdots\ne i_k\le m} g_{i_1\cdots i_k}\left(X_{i_1}^1,\dots,X_{i_k}^k\right)\varepsilon_{i_1}^1\cdots\varepsilon_{i_k}^k\right\|\right).$$

PROOF. The result follows directly from Theorem 2 by letting $Y_i^j = (X_i^j, \varepsilon_i^j)$ and taking $f_{i_1\dots i_k}^m(Y_{i_1},\dots,Y_{i_k}) = g_{i_1,\dots,i_k}(X_{i_1},\dots,X_{i_k})\varepsilon_{i_1}\cdots\varepsilon_{i_k}$. □

In Corollaries 2 and 3, we will make the assumption that the kernels f_{ij} are symmetric (as defined in Theorem 1) and degenerate, that is, we require that

$$E\left(f_{ij}(X_i, X_j)\mid X_i\right) = E\left(f_{ij}(X_i, X_j)\mid X_j\right) = 0.$$

A special case of Corollaries 2 and 3 involving quadratic forms was introduced in de la Peña and Klass (1990).

The following is an extension of Lévy's inequality.

COROLLARY 2. *In the notation of Theorem 1,*

$$E\Phi\left(\max_{1\le m\le n}\left\|\sum_{1\le i\ne j\le m} f_{ij}(X_i, X_j)\right\|\right) \le 2E\Phi\left(32\left\|\sum_{1\le i\ne j\le n} f_{ij}(X_i, X_j)\right\|\right).$$

PROOF. From Theorem 1 and Theorem 2 it follows that

$$E\Phi\left(\max_{1\le m\le n}\left\|\sum_{1\le i\ne j\le m} f_{ij}(X_i, X_j)\right\|\right)$$
$$\le E\Phi\left(4\max_{1\le m\le n}\left\|\sum_{1\le i\ne j\le m} f_{ij}\left(X_i, \tilde{X}_j\right)\right\|\right)$$

[by conditioning on $X_1, \dots, X_n$ and using Lemma A2]

$$\le 2E\Phi\left(8\left\|\sum_{1\le i\ne j\le n} f_{ij}\left(X_i, \tilde{X}_j\right)\right\|\right) \quad \text{[by Theorems 1 and 2 again]}$$
$$\le 2E\Phi\left(32\left\|\sum_{1\le i\ne j\le n} f_{ij}(X_i, X_j)\right\|\right).$$

REMARK. The case $\Phi(x) = |x|$ is of special interest, since the preceeding result is a strict improvement over the use of Doob's inequality to bound the L^1 norm of the maximum of a martingale when applied to this problem.

COROLLARY 3 (Khintchine's inequality for degenerate U-statistics). *Let $\{X_i\}$ be a sequence of independent random variables, $\{f_{ij}\}$ a sequence of real-valued functions such that $f_{ij} = f_{ji}$ and $f_{ij}(X_i, X_j) = f_{ij}(X_j, X_i)$. Let $\Phi(x)$ be any convex increasing function for $x \geq 0$ such that $\Phi(2x) \leq 2^\alpha \Phi(x)$ for some $\alpha \geq 1$. If $E(f_{ij}(X_i, X_j) \mid X_i) = E(f_{ij}(X_i, X_j) \mid X_j) = 0$, then*

$$C_{1\alpha} E\Phi\left(\sqrt{\sum_{1 \leq i \neq j \leq n} \left(f_{ij}(X_i, X_j)\right)^2}\right) \leq E\Phi\left(\left|\sum_{1 \leq i \neq j \leq n} f_{ij}(X_i, X_j)\right|\right)$$

$$\leq C_{2\alpha} E\Phi\left(\sqrt{\sum_{1 \leq i \neq j \leq n} \left(f_{ij}(X_i, X_j)\right)^2}\right),$$

where $0 < C_{1\alpha}, C_{2\alpha} < \infty$ depend only on α.

