STABILITY IN PROBABILITY
BANACH CENTER PUBLICATIONS, VOLUME 90
INSTITUTE OF MATHEMATICS
POLISH ACADEMY OF SCIENCES
WARSZAWA 2010

ASYMPTOTIC RATE OF CONVERGENCE IN THE DEGENERATE U-STATISTICS OF SECOND ORDER

OLGA YANUSHKEVICHIENE

Vilnius Pedagogical University
Studentu 39, Vilnius LT-08106, Lithuania, and
Institute of Mathematics and Informatics
Akademijos, 4, Vilnius LT-08663, Lithuania
E-mail: olgjan@zebra.lt

Abstract. Let $X, X_1, \dots, X_n$ be independent identically distributed random variables taking values in a measurable space $(\Theta, \Re)$. Let $h(x, y)$ and $g(x)$ be real valued measurable functions of the arguments $x,\ y \in \Theta$ and let $h(x, y)$ be symmetric. We consider U-statistics of the type

$$T(X_1, \dots, X_n) = n^{-1} \sum_{1 \le i < k \le n} h(X_i, X_k) + n^{-1/2} \sum_{1 \le i \le n} g(X_i).$$

Let q_i $(i \ge 1)$ be eigenvalues of the Hilbert-Schmidt operator associated with the kernel $h(x, y)$, and q_1 be the largest in absolute value one. We prove that

$$\Delta_n = \rho(T(X_1, \dots, X_n), T(G_1, \dots, G_n)) \le \frac{c\beta'^{1/6}}{\sqrt{|q_1|} n^{1/12}},$$

where $G_i,\ 1 \le i \le n$, are i.i.d. Gaussian random vectors, ρ is the Kolmogorov (or uniform) distance and $\beta' := \mathbf{E}\,|h(X, X_1)|^3 + \mathbf{E}\,|h(X, X_1)|^{18/5} + \mathbf{E}\,|g(X)|^3 + \mathbf{E}\,|g(X)|^{18/5} + 1 < \infty$.

1. Introduction and formulation of the main result. Let $X, X_1, \dots, X_n$ be independent identically distributed random variables taking values in a measurable space $(\Theta, \Re)$. Let $h : \Theta^2 \to \mathbf{R}$ and $g : \Theta \to \mathbf{R}$ be real-valued measurable functions. Let h be symmetric, that is, $h(x, y) = h(y, x)$ for all $x, y \in \Theta$. Assume that $\mathbf{E}g(X) = 0,\ \mathbf{E}h(x, X) = 0$, for all $x \in \Theta$. Let us consider the U-statistic

$$T = T(X_1, \dots, X_n) = n^{-1} \sum_{1 \le i < k \le n} h(X_i, X_k) + n^{-1/2} \sum_{1 \le i \le n} g(X_i). \tag{1}$$

2010 *Mathematics Subject Classification*: Primary 60F99; Secondary 46A32.
Key words and phrases: second order U-statistic, limit distribution, Kolmogorov distance, eigenvalues, Hilbert-Schmidt operator.
The paper is in final form and no version of it will be published elsewhere.

DOI: 10.4064/bc90-0-17

Let $Z, Z_1, \ldots, Z_n$ be independent identically distributed random variables taking values in a measurable space $(\Theta, \Re)$, for which the conditions

$$\begin{gathered}
\mathbf{E}g(Z) = 0,\ \mathbf{E}h(x, Z) = 0,\ \mathbf{E}g^2(Z) = \mathbf{E}g^2(X), \\
\mathbf{E}h(x, Z)h(y, Z) = \mathbf{E}h(x, X)h(y, X), \\
\mathbf{E}h(x, Z)g(Z) = \mathbf{E}h(x, X)g(X) \text{ for all } x \in \Theta
\end{gathered} \tag{2}$$

are satisfied. Denote by $\mathcal{L}(Y)$ the distribution of a random variable Y. We shall prove in this work that $\mathcal{L}(T(X_1, \ldots, X_n))$ approaches $\mathcal{L}(T(Z_1, \ldots, Z_n))$ as $n \to \infty$ if conditions (2), (3) and (4) are satisfied. We also find the rate of convergence. Thus it is proved that in view of the conditions mentioned above, the limit distribution of the second-order U-statistic possesses the property of stability in the sense that it is independent of the distribution of the initial random variable X. This result can be regarded as being intermediate in determining the rate of convergence of T-statistic to its own limit distribution. Note that the rate of convergence of second-order random polynomials was investigated by Yanushkevichiene (1998, 2006).

