Chin. Ann. of Math.
19B: 2(1998), 157-166.

BERNSTEIN TYPE INEQUALITIES FOR DEGENERATE *U*-STATISTICS WITH APPLICATIONS**

WANG QIYING*

Abstract

The author discusses Bernstein type inequalities for degenerate U-statistics. As applications of these results, Cramer type large deviations for studentized U-statistics are obtained under mild conditions.

Keywords U-statistics, Large deviation, Bernstein type inequality, Studentized U-statistics
1991 MR Subject Classification 60F15, 62E20
Chinese Library Classification O211.4, O212.2

§1. Introduction and Main Results

Let $X_1, X_2, \cdots, X_n, \cdots$ be independent and identically distributed (i.i.d.) random variables with common distribution function P. Let $h : R^m \to R$ be measurable and symmetric in its arguments. The U-statistics corresponding to the kernel h is defined by

$$U_n^m(h) = \frac{(n-m)!}{n!} \sum_{(i_1,\cdots,i_m)\in I_m^n} h(X_{i_1}, \cdots, X_{i_m}), \tag{1.1}$$

where $I_m^n = \{(i_1, \cdots, i_m) : i_j \in N, 1 \le i_j \le n, i_j \ne i_k, \text{ if } j \ne k\}$.

The kernel $h(\cdot)$ is called r-degenerate for P, $0 \le r \le m-1$, if

$$\max\{l \ge 0; E(h(X_1, \cdots, X_m)|G_l) = 0 \text{ a.s.}\} = r, \tag{1.2}$$

where G_l is the field generated by $X_1, \cdots, X_l$; G_0 is the trivial field. For $r = 0$ we deal with so-called centered non-degenerate kernel; $m-1$ degenerate kernel is called also P-canonical (see [1]).

Exponential bounds for $P(U_n^m(h) \ge x)$ are studied by many authors under different set of assumations. A far more comeplete list of papers containing results of this kind includes Hoeffding[10], Borisov[4], Christofides[8], Arcones and Gine[1], Miao and Zhang[12], Cheng[7], as well as Arcones[16].

In the previous results, for degenerate U-statistics Arcones and Gine[1] got:

Manuscript received September 19, 1995.
*Institute of Mathematics, Nanjing University, Nanjing 210093, China.
**Project supported by the National Natural Science Foundation of China and the Natural Science Foundation of Jiangsu Province.

Theorem A. *Let $h(\cdot)$ be P-canonical with $|h| \le A$. Then, there exist constants C_1, C_2 depending only on m such that for all $n \ge m$,*

$$P\left(\binom{n}{m}^{1/2} |U_n^m(h)| \ge x\right) \le C_1 \exp\left(-C_2 (x/A)^{2/m}\right). \tag{1.3}$$

The followimg Theorem is due to Cheng[7].

Theorem C. *Let $h(\cdot)$ be P-canonical with $B \equiv E\exp(|h|) < \infty$. Then for all $n \ge m$,*

$$P(|U_n^m(h)| \ge x) \le 2\exp\Big(0.55 \frac{n}{(n-1)\cdots(n-j+1)} \frac{(j+1)!}{j} - \frac{nx}{2^{j+3}\log n}\Big) + B\exp\Big(-\frac{1}{2^{j+2}}\Big[\frac{n}{j}\Big]x\Big). \tag{1.4}$$

In this paper, we obtain the exact estimate of constants C_1, C_2 in Theorem A. Also, we discuss Bernstein type inequalities for degenerate U-statistics under $E\exp(|h|^{2\delta}) < \infty$ for some $o < \delta \le 1$, which extend and sharpen Theorem C with many other similar results. To establish our main results, the upper bounds of $E(U_n^m(h))^p$ for any real number $p \ge 2$ are better estimated (see the following Lemma 1.1) which are interesting in itself. In the third section, we discuss Cramer type deviation for studentized U-statistics by using these exponential inequalities, which improve the similar results given in [15].

Theorem 1.1. *Let $h(\cdot)$ be P-canonical with $|h| \le C$. Then for any $x \ge A_0 e^{m/2}, n \ge m$,*

$$P\left(\binom{n}{m}^{1/2} |U_n^m(h)| \ge x\right) \le \exp\Big(-\frac{m}{2}\Big[1 + \frac{1}{e}\Big(\frac{x}{A_0}\Big)^{2/m}\Big]\Big) \tag{1.5}$$

where $A_0 = \frac{\sqrt{\pi}}{2} C \left(\frac{8}{e} m^2\right)^m$.

In the following Theorem 1.2 we relax the conditions of Theorem 1.1.

