On Large Deviations of U-Statistics

Yuri V. Borovskikh
Transport University
St Petersburg

Neville C. Weber
University of Sydney
Sydney

January 11, 2000

Large deviation results with explict order terms and Cramér's series are developed for non-degenerate U-statistics of degree m with bounded kernels using the contraction technique of Keener, Robinson and Weber (1998) and a partitioning method to bound the degenerate remainder term.

1 Introduction and Results.

Let $X_1, \cdots, X_n$ be independent and identically distributed random variables taking values in the measurable space $(\mathbf{X}, \mathcal{X})$. Consider a U-statistic $U_n = U_n(X_1, \cdots, X_n)$ of the form

$$U_n = \frac{(m-1)\cdots 2 \cdot 1}{(n-1)\cdots(n-m+1)} \sum_{1 \le i_1 < \cdots < i_m \le n} h(X_{i_1}, \ldots, X_{i_m})$$

where $h : \mathbf{X}^m \to R$ is an integrable, symmetric function. According to the Hoeffding representation

$$U_n = S_n - \frac{n(m-1)}{m}\mu + v_n,$$

where

$$\mu = Eh(X_1, \cdots, X_m),$$

AMS 1991 subject classification. Primary 60F10. *Key words and phrases.* Large deviations, U-statistics, Cramer's series.

$$
\begin{aligned}
S_n &= \sum_{i=1}^{n} g(X_i), \\
g(x) &= E(h(X_1, \cdots, X_{m-1}, X_m)|X_m = x),
\end{aligned}
$$

and

$$
\begin{aligned}
v_n &= \frac{m-1}{n-1} \sum_{1 \le i_1 < i_2 \le n} g_2(X_{i_1}, X_{i_2}) + \cdots \\
&\quad + \frac{(m-1)\cdots 2 \cdot 1}{(n-1)\ldots(n-m+1)} \sum_{1 \le i_1 < \cdots < i_m \le n} g_m(X_{i_1}, \ldots, X_{i_m}),
\end{aligned}
$$

where $g_c : \mathbf{X}^c \to R,\ c = 2, \ldots, m$ are canonical functions possessing the degeneracy property $Eg_c(x_1, \ldots, x_{c-1}, X_c) = 0$ a.s., $c = 2, \ldots, m$.

Let $\sigma^2 = E(g(X_1) - \mu)^2$ and assume $0 < \sigma^2 < \infty$. By the central limit theorem for non-degenerate U-statistics,

$$
P\left(\frac{U_n - n\mu/m}{\sigma\sqrt{n}} \le x\right) \to \Phi(x),
$$

uniformly in x as $n \to \infty$, provided

$$
t^{\gamma_c} P(|g_c(X_1, \ldots, X_c)| > t) \to 0 \text{ as } \quad t \to \infty
$$

for each $c = 2, \ldots, m$, where $\gamma_c = 2c/(2c-1)$ and $\Phi(x)$ denotes the standard normal distribution function.

We are interested in the relative error $o(1)$ in

$$
P\left(\frac{U_n - n\mu/m}{\sigma\sqrt{n}} > x\right) = (1 - \Phi(x))(1 + o(1))
$$

as x tends to infinity together with n . In what follows, without loss of generality, we take $\mu = 0$.

We will develop the new contraction technique proposed by Keener, Robinson and Weber (1998) for U-statistics of degree 2 to establish Cramér's series for large deviation probabilities for U-statistics of any degree m. We find the explicit solution of the contraction equation in the form of a convergent power series in t with coefficients depending on $x \in \mathbf{X}$. This enables the construction of the Cramér's series with coefficients in explicit form. Further we use a new partitioning method to handle the error term. The degenerate part of the U-statistic is expressed as a sum of separate terms. This partitioning enables us to bound the required expectation by the sum of several terms where in each term we can identify a sum that is independent of all other components in that term. We then use this independence to bound the necessary expectations.This new result also enables the range of x values for which the expansion in Keener *et al.* holds to be extended to $x = o(\sqrt{n})$. Aleshkevichiene (1990) established large deviation results for U-statistics of degree 2 with bounded kernels with the help of the cumulant method. In their Theorem 3 instead of the full Cramér's series they have only a truncated version.

Let $\mathcal{C} = \{g : \| g \|_\infty \leq c\}$, where $\| \cdot \|_\infty$ denotes the essential supremum. If $d(g_0, g_1) = \| g_0 - g_1 \|_\infty$, then $(\mathcal{C}, d)$ is a complete metric space. For $g, h \in \mathcal{C}$ define the operator $\mathcal{H}_t^{m-1} : g \to \mathcal{H}_t^{m-1} g$, by

$$\mathcal{H}_t^{m-1} g = \frac{E\{\prod_{s=1}^{m-1} e^{tg(X_s)} h(X_1, \cdots, X_{m-1}, \cdot)\}}{(Ee^{tg(X_1)})^{m-1}}, \ t \in R,$$

for any fixed integer $m = 2, 3, \ldots$.

As in Keener *et al.* (1998) we can to show that the operator $\mathcal{H}_t^{m-1}$ is a contraction if $2(m-1)c|t| \exp(2(m-1)c|t|) < 1$. Therefore from the contraction mapping theorem (Lemma 14.1.1 in Lang (1994)) it follows that the contraction equation $g = \mathcal{H}_t^{m-1} g$ has an essentially unique solution $g_t \in \mathcal{C}$, i.e.

$$g_t = \mathcal{H}_t^{m-1} g_t. \tag{1}$$

Our first result relates to the equation (1). We give the contraction solution $g_t(x)$ of (1) in the form of an analytic function in the parameter t. To formulate this result we need some notation. For $x \in \mathbf{X}$ and $k = 1, 2, \ldots$ let

$$c_k(x) = \sum_{l=0}^{k-1} a_l(x) \ b_{k-1-l}, \tag{2}$$

where,

$$\begin{aligned} a_0(x) &= g(x), \\ a_l(x) &= \sum_{\nu_1+2\nu_2+\ldots+l\nu_l=l} Eh(X_1, \ldots, X_{m-1}, x) \prod_{s=1}^{l} \frac{1}{\nu_s!} \left(\sum_{j=1}^{m-1} c_s(X_j) \right)^{\nu_s}, \quad l \geq 1 \end{aligned}$$

and

$$\begin{aligned} b_0 &= 1, \\ b_l &= \sum_{\nu_1+2\nu_2+\ldots+l\nu_l=l} (\nu_1 + \ldots + \nu_l)! \prod_{s=1}^{l} \frac{(-1)^{\nu_s}}{\nu_s!} \beta_s^{\nu_s}, \quad l \geq 1, \end{aligned}$$

where
$$\beta_l = \sum_{\nu_1+2\nu_2+\ldots+l\nu_l=l} E\prod_{s=1}^{l} \frac{1}{\nu_s!}\left(\sum_{j=1}^{m-1} c_s(X_j)\right)^{\nu_s}, \quad l \geq 1.$$
The summation in $a_l(x)$, b_l and β_l is carried out over all integer-valued vectors $(\nu_1, \ldots, \nu_l)$ such that $\nu_1 + 2\nu_2 + \ldots + l\nu_l = l$, $\nu_j \geq 0$, $j = 1, \ldots, l$. Thus, for example,

$$\begin{aligned}
a_1(x) &= (m-1)Eh\left(X_1, \ldots, X_{m-1}, x\right) c_1\left(X_1\right), \\
a_2(x) &= (m-1)Eh\left(X_1, \ldots, X_{m-1}, x\right) c_2\left(X_1\right) + \\
&+\frac{1}{2}(m-1)Eh\left(X_1, \ldots, X_{m-1}, x\right) c_1^2\left(X_1\right) + \\
&+\binom{m-1}{2} Eh\left(X_1, \ldots, X_{m-1}, x\right) c_1\left(X_1\right) c_1\left(X_2\right),
\end{aligned}$$

and

$$\begin{aligned}
b_1 &= -\beta_1, \\
b_2 &= \beta_1^2 - \beta_2,
\end{aligned}$$

with

$$\begin{aligned}
\beta_1 &= (m-1)Ec_1\left(X_1\right), \\
\beta_2 &= (m-1)Ec_2\left(X_1\right) + \frac{1}{2}(m-1)Ec_1^2\left(X_1\right) + \binom{m-1}{2}\left(Ec_1(X_1)\right)^2.
\end{aligned}$$

Here, by definition $\binom{n}{k} = n\,(n-1)\,\ldots\,(n-k+1)\,/\,k!$, if $n \geq k$ and $\binom{n}{k} = 0$, if $n < k$. The expression on the right in (2) depends only on $c_1(x), \ldots, c_{k-1}(x)$. Hence (2) is a recurrence relation defining $c_k(x)$, $k = 1, 2, \ldots$. For example,

$$\begin{aligned}
c_1(x) &= g(x), \\
c_2(x) &= (m-1)Eh(X_1, \ldots, X_{m-1}, x)g(X_1), \\
c_3(x) &= \binom{m-1}{2} Eh(X_1, \ldots, X_{m-1}, x)g(X_1)g(X_2) \\
&+(m-1)^2 Eh(X_1, \ldots, X_{m-1}, x)h(X_1, X_m \ldots, X_{2m-2})g(X_m) \\
&+\frac{1}{2}(m-1)Eh(X_1, \ldots, X_{m-1}, x)g^2(X_1) \\
&-\left((m-1)^2 + \frac{1}{2}(m-1)\right)\, g(x)Eg^2(X_1).
\end{aligned}$$