We are going to use the notation

$$\Delta_n = \rho(\mathcal{L}(T(X_1, \ldots, X_n)), \mathcal{L}(T(Z_1, \ldots, Z_n))),$$

where ρ is the Kolmogorov (or uniform) metric.

Let $q_1, q_2, \ldots$ be the eigenvalues of the Hilbert-Schmidt operator Q, associated with the kernel h and measure $\mu = \mathcal{L}(X)$ (see Section 2 for detailed definitions), and $\bar{q_1}, \bar{q_2}, \ldots$ be eigenvalues of the Hilbert-Schmidt operator Q, associated with the kernel h and measure $\bar{\mu} = \mathcal{L}(Z)$. Without loss of generality, we assume that $|q_1| \geq |q_2| \geq \ldots$ and $|\bar{q_1}| \geq |\bar{q_2}| \geq \ldots$.

Assume also that

$$q_i = \bar{q_i},\ i = 1, 2, \ldots, \tag{3}$$

and denote

$$\gamma_s = \mathbf{E}\,|g(X)|^s,\ \bar{\gamma_s} = \mathbf{E}\,|g(Z)|^s,\ \beta_s = \mathbf{E}\,|h(X, X_1)|^s,\ \bar{\beta_s} = \mathbf{E}\,|h(Z, Z_1)|^s,$$

where $s > 0$,

$$\beta' := \beta_3 + \beta_{18/5} + \gamma_3 + \gamma_{18/5} + 1,\ \bar{\beta} = \bar{\beta}_3 + \bar{\beta}_{18/5} + \bar{\gamma_3} + \bar{\gamma}_{18/5} + 1,\ \beta := \max(\beta', \bar{\beta}),$$

and suppose that

$$\beta < \infty, \quad \beta_2 > 0, \quad \bar{\beta}_2 > 0. \tag{4}$$

The conditions $\beta_2 > 0,\ \bar{\beta}_2 > 0$ ensure that the quadratic part of the statistic T is not asymptotically negligible and therefore T is not asymptotically normal.

In the sequel, we denote by c some positive absolute constants which may differ from line to line or from formula to formula. The following theorem is our main result.

Theorem 1.1. *If conditions (2), (3) and (4) are satisfied then*

$$\Delta_n \leq c\beta^{1/6}|q_1|^{-1/2}n^{-1/12}. \tag{5}$$

The order of this upper bound cannot be improved, since Senatov (1996) has shown that, in the CLT in multidimensional Euclidian space for balls, whose center is not at zero, the respective rates of convergence are defined by the expression $O(\frac{n^{-k/12}}{(q_1 \ldots q_k)^{1/2}})$, under the condition that $|q_k| > 0$ and $k \le 6$. In our case, we have only some information about the first eigenvalue, therefore the order of the bound is equal to $1/12$.

Example. Let

$$T = T(X_1, \dots, X_n) = n^{-1} \sum_{1 \le i < k \le n} X_i X_k + n^{-1/2} \sum_{1 \le i \le n} X_i,$$

and $X, X_1, \dots, X_n$ be independent identically distributed one-dimensional random variables with $\mathbf{E}X = 0,\ \mathbf{E}X^2 = 1$. In this case conditions (2) are satisfied and the Theorem 1.1 holds for all one-dimensional Z with moments $\mathbf{E}Z = 0,\ \mathbf{E}Z^2 = 1$.

2. Special representations of bivariate U-statistics. The following representation of bivariate U-statistics goes back to Dunford and Schwartz (1963). Consider the measurable space $(\Theta, \Re, \mu)$ with measure $\mu = \mathcal{L}(X)$ where $X \in \Theta$. Let $L^2 = L^2(\Theta, \Re, \mu)$ denote the real Hilbert space of square integrable real functions. The Hilbert-Schmidt operator Q is defined via

$$Qf(x) = \int_\Theta h(x, y) f(y) \mu(dy), \qquad f \in L^2.$$

Let $\{e_j : j \ge 1\}$ denote an orthonormal complete system of eigenfunctions of Q ordered by decreasing absolute values of the corresponding eigenvalues $|q_1| \ge |q_2| \ge \dots$. Then

$$\mathbf{E}\, h^2(X, X_1) = \sum_{j \ge 1} q_j^2 < \infty,\ h(x, y) = \sum_{j \ge 1} q_j e_j(x) e_j(y). \tag{6}$$

Consider the subspace $L^2(g, h) \subset L^2$ generated by g, h and eigenfunctions e_j corresponding to nonzero eigenvalues $q_j \ne 0$. Introducing, if necessary, a normalized eigenfunction, say e_0, such that $Qe_0 = 0$, we can assume that $e_0, e_1, \dots$ is an orthonormal basis of $L^2(g, h)$. Thus, we can write

$$g(X) = \sum_{j \ge 0} a_j e_j(X), \qquad \gamma_2 = \mathbf{E}\, g^2(X) = \sum_{j \ge 0} a_j^2, \tag{7}$$

with $a_j = \mathbf{E}\, g(X) e_j(X)$. It is easy to see that $\mathbf{E}\, e_j(X) = 0$, for all j. Therefore $(e_j(X))_{j \ge 0}$ is an orthonormal system of mean zero random variables.