Theorem 1.2 *Let $h(\cdot)$ be P-canonical with*

$$\alpha \equiv E\exp(|h|^{2\delta}) < \infty, \qquad \text{for some } 0 < \delta \le 1. \tag{1.6}$$

Then for all $x \ge [2e(C_0+1)]^{m/2}$, $n \ge n_0$,

$$P\left(\binom{n}{m}^{1/2} |U_n^m(h)| \ge xA_1\right) \le \exp\left(-\frac{(2e)^{-1} m x^{2/m}}{C_0 + \left(Bx^{1/m}[\frac{n}{m}]^{-\delta/2}\right)^{2/(m\delta+1)}}\right) \tag{1.7}$$

where $A_1 = \frac{\sqrt{\pi}}{2} \left(\frac{8}{e} m^2\right)^m$, $B^{-1} = \min[\sqrt{2eC^\delta \delta}, 2e(C_0+1)C^\delta \delta], n_0, C_0, C$ are chosed as those of Lemma 2.2.

Remark 1.1. Let $h(\cdot)$ be P-canonical with $E\exp(|h|) < \infty$. It follows from Theorem C that there exists constant $C > 0$ such that

$$P\left(\binom{n}{m}^{1/2} |U_n^m(h)| \ge x\right) \le \exp\Big(-\frac{Cn^{1-m/2}x}{\log n}\Big), \quad \text{for all } x > 0. \tag{1.8}$$

It is easy to see that inequality (1.7) is better than (1.8) in the range $x^{2/m} \le n^{1/2}$.

In the following, we define for $k = 1, \cdots, m$

$$\prod_{k,m} h(x_1, \cdots, x_k) = (\delta_{x_1} - P)\cdots(\delta_{x_k} - P)\, P^{m-k} h, \tag{1.9}$$

where $Q_1 \cdots Q_m h = \int \cdots \int h(x_1, \cdots x_m) dQ_1(x_1) \cdots dQ_m(x_m)$. Then

$$U_n^m(h) = \sum_{k=0}^{m} \binom{m}{k} U_n^k \left(\prod_{k,m} h \right), \tag{1.10}$$

where $\prod_{k,m} h(\,\cdot\,)$ is a P-canonical function of k-variable,

$$U_n^k \Big(\prod_{k,m} h \Big) = \frac{(n-k)!}{n!} \sum_{(i_1,\cdots,i_k)\in I_k^n} \prod_{k,m} h(X_{i_1}, \cdots, X_{i_k}) \tag{1.11}$$

and the first r summands of the sum in (1.11) vanish if $h(\,\cdot\,)$ is r-degenerate for P.

It is well-known that $|h| \le A$ implies $|\Pi_{k,m} h| \le A$, and conditions (1.6) imply

$$E \exp \Big(\Big| \prod_{k,m} h \Big|^{2\delta} \Big) < \infty, \quad \text{for some } \ 0 < \delta \le 1 \tag{1.12}$$

by using Jensen's inequality. Therefore, combining (1.10)–(1.12) and Theorems 1.1 and 1.2, we obtain easily the following results.

Corollary 1.1. *Let $h(\,\cdot\,)$ be r-degenerate for P, $0 \le r \le m-1$ with $|h| \le C$ and $P^m h = 0$ if r=0. Then for all $x > 0$, $n \ge m$,*

$$P \left(n^{(r+1)/2} \left| U_n^{r+1}(h) \right| \ge x \right) \le \sum_{k=r+1}^{m} \exp \left(-C(m,k) x^{2/k} n^{(k-r-1)/k} \right), \tag{1.13}$$

where positive constants C(m,k) depend on m, k only.

Corollary 1.2. *Let h be r-degenerate $(1 \le r \le m-1)$ for P and (1.6) hold. Then for all $x > 0$, $n \ge m$*

$$\begin{aligned} &P \left(n^{(r+1)/2} \left| U_n^{r+1}(h) \right| \ge x \right) \\ &\le \sum_{k=r+1}^{m} \exp \Big(-C_1(m,k) \frac{x^{2/k} n^{(k-r-1)/k}}{1 + \big(x^{1/k} n^{(k-r-1-\delta)/2} \big)^{2/(k\delta+1)}} \Big), \end{aligned} \tag{1.14}$$

where positive constants $C_1(m,k)$ depend on m, k only.

Remark 1.2. Inequality (1.14) generalizes the main results given by Borisov[4], where the author established (1.14) for degenerate U-statistics satisfying

$$|h(x_1, x_2, \cdots, x_m)| \le g(x_1) \cdots g(x_m).$$

§2. Lemmas and Proofs of Main Results

To prove the Theorems, we need the following lemmas.