Theorem 1. *Let* $\| h \|_\infty = c < \infty$. *Then for all* $x \in \mathbf{X}$ *and* $k = 1, 2, \ldots$

$$(3) \qquad | c_k(x) | \leq E \mid h(X_1, \ldots, X_{m-1}, x) \mid (c(m-1))^{k-1} \cdot 10^k$$

and the function

$$(4) \qquad g_t(x) = \sum_{k=1}^{\infty} c_k(x) t^{k-1}, \ x \in \mathbf{X}$$

is the unique analytical solution of the contraction equation (1) in the region $0 \leq t < (10c(m-1))^{-1}$. *Further*

$$(5) \, | g_t(x) - g(x) | \leq E \mid h(X_1, \ldots, X_{m-1}, x) \mid (10c(m-1)t) \, (1 - 10c(m-1)t)^{-1}.$$

Note that the contraction function $g_t(x)$ in (4) can be written as

$$g_t(x) = g(x) + \sum_{k=1}^{\infty} c_{k+1}(x) t^k$$

and, hence, the estimate (5) is true by virtue of (3). Moreover, throughout what follows we assume that $m \geq 2$ since in the case $m = 1$ we have $g_t(x) = g(x)$ for all $t > 0$.

Theorem 2. *Let* $\| h \|_\infty = c < \infty$ *and* $\sigma^2 = Eg^2(X_1) > 0$. *Suppose that* $x = o(\sqrt{n}), \ n \to \infty$. *Then*

$$(6) \quad P\left(U_n > x\sigma\sqrt{n}\right) / (1 - \Phi(x)) = \exp\left\{ \frac{x^3}{\sqrt{n}} \, \lambda_m\left(\frac{x}{\sqrt{n}}\right) \right\} \left(1 + O\left(\frac{1+x}{\sqrt{n}}\right)\right)$$

with Cramér's series

$$(7) \qquad \lambda_m(u) = \sum_{k=0}^{\infty} \lambda_{mk} u^k,$$

where

$$|\lambda_{mk}| \leq 2\delta^{k+3}, \quad k = 0, 1, \ldots$$

with

$$\delta = \left(400em(\frac{ce}{\sigma})^3\right)(1 + 20ecm).$$

Explicit formulae for evaluation of the coefficients λ_{mk} for any $k = 0, 1, \ldots$ are given at (25). The power series in (7) is convergent for $|u| < \delta$.

2 Proofs.

Proof of Theorem 1. The following lemma provides the inequalities needed to establish (3).

Lemma 1. *If* $\| h \|_\infty = c < \infty$, *then for all* $k = 1, 2, \ldots$

$$(8) \qquad | c_k(x) | \le c(x) \left(c(m-1)\right)^{k-1} \gamma_k, \quad x \in \mathbf{X}$$

where

$$c(x) = E \mid h(X_1, \ldots, X_{m-1}, x) \mid,$$

$$\gamma_{k+1} = \frac{1}{k+1} \sum_{\nu_1 + 2\nu_2 + \ldots + k\nu_k = k} \prod_{s=1}^{k} \frac{1}{\nu_s!} ((k+1)\alpha_s)^{\nu_s}, \quad k \ge 1,$$

$\gamma_1 = 1$, $\alpha_1 = 2$ *and*

$$\alpha_k = \sum_{\nu_1 + 2\nu_2 + \ldots + k\nu_k = k} (\nu_1 + \ldots + \nu_k - 1)! \prod_{s=1}^{k} \frac{1}{\nu_s!} \left(\frac{1}{s!}\right)^{\nu_s}, \quad k \ge 2.$$

Proof. We will use induction on k. Clearly

$$| c_1(x) | \le c(x)\gamma_1,$$
$$| c_2(x) | \le c(x) c(m-1)\gamma_2$$

since $\gamma_1 = 1$ and $\gamma_2 = 2$.

Now suppose that the inequalities (8) are valid for $s = 1, \ldots, k-1$, that is,

$$| c_s(x) | \le c(x) \left(c(m-1)\right)^{s-1} \gamma_s, \quad s = 1, \ldots, k-1.$$

Further we write by virtue of (2)

$$| c_k(x) | \le \sum_{l=0}^{k-1} | a_l(x) | \cdot | b_{k-l-1} |,$$

where by the induction hypothesis

$$\begin{aligned} | a_l(x) | &\le c(x) \sum_{\nu_1 + 2\nu_2 + \ldots + l\nu_l = l} \prod_{s=1}^{l} \frac{1}{\nu_s!} \left\{ \left(c(m-1)\right)^s \gamma_s \right\}^{\nu_s} \\ &= c(x) \left(c(m-1)\right)^l \sum_{\nu_1 + 2\nu_2 + \ldots + l\nu_l = l} \prod_{s=1}^{l} \frac{1}{\nu_s!} \gamma_s^{\nu_s} \\ &= c(x) \left(c(m-1)\right)^l \bar{a}_l, \end{aligned}$$

and

$$\begin{aligned}
|\beta_l| &\leq \sum_{\nu_1+2\nu_2+\ldots+l\nu_l=l} \prod_{s=1}^{l} \frac{1}{\nu_s!}\Big\{(c(m-1))^s\,\gamma_s\Big\}^{\nu_s} \\
&= (c(m-1))^l \sum_{\nu_1+2\nu_2+\ldots+l\nu_l=l} \prod_{s=1}^{l} \frac{1}{\nu_s!}\gamma_s^{\nu_s} \\
&= (c(m-1))^l\,\bar{a}_l,
\end{aligned}$$

and also

$$\begin{aligned}
|\,b_l\,| &\leq \sum_{\nu_1+2\nu_2+\ldots+l\nu_l=l} (\nu_1+\ldots+\nu_l)! \prod_{s=1}^{l} \frac{1}{\nu_s!}\Big\{(c(m-1))^s\,\bar{a}_s\Big\}^{\nu_s} \\
&= (c(m-1))^l \sum_{\nu_1+2\nu_2+\ldots+l\nu_l=l} (\nu_1+\ldots+\nu_l)! \prod_{s=1}^{l} \frac{1}{\nu_s!}\bar{a}_s^{\nu_s} \\
&= (c(m-1))^l\,\bar{b}_l,
\end{aligned}$$

setting $\bar{a}_0 = 1$ and $\bar{b}_0 = 1$.

Hence,

$$|\,c_k(x)\,| \leq c(x)\,(c(m-1))^{k-1} \sum_{l=0}^{k-1} \bar{a}_l\ \bar{b}_{k-l-1}. \tag{9}$$

It is sufficient to show that in (9)

$$\gamma_k = \sum_{l=0}^{k-1} \bar{a}_l\,\bar{b}_{k-l-1}. \tag{10}$$

Let

$$\gamma(z) = \sum_{k=1}^{\infty} \gamma_k\,z^k$$

be the generating function of the sequence $\{\gamma_k\}$. If (10) holds then

$$\begin{aligned}
\gamma(z) &= \sum_{k=1}^{\infty}\left(\sum_{l=0}^{k-1}\bar{a}_l\,\bar{b}_{k-l-1}\right) z^k = z\sum_{k=0}^{\infty}\left(\sum_{l=0}^{k}\bar{a}_l\,\bar{b}_{k-l}\right) z^k \\
&= z\left(\sum_{k=0}^{\infty}\bar{a}_k\,z^k\right)\cdot\left(\sum_{k=0}^{\infty}\bar{b}_k\,z^k\right)
\end{aligned}$$

where

$$\sum_{k=0}^{\infty}\bar{a}_k\,z^k = 1 + \sum_{k=1}^{\infty}\left(\sum_{\nu_1+2\nu_2+\ldots+k\nu_k=k}\prod_{s=1}^{k}\frac{1}{\nu_s!}\gamma_s^{\nu_s}\right) z^k = e^{\gamma(z)}$$

and

$$\begin{aligned}
\sum_{k=0}^{\infty}\bar{b}_k\,z^k &= 1 + \sum_{k=1}^{\infty}\left(\sum_{\nu_1+2\nu_2+\ldots+k\nu_k=k}(\nu_1+\ldots+\nu_k)!\prod_{s=1}^{k}\frac{1}{\nu_s!}\bar{a}_s^{\nu_s}\right) z^k \\
&= \left(1-\sum_{k=1}^{\infty}\bar{a}_k\,z^k\right)^{-1} = \frac{1}{2-e^{\gamma(z)}}.
\end{aligned}$$

Therefore from (10) we obtain the equation for the generating function

$$\gamma(z) = z \, \frac{e^{\gamma(z)}}{2 - e^{\gamma(z)}}, \quad \gamma(0) = 0. \tag{11}$$

