The Annals of Probability
1995, Vol. 23, No. 2, 806–816

DECOUPLING INEQUALITIES FOR THE TAIL PROBABILITIES OF MULTIVARIATE U-STATISTICS

BY VICTOR H. DE LA PEÑA[1] AND S. J. MONTGOMERY-SMITH[2]

Columbia University and University of Missouri, Columbia

In this paper we present a decoupling inequality that shows that multivariate U-statistics can be studied as sums of (conditionally) independent random variables. This result has important implications in several areas of probability and statistics including the study of random graphs and multiple stochastic integration. More precisely, we get the following result: Let $\{X_j\}$ be a sequence of independent random variables on a measurable space $(\mathscr{S}, S)$ and let $\{X_i^{(j)}\}$, $j = 1, \dots, k$, be k independent copies of $\{X_i\}$. Let $f_{i_1 i_2 \cdots i_k}$ be families of functions of k variables taking $(S \times \cdots \times S)$ into a Banach space $(B, \|\cdot\|)$. Then, for all $n \geq k \geq 2$, $t > 0$, there exist numerical constants C_k depending on k only so that

$$P\left(\left\| \sum_{1 \leq i_1 \neq i_2 \neq \cdots \neq i_k \leq n} f_{i_1 \cdots i_k}\left(X_{i_1}^{(1)}, X_{i_2}^{(1)}, \dots, X_{i_k}^{(1)}\right) \right\| \geq t\right)$$

$$\leq C_k P\left(C_k \left\| \sum_{1 \leq i_1 \neq i_2 \neq \cdots \neq i_k \leq n} f_{i_1 \cdots i_k}\left(X_{i_1}^{(1)}, X_{i_2}^{(2)}, \dots, X_{i_k}^{(k)}\right) \right\| \geq t\right).$$

The reverse bound holds if, in addition, the following symmetry condition holds almost surely:

$$f_{i_1 i_2 \cdots i_k}\left(X_{i_1}, X_{i_2}, \dots, X_{i_k}\right) = f_{i_{\pi(1)} i_{\pi(2)} \cdots i_{\pi(k)}}\left(X_{i_{\pi(1)}}, X_{i_{\pi(2)}}, \dots, X_{i_{\pi(k)}}\right),$$

for all permutations π of $(1, \dots, k)$.

1. Introduction. In this paper we provide the multivariate extension of the tail probability decoupling inequality for generalized U-statistics of order 2 and quadratic forms presented in de la Peña and Montgomery-Smith (1993). This type of inequality permits the transfer of some results for sums of independent random variables to the case of U-statistics. Our work builds mainly on recent work of Kwapień and Woyczynski (1992) as well as on results for U-statistics from Giné and Zinn (1992) and papers dealing with inequalities for multilinear forms of symmetric and hypercontractive random variables in de la Peña, Montgomery-Smith and Szulga (1992) and de la Peña (1992). It is to be remarked that the decoupling inequalities for multilinear forms introduced in McConnell and Taqqu (1986) provided us with our first exposure to this decoupling problem. For an expanded list of references on the subject, see, for example, Kwapień and Woyczynski (1992).

Received August 1993; revised August 1994.

[1] Supported in part by NSF grants.

[2] Supported by the University of Missouri Research Board.

AMS 1991 *subject classifications.* Primary 60E15; secondary 60D05.
Key words and phrases. U-statistics, decoupling.

2. Main result.

THEOREM 1. *Let* $\{X_i\}$ *be a sequence of independent random variables on a measurable space* $(\mathscr{S}, S)$ *and let* $\{X_i^{(j)}\}$, $j = 1, \dots, k$, *be* k *independent copies of* $\{X_i\}$. *Let* $f_{i_1 i_2 \cdots i_k}$ *be families of functions of* k *variables taking* $(S \times \cdots \times S)$ *into a Banach space* $(B, \|\cdot\|)$. *Then, for all* $n \geq k \geq 2$, $t > 0$, *there exist numerical constants* C_k, $\tilde{C}_k$ *depending on* k *only so that*

$$P\left(\left\| \sum_{1 \leq i_1 \neq i_2 \neq \cdots \neq i_k \leq n} f_{i_1 \cdots i_k}\left(X_{i_1}^{(1)}, X_{i_2}^{(1)}, \dots, X_{i_k}^{(1)}\right) \right\| \geq t\right)$$

$$\leq C_k P\left(C_k \left\| \sum_{1 \leq i_1 \neq i_2 \neq \cdots \neq i_k \leq n} f_{i_1 \cdots i_k}\left(X_{i_1}^{(1)}, X_{i_2}^{(2)}, \dots, X_{i_k}^{(k)}\right) \right\| \geq t\right).$$