Throughout the rest of the paper we shall assume that all random variables and vectors are totally independent, if the contrary is not clear from the context.

Bentkus and Götze (1999) modified the representation of Dunford and Schwartz in the following way. Let $\mathbf{R}^\infty$ denote the space of all real sequences $x = (x_0, x_1, x_2, \dots)$, $x_j \in \mathbf{R}$. The Hilbert space $l_2 \subset \mathbf{R}^\infty$ consists of $x \in \mathbf{R}^\infty$ such that

$$|x|^2 \overset{def}{=} \langle x, x \rangle,\ |x| < \infty,\ \langle x, y \rangle = \sum_{j \ge 0} x_j y_j.$$

Consider a random vector

$$\mathbf{X} \overset{def}{=} (e_0(X), e_1(X), \dots),$$

which takes values in $\mathbf{R}^\infty$. Since $\{e_j(X)\}_{j \ge 0}$ is a system of mean zero uncorrelated random variables with variances 1, the random vector $\mathbf{X}$ has identity covariance and mean zero. Using (6) and (7), we can write

$$h(X, X_1) = \langle Q\mathbf{X}, \mathbf{X}_1 \rangle,\ g(X) = \langle a, \mathbf{X} \rangle, \tag{8}$$

where we define $Qx = (0, q_1x_1, q_2x_2, \dots)$, for $x \in \mathbf{R}^\infty$, and $a = (a_j)_{j\geq 0} \in \mathbf{R}^\infty$. Equalities (8) allow us to assume that the measurable space Θ is $\mathbf{R}^\infty$, the random variable X is a random vector taking values in $\mathbf{R}^\infty$ with mean zero and identity covariance, and that

$$h(X, X_1) = \langle QX, X_1\rangle,\ \ g(X) = \langle a, X\rangle.$$

In particular, without loss of generality, we shall assume throughout that the kernels $h(x, y)$ and $g(x)$ are linear functions in each of their arguments.

Let G, G_i, $1 \leq i \leq n$ be i.i.d. Gaussian random vectors $G_i = (G_{1,i}, G_{2,i}, \dots)$ with values in $\mathbf{R}^\infty$, where $G_{1,i}, G_{2,i}, \dots$ denote i.i.d. standard normal random variables. We assume throughout the rest of the paper that

$$\mathbf{E}g(G) = 0,\ \ \mathbf{E}h(x, G) = 0,\ \ \mathbf{E}g^2(G) = \mathbf{E}g^2(X), \tag{9}$$

$$\mathbf{E}h(x, G)h(y, G) = \mathbf{E}h(x, X)h(y, X),$$

$$\mathbf{E}h(x, G)g(G) = \mathbf{E}h(x, X)g(X), \quad x \in \Theta.$$

Note that the possibility of selecting Gaussian random variables in such way that equalities (9) are satisfied was proven by Bentkus and Götze (1999).

3. Lemma. To prove the theorem we need the following lemma.

LEMMA 3.1. *Let η be a standard Gaussian random variable and $q_1 > 0$. Then the distribution function $H(x) = \mathbf{P}\{q_1\eta^2 < x\}$ satisfies the Lipschitz condition*

$$\sup_x |H(x + \varepsilon) - H(x)| \leq c q_1^{-1/2}\sqrt{\varepsilon}, \tag{10}$$

where $\varepsilon > 0$.

Proof of the lemma. The distribution of η^2 has the density

$$f(x) = \frac{1}{\sqrt{2}\Gamma(1/2)\sqrt{x}} e^{-x/2} \text{ for } x > 0, \quad \text{and} \quad f(x) = 0 \text{ for } x \leq 0.$$

Denote by $H_1(x)$ the distribution function of the random variable η^2. Then

$$H_1(x + \varepsilon) - H_1(x) = \int_x^{x+\varepsilon} f(u)du.$$

Analyzing the first derivative H_1' of this function, it is easy to see that the latter does not increase. Therefore we can write

$$|H_1(x + \varepsilon) - H_1(x)| \leq c \int_0^\varepsilon \frac{e^{-u}}{\sqrt{u}} du \leq c\sqrt{\varepsilon}.$$

Passing from the distribution function H_1 to the distribution function H we arrive at the assertion of the lemma.