Denote

$$\phi(y) = y - \ln(2 - e^{y}), \quad |y| < \ln 2. \tag{12}$$

Then equation (11) is reduced to the form

$$\gamma(z) = z \, \exp\{\phi(\gamma(z))\}. \tag{13}$$

Here

$$\phi(y) = \sum_{k=1}^{\infty} \alpha_k y^k, \quad |y| < \ln 2, \tag{14}$$

where

$$\begin{aligned} \alpha_1 &= 2, \\ \alpha_k &= \sum_{\nu_1 + 2\nu_2 + \ldots + k\nu_k = k} (\nu_1 + \ldots + \nu_k - 1)! \prod_{s=1}^{k} \frac{1}{\nu_s!} \left(\frac{1}{s!}\right)^{\nu_s}, \quad k \geq 2. \end{aligned}$$

This expansion follows from

$$\phi(y) = y - \ln\left(1 - \sum_{k=1}^{\infty} \frac{1}{k!} y^k\right) = \sum_{k=1}^{\infty} \alpha_k y^k.$$

Using the Lagrange expansion (see, Whittaker and Watson(1927)) we find the solution of equation (13) is

$$\gamma(z) = \sum_{k=1}^{\infty} \gamma_k \, z^k \tag{15}$$

with $\gamma_1 = 1$ and

$$\gamma_{k+1} = \frac{1}{k+1} \sum_{\nu_1 + 2\nu_2 + \ldots + k\nu_k = k} \prod_{s=1}^{k} \frac{1}{\nu_s!} \, ((k+1)\alpha_s)^{\nu_s}, \ k \geq 1.$$

According to (10) - (15) this proves Lemma 1.

Now we show that

$$\gamma_k \leq 10^k, \quad k \geq 1. \tag{16}$$

Write for all $k \geq 1$

$$\begin{aligned}
\gamma_{k+1} &= \frac{2^k}{k+1} \sum_{\nu_1+2\nu_2+\ldots+k\nu_k=k} \prod_{s=1}^{k} \frac{1}{\nu_s!} \left((k+1)\frac{\alpha_s}{2^s} \right)^{\nu_s} \\
&\leq 2^k \prod_{s=1}^{k} \left\{ \sum_{\nu_s=0}^{\infty} \frac{1}{\nu_s!} \left((k+1)\frac{\alpha_s}{2^s} \right)^{\nu_s} \right\} \\
&= 2^k \prod_{s=1}^{k} \exp\left\{ (k+1)\frac{\alpha_s}{2^s} \right\} \\
&\leq 2^k \exp\left\{ (k+1)\phi(\frac{1}{2}) \right\} \\
&\leq \left(\frac{2\sqrt{e}}{2-\sqrt{e}} \right)^{k+1} \leq 10^{k+1}.
\end{aligned}$$

From (8) and (16) we get the estimate (3).

Lemma 2. *If* $0 \leq t < (10c(m-1))^{-1}$ *then the function* $g_t(x)$ *defined in (4) is the solution of the contraction equation (1).*

Proof. Denote $z_t(x) = t\, g_t(x)$ and rewrite the contraction equation (1) in the form

$$z_t(x) = t \cdot \frac{E \prod_{s=1}^{m-1} e^{z_t(X_s)} h\left(X_1, \ldots, X_{m-1}, x\right)}{\left(E e^{z_t(X_1)}\right)^{m-1}}. \tag{17}$$

We must prove that the analytical function

$$z_t(x) = \sum_{k=1}^{\infty} c_k(x) t^k \tag{18}$$

satisfies equation (17). The numerator in (17) has the expansion

$$\begin{aligned}
&E \prod_{s=1}^{m-1} e^{z_t(X_s)} h\left(X_1, \ldots, X_{m-1}, x\right) \\
&= E \exp\Big\{ \sum_{k=1}^{\infty} \Big(\sum_{j=1}^{m-1} c_k(X_j) \Big) t^k \Big\} h\left(X_1, \ldots, X_{m-1}, x\right) \\
&= \sum_{k=0}^{\infty} \Big\{ \sum_{\nu_1+2\nu_2+\ldots+k\nu_k=k} E h\left(X_1, \ldots, X_{m-1}, x\right) \prod_{s=1}^{k} \frac{1}{\nu_s!} \Big(\sum_{j=1}^{m-1} c_s(X_j) \Big)^{\nu_s} \Big\} t^k \\
&= \sum_{k=0}^{\infty} a_k(x)\, t^k, \quad a_0(x) = g(x).
\end{aligned}$$

Similarly the denominator in (17) has the expansion

$$\begin{aligned}
&\left(E e^{z_t(X_1)}\right)^{-(m-1)} \\
&= \left(E \exp\Big\{ \sum_{k=1}^{\infty} \Big(\sum_{j=1}^{m-1} c_k(X_j) \Big) t^k \Big\} \right)^{-1}
\end{aligned}$$

$$= \left(1+\sum_{k=1}^{\infty}\left\{\sum_{\nu_1+2\nu_2+\ldots+k\nu_k=k} E\prod_{s=1}^{k}\frac{1}{\nu_s!}\left(\sum_{j=1}^{m-1}c_s(X_j)\right)^{\nu_s}\right\}t^k\right)^{-1}$$

$$= (1+\sum_{k=1}^{\infty}\beta_k t^k)^{-1}$$

$$= 1+\sum_{k=1}^{\infty}\left\{\sum_{\nu_1+2\nu_2+\ldots+k\nu_k=k}(\nu_1+\ldots+\nu_k)!\prod_{s=1}^{k}\frac{(-1)^{\nu_s}}{\nu_s!}\beta_s^{\nu_s}\right\}t^k$$

$$= \sum_{k=0}^{\infty} b_k\, t^k, \quad b_0=1.$$

Hence

$$\frac{E\prod_{s=1}^{m-1}e^{z_t(X_s)}h\left(X_1,\ldots,X_{m-1},x\right)}{\left(Ee^{z_t(X_1)}\right)^{m-1}} = \sum_{k=0}^{\infty}\left\{\sum_{l=0}^{k}a_l(x)\,b_{k-l}\right\}t^k$$

and thus

$$t\cdot\frac{E\prod_{s=1}^{m-1}e^{z_t(X_s)}h\left(X_1,\ldots,X_{m-1},x\right)}{\left(Ee^{z_t(X_1)}\right)^{m-1}} = \sum_{k=1}^{\infty}\left\{\sum_{l=0}^{k-1}a_l(x)\,b_{k-l-1}\right\}t^k. \tag{19}$$

Comparing (17) - (19) we see that for all $k=1,2,\ldots$

$$c_k(x) \;=\; \sum_{l=0}^{k-1}a_l(x)\,b_{k-1-l}, \quad x\in\mathbf{X}.$$

This formula coincides with (2). As a result we obtain the proof of Lemma 2.

Theorem 1 follows from Lemma 1 and Lemma 2.

Proof of Theorem 2. Define F_t by

$$dF_t = e^{tg_t-\psi(t)}dF,$$

where

$$\psi(t) = \ln E\exp\{tg_t(X_1)\}$$

and F is the distribution function of X_1. Let E_t denote the expectation under P_t, where P_t is a probability measure under which $X_1, X_2, \ldots$ are i.i.d. from F_t.

In the following we will need in the solution of the saddle-point equation

$$\mu_t = \frac{m\sigma x}{\sqrt{n}}, \tag{20}$$

where $\mu_t = E_t g_t(X_1)$ and $xn^{-1/2}\to 0$.

Lemma 3. *Let $t=t\left(\frac{x}{\sigma\sqrt{n}}\right)$ be the solution of equation (20). Then*

$$t=\sum_{k=1}^{\infty}\pi_k\left(\frac{x}{\sigma\sqrt{n}}\right)^k \tag{21}$$

is a convergent power series with $\pi_1 = 1$ *and coefficients* $\pi_k,\ k \geq 2$ *specified at (23).*

Proof. By definition

$$\mu_t = \frac{E\exp\Big\{t\,\sum_{j=1}^m g_t(X_j)\Big\}h\,(X_1,\ldots,X_m)}{E\exp\Big\{t\,\sum_{j=1}^m g_t(X_j)\Big\}}.$$

With the help of (4) we obtain the expansions

$$E\exp\Big\{t\sum_{j=1}^m g_t(X_j)\Big\}h\,(X_1,\ldots,X_m) = \sum_{k=1}^\infty d_k(h)t^k, \quad d_1(h) = m\sigma^2$$

and

$$E\exp\Big\{t\sum_{j=1}^m g_t(X_j)\Big\} = 1 + \sum_{k=1}^\infty d_k(1)t^k,$$

with

$$d_k(h) = \sum_{\nu_1+2\nu_2+\ldots+k\nu_k=k} Eh\,(X_1,\ldots,X_m)\prod_{s=1}^k\frac{1}{\nu_s!}\Big\{\sum_{j=1}^m c_s(X_j)\Big\}^{\nu_s}, \quad k\geq 1$$

where due to (3)

$$\begin{aligned}
|d_k(h)| &\leq c(10c\,m)^k\sum_{\nu_1+2\nu_2+\ldots+k\nu_k=k}\prod_{s=1}^k\frac{1}{\nu_s!}\\
&\leq c(10c\,m)^k\prod_{s=1}^k\left\{\sum_{\nu_s=0}^\infty\frac{1}{\nu_s!}\right\} = c(10ecm)^k, \quad k\geq 1.
\end{aligned}$$