If, in addition, the following symmetry condition holds almost surely,

$$f_{i_1 i_2 \cdots i_k}\left(X_{i_1}, X_{i_2}, \dots, X_{i_k}\right) = f_{i_{\pi(1)} i_{\pi(2)} \cdots i_{\pi(k)}}\left(X_{i_{\pi(1)}}, X_{i_{\pi(2)}}, \dots, X_{i_{\pi(k)}}\right),$$

for all permutations π *of* $(1, \dots, k)$, *then*

$$P\left(\left\| \sum_{1 \leq i_1 \neq i_2 \neq \cdots \neq i_k \neq n} f_{i_1 \cdots i_k}\left(X_{i_1}^{(1)}, X_{i_2}^{(2)}, \dots, X_{i_k}^{(k)}\right) \right\| \geq t\right)$$

$$\leq \tilde{C}_k P\left(\tilde{C}_k \left\| \sum_{1 \leq i_1 \neq i_2 \neq \cdots \neq i_k \leq n} f_{i_1 \cdots i_k}\left(X_{i_1}^{(1)}, X_{i_2}^{(1)}, \dots, X_{i_k}^{(1)}\right) \right\| \geq t\right).$$

NOTE. In this paper we use the notation $\{i_1 \neq i_2 \neq \cdots \neq i_k\}$ to denote that all of $i_1, \dots, i_k$ are different.

3. Preliminary results. Throughout this paper we will be using two results found in earlier work. The first one comes from de la Peña and Montgomery-Smith (1993). For completeness we reproduce the proof here.

LEMMA 1. *Let* X, Y *be two i.i.d. random variables. Then*

$$P(\|X\| \geq t) \leq 3P\left(\|X + Y\| \geq \frac{2t}{3}\right). \tag{1}$$

PROOF. Let X, Y, Z be i.i.d. random variables. Then

$$\begin{aligned} &P(\|X\| \geq t) \\ &\quad = P(\|(X + Y) + (X + Z) - (Y + Z)\| \geq 2t) \\ &\quad \leq P(\|X + Y\| \geq 2t/3) + P(\|X + Z\| \geq 2t/3) + P(\|Y + Z\| \geq 2t/3) \\ &\quad = 3P(\|X + Y\| \geq 2t/3). \end{aligned}$$

□

The second result comes from Kwapień and Woyczynski (1992) and can also be found in de la Peña and Montgomery-Smith (1993).

PROPOSITION 1. *Let Y be any mean zero random variable with values in a Banach space $(B, \|\cdot\|)$. Then, for all $a \in B$,*

$$P(\|a + Y\| \geq \|a\|) \geq \frac{\kappa}{4}, \tag{2}$$

where $\kappa = \inf_{x' \in B'}((E|x'(Y)|)^2/E(x'(Y))^2)$. (Here B' denotes the family of linear functionals on B.)

PROOF. Note first that if ξ is a random variable for which $E\xi = 0$, then $P(\xi \geq 0) \geq 1/4((E|\xi|)^2/E(\xi^2))$. From this we deduce that $P(x'(Y) \geq 0) \geq 1/4((E|x'(Y)|)^2/E(x'(Y))^2)$. The result then follows, because if $x' \in B'$ is such that $\|x'\| = 1$ and $x'(a) = \|a\|$, then $\{\|a + Y\| \geq \|a\|\}$ contains $\{x'(a + Y) \geq x'(a)\} = \{x'(Y) \geq 0\}$. □

LEMMA 2. *Let $x, a_{i_1}, a_{i_1 i_2}, \dots, a_{i_1 i_2 \cdots i_k}$ belong to a Banach space $(B, \|\cdot\|)$. Let $\{\varepsilon_i\}$ be a sequence of symmetric Bernoulli random variables. Then*

$$P\left(\left\|x + \sum_{r=1} \sum_{1 \leq i_1 \neq i_2 \neq \cdots \neq i_r \leq n} a_{i_1 \cdots i_r} \varepsilon_{i_1} \cdots \varepsilon_{i_r}\right\| \geq \|x\|\right) \geq c_k^{-1}$$

for a universal constant $1 < c_k < \infty$ depending on k only.

PROOF. Suppose that $x, a_{i_1}, a_{i_1 i_2}, \dots, a_{i_1 i_2 \cdots i_k}$ are in R. Then since the ε's are hypercontractive, by (1.4) of Kwapień and Szulga (1991) and the easy argument of the proof of Lemma 3 in de la Peña and Montgomery-Smith (1993), for some $\sigma > 0$, we get

$$\begin{aligned}
&\left(E\left|\sum_{r=1}^{k} \sum_{1 \leq i_1 \neq \cdots \neq i_r \neq n} a_{i_1 \cdots i_r} \varepsilon_{i_1} \cdots \varepsilon_{i_r}\right|^4\right)^{1/4} \\
&\quad = \left(E\left|\sum_{r=1}^{k} \sum_{1 \leq i_1 < \cdots < i_r \leq n} b_{i_1 \cdots i_r} \varepsilon_{i_1 r} \varepsilon_{i_r}\right|^4\right)^{1/4} \\
&\quad \leq \sigma^{-k} \left(E\left|\sum_{r=1}^{k} \sum_{1 \leq i_1 < \cdots < i_k \leq n} b_{i_1 \cdots i_r} \varepsilon_{i_1} \cdots \varepsilon_{i_r}\right|^2\right)^{1/2} \\
&\quad = \sigma^{-k} \left(E\left|\sum_{r=1}^{k} \sum_{1 \leq i_1 \neq \cdots \neq i_k \leq n} a_{i_1 \cdots i_r} \varepsilon_{i_1} \cdots \varepsilon_{i_r}\right|^2\right)^{1/2},
\end{aligned}$$

where $b_{i_1 \cdots i_r} = \sum_{\pi \in S_r} a_{i_{\pi(1)} \cdots i_{\pi(r)}}$ and S_r denotes the set of all permutations of $\{1, \dots, r\}$.