4. Proof of the Theorem. We have to prove inequality (5). Using the triangle inequality we can write

$$\Delta_n \leq \rho(\mathcal{L}(T(X_1, \dots, X_n)), \mathcal{L}(T(G_1, \dots, G_n))) + \rho(\mathcal{L}(T(G_1, \dots, G_n)), \mathcal{L}(T(Z_1, \dots, Z_n))). \tag{11}$$

It has been shown by Bentkus and Götze (1999) that it is possible to represent the statistic $T(G_1,\dots,G_n)$ in the form

$$T(G_1,\dots,G_n)=n^{-1}\sum_{1\le i<k\le n}\langle QG_i,G_k\rangle+n^{-1/2}\sum_{1\le i\le n}\langle a,G_i\rangle,$$

where $a=(a_i)_{i\ge1}$ is some constant. Let $G_{i,j},\ i\ge1$ be the components of the vector G_j. Then we can rewrite the above expression in the form

$$T(G_1,\dots,G_n)=\sum_{i\ge1}\Big(n^{-1}q_i\sum_{1\le j<k\le n}G_{i,j}G_{i,k}+n^{-1/2}\sum_{1\le j\le n}a_iG_{i,j}\Big),$$

where $G_{i,j}$ are independent normally distributed random variables with the first two moments equal to 0 and 1 respectively. It is easy to see that

$$2n^{-1}q_i\sum_{1\le j<k\le n}G_{i,j}G_{i,k}=(n-1)q_i\overline{G}_i^2-q_iS_i^2,$$

where

$$\overline{G}_i=n^{-1}\sum_{1\le j\le n}G_{i,j}\ \text{and}\ S_i^2=n^{-1}\sum_{1\le j\le n}G_{i,j}^2-\overline{G}_i^2.$$

It is also well known (see Cramer (1946)) that random variables $\overline{G}_i$ and S_i are independent. The random variables $\overline{G}_i$ can be written in the form $\overline{G}_i=\eta_i/\sqrt{n}$, so we can write

$$n^{-1}q_i\sum_{1\le j<k\le n}G_{i,j}G_{i,k}=\frac{n-1}{n}q_i\eta_i^2-q_iS_i^2.$$

Then

$$\begin{aligned}T(G_1,\dots,G_n)&=\sum_{i\ge1}\left(\frac{n-1}{2n}q_i\eta_i^2-\frac{q_i}{2}S_i^2+a_i\eta_i\right)\\&=\sum_{i\ge1}\left(\frac12q_i(\eta_i^2-1)-\frac{1}{2n}q_i\eta_i^2-\frac{q_i}{2}(S_i^2-1)+a_i\eta_i\right).\end{aligned}$$

Using the independence of $\eta_i,\ \eta_j,\ S_i,\ S_j,\ i\ne j$ and Lemma 3.1 it is easy to notice that the distribution function of $T(G_1,\dots,G_n)$ satisfies the Lipschitz condition with exponent 1/2. Now we can write

$$\begin{aligned}&\sup_x|\mathbf{P}(T(G_1,\dots,G_n)\le x+\varepsilon)-\mathbf{P}(T(G_1,\dots,G_n)\le x)|\\&=\sup_x\left|\mathbf{P}(q_1\eta_1^2-\frac1nq_1\eta_1^2+2a_1\eta_1\le x+2\varepsilon)-\mathbf{P}\left(q_1\eta_1^2-\frac1nq_1\eta_1^2+2a_1\eta_1\le x\right)\right|\\&=\sup_x\left|\mathbf{P}\left(\frac{q_1(n-1)}{n}\left(\eta_1+\frac{a_1n}{q_1(n-1)}\right)^2\le x+2\varepsilon\right)-\mathbf{P}\left(\frac{q_1(n-1)}{n}\left(\eta_1+\frac{a_1n}{q_1(n-1)}\right)^2\le x\right)\right|\\&\le\frac{c'}{\sqrt{|q_1|}}\sqrt{\varepsilon}.\end{aligned}\tag{12}$$

Here and throughout, $c',c_1,c_2,\dots$ are some positive constants.