Lemma 2.1. Let $h(\,\cdot\,)$ be P-canonical. Then for all real $p \ge 2$,

$$E \left| \binom{n}{m}^{1/2} U_n^m(h) \right|^p \le C_{m,p} E \left[\frac{1}{[\frac{n}{m}]} \sum_{j=1}^{[\frac{n}{m}]} h^2(X_{jm+1}, \cdots, X_{(j+1)m}) \right]^{p/2} \tag{2.1}$$

$$\le C_{m,p} E |h(X_1, \cdots, X_m)|^p \tag{2.2}$$

where $C_{m,p} = \left(\frac{\sqrt{\pi}}{2} \right)^p \left(\frac{8}{e} m^2 \right)^{mp} (p-1)^{mp/2}$.

Remark 2.1. The upper bounds of $E|\binom{n}{m} U_n^m(h)|^p$ are studied by many authors (see, for example, [6, 7, 9]). But, inequality (2.1) seems to be new and inequality (2.2) is more accurate than the previous results

Proof. Let $a_{i_1,\cdots,i_m}$ be elements in a Banach space $(B,\|.\|)$,

$$Y=\sum_{1\le i_1<\cdots<i_m\le n}\epsilon_{i_1}\cdots\epsilon_{i_m}a_{i_1,\cdots,i_m},$$

where $\epsilon_j, j\ge 1$ are a squence of i.i.d. Rademacher random variables. In View of Khinchin-type inequalities for the Rademacher chaos (see, for example, [3]), it is evident that for any real fixed $p\ge 2$

$$E|Y|^p\le (p-1)^{mp/2}\Big(\sum_{1\le i_1<\cdots<i_m\le n}a^2_{i_1,\cdots,i_m}\Big)^{p/2}. \tag{2.3}$$

Now we apply Theorem 1.2 and Corollary 1 in [17] for a convex function $\phi(x)=|x|^p$, and then (2.3) with $a_{i_1,\cdots,i_m}=h(X_{i_1},\cdots,X_{i_m})$, to obtain

$$E\Big|\sum_{1\le i_1<\cdots<i_m\le n}h(X_{i_1},\cdots,X_{i_m})\Big|^p\le m^{mp}E\Big|\sum_{1\le i_1<\cdots<i_m\le n}h(X^1_{i_1},\cdots,X^m_{i_m})\Big|^p$$

(where $\{X^k_j\}^n_{j=1}, k=1,\cdots,m$ are independent copies of $\{X_j\}^n_{j=1}$)

$$\le (2m)^{mp}E\Big|\sum_{1\le i_1<\cdots<i_m\le n}h(X^1_{i_1},\cdots,X^m_{i_m})\epsilon^1_{i_1}\cdots\epsilon^m_{i_m}\Big|^p$$

(by symmetrization, cf. Ledoux and Talagrand (1991), where$\{\epsilon^k_j\}^n_{j=1}, k=1,\cdots,m$ are independent copies of $\{\epsilon_j\}^n_{j=1}$)

$$\le (2m)^{mp}[4^{m-1}(m-1)!]^pE\Big|\sum_{1\le i_1<\cdots<i_m\le n}h(X_{i_1},\cdots,X_{i_m})\epsilon_{i_1}\cdots\epsilon_{i_m}\Big|^p$$

$$\le C_{m,p}E\Big(\sum_{1\le i_1<\cdots<i_m\le n}h^2(X_{i_1},\cdots,X_{i_m})\Big)^{p/2} \tag{2.4}$$

(using that $k!\le\sqrt{2\pi k}k^ke^{-k}\exp(\frac{1}{12k})$).

Since $\binom{n}{m}^{-1}\sum_{1\le i_1<\cdots<i_m\le n}h^2(X_{i_1},\cdots,X_{i_m})$ is the average of $W(X_{i_1},\cdots,X_{i_n})$ over all the permutations $i_1,\cdots,i_n$ of $1,\cdots,n$ with

$$W(x_1,\cdots,x_n)=k^{-1}\sum_{j=0}^{k-1}h^2(X_{jm+1},\cdots,X_{(j+1)m})$$

and $k=[\frac{n}{m}]$ (see [14]), we have, by convexity of $x^{p/2}$ $(x>0)$,

$$E\Big|\binom{n}{m}^{1/2}U^m_n(h)\Big|^p$$

$$\le C_{m,p}E\left(\binom{n}{m}^{-1}\sum_{1\le i_1<\cdots<i_m\le n}h^2(X_{i_1},\cdots,X_{i_m})\right)^{p/2}$$

$$\le C_{m,p}E\Big[\frac{1}{[\frac{n}{m}]}\sum_{j=1}^{[\frac{n}{m}]}h^2(X_{jm+1},\cdots,X_{(j+1)m})\Big]^{p/2}. \tag{2.5}$$

Inequality (2.2) follows from Hölder's inequality. This proves Lemma 2.1.