Next, observe that $\|\xi\|_4 \leq \sigma^{-2}\|\xi\|_2$ implies that $\|\xi\|_2 \leq \sigma^{-4}\|\xi\|_1$. Take $x' \in B'$ so that $\|x'\| = 1$ and $x'(x) = \|x\|$. Then

$$\begin{aligned}
&P\left(\left\| x + \sum_{r=1}^{k} \sum_{1 \leq i_1 \neq i_2 \neq \cdots \neq i_r \leq n} a_{i_1 \cdots i_r} \varepsilon_{i_1} \cdots \varepsilon_{i_r} \right\| \geq \|x\| \right) \\
&\quad \geq P\left(x'(x) + \sum_{r=1}^{k} \sum_{1 \leq i_1 \neq i_2 \neq \cdots \neq i_r \leq n} x'(a_{i_1 \cdots i_r}) \varepsilon_{i_1} \cdots \varepsilon_{i_r} \geq x'(x) \right) \\
&\quad = P\left(\sum_{r=1}^{k} \sum_{1 \leq i_1 \neq i_2 \neq \cdots \neq i_r \leq n} x'(a_{i_1 \cdots i_r}) \varepsilon_{i_1} \cdots \varepsilon_{i_r} \geq 0 \right) \geq c_k^{-1}. \qquad \square
\end{aligned}$$

NOTE. Throughout this paper we will use c_k and C_k to denote numerical constants that depend on k only and may change from application to application.

4. Proof of the upper bound.

Our proof of this result is obtained by applying the argument used in the proof of the upper bound in the bivariate case plus an inductive argument. Let $\{\sigma_i\}$ be a sequence of independent symmetric Bernoulli random variables; that is, $P(\sigma_i = 1) = \frac{1}{2}$ and $P(\sigma_i = -1) = \frac{1}{2}$. Consider random variables $(Z_i^{(1)}, Z_i^{(2)})$ such that $(Z_i^{(1)}, Z_i^{(2)}) = (X_i^{(1)}, X_i^{(2)})$ if $\sigma_i = 1$ and $(Z_i^{(1)}, Z_i^{(2)}) = (X_i^{(2)}, X_i^{(1)})$ if $\sigma_i = -1$. Then $(1 + \sigma_i)$ and $(1 - \sigma_i)$ are either 0 or 2 and these random variables can be used to transform the problem from one involving X's to one involving Z's. Let us first illustrate the argument in the case that $k = 3$:

$$\begin{aligned}
&2^3 f_{i_1 i_2 i_3}\left(Z_{i_1}^{(1)}, Z_{i_2}^{(1)}, Z_{i_3}^{(2)}\right) \\
&\quad = \Big\{(1 + \sigma_{i_1})(1 + \sigma_{i_2})(1 + \sigma_{i_3}) f_{i_1 i_2 i_3}\left(X_{i_1}^{(1)}, X_{i_2}^{(1)}, X_{i_3}^{(2)}\right) \\
&\qquad + (1 + \sigma_{i_1})(1 + \sigma_{i_2})(1 - \sigma_{i_3}) f_{i_1 i_2 i_3}\left(X_{i_1}^{(1)}, X_{i_2}^{(1)}, X_{i_3}^{(1)}\right) \\
&\qquad + (1 + \sigma_{i_1})(1 - \sigma_{i_2})(1 + \sigma_{i_3}) f_{i_1 i_2 i_3}\left(X_{i_1}^{(1)}, X_{i_2}^{(2)}, X_{i_3}^{(2)}\right) \\
&\qquad + (1 - \sigma_{i_1})(1 + \sigma_{i_2})(1 + \sigma_{i_3}) f_{i_1 i_2 i_3}\left(X_{i_1}^{(2)}, X_{i_2}^{(1)}, X_{i_3}^{(2)}\right) \\
&\qquad + (1 + \sigma_{i_1})(1 - \sigma_{i_2})(1 - \sigma_{i_3}) f_{i_1 i_2 i_3}\left(X_{i_1}^{(1)}, X_{i_2}^{(2)}, X_{i_3}^{(1)}\right) \\
&\qquad + (1 - \sigma_{i_1})(1 + \sigma_{i_2})(1 - \sigma_{i_3}) f_{i_1 i_2 i_3}\left(X_{i_1}^{(2)}, X_{i_2}^{(1)}, X_{i_3}^{(1)}\right) \\
&\qquad + (1 - \sigma_{i_1})(1 - \sigma_{i_2})(1 + \sigma_{i_3}) f_{i_1 i_2 i_3}\left(X_{i_1}^{(2)}, X_{i_2}^{(2)}, X_{i_3}^{(2)}\right) \\
&\qquad + (1 - \sigma_{i_1})(1 - \sigma_{i_2})(1 - \sigma_{i_3}) f_{i_1 i_2 i_3}\left(X_{i_1}^{(2)}, X_{i_2}^{(2)}, X_{i_3}^{(1)}\right)\Big\},
\end{aligned} \tag{3}$$