Further

$$\begin{aligned}
\sum_{k=1}^\infty d_k(h)t^k &= tm\sigma^2\left(1+\sum_{k=1}^\infty\left(\frac{d_{k+1}(h)}{m\sigma^2}\right)t^k\right)\\
&= tm\sigma^2\exp\Big\{\ln\left(1+\sum_{k=1}^\infty\left(\frac{d_{k+1}(h)}{m\sigma^2}\right)t^k\right)\Big\}\\
&= tm\sigma^2\exp\Big\{\sum_{k=1}^\infty\mathcal{A}_k t^k\Big\},
\end{aligned}$$

where

$$\mathcal{A}_k = \sum_{\nu_1+2\nu_2+\ldots+k\nu_k=k}(\nu_1+\ldots+\nu_k-1)!\prod_{s=1}^k\frac{(-1)^{\nu_s-1}}{\nu_s!}\left(\frac{d_{s+1}(h)}{m\sigma^2}\right)^{\nu_s},$$

in addition

$$\begin{aligned}
|\mathcal{A}_k| &\leq \sum_{\nu_1+2\nu_2+\ldots+k\nu_k=k}(\nu_1+\ldots+\nu_k)!\prod_{s=1}^k\frac{1}{\nu_s!}\left(\frac{c}{m\sigma^2}(10ecm)^{s+1}\right)^{\nu_s}\\
&= (10ecm)^k\sum_{\nu_1+2\nu_2+\ldots+k\nu_k=k}(\nu_1+\ldots+\nu_k)!\prod_{s=1}^k\frac{1}{\nu_s!}\left(\frac{10ec^2}{\sigma^2}\right)^{\nu_s}\\
&= (10ecm)^k\left(\frac{10ec^2}{\sigma^2}\right)\left(1+\frac{10ec^2}{\sigma^2}\right)^{k-1}, \quad k\geq 1.
\end{aligned}$$

Further,

$$\left(1+\sum_{k=1}^{\infty} d_k(1)t^k\right)^{-1} = \exp\Big\{-\ln\left(1+\sum_{k=1}^{\infty} d_k(1)t^k\right)\Big\} = \exp\Big\{-\sum_{k=1}^{\infty} \mathcal{B}_k t^k\Big\},$$

where

$$\mathcal{B}_k = \sum_{\nu_1+2\nu_2+\ldots+k\nu_k=k} (\nu_1+\ldots+\nu_k-1)!\prod_{s=1}^{k} \frac{(-1)^{\nu_s-1}}{\nu_s!}(d_s(1))^{\nu_s},$$

in addition

$$\begin{aligned}
|\mathcal{B}_k| &\le \sum_{\nu_1+2\nu_2+\ldots+k\nu_k=k} (\nu_1+\ldots+\nu_k)!\prod_{s=1}^{k} \frac{1}{\nu_s!}(10ecm)^{s\nu_s} \\
&= (10ecm)^k \sum_{\nu_1+2\nu_2+\ldots+k\nu_k=k} (\nu_1+\ldots+\nu_k)!\prod_{s=1}^{k} \frac{1}{\nu_s!} \\
&= (10ecm)^k 2^{k-1}, \quad k \ge 1.
\end{aligned}$$

Hence the equation (20) is reduced to the form

$$t = \frac{x}{\sigma\sqrt{n}} \exp\{\mathcal{D}(t)\}, \tag{22}$$

where

$$\mathcal{D}(t) = \sum_{k=1}^{\infty} \mathcal{D}_k t^k,$$

with the coefficients $\mathcal{D}_k = \mathcal{B}_k - \mathcal{A}_k$, and

$$|\mathcal{D}_k| \le 2^k (10ecm)^k (10e(\frac{c}{\sigma})^2)^k, \quad k \ge 1.$$

By the Lagrange expansion we have the solution of equation (22)

$$t\left(\frac{x}{\sigma\sqrt{n}}\right) = \sum_{k=1}^{\infty} \pi_k \left(\frac{x}{\sigma\sqrt{n}}\right)^k,$$

where $\pi_1 = 1$ and

$$\pi_{k+1} = \frac{1}{k+1} \sum_{\nu_1+2\nu_2+\ldots+k\nu_k=k} \prod_{s=1}^{k} \frac{1}{\nu_s!}((k+1)\mathcal{D}_s)^{\nu_s}, \quad k \ge 1, \tag{23}$$

and due to (3)

$$\begin{aligned}
|\pi_{k+1}| &\le \frac{1}{k+1} \sum_{\nu_1+2\nu_2+\ldots+k\nu_k=k} \prod_{s=1}^{k} \frac{1}{\nu_s!}\left((k+1)(200cm(\frac{ce}{\sigma})^2)^s\right)^{\nu_s} \\
&= \frac{1}{k+1}(400cm(\frac{ce}{\sigma})^2)^k \sum_{\nu_1+2\nu_2+\ldots+k\nu_k=k} \prod_{s=1}^{k} \frac{1}{\nu_s!}\left((k+1)(\frac{1}{2})^s\right)^{\nu_s}
\end{aligned}$$

$$
\begin{aligned}
&\le \frac{1}{k+1}(400cm(\frac{ce}{\sigma})^2)^k \prod_{s=1}^{k}\left\{\sum_{\nu_s=0}^{\infty}\frac{1}{\nu_s!}\left((k+1)(\frac{1}{2})^s\right)^{\nu_s}\right\} \\
&\le \frac{1}{k+1}(400cm(\frac{ce}{\sigma})^2)^k \prod_{s=1}^{k}\exp\left\{(k+1)(\frac{1}{2})^s\right\} \\
&\le \frac{1}{k+1}(400cm(\frac{ce^2}{\sigma})^2)^k, \quad k \ge 1.
\end{aligned}
$$

This proves Lemma 3.

Let $U_n(t)$ denote the U-statistic U_n, where $X_1, X_2, \ldots$ are i.i.d. with the distribution F_t. Using the Hoeffding representation under P_t

$$U_n(t) = S_n(t) - \frac{n(m-1)}{m}\mu_t + v_n(t),$$

where

$$
\begin{aligned}
S_n(t) &= \sum_{i=1}^{n} g_t(X_i), \\
v_n(t) &= \frac{m-1}{n-1}\sum_{1\le i_1<i_2\le n} g_{2t}(X_{i_1}, X_{i_2}) + \cdots \\
&\quad + \frac{(m-1)\cdots 2\cdot 1}{(n-1)\ldots(n-m+1)}\sum_{1\le i_1<\cdots<i_m\le n} g_{mt}(X_{i_1},\ldots,X_{i_m}), \\
g_t(x) &= E_t(h(X_1,\cdots,X_{m-1},X_m)|X_m = x),
\end{aligned}
$$

where $g_{ct} : \mathbf{X}^c \to R,\ c = 2,\cdots,m$, are canonical functions satisfying $E_t g_{ct}(x_1,\cdots,x_{c-1},X_c) = 0$, a.s. and $X_1, X_2, \cdots$ are independent random variables all with an identical distribution function F_t .

In what follows t is the solution of the equation (20), that is t has the expansion (21). By virtue of (20)

$$U_n(t) - x\sigma\sqrt{n} = S_n(t) - n\mu_t + v_n(t)$$

and hence

$$
\begin{aligned}
&P(U_n > x\sigma\sqrt{n}) \\
&= \int\cdots\int I(U_n(y_1,\cdots,y_n) > x\sigma\sqrt{n})\prod_{i=1}^{n} dF(y_i)
\end{aligned}
$$

$$
\begin{aligned}
&= \int \cdots \int \exp\{\sum_{i=1}^{n}(\psi(t) - tg_t(y_i))\}\ I(U_n(y_1, \cdots, y_n) > x\sigma\sqrt{n}) \prod_{i=1}^{n} dF_t(y_i) \\
&= E_t[e^{n\psi(t) - tS_n(t)} I(U_n(t) > x\sigma\sqrt{n})] \\
(24) \quad &= e^{n\psi(t) - nt\mu_t} \cdot \mathcal{T}_n(t),
\end{aligned}
$$

where

$$\mathcal{T}_n(t) = E_t[e^{-t(S_n(t) - n\mu_t)} I(S_n(t) - n\mu_t + v_n(t) > 0)].$$