Lemma 2.2. *Let $X, X_j, j\ge 1$ be a sequence of nonnegative i.i.d. random variables with $A\equiv E\exp(|X|^\delta)<\infty$ for some $0<\delta\le 1$. Then,*

(1) *for all* $y \geq C_0$ *and* $n \geq n_0$,

$$P\Big(\frac{1}{n}\sum_{j=1}^{n} X_j \geq y\Big) \leq \exp[-(cny)^\delta], \tag{2.6}$$

(2) *for all real* $p \geq 1$,

$$E\Big(\frac{1}{n}\sum_{j=1}^{n} X_j\Big)^p \leq C_0^p + (Cn)^{-p}\Big(\frac{p}{\delta}\Big)^{p/\delta}, \tag{2.7}$$

where

$$C = \begin{cases} (1-\delta)/2^{1/\delta+2}, & \text{if } 0<\delta<1,\\ 1/4(2A+1), & \text{if } \delta = 1,\end{cases}$$
$$C_0 = \begin{cases} \sqrt{A}2^{2/(\delta+1)}\ \Gamma(1/\delta+1), & \text{if } 0<\delta<1,\\ \max(2EX_1, 1+EX_1), & \text{if } \delta = 1,\end{cases}$$
$$n_0 = \begin{cases} \frac{4}{A}(\delta^{1/\delta}(1-\delta)\ \Gamma(1/\delta+1))^{-2}, & \text{if } 0<\delta<1,\\ 2, & \text{if } \delta = 1.\end{cases}$$

Proof. Since $A \equiv E\exp(X) < \infty$ is equivalent to $EX^k \leq Ak!$ for all integer $k \geq 2$, inequality (2.6) follows from Bernstein's inequality (for $\delta = 1$) and Schmuckenschlaeger's inequality[13].

It follows from (2.6) that for all real $p \geq 1$,

$$\begin{aligned} E\Big(\frac{1}{n}\sum_{j=1}^{n} X_j\Big)^p &= \int_0^\infty P\Big(\frac{1}{n}\sum_{j=1}^{n} X_j \geq y\Big)dy^p \\ &\leq C_0^p + \int_{c_0}^\infty \exp[-(cny)^\delta]dy^p \\ &\leq C_0^p + (Cn)^{-p}\int_0^\infty \exp(-t)dt^{p/\delta} \\ &\leq C_0^p + (Cn)^{-p}\Big(\frac{p}{\delta}\Big)^{p/\delta}. \end{aligned} \tag{2.8}$$

Lemma 2.2 is complete.

Now we prove the theorems.

Proof of Theorem 1.1. It follows from Lemma 2.1 and Marcov's inequality that for all real $p \geq 2$,

$$\begin{aligned} P\left(\binom{n}{m}^{1/2}|U_n^m(h)| \geq x\right) &\leq x^{-p}(p-1)^{mp/2}A_0^p \\ &\leq \exp\Big(-p\log x + \frac{mp}{2}\log(p-1) + p\log A_0\Big). \end{aligned} \tag{2.9}$$

Inequality (1.5) follows by minimizing the exponent with respect p in (2.9) and noting $p \geq 2$.

Proof of Theorem 1.2. Let

$$X_j = h^2(X_{jm+1}, \cdots, X_{(j+1)m}).$$

Then $X_j, j \geq 1$ are a sequence of i.i.d. random variables with $E\exp(|X_1|^\delta) < \infty$, for some $0 < \delta \leq 1$. Therefore, combining (2.1), (2.7) and Marcov's inequality, we obtain that for all real $p \geq 1$

$$P\left(\binom{n}{m}^{1/2}|U_n^m(h)| \geq xA_1\right) \leq A_1^{-2p}x^{-2p}E\left|\binom{n}{m}^{1/2}U_n^m(h)\right|^{2p}$$
$$\leq x^{-2p}(2p)^{mp}\Big[C_0^p + \Big(C\Big[\frac{n}{m}\Big]\Big)^{-p}\Big(\frac{p}{\delta}\Big)^{p/\delta}\Big]. \tag{2.10}$$

Now we take

$$p = \begin{cases} \frac{(2e)^{-1}x^{2/m}}{C_0+1}, & \text{if } x^{2/m} \leq [\frac{n}{m}]^\delta B^{-1}, \\ \frac{(2e)^{-1}x^{2/m}}{C_0+(Bx^{1/m}[\frac{n}{m}]^{-\delta/2})^{2/(m\delta+1)}}, & \text{if } x^{2/m} \geq [\frac{n}{m}]^\delta B^{-1}. \end{cases} \tag{2.11}$$