where the "+" sign is chosen if the superscript of X_i agrees with that of Z_i and "−" otherwise. Next, set

$$\begin{aligned}T_{n,3} = \sum_{1\le i_1\ne i_2\ne i_3\le n} \Big\{ & f_{i_1i_2i_3}\big(X_{i_1}^{(1)}, X_{i_2}^{(1)}, X_{i_3}^{(2)}\big) + f_{i_1i_2i_3}\big(X_{i_1}^{(1)}, X_{i_2}^{(1)}, X_{i_3}^{(1)}\big) \\ & + f_{i_1i_2i_3}\big(X_{i_1}^{(1)}, X_{i_2}^{(2)}, X_{i_3}^{(2)}\big) + f_{i_1i_2i_3}\big(X_{i_1}^{(2)}, X_{i_2}^{(1)}, X_{i_3}^{(2)}\big) \\ & + f_{i_1i_2i_3}\big(X_{i_1}^{(1)}, X_{i_2}^{(2)}, X_{i_3}^{(1)}\big) + f_{i_1i_2i_3}\big(X_{i_1}^{(2)}, X_{i_2}^{(1)}, X_{i_3}^{(1)}\big) \\ & + f_{i_1i_2i_3}\big(X_{i_1}^{(2)}, X_{i_2}^{(2)}, X_{i_3}^{(2)}\big) + f_{i_1i_2i_3}\big(X_{i_1}^{(2)}, X_{i_2}^{(2)}, X_{i_3}^{(1)}\big)\Big\}.\end{aligned}$$

Letting $\mathscr{G}_2 = \sigma(X_i^{(1)}, X_i^{(2)}, i = 1,\dots,n)$ we get

$$T_{n,3} = 2^3 \sum_{1\le i_1\ne i_2\ne i_3\le n} E\Big(f_{i_1i_2i_3}\big(Z_{i_1}^{(1)}, Z_{i_2}^{(1)}, Z_{i_3}^{(2)}\big) \mid \mathscr{G}_2\Big).$$

More generally, for any $1 \le l_1,\dots,l_k \le 2$, one can obtain the expansion

$$\begin{aligned} & 2^k f_{i_1\cdots i_k}\big(Z_{i_1}^{(l_1)},\dots,Z_{i_k}^{(l_k)}\big) \\ & \quad = \sum_{1\le j_1,\dots,j_k\le 2} (1 \pm \sigma_{i_1})\cdots(1 \pm \sigma_{i_k}) f_{i_1\cdots i_k}\big(X_{i_1}^{(j_1)},\dots,X_{i_k}^{(j_k)}\big).\end{aligned} \tag{4}$$

The appropriate extension of $T_{n,3}$ is

$$T_{n,k} = \sum_{1\le i_1\ne \cdots \ne i_k\le n} \ \sum_{1\le j_1,\dots,j_k\le 2} f_{i_1\cdots i_k}\big(X_{i_1}^{(j_1)},\dots,X_{i_k}^{(j_k)}\big).$$

Again,

$$T_{n,k} = 2^k \sum_{1\le i_1\ne \cdots \ne i_k\le n} E\Big(f_{i_1\cdots i_k}\big(Z_{i_1}^{(l_1)},\dots,Z_{i_k}^{(l_k)}\big) \mid \mathscr{G}_2\Big).$$

From Lemma 1 we get

$$\begin{aligned} & P\left(\left\| \sum_{1\le i_1\ne i_2\ne \cdots \ne i_k\le n} f_{i_1\cdots i_k}\big(X_{i_1}^{(1)},\dots,X_{i_k}^{(1)}\big)\right\| \ge t\right) \\ & \quad \le 3P\left(3\left\| \sum_{1\le i_1\ne \cdots \ne i_k\le n} \Big\{f_{i_1\cdots i_k}\big(X_{i_1}^{(1)},\dots,X_{i_k}^{(1)}\big) \right.\right. \\ & \qquad\qquad \left.\left. + f_{i_1\cdots i_k}\big(X_{i_1}^{(2)},\dots,X_{i_k}^{(2)}\big)\Big\}\right\| \ge 2t\right)\end{aligned}$$

$$= 3P\left(3\left\|T_{n,k} + \sum_{1\le i_1\ne\cdots\ne i_k\le n} f_{i_1\cdots i_k}\left(X_{i_1}^{(1)},\dots,X_{i_k}^{(1)}\right) + f_{i_1\cdots i_k}\left(X_{i_1}^{(2)},\dots,X_{i_k}^{(2)}\right) - T_{n,k}\right\| \ge 2t\right)$$