Now, we prove that, for any $\varepsilon>0$, we have

$$\rho(\mathcal{L}(T(X_1,\dots,X_n)),\mathcal{L}(T(G_1,\dots,G_n)))\le\frac{c'}{\sqrt{|q_1|}}\sqrt{\varepsilon}+\Delta,\tag{13}$$

where

$$\Delta = \max_{\varphi} |\mathbf{E}\varphi(T(X_1, \dots, X_n)) - \mathbf{E}\varphi(T(G_1, \dots, G_n))|,$$

and the maximum is taken over all infinitely differentiable φ such that $|\varphi^{(k)}(u)| \le c_1/\varepsilon^k$, $k = 1, 2, 3$ and $0 \le \varphi(u) \le 1$, moreover, either

$$\varphi(u) = 1 \text{ for } u \le x - \varepsilon, \text{ and } \varphi(u) = 0 \text{ for } u \ge x, \tag{14}$$

or

$$\varphi(u) = 1 \text{ for } u \le x, \text{ and } \varphi(u) = 0 \text{ for } u \ge x + \varepsilon. \tag{15}$$

To do so, write $\delta^* = \mathbf{P}\{T(X_1, \dots, X_n) \le x\} - \mathbf{P}\{T(G_1, \dots, G_n) \le x\}$. We start with the proof of (13) in the case $\delta^* \ge 0$. Let us take a function φ such that (15) holds. Then

$$\begin{aligned}
\delta^* &= \mathbf{EI}\{T(X_1, \dots, X_n) \le x\} - \mathbf{P}\{T(G_1, \dots, G_n) \le x\} \\
&\le |\mathbf{E}\varphi(T(X_1, \dots, X_n)) - \mathbf{E}\varphi(T(G_1, \dots, G_n))| \\
&\quad + |\mathbf{E}\varphi(T(G_1, \dots, G_n)) - \mathbf{P}\{T(G_1, \dots, G_n)x\}| \\
&\le \Delta + \mathbf{P}\{x \le T(G_1, \dots, G_n) \le x + \varepsilon\}.
\end{aligned}$$

Using the Lipschitz condition, we get (13).

If $\delta^* < 0$, the proof of (13) is similar. One has to take a function φ that satisfies (14).

Denote

$$\Delta^*(\varphi) = |\mathbf{E}\varphi(T(X_1, \dots, X_n)) - \mathbf{E}\varphi(T(G_1, \dots, G_n))|.$$

Obviously

$$\begin{aligned}
\Delta^*(\varphi) &\le |\mathbf{E}\varphi(T(X_1, \dots, X_n)) - \mathbf{E}\varphi(T(X_1, \dots, X_{n-1}, G_n))| \\
&\quad + |\mathbf{E}\varphi(T(X_1, \dots, X_{n-1}, G_n)) - \mathbf{E}\varphi(T(X_1, \dots, X_{n-2}, G_{n-1}, G_n))| \\
&\quad + \dots + |\mathbf{E}\varphi(T(X_1, G_2, \dots, G_n)) - \mathbf{E}\varphi(T(G_1, \dots, G_n))| \\
&= \Delta^*_{1,n} + \dots + \Delta^*_{n,n}.
\end{aligned} \tag{16}$$

Let us prove now that

$$\rho(\mathcal{L}(T(X_1, \dots, X_n)), \mathcal{L}(T(G_1, \dots, G_n))) \le c_2 |q_1|^{-1/2} \beta'^{1/6} n^{-1/12}. \tag{17}$$

We shall use induction on n. Assume that for all $m \le n - 1$, the inequality

$$\rho(\mathcal{L}(T(X_1, \dots, X_{i-1}, G_i, \dots, G_m)), \mathcal{L}(T(G_1, \dots, G_m))) \le c_2 |q_1|^{-1/2} \beta'^{1/6} m^{-1/12} \tag{18}$$

holds true for all $2 \le i \le m$ and all functions h and g for which the conditions of Section 1 are satisfied. Here β' is the moment of the functions h and g, defined in Section 1.

It is easy to see that bounds (18) hold for $m = 2$. Indeed, let $m = 2$. As $\rho \le 1$, it suffices to show that $|q_1|^{-1/2} \beta'^{1/6} \ge 1$. We have

$$|q_1|^{-1/2} \beta'^{1/6} \ge |q_1|^{-1/2} (\mathbf{E}\,|h|^3)^{1/6} = (q_1^{-2} (\mathbf{E}\,|h|^3)^{2/3})^{1/4}.$$

Using (6), we obtain

$$|q_1|^{-1/2} \beta'^{1/6} \ge (q_1^{-2} \mathbf{E}\,|h|^2)^{1/4} = (q_1^{-2} (q_1^2 + q_2^2 + \dots))^{1/4} \ge 1.$$

Hence, the statement is proved for $m = 2$.