Recall that

$$B^{-1} = \min[\sqrt{2eC^\delta\delta}, 2e(C_0+1)C^\delta\delta].$$

Elementary calculation shows that if $x^{2/m} \leq [\frac{n}{m}]^\delta B^{-1}$,

$$C_0^p + \Big(C\Big[\frac{n}{m}\Big]\Big)^{-p}\Big(\frac{p}{\delta}\Big)^{p/\delta} \leq C_0^p + 1 \leq (C_0+1)^p; \tag{2.12}$$

if $x^{2/m} \geq [\frac{n}{m}]^\delta B^{-1}$ (note that $C_0 > 1$),

$$p \leq (2e)^{-1}B^{-\frac{2}{(m\delta+1)}}x^{\frac{2\delta}{m\delta+1}}[\frac{n}{m}]^{\frac{\delta}{m\delta+1}},$$

$$C_0^p + \Big(C\Big[\frac{n}{m}\Big]\Big)^{-p}\Big(\frac{p}{\delta}\Big)^{p/\delta} \leq C_0^p + C^{-p}\delta^{-\frac{p}{\delta}}(2e)^{-\frac{p}{\delta}}B^{-\frac{2p}{\delta}}\Big(Bx^{\frac{1}{m}}\Big[\frac{n}{m}\Big]^{-\frac{\delta}{2}}\Big)^{\frac{2mp}{m\delta+1}}$$
$$\leq C_0^p + \Big(Bx^{\frac{1}{m}}[\frac{n}{m}]^{-\frac{\delta}{2}}\Big)^{\frac{2mp}{m\delta+1}}$$
$$\leq \Big[C_0 + \Big(Bx^{\frac{1}{m}}[\frac{n}{m}]^{-\frac{\delta}{2}}\Big)^{\frac{2}{m\delta+1}}\Big]^{mp}. \tag{2.13}$$

Inequality (1.7) follows from (2.10)–(2.13) immediately. Theorem 1.2 is complete.

§3. Cramer Type Large Deviations for Studentized U-Statistics

Cramer type deviations for non-degenerate U-statistics were studied by many authors (see [2] for details). In this section, we discuss Cramer type deviations for studentized U-statistics. Explicitly, we establish the following theorem.

Theorem 3.1. *Let $h(\,\cdot\,)$ be non-degenerate kernel of degree 2 (i.e., in (1.1) $m = 2$) with* $\sigma_g^2 \equiv \mathrm{Var}[E(h(X_1, X_2)|X_1)] > 0$,

$$E\exp(|h|^{2\delta}) < \infty, \ \textit{for some } 0 < \delta \leq \frac{1}{2}. \tag{3.1}$$

Then

$$P\Big(\frac{\sqrt{n}(U_n^2(h) - Eh(X_1, X_2))}{S_n} \geq x\Big) = (1 - \Phi(x))(1 + o(1)) \tag{3.2}$$

uniformly in the range $0 \leq x \leq \rho(n)n^{\delta/2(2-\delta)}$ *with* $\rho(n) \to 0$, *where*

$$S_n^2 = \frac{4(n-1)}{(n-2)^2}\sum_{j=1}^n\left[(n-1)^{-1}\sum_{\substack{i=1\\i\neq j}}^n h(X_i, X_j) - U_n^2(h)\right]^2,$$

$\Phi(x)$ *denotes the standard normal distribution function.*

Remark 3.1. Condition (3.1) is equivalent to the condition that there exists a constant $A > 0$ such that for all $k = 1, 2, \cdots$

$$E|h(X_1, X_2) - E(X_1, X_2)|^k \le Ak^{k/2\delta}. \tag{3.3}$$

Vandemaele and Veraverbeke[15] proved that under the condition (3.3) relation (3.2) holds uniformly in the range $0 \le x \le \rho(n)n^\delta$ with

$$\alpha = \begin{cases} \frac{\delta}{2(3\delta+1)}, & \text{if } \delta \ge \frac{1}{4}, \\ \frac{\delta}{2(2-\delta)}, & \text{if } \delta \le \frac{1}{4}. \end{cases}$$

Corollary 3.1. *Let $h(\,\cdot\,)$ be a non-degenerate kernel of degree 2 with $\sigma_g^2 > 0, E\exp(|h|) < \infty$. Then, (3.2) holds uniformaly in the range $0 \le x \le \rho(n)n^{1/6}$ with $\rho(n) \to 0$.*

Remark 3.2. Condition $E\exp(|h|) < \infty$ is not the optimal. But, it is difficult to remove for studentized U-statistics.

For the proof of Theorem 3.1, we need the following Lemma 3.1.