$$\le \left\{3P(3\|T_{n,k}\| \ge t) + 3P\left(3\left\|\sum_{1\le i_1\ne\cdots\ne i_k\le n}\ \sum_{\substack{1\le j_1,\dots,j_k\le 2\\ \text{not all } j\text{'s equal}}} f_{i_1\cdots i_k}\left(X_{i_1}^{(j_1)},\dots,X_{i_k}^{(j_k)}\right)\right\| \ge t\right)\right\} \tag{5}$$

$$\le \left\{3P(3\|T_{n,k}\| \ge t) + \sum_{\substack{1\le j_1,\dots,j_k\le 2\\ \text{not all } j\text{'s equal}}} C_k P\left(C_k\left\|\sum_{1\le i_1\ne\cdots\ne i_k\le n} f_{i_1\cdots i_k}\left(X_{i_1}^{(j_1)},\dots,X_{i_k}^{(j_k)}\right)\right\| \ge t\right)\right\}.$$

(Recall that C_k, c_k are numerical constants that depend on k only and may change from application to application.)

Observe also that using (4) and the fact that the σ's are independent of the X's, Lemma 2 with $x = T_{n,k}$ gives for any fixed $1 \le l_1,\dots,l_k \le 2$,

$$P\left(2^k\left\|\sum_{1\le i_1\ne\cdots\ne i_k\le n} f_{i_1\cdots i_k}\left(Z_{i_1}^{(l_1)},\dots,Z_{i_k}^{(l_k)}\right)\right\| \ge \|T_{n,k}\| \mid \mathscr{G}_2\right) \ge c_k^{-1}. \tag{6}$$

Integrating over $(\|T_{n,k}\| \ge t\}$ and using the fact that $\{(X_i^{(1)}, X_i^{(2)})\colon i = 1,\dots,n\}$ has the same joint distribution as $\{(Z_i^{(1)}, Z_i^{(2)})\colon i = 1,\dots,n\}$ we obtain that

$$\begin{aligned} &P\left(2^k\left\|\sum_{1\le i_1\ne\cdots\ne i_k\le n} f_{i_1\cdots i_k}\left(X_{i_1}^{(l_1)},\dots,X_{i_k}^{(l_k)}\right)\right\| \ge t\right)\\ &\quad= P\left(2^k\left\|\sum_{1\le i_1\ne\cdots\ne i_k\le n} f_{i_1\cdots i_k}\left(Z_{i_1}^{(l_1)},\dots,Z_{i_k}^{(l_k)}\right)\right\| \ge t\right)\\ &\quad\ge c_k^{-1}P\left(\|T_{n,k}\| \ge t\right). \end{aligned} \tag{7}$$

It is obvious that the upper bound decoupling inequality holds for the case of U-statistics of order 1. Assume that it holds for U-statistics of orders $2,\dots,k-1$. Putting (5) and (7) together with $1 \le l_1,\dots,l_k \le 2$, not all l's equal, we get

$$P\left(\left\|\sum_{1\le i_1\ne\cdots\ne i_k\le n} f_{i_1\cdots i_k}\left(X_{i_1}^{(1)},\ldots,X_{i_k}^{(1)}\right)\right\|\ge t\right)$$

$$\le \left\{3P(3\|T_{n,k}\|\ge t)\right.$$

$$\left.+\sum_{\substack{1\le j_1,\ldots,j_k\le 2\\ \text{not all } j\text{'s equal}}} C_k P\left(C_k\left\|\sum_{1\le i_1\ne\cdots\ne i_k\le n} f_{i_1\cdots i_k}\left(X_{i_1}^{(j_1)},\ldots,X_{i_k}^{(j_k)}\right)\right\|\ge t\right)\right\}$$

$$\le \sum_{\substack{1\le j_1,\ldots,j_k\le 2\\ \text{not all } j\text{'s equal}}} C_k P\left(C_k\left\|\sum_{1\le i_1\ne\cdots\ne i_k\le n} f_{i_1\cdots i_k}\left(X_{i_1}^{(j_1)},\ldots,X_{i_k}^{(j_k)}\right)\right\|\ge t\right)$$

$$\le C_k P\left(C_k\left\|\sum_{1\le i_1\ne\cdots\ne i_k\le n} f_{i_1\cdots i_k}\left(X_{i_1}^{(1)},\ldots,X_{i_k}^{(k)}\right)\right\|\ge t\right),$$

where again, the last line follows by the decoupling result for U-statistics of orders $2,\ldots,k-1$ of the inductive hypothesis. Since the statement "not all j's equal" means that there are less than k j's which equal, the variables whose j's are equal can be decoupled using (conditionally on the other variables) the decoupling inequalities for U-statistics of order $2,\ldots,k-1$. □

Next we give the proof of the lower bound.

5. Proof of the lower bound.

In order to show the lower bound we require the following result.