Due to (4) we have in (24) the following expansions for sufficiently small t

$$\psi(t) = \ln(1 + \sum_{k=2}^{\infty} \mathcal{G}_k t^k) = \sum_{k=2}^{\infty} \mathcal{H}_k t^k,$$

where

$$\mathcal{G}_k = \sum_{\nu_1 + 2\nu_2 + \ldots + k\nu_k = k} E \prod_{s=1}^{k} \frac{1}{\nu_s!} c_s^{\nu_s}(X_1),$$

$$\mathcal{H}_k = \sum_{\nu_1 + 2\nu_2 + \ldots + k\nu_k = k} (\nu_1 + \cdots + \nu_k - 1)! \prod_{s=1}^{k} \frac{(-1)^{\nu_s - 1}}{\nu_s!} \mathcal{G}_s^{\nu_s},$$

and by virtue of (3) for $k \geq 1$

$$|\mathcal{G}_k| \leq (10ecm)^k \quad \text{and} \quad |\mathcal{H}_k| \leq (10ecm)^k 2^{k-1}.$$

Since t is the solution of the saddle-point equation (20) then

$$\psi(t) = \sum_{k=2}^{\infty} \psi_k (\frac{x}{\sigma\sqrt{n}})^k$$

with coefficients

$$\psi_k = \sum_{\nu_1 + 2\nu_2 + \ldots + k\nu_k = k} (\nu_1 + \cdots + \nu_k)! \mathcal{H}_{\nu_1 + \cdots + \nu_k} \prod_{s=1}^{k} \frac{1}{\nu_s!} \pi_s^{\nu_s},$$

and, moreover, due to (3)

$$
\begin{aligned}
|\psi_k| &\leq \frac{1}{2} \sum_{\nu_1 + 2\nu_2 + \ldots + k\nu_k = k} (\nu_1 + \cdots + \nu_k)! (20ecm)^{\nu_1 + \cdots + \nu_k} \prod_{s=1}^{k} \frac{1}{\nu_s!} ((400cm(\frac{ce^2}{\sigma})^2)^s)^{\nu_s} \\
&= \frac{1}{2} (400cm(\frac{ce^2}{\sigma})^2)^k \sum_{\nu_1 + 2\nu_2 + \ldots + k\nu_k = k} (\nu_1 + \cdots + \nu_k)! \prod_{s=1}^{k} \frac{1}{\nu_s!} (20ecm)^{\nu_s} \\
&\leq (400cm(\frac{ce^2}{\sigma})^2)^k (20ecm)(1 + 20ecm)^{k-1}, \quad k \geq 1.
\end{aligned}
$$

Thus we have established that if $g_t(x)$ is the contraction function of the form (4) and $t = t\left(\frac{x}{\sigma\sqrt{n}}\right)$ is the solution of the saddle-point equation (20) then in (24) the following expansion is true

$$\begin{aligned} n(\psi(t) - t\mu_t) &= n\psi\left(t\left(\frac{x}{\sigma\sqrt{n}}\right)\right) - mt\left(\frac{x}{\sigma\sqrt{n}}\right)x\sigma\sqrt{n} \\ &= -\frac{x^2}{2} + \frac{x^3}{\sqrt{n}}\sum_{k=0}^{\infty}\lambda_{mk}\left(\frac{x}{\sqrt{n}}\right)^k = -\frac{x^2}{2} + \frac{x^3}{\sqrt{n}}\lambda_m\left(\frac{x}{\sqrt{n}}\right), \end{aligned} \tag{25}$$

defining Cramér's series (7) with the coefficients

$$\lambda_{mk} = \frac{1}{\sigma^{k+3}}\psi_{k+3} - \frac{m}{\sigma^{k+1}}\pi_{k+2}, \quad k = 0, 1, \ldots,$$

which satisfy the inequality

$$|\lambda_{mk}| \leq 2\left(400em(\frac{ce}{\sigma})^3(1 + 20ecm)\right)^{k+3}.$$

Next we need to estimate $\mathcal{T}_n(t)$ in (24). Denote

$$v_{ns}(t) = \frac{(m-1)\cdots(m-s+1)}{(n-1)\ldots(n-s+1)}\sum_{1\leq i_1<\cdots<i_s\leq n} g_{st}(X_{i_1}, \ldots, X_{i_s}).$$

Then

$$v_n(t) = v_{n2}(t) + \cdots + v_{nm}(t).$$

Lemma 4. *For all integers $n \geq m$, $m \geq 2$, all distribution functions, F_t, and all kernels h such that $\| h \|_\infty \leq c$*

$$E_t \exp\{\delta|v_n(t)|\} \leq d$$

for some positive constants δ and d depending only on m and c.

Proof. Write

$$\begin{aligned} E_t \exp\{\delta|v_n(t)|\} &\leq E_t \prod_{s=2}^{m} \exp\{\delta|v_{ns}(t)|\} \\ &\leq \sum_{s=2}^{m} q_s^{-1} E_t \exp\{\delta q_s|v_{ns}(t)|\}, \end{aligned}$$

since for any $x_i \in R$ and $q_2^{-1} + \cdots + q_m^{-1} = 1$,

$$|x_2 \cdots x_m| \leq q_2^{-1}|x_2|^{q_2} + \cdots + q_m^{-1}|x_m|^{q_m}. \tag{26}$$

By the inequality of Arcones and Giné (1993) for any $x > 0$ there are constants γ_i depending only on s such that for all kernels h with $\| h \|_\infty \leq c$

$$P_t(|n^{-\frac{s}{2}} \sum_{1 \leq i_1 < \cdots < i_s \leq n} g_{st}(X_{i_1}, \ldots, X_{i_s})| > x) \leq \gamma_1 \exp(-\gamma_2 (x/c)^{2/s}).$$

Hence,

$$E_t \exp\{\delta n^{-1} | \sum_{1 \leq i_1 < \cdots < i_s \leq n} g_{st}(X_{i_1}, \ldots, X_{i_s})|^{2/s}\} \leq d < \infty,$$

for some positive constants δ and d. Further, for all $s \geq 2$ and $n \geq m$

$$\begin{aligned} |v_{ns}(t)| &\leq \delta(m) |n^{-s+1} \sum_{1 \leq i_1 < \cdots < i_s \leq n} g_{st}(X_{i_1}, \ldots, X_{i_s})| \\ &= \delta(m) [n^{-s+2} | \sum_{1 \leq i_1 < \cdots < i_s \leq n} g_{st}(X_{i_1}, \ldots, X_{i_s})|^{(s-2)/s}] \\ &\quad \times [n^{-1} | \sum_{1 \leq i_1 < \cdots < i_s \leq n} g_{st}(X_{i_1}, \ldots, X_{i_s})|^{2/s}] \\ &\leq \delta_1(m) c^{(s-2)/s} [n^{-1} | \sum_{1 \leq i_1 < \cdots < i_s \leq n} g_{st}(X_{i_1}, \ldots, X_{i_s})|^{2/s}]. \end{aligned}$$

This proves the desired result.

Now we will apply Lemma 4 to the estimation of $\mathcal{T}_n(t)$. We introduce the following notation where the dependence on t has been suppressed. Let

$$\begin{aligned} \tilde{S} &= \frac{1}{\sigma_t \sqrt{n}} \sum_{i=1}^{n} (g_t(X_i) - \mu_t), \\ \tilde{v}_n &= \frac{1}{\sigma_t \sqrt{n}} v_n(t), \\ \Phi_n(x) &= P_t(\tilde{S} + \tilde{v}_n \leq x), \ x \in R, \end{aligned}$$

$$\sigma_t^2 = E_t (g_t(X_1) - \mu_t)^2, \ \ x_t = \sqrt{n} t \sigma_t.$$

Then $\mathcal{T}_n(t)$ can be written as

$$\mathcal{T}_n(t) = E_t e^{-x_t \tilde{S}} \ I(\tilde{S} + \tilde{v}_n > 0).$$

Further we will use the formula

$$e^z = \sum_{k=0}^{r-1} \frac{z^k}{k!} + \frac{z^r}{(r-1)!} \int_0^1 (1-u)^{r-1} e^{uz} du, \tag{27}$$

which is valid for all complex z and positive integers $r = 1, 2, \ldots$.

With the help of (27) we have the following representation

$$\begin{aligned}
\mathcal{T}_n(t) &= E_t e^{x_t \tilde{v}_n} \cdot \left\{ e^{-x_t(\tilde{S}+\tilde{v}_n)} \; I(\tilde{S} + \tilde{v}_n > 0) \right\} \\
&= E_t \left\{ e^{-x_t(\tilde{S}+\tilde{v}_n)} \; I(\tilde{S} + \tilde{v}_n > 0) \right\} \\
&\quad + x_t \int_0^1 E_t \tilde{v}_n e^{u x_t \tilde{v}_n} \cdot \left\{ e^{-x_t(\tilde{S}+\tilde{v}_n)} \; I(\tilde{S} + \tilde{v}_n > 0) \right\} du \\
&= I_n(t) + t \int_0^1 I_n(u,t) du, \qquad (28)
\end{aligned}$$

where

$$I_n(t) = E_t \left\{ e^{-x_t(\tilde{S}+\tilde{v}_n)} \; I(\tilde{S} + \tilde{v}_n > 0) \right\},$$

$$I_n(u,t) = E_t v_n(t) e^{u t v_n(t)} \cdot \left\{ e^{-x_t(\tilde{S}+\tilde{v}_n)} \; I(\tilde{S} + \tilde{v}_n > 0) \right\}.$$

Estimate $\mathbf{I_n(t)}$.