Lemma 3.1. *Let A_n, B_n, C_n be sequences of random variables. If for some $\alpha > 0$,*

$$P(A_n \ge x) = (1 - \Phi(x))(1 + o(1)) \tag{3.4}$$

uniformly in the range $0 \le x \le \rho(n)n^\alpha$ with $\rho(n) \to 0$, and if

$$P(B_n \ge n^{-\alpha}) = o(1 - \Phi(x)), \tag{3.5}$$

$$P(|C_n^2 - 1| \ge n^{-2\alpha}) = o(1 - \Phi(x)) \tag{3.6}$$

uniformly in the range $0 \le x \le \rho(n)n^\alpha$ with $\rho(n) \to 0$,then

$$P(C_n^{-1}A_n + B_n \ge x) = (1 - \Phi(x))(1 + o(1)) \tag{3.7}$$

uniformly in the range $0 \le x \le \rho(n)n^\alpha$ with $\rho(n) \to 0$.

Proof. Since $|C_n^2 - 1| \le n^{-2\alpha}$ is equivalent to

$$(1 - n^{-2\alpha})^{1/2} \le C_n \le (1 + n^{-2\alpha})^{1/2},$$

similar to the proof of Lemma 4 in [15], we have

$$P(C_n^{-1}A_n \ge x) = (1 - \Phi(x))(1 + o(1)) \tag{3.8}$$

uniformly in the range $0 \le x \le \rho(n)n^\alpha$ with $\rho(n) \to 0$. Relation (3.7) follows from (3.5),(3.8) and classical methods (see, for example, [11]).

Proof of Theorem 3.1. From [2], we have that

$$P\Big(\frac{\sqrt{n}(U_n^2(h) - Eh(X_1, X_2))}{2\sigma_g} \ge x\Big) = (1 - \Phi(x))(1 + o(1)) \tag{3.9}$$

uniformly in the range $0 \le x \le \rho(n)n^{\delta/2(2-\delta)}$ with $\rho(n) \to 0$. By using Lemma 3.1, to prove (3.2) it is enough to show that

$$P\Big(\Big|\frac{S_n^2}{4\sigma_g^2} - 1\Big| \ge n^{-\delta/(2-\delta)}\Big) = o(1 - \Phi(x)) \tag{3.10}$$

uniformly in the range $0 \le x \le \rho(n)n^{\delta/2(2-\delta)}$ with $\rho(n) \to 0$. Let

$$\begin{aligned} g(X_j) &= E(h(X_i, X_j)|X_j), \\ \tilde{g}(X_j) &= E[\psi(X_i, X_j)g(X_i)|X_j], \\ \psi(X_i, X_j) &= h(X_i, X_j) - g(X_i) - g(X_j) + Eh(X_i, X_j) \\ f(X_j) &= 4(g^2(X_j) - \sigma_g^2) + 8\tilde{g}(X_j). \end{aligned}$$

It follows from [5] that

$$S_n^2 = 4\sigma_g^2 + T_n + R_n, \tag{3.11}$$

where

$$T_n = \frac{1}{n}\sum_{j=1}^{n} f(X_j), \qquad R_n = \sum_{j=1}^{n} R_{nj} \tag{3.12}$$

with

$$R_{n1} = -4\binom{n}{2}^{-1}\sum_{i<j} g(X_i)g(X_j),$$

$$R_{n2} = 4\binom{n}{2}^{-1}\sum_{i<j}[(g(X_i)+g(X_j))\psi(X_i,X_j) - \tilde{g}(X_i) - \tilde{g}(X_j)],$$

$$R_{n3} = -\frac{8}{n}\sum_{i=1}^{n}\left[g(x_i)\binom{n-1}{2}^{-1}\sum_{r<m}^{(i)}\psi(X_r,X_m)\right],$$

$$R_{n4} = \frac{4}{n-2}\sum_{i=1}^{n}\left[\binom{(n-1)}{2}^{-1}\sum_{r<m}^{(i)}\psi(X_i,X_j)\psi(X_i,X_m)\right],$$

$$R_{n5} = -\frac{4n(n-1)}{(n-2)^2}\left[\binom{n}{2}^{-1}\sum_{i<j}\psi(X_i,X_r)\right]^2,$$

$$R_{n6} = \frac{4n}{(n-2)^2}\binom{n}{2}^{-1}\sum_{i<j}\psi^2(X_i,X_j).$$

Remark that for above notation we write $\sum\limits_{r<m}^{(k)}$ for $\sum\limits_{\substack{1\le r<m\le n\\ r\ne k, m\ne k}}$.