Now we will prove that (17) and (18) hold for $m = n$. We have

$$\begin{aligned}T(X_1,\dots,X_n) = \frac{1}{n}\{&h(X_1,X_2)+\sqrt{n}g(X_1)+\cdots+h(X_1,X_{n-1})+\sqrt{n}g(X_{n-1})\\ &+h(X_1,X_n)+\sqrt{n}g(X_n)+h(X_2,X_3)+\cdots+h(X_2,X_n)\\ &+\cdots+h(X_{n-2},X_{n-1})+h(X_{n-2},X_n)+h(X_{n-1},X_n)\}.\end{aligned}$$

We denote by w the sum of all summands of the above expression for $T(X_1,\dots,X_n)$, which do not involve X_n:

$$w = \frac{1}{n}\{h(X_1,X_2)+\sqrt{n}g(X_1)+\cdots+h(X_1,X_{n-1})+\sqrt{n}g(X_{n-1})+\cdots+h(X_{n-2},X_{n-1})\}$$

and by l_n the sum of all summands involving X_n:

$$l_n = \frac{1}{n}\{h(X_1,X_n)+\sqrt{n}g(X_n)+\cdots+h(X_{n-2},X_n)+h(X_{n-1},X_n)\}.$$

Replacing X_n by G_n we get

$$\begin{aligned}T(X_1,\dots,X_{n-1},G_n) &= w+\frac{1}{n}\{h(X_1,G_n)+\sqrt{n}g(G_n)+\cdots+h(X_{n-2},G_n)+h(X_{n-1},G_n)\}\\ &= w+l_n^*,\end{aligned}$$

where l_n^* is obtained from l_n replacing X_n by G_n. We expand into the Taylor series

$$\varphi(x+y)=\varphi(x)+\varphi'(x)y+\frac{1}{2}\varphi''(x)y^2+\frac{1}{2}\mathbf{E}\varphi'''(x+\tau y)(1-\tau)^2y^3.$$

Here τ is a random variable uniformly distributed in $[0,1]$ and independent of all the other random variables. Let us apply the expansion to $x = w$ and $y = l_n$. Write

$$\begin{aligned}\Delta_{1,n}^* &= |\mathbf{E}\varphi(T(X_1,\dots,X_n))-\mathbf{E}\varphi(T(X_1,\dots,X_{n-1},G_n))|\\ &= \Big|\mathbf{E}\varphi(w)-\mathbf{E}\varphi(w)+\mathbf{E}\varphi'(w)l_n-\mathbf{E}\varphi'(w)l_n^*+\frac{1}{2}\mathbf{E}\varphi''(w)l_n^2-\frac{1}{2}\mathbf{E}\varphi''(w)(l_n^*)^2\\ &\quad+\frac{1}{2}\mathbf{E}\varphi'''(w+l_n\tau)l_n^3(1-\tau)^2-\frac{1}{2}\mathbf{E}\varphi'''(w+l_n^*\tau)(l_n^*)^3(1-\tau)^2\Big|.\end{aligned}$$

Let $\mathbf{E}_{X_1,\dots,X_{n-1},\tau}$ be the expectation with respect to the random variables $X_1,\dots,X_{n-1},\tau$, $\mathbf{E}_{X_n}$ be the expectation with respect to the random variable X_n and so on. Write

$$\begin{aligned}\Delta_{1,n}^* = \Big|&\mathbf{E}_{X_1,\dots,X_{n-1},\tau}\varphi'(w)\mathbf{E}_{X_n}l_n-\mathbf{E}_{X_1,\dots,X_{n-1},\tau}\varphi'(w)\mathbf{E}_{G_n}l_n^*\\ &+\frac{1}{2}\mathbf{E}_{X_1,\dots,X_{n-1},\tau}\varphi''(w)\mathbf{E}_{X_n}l_n^2-\frac{1}{2}\mathbf{E}_{X_1,\dots,X_{n-1},\tau}\varphi''(w)\mathbf{E}_{G_n}(l_n^*)^2\\ &+\frac{1}{2}\mathbf{E}\varphi'''(w+l_n\tau)l_n^3(1-\tau)^2-\frac{1}{2}\mathbf{E}\varphi'''(w+l_n^*\tau)(l_n^*)^3(1-\tau)^2\Big|.\end{aligned}$$

Using (9), we get

$$\Delta_{1,n}^* = \left|\frac{1}{2}\mathbf{E}\varphi'''(w+l_n\tau)l_n^3(1-\tau)^2-\frac{1}{2}\mathbf{E}\varphi'''(w+l_n^*\tau)(l_n^*)^3(1-\tau)^2\right|. \tag{19}$$