LEMMA 3. *Let $1\le l\le k$. Then there is a constant C_k such that*

$$P\left(\left\|\sum_{1\le i_1\ne i_2\ne\cdots\ne i_k\le n} f_{i_1\cdots i_k}\left(X_{i_1}^{(1)},X_{i_2}^{(1)},\ldots,X_{i_k}^{(1)}\right)\right\|\ge t\right)$$

$$\ge C_k^{-1}P\left(\left\|\sum_{1\le i_1\ne i_2\ne\cdots\ne i_k\le n}\ \sum_{1\le j_1,\ldots,j_k\le l} f_{i_1\cdots i_k}\left(X_{i_1}^{(j_1)},X_{i_2}^{(j_2)},\ldots,X_{i_k}^{(j_k)}\right)\right\|\ge C_k t\right).$$

PROOF. Let $\{\delta_r\}$, $r=1,\ldots,l$, be a sequence of random variables for which $P(\delta_r=1)=1/l$ and $P(\delta_r=0)=1-1/l$, and $\sum_{r=1}^{l}\delta_r=1$. Set $\varepsilon_r=\delta_r-1/l$

for $r = 1, \dots, l$. Then it is easy to see that there exists $\sigma_l > 0$ depending only upon l such that for any real number x_0 and any sequence of real constants $\{a_i\}$,

$$\left\| x_0 + \sum_{r=1}^{l} a_r \varepsilon_r \right\|_4 \le \left\| x_0 + \sigma_l^{-1} \sum_{r=1}^{l} a_r \varepsilon_r \right\|_2 . \tag{8}$$

One can also use the results of Section 6.9 of Kwapień and Woyczynski [(1992), pages 180 and 181] to assert this since the ε's satisfy Conditions 1 through 3 stated there. □

Let $\{(\delta_{i1}, \dots, \delta_{il}),\ i = 1, \dots, n\}$ be n independent copies of $(\delta_1, \dots, \delta_l)$. As before, we define

$$\varepsilon_{ij} = \delta_{ij} - \frac{1}{l}. \tag{9}$$

Since the vectors $\mathscr{E}_i = (\varepsilon_{i1}, \dots, \varepsilon_{il})$ are independent, by an argument given in Kwapień and Szulga (1991), for $i = 1, \dots, n$, for all constants x_0, a_{ij} in R,

$$\begin{aligned} \left\| x_0 + \sum_{i=1}^{n} \sum_{r=1}^{l} a_{ir} \varepsilon_{ir} \right\|_4 &\le \left\| x_0 + \sigma_l^{-1} \sum_{i=1}^{n} \sum_{r=1}^{l} a_{ir} \varepsilon_{ir} \right\|_2 \\ &\le \sigma_l^{-1} \left\| x_0 + \sum_{i=1}^{n} \sum_{r=1}^{l} a_{ir} \varepsilon_{ir} \right\|_2 , \end{aligned} \tag{10}$$

and recentering, we obtain

$$\left\| x_0 + \sum_{i=1}^{n} \sum_{r=1}^{l} a_{ir} \delta_{ir} \right\|_4 \le \sigma_l^{-1} \left\| x_0 + \sum_{i=1}^{n} \sum_{r=1}^{l} a_{ir} \delta_{ir} \right\|_2 . \tag{11}$$

Next we use the sequence $\mathscr{E}_i$, $i = 1, \dots, n$, in defining the analogue of the Z's used in our proof of the upper bound.

For each i, let $Z_i = X_i^{(j)}$ if $\delta_{ij} = 1$. Then $\{Z_i,\ i = 1, \dots, n\}$ has the same joint distribution as $\{X_i^{(1)},\ i = 1, \dots, n\}$ and

$$f_{i_1 \cdots i_k}(Z_{i_1}, \dots, Z_{i_k}) = \sum_{1 \le j_1, j_2, \dots, j_k \le l} \delta_{i_1 j_1} \cdots \delta_{i_k j_k} f_{i_1 \cdots i_k}\big(X_{i_1}^{(j_1)}, \dots, X_{i_k}^{(j_k)}\big).$$

The fact that $E\delta_{i_r j_r} = 1/l$ for all i_r, j_r gives

$$E\big(f_{i_1 \cdots i_k}(Z_{i_1}, \dots, Z_{i_k}) \mid \mathscr{G}_l\big) = \left(\frac{1}{l}\right)^k \sum_{1 \le j_1, \dots, j_n \le l} f_{i_1 \cdots i_k}\big(X_{i_1}^{(j_1)}, \dots, X_{i_k}^{(j_k)}\big),$$

where $\mathscr{G}_l = \sigma((X_i^{(1)}, \dots, X_i^{(l)}),\ i = 1, \dots, n)$.

Let

$$\begin{aligned} U_n &= \sum_{1 \le i_1 \ne i_2 \ne \cdots \ne i_k \le n} f_{i_1 i_2 \cdots i_k}(Z_{i_1}, \dots, Z_{i_k}) \\ &= \sum_{1 \le i_1 \ne i_2 \ne \cdots \ne i_k \le n} \ \sum_{1 \le j_1, \dots, j_k \le l} \delta_{i_1 j_1} \cdots \delta_{i_k j_k} f_{i_1 \cdots i_k}\big(X_{i_1}^{(j_1)}, \dots, X_{i_k}^{(j_k)}\big). \end{aligned}$$