Elementary calculus shows that for any positive constant $C>0$,

$$\begin{aligned}\exp(-cn^{\delta/(2-\delta)}) &\le \exp\Big(-\frac{\rho(n)}{2}n^{\delta/(2-\delta)}\Big)\\ &\le 1-\Phi(\rho^{1/2}(n)n^{\delta/2(2-\delta)})\\ &= o(1-\Phi(x)), as\ n\to\infty\end{aligned} \tag{3.13}$$

uniformly in the range $0\le x\le \rho(n)n^{\delta/2(2-\delta)}$ with $\rho(n)\to 0$.

In view of (3.11)–(3.13), relation (3.10) follows from the following lemma.

Lemma 3.2. *Assume* (3.1). *Then,*

(1)

$$P\left(T_n \ge n^{-\delta/(2-\delta)}\right) = o(1-\Phi(x)) \tag{3.14}$$

uniformly in the range $0\le x\le \rho(n)n^{\delta/2(2-\delta)}$ *with* $\rho(n)\to 0$.

(2) *there exist constants* $C, C_1>0$ (*depend on* δ *only*) *such that for* $j=1,\cdots,6$,

$$P\left(|R_{nj}|\ge n^{-\delta/(2-\delta)}\right)\le C_1\exp(-Cn^{\delta/(2-\delta)}). \tag{3.15}$$

Proof of Lemma 3.2. In view of (3.3), relation (3.14) follows from (14) in [15].

Next we prove (3.15). Let $A(X_i, X_j)$ denote the one of the following variables

$$g(X_i)g(X_j), \quad (g(X_i)+g(X_j))\psi(X_i,X_j)-\tilde g(X_i)-\tilde g(X_j), \quad \psi(X_k,X_i)\psi(X_k,X_j),\ i,j\neq k.$$

It is easy to show that for all $i \neq j$,

$$E(A(X_i,X_j)|X_i)=E(A(X_i,X_j)|X_j)=0 \tag{3.16}$$

and for some $0<\delta\le 1/2$,

$$E\exp(|A(X_1,X_2)|^{\delta})<\infty, \tag{3.17}$$

$$E\exp(|g(X_1)|^{2\delta})<\infty, \tag{3.18}$$

$$E\exp(|\psi(X_1,X_2)|^{2\delta})<\infty \tag{3.19}$$

by using Jensen's inequality.

It follows from (3.16)–(3.19) and Corollary 1.2 that there exists a constant $C>0$ such that if $0<\delta\le 1/2$,

$$\begin{aligned} P(|R_{nk}|\ge n^{-\delta/(2-\delta)}) &\le P(n|R_{nk}|\ge Cn^{2(1-\delta)/(2-\delta)}) \\ &\le \exp\left(-Cn^{(6-5\delta)\delta/2(\delta+1)(2-\delta)}\right) \\ &\le \exp(-Cn^{\delta/(2-\delta)}), \quad \text{for } k=1,2. \end{aligned} \tag{3.20}$$

Similarly, we obtain that if $0<\delta\le 1/2$,

$$\begin{aligned} P(|R_{n5}|\ge n^{-\delta/(2-\delta)}) &\le P\left(\binom{n}{2}^{-1}\Big|\sum_{i<j}\psi(X_i,X_j)\Big|\ge n^{-\delta/(2-\delta)}\right) \\ &\le \exp(-Cn^{\delta/(2-\delta)}), \end{aligned} \tag{3.21}$$

$$\begin{aligned} &P(|R_{n4}|\ge n^{-\delta/(2-\delta)}) \\ &\le \sum_{j=1}^{n} P\left(\binom{n-1}{2}^{-1}\Big|\sum_{r<m}^{(j)}\psi(X_j,X_r)\psi(X_j,X_m)\Big|\ge Cn^{-\delta/(2-\delta)}\right) \\ &\le \exp\left(-Cn^{(6-5\delta)\delta/2(\delta+1)(2-\delta)}\right) \\ &\le n\exp(-Cn^{\delta/(2-\delta)}), \end{aligned} \tag{3.22}$$

$$\begin{aligned} &P(|R_{n3}|\ge n^{-\delta/(2-\delta)} \\ &\le \sum_{j=1}^{n} P\left(\binom{n-1}{2}^{-1}\Big|g(X_j)\sum_{r<m}^{(j)}\psi(X_r,X_m)\Big|\ge Cn^{-\delta/(2-\delta)}\right) \\ &\le nP(|g(X_1)|\ge n^{\frac{1}{2(2-\delta)}})+nP\left(\binom{n-1}{2}^{-\frac12}\Big|\sum_{r<m}^{(1)}\psi(X_r,X_m)\Big|\ge Cn^{\frac{3-4\delta}{2(2-\delta)}}\right) \\ &\le n\exp(-n^{\frac{\delta}{2-\delta}})E\exp(|g(X_1)|^{2\delta})+nexp\left(-Cn^{\frac{5(1-\delta)\delta}{(2\delta+1)(2-\delta)}}\right) \\ &\le C_1\exp(-Cn^{\delta/(2-\delta)}). \end{aligned} \tag{3.23}$$