By definition

$$\begin{aligned}
I_n(t) &= \int e^{-x_t y} \; I(y > 0) d\Phi_n(y) \\
&= \int_0^\infty e^{-x_t y} d\Phi_n(y) \\
&= \int_0^\infty e^{-x_t y} d\Phi(y) + \int_0^\infty e^{-x_t y} d(\Phi_n(y) - \Phi(y)). \qquad (29)
\end{aligned}$$

Here by (4), (21) and (27)

$$\begin{aligned}
\int_0^\infty e^{-x_t y} d\Phi(y) &= \int_0^\infty e^{-x(1+t\cdot O(1))y} d\Phi(y) \\
&= \int_0^\infty e^{-xy} d\Phi(y) \\
&\quad - \int_0^1 \left\{ xt \cdot O(1) \int_0^\infty y e^{-xy - xtyu \cdot O(1)} d\Phi(y) \right\} du.
\end{aligned}$$

Further,

$$\int_0^\infty e^{-xy} d\Phi(y) = e^{x^2/2} \cdot (1 - \Phi(x))$$

and since $t = \frac{x}{\sigma\sqrt{n}} \left(1 + O\left(\frac{x}{\sqrt{n}}\right)\right)$ and $e^{-y^2/2} \leq 1$

$$xt \int_0^\infty y e^{-xy - xtyu \cdot O(1)} d\Phi(y) \leq \frac{x^2}{\sigma\sqrt{n}} (1 + O(1)) \int_0^\infty y e^{-(x + xtuO(1))y} dy.$$

After the substitution $z = (x + xtu \cdot O(1))y$

$$xt \int_0^\infty y e^{-xy - xtyu \cdot O(1)} d\Phi(y)$$

$$
\begin{aligned}
&\leq \frac{x^2}{\sigma\sqrt{n}}(1+O(1))(x+xtu\cdot O(1))^{-2}\int_0^\infty e^{-z}zdz,\\
&= e^{x^2/2}(1-\Phi(x))\cdot O\left(\frac{1+x}{\sqrt{n}}\right).
\end{aligned}
$$

Consequently,

$$
\int_0^\infty e^{-x_t y}d\Phi(y) = e^{x^2/2}(1-\Phi(x))\left(1+O\left(\frac{1+x}{\sqrt{n}}\right)\right). \tag{30}
$$

Then we have in (29)

$$
\int_0^\infty e^{-x_t y}d(\Phi_n(y)-\Phi(y)) = -(\Phi_n(0)-\Phi(0)) + x_t\cdot\int_0^\infty(\Phi_n(y)-\Phi(y))e^{-x_t y}dy.
$$

By the normal approximation theorem for U-statistics (see Borovskikh (1996), (1998))

$$
\sup_y|\Phi_n(y)-\Phi(y)| \leq A_m(a_t+b_t)\frac{1}{\sqrt{n}}, \tag{31}
$$

where A_m is a positive constant depending only on m, and

$$
\begin{aligned}
a_t &= \sigma_t^{-3/2}\cdot E_t|g_t(X_1)-\mu_t|^3,\\
b_t &= \sum_{s=2}^m \sigma_t^{-\delta_s}\cdot E_t|g_{st}(X_1,\ldots,X_s)|^{\delta_s},
\end{aligned}
$$

with $\delta_s = (2s+1)/(2s-1)$.

By applying (4), (21) and (31) we have

$$
\begin{aligned}
a_t &\leq 2^3c^3\sigma^{-3/2}(1+t\cdot O(1)),\\
b_t &\leq \sum_{s=2}^m (2^{s+1}c)^{\delta_s}\sigma^{-\delta_s}(1+t\cdot O(1)).
\end{aligned}
$$

Hence we get the following estimate

$$
|\int_0^\infty e^{-x_t y}d(\Phi_n(y)-\Phi(y))| \leq e^{x^2/2}(1-\Phi(x))\cdot O\left(\frac{1+x}{\sqrt{n}}\right). \tag{32}
$$

From (29), (30) and (32) we see that

$$
I_n(t) = e^{x^2/2}(1-\Phi(x))\cdot\left(1+O\left(\frac{1+x}{\sqrt{n}}\right)\right). \tag{33}
$$

Estimate $\mathbf{I_n(u,t)}$.

First we will prove that for any $m \geq 2$

$$
|I_n(u,t)| \leq c_1\mathcal{R}(t) + c_2\frac{1}{\sqrt{n}}, \tag{34}
$$

where

$$\mathcal{R}(t) = E_t\{e^{c_3|v_{n2}(t)|}[e^{-x_t(\tilde{S}+\tilde{v}_{n2})} \cdot I(\tilde{S} + \tilde{v}_{n2} + \varepsilon > 0)]\}$$

for some positive constants c_i and $\varepsilon = n^{-\frac{1}{2}}$. For $0 < u < 1$, since $x \leq \exp(x),\ x > 0$,

$$|I_n(u,t)| \leq c_1^{-1} E_t\{e^{(c_1+t)(|v_{n2}(t)|+|r_n|)}\ [e^{-x_t(\tilde{S}+\tilde{v}_{n2})} I(\tilde{S} + \tilde{v}_{n2} + \tilde{r}_n > 0)]\}$$

where $r_n = v_{n3}(t) + \cdots + v_{nm}(t)$ and $\tilde{r}_n = r_n/(\sigma_t\sqrt{n})$.

Applying the inequalities $e^x - 1 \leq xe^x,\ x \geq 0$ with $x = (c_1 + t)|r_n|$ and $\tilde{r}_n > -(\tilde{S} + \tilde{v}_{n2})$ we get

$$\begin{aligned}
|I_n(u,t)| &\leq c_1^{-1} E_t\{e^{(c_1+t)|v_{n2}(t)|}[e^{-x_t(\tilde{S}+\tilde{v}_{n2})} I(\tilde{S} + \tilde{v}_{n2} + \tilde{r}_n > 0)]\} \\
&+ c_1^{-1}(c_1 + t) E_t\{|r_n| \exp((c_1 + t)|v_{n2}(t)| + (c_1 + 2t)|r_n|)\}.
\end{aligned}$$

By Hölder's inequality and Lemma 4

$$E_t\{|r_n| \exp((c_1 + t)|v_{n2}(t)| + (c_1 + 2t)|r_n|)\} \leq (E_t r_n^2)^{\frac{1}{2}} \sqrt{d}.$$

Here

$$\begin{aligned}
E_t r_n^2 &\leq (m-2)\delta^2(m) n^{-1} \sum_{s=3}^{m} E_t |n^{-s/2} \sum_{1 \leq i_1 < \cdots < i_s \leq n} g_{st}(X_{i_1}, \ldots, X_{i_s})|^2 \\
&= (m-2)\delta^2(m) n^{-1} \sum_{s=3}^{m} \binom{n}{s} n^{-s} E_t g_{st}^2(X_1, \ldots, X_s).
\end{aligned}$$

Further write

$$\begin{aligned}
&E_t\{e^{(c_1+t)|v_{n2}(t)|}[e^{-x_t(\tilde{S}+\tilde{v}_{n2})} I(\tilde{S} + \tilde{v}_{n2} + \tilde{r}_n > 0)]\} \\
&\quad = E_t\{e^{(c_1+t)|v_{n2}(t)|}[e^{-x_t(\tilde{S}+\tilde{v}_{n2})} I(\tilde{S} + \tilde{v}_{n2} + \tilde{r}_n > 0) I(|\tilde{r}_n| \leq \varepsilon)]\} \\
&\qquad + E_t\{e^{(c_1+t)|v_{n2}(t)|}[e^{-x_t(\tilde{S}+\tilde{v}_{n2})} I(\tilde{S} + \tilde{v}_{n2} + \tilde{r}_n > 0) I(|\tilde{r}_n| > \varepsilon)]\} \\
&\quad \leq \mathcal{R}(t) + E_t\{e^{(c_1+t)|v_{n2}(t)|+t|r_n|} I(|\tilde{r}_n| > \varepsilon)\} \\
&\quad \leq \mathcal{R}(t) + e^{-\sigma_t^{2/3} n^{1/3}} E_t\{e^{(c_1+t)|v_{n2}(t)|+t|r_n|+(n^{1/2}|r_n|)^{2/3}}\}.
\end{aligned}$$

Since

$$(n^{1/2}|r_n|)^{2/3} \leq \delta_2(m) \sum_{s=3}^{m} (n^{-1} | \sum_{1 \leq i_1 < \cdots < i_s \leq n} g_{st}(X_{i_1}, \ldots, X_{i_s})|^{2/s})$$

(34) follows from Lemma 4.