Let $\mathscr{D}_i = (\delta_{i1}, \dots, \delta_{il})$. Since the $\mathscr{D}$'s are independent of the X's, if we let

$$\begin{aligned} &g_{i_1 \cdots i_k}\left(\mathscr{D}_{i_1}, \dots, \mathscr{D}_{i_k}\right) \\ &\quad = \sum_{1 \le j_1, \dots, j_k \le l} \delta_{i_1 j_1} \cdots \delta_{i_k j_k} f_{i_1 \cdots i_k}\left(X_{i_1}^{(j_1)}, \dots, X_{i_k}^{(j_k)}\right). \end{aligned}$$

then, since

$$f_{i_1 \cdots i_k}\left(X_{i_1}, \dots, X_{i_k}\right) = f_{i_{(\pi(1))} \cdots i_{(\pi(k))}}\left(X_{i_{\pi(1)}}, \dots, X_{i_{\pi(k)}}\right),$$

we have that

$$g_{i_1 \cdots i_k}\left(\mathscr{D}_{i_1}, \dots, \mathscr{D}_{i_k}\right) = g_{i_{(\pi(1))} \cdots i_{(\pi(k))}}\left(\mathscr{D}_{i_{\pi(1)}}, \dots, \mathscr{D}_{i_{\pi(k)}}\right).$$

Therefore, the two-sided decoupling inequality in de la Peña (1992) can be applied and for every convex increasing function Φ, every $\mathscr{G}_l$-measurable function T and every set of k independent copies $\mathscr{D}_i^{(r)}$, $r = 1, \dots, k$, of $\mathscr{D}_i$ there exist numerical constants A_k, B_k so that

$$\begin{aligned} &E\left(\Phi\left(A_k \left\| T + \sum_{1 \le i_1 \ne i_2 \ne \cdots \ne i_k \le n} g_{i_1 \cdots i_k}\left(\mathscr{D}_{i_1}, \dots, \mathscr{D}_{i_k}\right)\right\|\right) \mid \mathscr{G}_l\right) \\ &\quad \le E\left(\Phi\left(\left\| T + \sum_{1 \le i_1 \ne i_2 \ne \cdots \ne i_k \le n} g_{i_1 \cdots i_k}\left(\mathscr{D}_{i_1}^{(1)}, \dots, \mathscr{D}_{i_k}^{(k)}\right)\right\|\right) \mid \mathscr{G}_l\right) \\ &\quad \le E\left(\Phi\left(B_k \left\| T + \sum_{1 \le i_1 \ne i_2 \ne \cdots \ne i_k \le n} g_{i_1 \cdots i_k}\left(\mathscr{D}_{i_1}, \dots, \mathscr{D}_{i_k}\right)\right\|\right) \mid \mathscr{G}_l\right). \end{aligned}$$

This result with (11) shows that, conditionally on $\mathscr{G}_l$,

$$\|U_n - T_n\|_4 \le \sigma_l^{-k} \frac{B_k}{A_k} \|U_n - T_n\|_2, \tag{12}$$

where

$$\begin{aligned} T_n &= E(U_n \mid \mathscr{G}_l) \\ &= \left(\frac{1}{l}\right)^k \sum_{1 \le i_1 \ne i_2 \ne \cdots \ne i_k \le n} \sum_{1 \le j_1, \dots, j_k \le l} f_{i_1 \cdots i_k}\left(X_{i_1}^{(j_1)}, X_{i_2}^{(j_2)}, \dots, X_{i_k}^{(j_k)}\right). \end{aligned}$$

[See also the proofs of Lemma 2 and Lemma 6.5.1 of Kwapień and Woyczynski (1992).]

Thus we have that

$$P(\|U_n\| \ge \|T_n\| \mid \mathscr{G}_l) \ge c_k^{-1}. \tag{13}$$

This follows from the use of (12) and Proposition 1 with $a = T_n$ and $Y = U_n - T_n$. We also use the fact that for any random variable ξ and positive constant c, $\|\xi\|_4 \le c\|\xi\|_2$ implies that $\|\xi\|^2 \le c^2\|\xi\|_1$ (see also the proof of Lemma 2 for the approach to transfer the problem from one on Banach space-valued random variables to one on real-valued).

Integrating (13) over the set $\{\|T_n\| \geq t\}$, we get

$$P\left(\left\| \sum_{1 \leq i_1 \neq i_2 \neq \cdots \neq i_k \leq n} f_{i_1 \cdots i_k}\left(X_{i_1}^{(1)}, X_{i_2}^{(1)}, X_{i_k}^{(1)}\right) \right\| \geq t\right)$$

$$= P\left(\left\| \sum_{1 \leq i_1 \neq i_2 \neq \cdots \neq i_k \leq n} f_{i_1 \cdots i_k}\left(Z_{i_1}, Z_{i_2}, \dots, Z_{i_k}\right) \right\| \geq t\right)$$

$$\geq c_k^{-1} P\left(C_k \left\| \sum_{1 \leq i_1 \neq i_2 \neq \cdots \neq i_k \leq n} \sum_{1 \leq j_1, \dots, j_k \leq l} f_{i_1 \cdots i_k}\left(X_{i_1}^{(j_1)}, X_{i_2}^{(j_2)}, \dots, X_{i_k}^{(j_k)}\right) \right\| \geq t\right)$$

and Lemma 3 is proved. □

The end of the proof of the lower bound follows by using induction and the iterative procedure introduced to obtain the proof of the lower bound multivariate decoupling inequality in de la Peña (1992). We give a different expression of the same proof, motivated by ideas from de la Peña, Montgomery-Smith and Szulga (1992). We will use S_k to denote the set of permutations of $\{1, \dots, k\}$.