Next we prove (3.15) for $j=6$. Note that (3.19) implies there exists a constant $A>0$ such that for all real $p\ge 1$,

$$E|\psi(X_1,X_2)|^{2p}\le Ap^{p/\delta}. \tag{3.24}$$

Therefore, similar to the proof of (2.5), we obtain for all real $p \geq 2$

$$\begin{aligned} P(|R_{n6}| \geq n^{-\delta/(2-\delta)}) &\leq P\left(\binom{n}{2}^{-1}\sum_{i<j}\psi^2(X_i,X_j) \geq Cn^{2(1-\delta)/(2-\delta)}\right) \\ &\leq C^{-p}n^{-\frac{2(1-\delta)p}{2-\delta}}E\left[\binom{n}{2}^{-1}\sum_{i<j}\psi^2(X_i,X_j)\right]^p \\ &\leq C^{-p}n^{-\frac{2(1-\delta)p}{2-\delta}}E|\psi(X_1,X_2)|^{2p} \\ &\leq An^{-\frac{2(1-\delta)p}{2-\delta}}p^{p/\delta}C^{-p}. \end{aligned} \tag{3.25}$$

Choosing $p = e^{-1}n^{-\frac{2(1-\delta)\delta}{2-\delta}}C^{\delta}$, it follows from (3.25) that if $0 < \delta \leq 1/2$,

$$P(|R_{n6}| \geq n^{-\delta/(2-\delta)}) \leq A\exp(-Cn^{\delta/(2-\delta)}). \tag{3.26}$$

Combining (3.21)–(3.25), we prove Lemma 3.2. The proof of Theorem 3.1 is complete.

References

[1] Arcones, M. A. & Gine, E., Limit theorems for U-processes, *Ann. Prob.,* **23** (1993), 1494–1542.

[2] Aleshkyavichene, A. K., Probability of large deviation for U-statistics, *Lithuanian Mathematical Journal,* **32**(1992), 4–14.

[3] Borel, C., On the integrability of Banach space valued Walsh polymomiales, Lecture Notes in Math., 721, 1–3, Springer, New York, 1979.

[4] Borisov, J. S., Exponential inequalities for the distributions of Von-Mises and U-Statistics, *Prob.Theory and Math. Stat.,* **I** (Vilnius)(1990), 168–178.

[5] Callaert, H. & Veraverbeke, N., The order of the normal approximation for a studentized U-statistics, *Ann. Stat.,* **9**(1981), 194–200.

[6] Chen Xiru, The limit properties for the Von-mises and U-statistics, *Scientia Sinica, Series A,* **10**(1980), 522–532 (in Chinese).

[7] Cheng Ping, Some inequalities on U-statistics, *Chinese Science Bulletin,* **20**(1996), 1835–1837.

[8] Christofides, T. C., Probability inequalities with exponential bounds for U-statistics, *Stat. Prob. Letters,* **12**(1991), 257–261.

[9] Dehling, H., Denker, M. & Philipp, W., The almost sure invariance principle for the empirical process of U-statistics structure, *Ann. Inst. Henri Poincare,* **23**(1987), 121–133.

[10] Hoeffding, W., Probability inequalities for sums of bounded random variables, *J. Amer. Statist. Math.,* **18**(1963), 277–287.

[11] Inglot, T., Kallenberg, W. C. M. & Ledwina, T., Strong moderate deviation theorems, *Ann. Prob.,* **20**(1992), 987–1000.

[12] Miao, B. & Zhang, S. G., The exponential convergence rate of the projection residuce of U-statistics and their application, *Chinese J.of Applied Probability and Statitics,* **10**(1994), 253–263.

[13] Schmuckenschlaeger, M., Bernstein inequalities for a class of random variables, *Proc. Amer. Math. Society,* **117**(1993), 1159–1163.

[14] Serfling, R. J., Approximation theorems of mathematicl statistics, Wiley, New York, 1980.

[15] Vardemaele, M. & Veraverbeke, N., Cramer type large deviations for studentized U-statistics, *Metrika,* **32**(1985), 165–180.

[16] Arcones, M. A., A Bernstein-type inequality for U-Statistics and U-processes, *Stat. Probab. Letters,* **22** (1995), 239–247.

[17] de la Peña, V., Decoupling and Khintchine's inequalities for U-Statistics, *Ann. Probab.,* **20** (1992), 1877–1892.