In what follows we shall estimate $\mathcal{R}(t)$. We do this by partitioning $v_{n2}(t)$ into separate terms. This leads to the desired estimate

$$\mathcal{R}(t) \le c_4 x^{-1}. \tag{35}$$

For $n \to \infty$ let $k = [n/2],\ k_1 = [k/2], k_2 = k - k_1,\ n_1 = [(n-k)/2]$ and $n_2 = n - k - n_1$, where $[z]$ denotes the integer part of z. Then

$$v_{n2}(t) = v_k^1 + v_{n-k}^2 + v_{k_1,n-k}^3 + v_{k_2,n-k}^4, \tag{36}$$

where

$$\begin{aligned}
v_k^1 &= \frac{m-1}{n-1} \sum_{1 \le i_1 < i_2 \le k} g_{2t}(X_{i_1}, X_{i_2}),\\
v_{n-k}^2 &= \frac{m-1}{n-1} \sum_{k+1 \le i_1 < i_2 \le n} g_{2t}(X_{i_1}, X_{i_2}),\\
v_{k_1,n-k}^3 &= \frac{m-1}{n-1} \sum_{i_1=1}^{k_1} \sum_{i_2=k+1}^{n} g_{2t}(X_{i_1}, X_{i_2}),\\
v_{k_2,n-k}^4 &= \frac{m-1}{n-1} \sum_{i_1=k_1+1}^{k} \sum_{i_2=k+1}^{n} g_{2t}(X_{i_1}, X_{i_2}).
\end{aligned}$$

Further let

$$v_{k_1,n-k}^3 = v_{k_1 n_1}^{31} + v_{k_1 n_2}^{32}, \tag{37}$$

where

$$\begin{aligned}
v_{k_1 n_1}^{31} &= \frac{m-1}{n-1} \sum_{i_1=1}^{k_1} \sum_{i_2=k+1}^{n_1} g_{2t}(X_{i_1}, X_{i_2}),\\
v_{k_1 n_2}^{32} &= \frac{m-1}{n-1} \sum_{i_1=1}^{k_1} \sum_{i_2=n_1+1}^{n} g_{2t}(X_{i_1}, X_{i_2}),
\end{aligned}$$

and

$$v_{k_2,n-k}^4 = v_{k_2 n_1}^{41} + v_{k_2 n_2}^{42}, \tag{38}$$

where

$$\begin{aligned}
v_{k_2 n_1}^{41} &= \frac{m-1}{n-1} \sum_{i_1=k_1+1}^{k} \sum_{i_2=k+1}^{n_1} g_{2t}(X_{i_1}, X_{i_2}),\\
v_{k_2 n_2}^{42} &= \frac{m-1}{n-1} \sum_{i_1=k_1+1}^{k} \sum_{i_2=n_1+1}^{n} g_{2t}(X_{i_1}, X_{i_2}).
\end{aligned}$$

We also define the corresponding partitioned terms for the normalized quantity $\tilde{v}_{n2}$. For any real a, b and c we have the following inequality

$$I(a+b>c) \leq I(2a>c) + I(2b>c). \tag{39}$$

Hence by (36) and four applications of the inequality (39)

$$\begin{aligned} I(\tilde{S}+\tilde{v}_{n2}+\varepsilon>0) &\leq I(\tilde{S}+2\tilde{v}_k^1>0) \\ &\quad +I(\tilde{S}+2\tilde{v}_{n-k}^2+2\tilde{v}_{k_1,n-k}^3+2\tilde{v}_{k_2,n-k}^4+2\varepsilon>0) \\ &\leq I(\tilde{S}+2\tilde{v}_k^1>0)+I(\tilde{S}+4\tilde{v}_{n-k}^2>0) \\ &\quad +I(\tilde{S}+4\tilde{v}_{k_1,n-k}^3+4\tilde{v}_{k_2,n-k}^4+4\varepsilon>0) \\ &\leq I(\tilde{S}+2\tilde{v}_k^1>0)+I(\tilde{S}+4\tilde{v}_{n-k}^2>0) \\ &\quad +I(\tilde{S}+8\tilde{v}_{k_1,n-k}^3>0)+I(\tilde{S}+16\tilde{v}_{k_2,n-k}^4>0) \\ &\quad +I(\tilde{S}+16\varepsilon>0). \end{aligned} \tag{40}$$

Using (40) we can bound $\mathcal{R}(t)$ as follows

$$\begin{aligned} \mathcal{R}(t) &\leq E_t\{e^{c_3|v_{n2}(t)|}[e^{-x_t(\tilde{S}+\tilde{v}_{n2})}I(\tilde{S}+2\tilde{v}_k^1>0)]\} \\ &\quad +E_t\{e^{c_3|v_{n2}(t)|}[e^{-x_t(\tilde{S}+\tilde{v}_{n2})}I(\tilde{S}+4\tilde{v}_{n-k}^2>0)]\} \\ &\quad +E_t\{e^{c_3|v_{n2}(t)|}[e^{-x_t(\tilde{S}+\tilde{v}_{n2})}I(\tilde{S}+8\tilde{v}_{k_1,n-k}^3>0)]\} \\ &\quad +E_t\{e^{c_3|v_{n2}(t)|}[e^{-x_t(\tilde{S}+\tilde{v}_{n2})}I(\tilde{S}+16\tilde{v}_{k_2,n-k}^4>0)]\} \\ &\quad +E_t\{e^{c_3|v_{n2}(t)|}[e^{-x_t(\tilde{S}+\tilde{v}_{n2})}I(\tilde{S}+16\varepsilon>0)]\}. \end{aligned} \tag{41}$$

In (41) each expectation has the same structure. Consider the first expectation in (41). By (26) and (36)

$$\begin{aligned} e^{c_3|v_{n2}(t)|}[e^{-x_t(\tilde{S}+\tilde{v}_{n2})}] &= e^{c_3|v_{n2}(t)|+tv_k^1-tv_{n-k}^2-tv_{k_1,n-k}^3-tv_{k_2,n-k}^4}e^{-x_t(\tilde{S}+2\tilde{v}_k^1)} \\ &\leq \sum_{i=1}^4 q_{i1}^{-1}e^{\delta_{i1}|v^i|}[e^{-x_t(\tilde{S}+2\tilde{v}^1)}], \end{aligned}$$

where for simplicity we write $v^1=v_k^1$, $v^2=v_{n-k}^2$ and so forth. Hence from (41) we obtain

$$\mathcal{R}(t) \leq \sum_{i=1}^4\sum_{j=1}^4 q_{ij}^{-1}Q_{ij}+\sum_{i=1}^4 q_i^{-1}Q_i, \tag{42}$$

where

$$
\begin{aligned}
Q_{ij} &= E_t\{e^{\delta_{ij}|v^i|}[e^{-x_t(\tilde{S}+2^j\tilde{v}^j)}I(\tilde{S}+2^j\tilde{v}^j>0)]\},\\
Q_i &= E_t\{e^{\delta_i|v^i|}[e^{-x_t\tilde{S}}I(\tilde{S}+16\varepsilon>0)]\},
\end{aligned}
$$

and $\delta_i,\ \delta_{ij}$ are positive constants and $\varepsilon = n^{-1/2}$.

We shall use (37) and (38) to further partition the expectations when i or j takes the value 3 or 4. Thus using (26), for some constants $q_i,\ \delta_j^i$,

$$
\begin{aligned}
Q_{31} &\le q_1^{-1}E_t\{e^{\delta_1^1|v^{31}|}[e^{-x_t(\tilde{S}+2\tilde{v}^1)}I(\tilde{S}+2\tilde{v}^1>0)]\}\\
&+ q_2^{-1}E_t\{e^{\delta_1^2|v^{32}|}[e^{-x_t(\tilde{S}+2\tilde{v}^1)}I(\tilde{S}+2\tilde{v}^1>0)]\},
\end{aligned}
\tag{43}
$$

with a similar expression for Q_{41}. For terms like Q_{13} and Q_{14} we have

$$
\begin{aligned}
Q_{13} &\le q_1^{-1}E_t\{e^{\delta_2^1|v^1|}[e^{-x_t(\tilde{S}+16\tilde{v}^{31})}I(\tilde{S}+16\tilde{v}^{31}>0)]\}\\
&+ q_2^{-1}E_t\{e^{\delta_2^2|v^1|}[e^{-x_t(\tilde{S}+16\tilde{v}^{32})}I(\tilde{S}+16\tilde{v}^{32}>0)]\}\\
&+ q_3^{-1}E_t\{e^{\delta_2^3|v^{31}|}[e^{-x_t(\tilde{S}+16\tilde{v}^{31})}I(\tilde{S}+16\tilde{v}^{31}>0)]\}\\
&+ q_4^{-1}E_t\{e^{\delta_2^4|v^{31}|}[e^{-x_t(\tilde{S}+16\tilde{v}^{32})}I(\tilde{S}+16\tilde{v}^{32}>0)]\},\\
&+ q_5^{-1}E_t\{e^{\delta_2^5|v^{32}|}[e^{-x_t(\tilde{S}+16\tilde{v}^{31})}I(\tilde{S}+16\tilde{v}^{31}>0)]\},\\
&+ q_6^{-1}E_t\{e^{\delta_2^6|v^{32}|}[e^{-x_t(\tilde{S}+16\tilde{v}^{32})}I(\tilde{S}+16\tilde{v}^{32}>0)]\},
\end{aligned}
\tag{44}
$$