The Mazur–Orlicz formula tells us that for any $1 \leq j_1, \dots, j_k \leq k$,

$$\sum_{0 \leq \delta_1, \dots, \delta_k \leq 1} (-1)^{k - \delta_1 - \cdots - \delta_k} \delta_{j_1} \cdots \delta_{j_k}$$

is 0 unless $j_1, \dots, j_k$ is a permutation of $1, \dots, k$, in which case it is 1. Hence

$$\sum_{\pi \in S_k} f_{i_1 \cdots i_k}\left(X_{i_1}^{(\pi(1))}, \dots, X_{i_k}^{(\pi(k))}\right)$$

$$= \sum_{0 \leq \delta_1, \dots, \delta_k \leq 1} (-1)^{k - \delta_1 - \cdots - \delta_k} \sum_{1 \leq j_1, \dots, j_k \leq k} \delta_{j_1} \cdots \delta_{j_k} f_{i_1 \cdots i_k}\left(X_{i_1}^{(j_1)}, \dots, X_{i_k}^{(j_k)}\right).$$

By the symmetry properties of f,

$$\sum_{1 \leq i_1 \neq \cdots \neq i_k \leq n} f_{i_1 \cdots i_k}\left(X_{i_1}^{(1)}, \dots, X_{i_k}^{(k)}\right)$$

$$= \frac{1}{k!} \sum_{1 \leq i_1 \neq \cdots \neq i_k \leq n} \sum_{0 \leq \delta_1, \dots, \delta_k \leq 1} (-1)^{k - \delta_1 - \cdots - \delta_k}$$

$$\times \sum_{1 \leq j_1, \dots, j_k \leq k} \delta_{j_1} \cdots \delta_{j_k} f_{i_1 \cdots i_k}\left(X_{i_1}^{(j_1)}, \dots, X_{i_k}^{(j_k)}\right).$$

Therefore,

$$
\begin{aligned}
&\Pr\left(\left\| \sum_{1 \le i_1 \ne \cdots \ne i_k \le n} f_{i_1 \cdots i_k}\left(X_{i_1}^{(1)}, \ldots, X_{i_k}^{(k)}\right) \right\| \ge t\right) \\
&\quad \le \sum_{0 \le \delta_1, \ldots, \delta_k \le 1} \Pr\left(\left\| \sum_{1 \le i_1 \ne \cdots \ne i_k \le n} \sum_{1 \le j_1, \ldots, j_k \le k} \right.\right. \\
&\qquad\qquad \left.\left. \delta_{j_1} \cdots \delta_{j_k} f_{i_1 \cdots i_k}\left(X_{i_1}^{(j_1)}, \ldots, X_{i_k}^{(j_k)}\right) \right\| \ge k!t/2^k\right) \\
&\quad = \sum_{l=1}^{k} \binom{k}{l} \Pr\left(\left\| \sum_{1 \le i_1 \ne \cdots \ne i_k \le n} \sum_{1 \le j_1, \ldots, j_k \le l} f_{i_1 \cdots i_k}\left(X_{i_1}^{(j_1)}, \ldots, X_{i_k}^{(j_k)}\right) \right\| \right. \\
&\qquad\qquad \left. \ge k!t/2^k\right),
\end{aligned}
$$

and this combined with Lemma 3 is sufficient to show the result. □

REFERENCES

DE LA PEÑA, V. H. (1992). Decoupling and Khintchine's inequalities for *U*-statistics. *Ann. Probab.* **20** 1877–1892.

DE LA PEÑA, V. H., MONTGOMERY-SMITH, S. J. and SZULGA, J. (1992). Contraction and decoupling inequalities for multilinear forms and *U*-statistics. Preprint.

DE LA PEÑA, V. H. and MONTGOMERY-SMITH, S. J. (1993). Bounds on the tail probability of *U*-statistics and quadratic forms. Preprint.

GINÉ, E. and ZINN, J. (1992). A remark on convergence in distribution of *U*-statistics. Preprint.

KWAPIEŃ, S. and SZULGA, J. (1991). Hypercontraction methods in moment inequalities for series of independent random variables in normed spaces. *Ann. Probab.* **19** 369–379.

KWAPIEŃ, S. and WOYCZYNSKI, W. (1992). *Random Series and Stochastic Integrals: Single and Multiple*. Birkhäuser, Boston.

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

DEPARTMENT OF STATISTICS
COLUMBIA UNIVERSITY
NEW YORK, NEW YORK

DEPARTMENT OF MATHEMATICS
MATHEMATICAL SCIENCES BUILDING
UNIVERSITY OF MISSOURI
COLUMBIA, MISSOURI 65211