Denote $|\mathbf{E}\varphi'''(w+l_n\tau)l_n^3(1-\tau)^2|$ by r and let us estimate its value. Write

$$l_n = l_n'+l_n''$$

with

$$l_n' = \frac{1}{n}\Big(\sqrt{n}g(X_n) + \sum_{1\le k\le[(n-1)/2]} h(X_k, X_n)\Big),\ l_n'' = \frac{1}{n}\sum_{[(n-1)/2]<k\le n-1} h(X_k, X_n),$$

where $[(n-1)/2]$ is the integer part of the number $(n-1)/2$. In a similar manner, we define $(l_n^*)'$ and $(l_n^*)''$. Using the inequality $(a+b)^3 \le 8(a^3+b^3)$ for $a>0,\ b>0$, we find

$$\begin{aligned} r &\le \frac{8c_1}{\varepsilon^3}(\mathbf{E}\mathbf{I}(x \le w + \tau l_n' + \tau l_n'' \le x+\varepsilon)|l_n'|^3 + \mathbf{E}\mathbf{I}(x \le w + \tau l_n' + \tau l_n'' \le x+\varepsilon)|l_n''|^3) \\ &= \frac{8c_1}{\varepsilon^3}(r_1 + r_2). \end{aligned} \tag{20}$$

We estimate only r_1, since the estimation of r_2 is similar. Now we fix $X_1, \dots, X_{[(n-1)/2]}$, X_n, τ. Let $\mathbf{E}_* = \mathbf{E}_{X_{[(n-1)/2]+1},\dots,X_{n-1}}$, then

$$r_1 = \mathbf{E}|l_n'|^3\mathbf{E}_*\mathbf{I}(x \le w + \tau l_n' + \tau l_n'' \le x+\varepsilon) \le \mathbf{E}\ |l_n'|^3 \sup_x \mathbf{P}\{x \le w + \tau l_n'' \le x+\varepsilon\}$$

(we use the independence assumption)

$$= \mathbf{E}\ |l_n'|^3 \sup_x \mathbf{P}\{x \le T^* \le x+\varepsilon\},$$

where $T^* = T^*(X_{[(n-1)/2]+1}, \dots, X_{n-1})$. We get T^* from $w + \tau l_n''$ and then we fix $X_1, \dots, X_{[(n-1)/2]}$, X_n, τ. The corresponding function $h^* = h$ is the same and g^* has the following form

$$g^*(X_j) = g(X_j) + n^{-1/2}\sum_{1\le k\le[(n-1)/2]} h(X_k, X_j) + n^{-1/2}\tau h(X_n, X_j),$$

where $j = [(n-1)/2]+1, \dots, n-1$. Using the inequalities (12), (18), we get

$$\sup_x \mathbf{P}(x \le T^* \le x+\varepsilon) \le \frac{c'}{\sqrt{|q_1|}}\sqrt{\varepsilon} + \frac{2c_2}{\sqrt{|q_1|}}\left(\frac{2\beta^{*2}}{n-1}\right)^{1/12}, \tag{21}$$

where $\beta^* = \beta_3 + \beta_{18/5} + \gamma_3^* + \gamma_{18/5}^* + 1,\ \gamma_s^* = \mathbf{E}_*\ |g^*(X)|^s$.

From Theorem 20 in Petrov (1987), we derive that

$$\begin{aligned} n^{18/5}\mathbf{E}|l_n'|^{18/5} &= \mathbf{E}\Big|n^{9/5}g(X_n) + \sum_{1\le k\le[(n-1)/2]} h(X_k, X_n)\Big|^{18/5} \\ &\le c_4 n^{9/5}(\gamma_{18/5} + \beta_{18/5}) \le c_4\beta' n^{9/5}. \end{aligned} \tag{22}$$

In Bentkus and Götze (1999) we can find the following inequalities

$$\mathbf{E}|g(G)|^s \le c_s\mathbf{E}|g(X)|^s, \quad \mathbf{E}|h(x,G)|^s \le c_s\mathbf{E}|h(x,X)|^s, \quad s \ge 2,$$

where c_s are some constants depending on s. Let us take c_4 so large that inequality (22) will be true for $(l_n^*)'$ as well. In a similar way we get

$$n^3\mathbf{E}|l_n'|^3 = \mathbf{E}\Big|n^{3/2}g(X_n) + \sum_{1\le k\le[(n-1)/2]} h(X_k, X_n)\Big|^3 \le c_3 n^{3/2}(\gamma_3 + \beta_3).$$