and for terms like Q_{34} and Q_{43} we have

$$
\begin{aligned}
Q_{43} &\le q_1^{-1}E_t\{e^{\delta_3^1|v^{41}|}[e^{-x_t(\tilde{S}+16\tilde{v}^{31})}I(\tilde{S}+16\tilde{v}^{31}>0)]\}\\
&+ q_2^{-1}E_t\{e^{\delta_3^2|v^{41}|}[e^{-x_t(\tilde{S}+16\tilde{v}^{32})}I(\tilde{S}+16\tilde{v}^{32}>0)]\}\\
&+ q_3^{-1}E_t\{e^{\delta_3^3|v^{42}|}[e^{-x_t(\tilde{S}+16\tilde{v}^{31})}I(\tilde{S}+16\tilde{v}^{31}>0)]\}\\
&+ q_4^{-1}E_t\{e^{\delta_3^4|v^{42}|}[e^{-x_t(\tilde{S}+16\tilde{v}^{32})}I(\tilde{S}+16\tilde{v}^{32}>0)]\}\\
&+ q_5^{-1}E_t\{e^{\delta_3^5|v^{31}|}[e^{-x_t(\tilde{S}+16\tilde{v}^{31})}I(\tilde{S}+16\tilde{v}^{31}>0)]\}\\
&+ q_6^{-1}E_t\{e^{\delta_3^6|v^{31}|}[e^{-x_t(\tilde{S}+16\tilde{v}^{32})}I(\tilde{S}+16\tilde{v}^{32}>0)]\}\\
&+ q_7^{-1}E_t\{e^{\delta_3^7|v^{32}|}[e^{-x_t(\tilde{S}+16\tilde{v}^{31})}I(\tilde{S}+16\tilde{v}^{31}>0)]\}\\
&+ q_8^{-1}E_t\{e^{\delta_3^8|v^{32}|}[e^{-x_t(\tilde{S}+16\tilde{v}^{32})}I(\tilde{S}+16\tilde{v}^{32}>0)]\}.
\end{aligned}
\tag{45}
$$

All of the terms we need to bound to obtain (35) have similar structure. Let

$$
Q = E_t\{e^{\delta|\Delta_n|}[e^{-x_t(\xi_n+\eta_n)}I(\xi_n+\eta_n>0)]\},
\tag{46}
$$

where

$$\xi_n = \frac{1}{\sigma_t\sqrt{n}}\sum_{j\in J} g_t(X_j) + \tilde{\Delta}_{n1},$$

$$\eta_n = \frac{1}{\sigma_t\sqrt{n}}\sum_{j\in \bar{J}} g_t(X_j) + \tilde{\Delta}_{n2},$$

Δ_n is part of the sum v_n, $\tilde{\Delta}_{n1}$ and $\tilde{\Delta}_{n2}$ are parts of the normalized sum $\tilde{v}_n$, and $\bar{J} = \{1,\dots,n\}\backslash J$. The set of indices J has cardinality $|J| \sim \lambda n,\ 0 < \lambda < 1$ as $n \to \infty$. Here the random variable ξ_n is independent of Δ_n and η_n.

The expectations Q_i in (42) have a simpler form than (46). All expectations (43) - (45) have the structure (46). The following lemma provides a bound for Q in (46).

Lemma 5. *Under the conditions of Theorem 2*

$$Q \le c_4 x^{-1} + c_5 n^{-1/2}$$

for some positive constants c_4 and c_5.

Proof. Let $\lambda_n = (|J|/n)^{1/2}$ and

$$\tilde{\xi}_n = \frac{1}{\sigma_t\sqrt{|J|}}\sum_{j\in J} g_t(X_j) + \lambda_n^{-1}\tilde{\Delta}_{n1}$$

and $\tilde{\Phi}_n(u) = P(\tilde{\xi}_n \le u),\ u \in R$. Further, in (46), since ξ_n is independent of Δ_n and η_n we can write

$$Q = E_t\{e^{\delta|\Delta_n| - x_t\eta_n} E^{\xi} e^{-x_t\xi_n} I(\xi_n + \eta_n > 0)\}, \tag{47}$$

where $E^{\xi}\{\cdots\}$ denotes the expectation with respect to ξ_n. Here (see the proof of (33)),

$$\begin{aligned} &E^{\xi} e^{-x_t\xi_n} I(\xi_n + \eta_n > 0) \\ &= \int_{-\infty}^{\infty} e^{-x_t\lambda_n u} I(\lambda_n u + \eta_n > 0) d\tilde{\Phi}_n(u) \end{aligned}$$

$$= \int_{-\eta_n\lambda_n^{-1}}^{\infty} e^{-x_t\lambda_n u} d(\tilde{\Phi}_n(u) - \Phi(u)) + \int_{-\eta_n\lambda_n^{-1}}^{\infty} e^{-x_t\lambda_n u} d\Phi(u). \tag{48}$$

In (48)

$$\int_{-\eta_n\lambda_n^{-1}}^{\infty} e^{-x_t\lambda_n u} d(\tilde{\Phi}_n(u) - \Phi(u))$$

$$= \quad -e^{x_t\eta_n}(\tilde{\Phi}_n(-\eta_n\lambda_n^{-1}) - \Phi(-\eta_n\lambda_n^{-1})) + x_t\lambda_n \int_{-\eta_n\lambda_n^{-1}}^{\infty} e^{-x_t\lambda_n u}(\tilde{\Phi}_n(u) - \Phi(u))du$$

and by the inequality (31)

$$\left|\int_{-\eta_n\lambda_n^{-1}}^{\infty} e^{-x_t\lambda_n u} d(\tilde{\Phi}_n(u) - \Phi(u))\right| \leq c_6 e^{x_t\eta_n}\frac{1}{\sqrt{n}}.$$

Further in (48)

$$\begin{aligned} \int_{-\eta_n\lambda_n^{-1}}^{\infty} e^{-x_t\lambda_n u} d\Phi(u) &= \frac{1}{\sqrt{2\pi}}\int_{-\eta_n\lambda_n^{-1}}^{\infty} e^{-x_t\lambda_n u}e^{-u^2/2}du \\ &\leq \int_{-\eta_n\lambda_n^{-1}}^{\infty} e^{-x_t\lambda_n u}du \\ &= (x_t\lambda_n)^{-1}e^{x_t\eta_n}, \end{aligned}$$

where $x_t \sim x,\ \lambda_n \sim \lambda$ as $n \to \infty$. Hence in (47)

$$Q \leq (c_6\frac{1}{\sqrt{n}} + c_7\frac{1}{x})Ee^{\delta|\Delta_n|}. \tag{49}$$

Note that for the expectation of Q_i in (42) we can replace (49) by the inequality

$$Q \leq (c_8\frac{1}{\sqrt{n}} + c_9\frac{1}{x})Ee^{\delta|\Delta_n| + \delta^1|s_n|},$$

where $s_n = (\sigma_t\sqrt{|\bar{J}|})^{-1}\sum_{j\in\bar{J}} g_t(X_j)$. Here Hölder's inequality, Lemma 4 and the well known inequality $E\exp(\delta^1|s_n|) \leq d_1 < \infty$ for bounded random variables $g_t(X_j),\ j \in \bar{J}$ provides the required bound for Q. From (49) and Lemma 4 we obtain the proof of Lemma 5, and hence (35). Combining (24) - (25), (28) and (33) - (35) completes the proof of Theorem 2.

REFERENCES

Aleshkevichiene, A.K. (1990). Probabilities of large deviations for U-statistics and von Mises functionals, *Theor. Ver. i Prim.* **35**, 1-14.

Arcones, M.A. and Giné, E. (1993). Limit theorems for U-processes. *Ann. Probab.* **21**, 1494-1542.

Borovskikh, Yu.V. (1996). *U-statistics in Banach Spaces.* VSP, Utrecht, The Netherlands, xii + 420 p.

Borovskikh, Yu.V. (1998). Sharp estimates of the rate of convergence for U-statistics. Report 98-21, School of Mathematics and Statistics, University of Sydney. +42p.

Keener, R.W., Robinson, J., and Weber, N.C. (1998). Tail probability approximations for U-statistics, *Statist. and Probab. Letters*, **37**, 59-65.

Lang, S. (1994). *Real and Functional Analysis.* Springer-Verlag, New York, 3rd edition.

Whittaker, E.T. and Watson, G.N. (1927). *A Course of Modern Analysis.* Cambridge at the University Press, 4th edition.

Department of Applied Mathematics
Transport University
Moskovsky Avenue, 9
190031, St.Petersburg
Russia
e-mail:borovski@spb.cityline.ru

School of Mathematics
and Statistics, FO7
University of Sydney
NSW 2006, Australia
e-mail:neville@maths.usyd.edu.au