PROOF. We will use the symbol $A_n \cong_\alpha B_n$ to denote that the ratio of adjacent quantities is bounded away from zero and infinity by positive constants depending on α only.

By Theorem 1,

$$E\Phi\left(\left|\sum_{1 \leq i \neq j \leq n} f_{ij}(X_i, X_j)\right|\right)$$

$$\cong_\alpha E\Phi\left(\left|\sum_{1 \leq i \neq j \leq n} f_{ij}\left(X_i, \tilde{X}_j\right)\right|\right) \quad \text{[Lemma A1 applied twice]}$$

$$\cong_\alpha E\Phi\left(\left|\sum_{1 \leq i \neq j \leq n} f_{ij}\left(X_i, \tilde{X}_j\right)\varepsilon_i \tilde{\varepsilon}_j\right|\right) \quad \text{[by Corollary 1]}$$

$$\cong_\alpha E\Phi\left(\left|\sum_{1 \leq i \neq j \leq n} f_{ij}(X_i, X_j)\varepsilon_i \varepsilon_j\right|\right)$$

$$\cong_\alpha E\Phi\left(\left|\sum_{1 \leq i \neq j \leq n} f_{ij}(X_i, X_j)\varepsilon_i \tilde{\varepsilon}_j\right|\right)$$

[by Theorem 1, conditioning on $X_1, \ldots, X_n$]

$$\cong_\alpha E\Phi\left(\sqrt{\sum_{1 \leq i \neq j \leq n} f_{ij}^2(X_i, X_j)}\right).$$

The last line follows by conditioning on $(X_1, \ldots, X_n)$ and using the regular Khintchine's inequality. [See McConnell and Taqqu (1986) for a very nice proof of the version of Khintchine's inequality we are using.] □

To make the paper more complete, we present the following results.

APPENDIX

Symmetrization lemma.

LEMMA A1. *Let $X_1, \dots, X_n$ be a sequence of independent random elements taking values in a Banach space $(B, \|\cdot\|)$ with $EX_i = 0$ for all i. Let $\{\varepsilon_i\}$ be a sequence of independent Bernoulli r.v's independent of $\{X_i\}$. Then, for any convex increasing function Φ,*

$$E\Phi\left(\frac{1}{2}\left\|\sum_{i=1}^{n} X_i\varepsilon_i\right\|\right) \le E\Phi\left(\left\|\sum_{i=1}^{n} X_i\right\|\right) \le E\Phi\left(2\left\|\sum_{i=1}^{n} X_i\varepsilon_i\right\|\right).$$

PROOF. This follows easily from known symmetrization results. See, for example, Araujo and Giné [(1980), Lemma 2.13]. □

The following is an easy consequence of Lévy's inequality.

LEMMA A2. *Let $X_1, \dots, X_n$ be independent mean-zero random elements in a Banach space $(B, \|\cdot\|)$. Let $\Phi(x)$, $x \ge 0$, be a convex increasing function. Then*

$$E\Phi\left(\max_{j\le n}\left\|\sum_{i=1}^{j} X_i\right\|\right) \le 2E\Phi\left(2\left\|\sum_{i=1}^{n} X_i\right\|\right).$$

PROOF. Let $\{\tilde{X}_i\}$ be an independent copy of $\{X_i\}$. Then $\{X_i - \tilde{X}_i\}$ is a sequence of independent symmetric random elements, and, by Jensen's inequality,

$$\begin{aligned}
E\Phi\left(\max_{j\le n}\left\|\sum_{i=1}^{j} X_i\right\|\right) &\le E\Phi\left(\max_{j\le n}\left\|\sum_{i=1}^{j}\left(X_i - \tilde{X}_i\right)\right\|\right)\\
&\le 2E\Phi\left(\left\|\sum_{i=1}^{n}\left(X_i - \tilde{X}_i\right)\right\|\right) \quad \text{[by Lévy's inequality]}\\
&\le 2\left(\frac{1}{2}E\Phi\left(2\left\|\sum_{i=1}^{n} X_i\right\|\right) + \frac{1}{2}E\Phi\left(2\left\|\sum_{i=1}^{n} \tilde{X}_i\right\|\right)\right)\\
&= 2E\Phi\left(2\left\|\sum X_i\right\|\right).
\end{aligned}$$

Measurability conditions. In what follows we provide the measurability conditions under which Theorems 1 and 2 of this paper hold, mainly summarizing D. Pollard's ideas in the subject [Pollard (1991)]. We are thankful to him for allowing us to include this summary that makes this paper self-contained.

In addition to the conditions of Theorem 1, assume that for each i, X_i, $\tilde{X}_i$ are the coordinates of the product measure space $(\Pi_{i=1}^{\infty} S_i \times \Pi_{i=1}^{\infty} S_i, \Pi_{i=1}^{\infty} P_i \times \Pi_{i=1}^{\infty} P_i)$ (the usual assumption in empirical process theory). For simplicity, let $N_n = \{n\}$ and consider Bochner integrable functions $\{f_{ij}\}$ with

$$\max_{1\le i\ne j\le n} E^*\Phi\left(\sup_{f_{ij}\in\Pi_i j}\|f_{ij}(X_i, X_j)\|\right) < \infty$$

where E^* stands for outer expectations. Then Theorem 1 can be restated as

$$E^*\Phi\left(\sup_{\mathbf{f}\in\Pi}\left\|\sum_{1\le i\ne j\le n} f_{ij}(X_i, X_j)\right\|\right) \le E^*\Phi\left(8\sup_{\mathbf{f}\in\Pi}\left\|\sum_{1\le i\ne j\le n} f_{ij}\left(X_i, \tilde{X}_j\right)\right\|\right).$$

We recall that $E^*f = Ef^*$, where f^* is the measurable support of f [see e.g., Dudley and Philipp (1983)]. To show this, one modifies Lemma 1 as follows: If $\{r_i\}_{i=1}^{\infty}$ is a Rademacher sequence [i.e., the r_i are independent and $P(r_i = 1) = P(r_i = -1) = \frac{1}{2}$] independent of $\{X_i\}$ and $\{\tilde{X}_i\}$, with all the variables defined as coordinates in a product probability space, and if we let $X_i^r = X_i$ if $r_i = 1$ and $X_i^r = \tilde{X}_i$ if $r_i = -1$, for $i = 1, \dots, m$, and likewise for $\tilde{X}_i^r$, then for $\mathscr{G} = \sigma(X_1, \dots, X_n; \tilde{X}_1, \dots, \tilde{X}_n)$,

$$\begin{aligned} E_{\mathscr{G}} f_{ij}(X_i^r, X_j^r) &= E_{\mathscr{G}} f_{ij}\left(X_i^r, \tilde{X}_j^r\right) \\ &= \tfrac{1}{4}\Big\{ f_{ij}(X_i, X_j) + f_{ij}\left(X_i, \tilde{X}_j\right) \\ &\qquad + f_{ij}\left(\tilde{X}_i, X_j\right) + f_{ij}\left(\tilde{X}_i, \tilde{X}_j\right)\Big\}. \end{aligned}$$

Then one uses this fact in the proof of Theorem 1 with only formal changes, using the following simple facts about outer expectations:

1. Fubini's inequality: If X_1 and X_2 are independent S_i variables and H: $S_1 \times S_2 \to R$, then $E_1^* E_2^* H(X_1, X_2) \le (E \times E)^* H(X_1, X_2)$, and
2. Jensen's inequality for E^*:

$$\Phi\left(\sup_{\mathbf{f}\in\Pi}\left\|E\sum_{1\le i\ne j\le n} f_{ij}(X_i, X_j)\right\|\right) \le E^*\Phi\left(\sup_{\mathbf{f}\in\Pi}\left\|\sum_{1\le i\ne j\le n} f_{ij}\left(X_i, \tilde{X}_j\right)\right\|\right).$$

(The first use of Jensen's inequality is obvious, and so is the second if one expresses the convex increasing function Φ as the sup of its supporting lines.) The same approach can be applied in the general case treated in Section 3.

REMARK. Even though Theorems 1 and 2 are already very general, it may be possible to extend them (using basically the same approach) by replacing the Banach space D by more general topological vector spaces. One may also want to let Φ be a function of the norm. Along this line of thought, Pisier (1990) has pointed out that our results can be viewed as decoupling results for point processes.

HISTORICAL NOTE. We are thankful to D. L. Burkholder for pointing us to an early decoupling result of D. L. Burkholder and T. R. McConnell [see Lemma 1 of Burkholder (1983)] dealing with problems involving martingale transforms of a Rademacher sequence. This result has been important in the study of integral operators on Lebesgue–Bochner spaces.

Acknowledgments. We are thankful to E. Giné for requesting the solution of a special case of Theorem 1. We are also thankful to M. J. Klass,

J. Winniki, D. Alemayehu and the referee for helpful comments and suggestions, and to J. Cvitanic for his help in obtaining the constants in the multivariate case.

REFERENCES

ARAUJO, A. and GINÉ, E. (1980). *The Central Limit Theorem for Real and Banach Valued Random Variables*. Wiley, New York.

BOURGAIN, J. and TZAFRIRI, L. (1987). Invertibility of "large" submatrices with applications to the geometry of Banach spaces and harmonic analysis. *Israel J. Math.* **57** (3) 137–224.

BURKHOLDER, D. L. (1983). A geometric condition that implies the existence of certain singular integrals of Banach-space-valued functions. In *Conference on Harmonic Analysis in Honor of Antoni Zygmund* (W. Beckner, A. P. Calderón, R. Fefferman and P. W. Jones, eds.). Wadsworth, Belmont, Calif.

DUDLEY, R. (1984). A course on empirical processes. *Ecole d'Eté de Probabilités de Saint-Flour, XII–1982. Lecture Notes in Math.* **1097** 1–142. Springer, New York.

DUDLEY, R. and PHILIPP, W. (1983). Invariance principles for sums of Banach space valued random elements and empirical processes. *Z. Warsch. Verw. Gebiete* **62** 509–552.

DE LA PEÑA, V. H. and KLASS, M. J. (1990). Order of magnitude bounds for expectations involving quadratic forms. Unpublished manuscript.

HITCZENKO, P. (1988). Comparison for moments of tangent sequences of random variables. *Probab. Theory Related Fields* **78** 223–230.

KWAPIEŃ, S. (1987). Decoupling inequalities for polynomial chaos. *Ann. Probab.* **15** 1062–1071.

MCCONNELL, T. R. and TAQQU, M. S. (1986). Decoupling inequalities for multilinear forms in independent symmetric random variables. *Ann. Probab.* **14** 943–954.

MCCONNELL, T. R. and TAQQU, M. S. (1987). Decoupling inequalities for Banach-valued multilinear forms in independent symmetric Banach-valued random variables. *Probab. Theory Related Fields* **75** 499–507.

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

NOLAN, D. and POLLARD, D. (1988). Functional limit theorems for *U*-processes. *Ann. Probab.* **16** 1291–1298.

PISIER, G. (1990). Personal communication.

POLLARD, D. (1991). Unpublished manuscript.

ZINN, J. (1985). Comparison of martingale difference sequences. *Probability in Banach Spaces V. Lecture Notes in Math.* **1153** 453–457. Springer, Berlin.

ZINN, J. (1989). Personal communication.

DEPARTMENT OF STATISTICS
COLUMBIA UNIVERSITY
NEW YORK, NEW YORK 10027