Obviously,

$$\mathbf{E}|l_n'|^3 \le c_3\beta' n^{-3/2}. \tag{23}$$

Let us estimate $\mathbf{E}\,(|l'_n|^3\beta^{*1/6})$. Using Hölder's inequality, we obtain

$$\begin{aligned}\mathbf{E}\,(|l'_n|^3\beta^{*1/6}) &\le (\mathbf{E}\,|l'_n|^{18/5})^{5/6}(\mathbf{E}\,\beta^*)^{1/6}\\ &= (\mathbf{E}\,|l'_n|^{18/5})^{5/6}(\mathbf{E}\,(\beta_3+\beta_{18/5}+|g^*(X)|^3+|g^*(X)|^{18/5}+1))^{1/6}.\end{aligned}$$

The following bounds are true

$$\mathbf{E}\,|g^*(X)|^{18/5} \le c_4(\gamma_{18/5}+\beta_{18/5}), \qquad \mathbf{E}\,|g^*(X)|^3 \le c_3(\gamma_3+\beta_3).$$

Denoting $c_5 = \max(c_4+1,\ c_3+1)$, we get

$$\mathbf{E}\,(|l'_n|^3\beta^{*1/6}) \le c_5^{1/6}\beta'^{1/6}(\mathbf{E}\,|l'_n|^{18/5})^{5/6} \le c_5\beta' n^{-3/2}. \tag{24}$$

Combining (21), (23), (24) and using that $\beta' > 1$, we get

$$r_1 \le \frac{\beta'}{\sqrt{|q_1|}n^{3/2}}(c'c_3\sqrt{\varepsilon}+3c_2c_5\beta'^{1/6}(n-1)^{-1/12}).$$

We can construct the same bounds for the second summand in (19). Finally we have

$$\Delta^*_{1,n} \le 16\frac{c_1c_5\beta'}{\sqrt{|q_1|\varepsilon^3 n^{3/2}}}(c'\sqrt{\varepsilon}+c_2\beta'^{1/6}(n-1)^{-1/12}).$$

Let $\varepsilon = \delta n^{-1/6}\beta'^{1/3}$, then we can write

$$\Delta^*_{1,n} \le 16c_1c_5(c'\sqrt{\delta}+3c_2)\delta^{-3}|q_1|^{-1/2}\beta'^{1/6}(n)^{-13/12}.$$

In view of (16) and (13), we obtain

$$\begin{aligned}&\rho(\mathcal{L}(T(X_1,\dots,X_n)),\mathcal{L}(T(G_1,\dots,G_n)))\\ &\qquad\le c'|q_1|^{-1/2}\sqrt{\varepsilon}+16c_1c_5(c'\sqrt{\delta}+3c_2)\delta^{-3}|q_1|^{-1/2}\beta'^{1/6}(n)^{-1/12}\\ &\qquad= (c'\sqrt{\delta}+16c_1c_5(c'\sqrt{\delta}+3c_2)\delta^{-3})|q_1|^{-1/2}\beta'^{1/6}(n)^{-1/12}.\end{aligned}$$

Choosing δ and c_2 so that the relations

$$112c_1c_3 \le \delta^3, \qquad c_2 \ge 2c'\sqrt{\delta},$$

hold, we achieve

$$\rho(\mathcal{L}(T(X_1,\dots,X_n)),\mathcal{L}(T(G_1,\dots,G_n))) \le c_2|q_1|^{-1/2}\beta'^{1/6}n^{-1/12} \le c_2|q_1|^{-1/2}\beta^{1/6}n^{-1/12}.$$

The second summand in (11) can be estimated similarly. This ends the proof of Theorem 1.1.

Acknowledgements. The author would like to thank V. Bentkus who formulated the problem under consideration as well as suggested an idea of the proof. The author is also much obliged to V. Senatov for advice related to the construction of the lower bounds.

References

V. Bentkus and F. Götze (1999), *Optimal bounds in non-Gaussian limit theorems for U-statistics*, The Annals of Probability 27, 454–521.

H. Cramer (1946), *Mathematical Methods of Statistics*, Stockholm.

N. Dunford and J. T. Schwartz (1963), *Linear Operators*, Wiley, New York.

V. V. Petrov (1987), *Limit Theorems for Sums of Random Variables*, Nauka, Moscow (in Russian).

V. V. Senatov (1996), *Qualitative effects in the estimates of the convergence rate in the central limit theorem in multidimensional spaces*, Proceedings of the Steklov Institute of Math. 215, 4.

O. Yanushkevichiene (1998), *On the rate of convergence of second-degree random polynomials*, Journal of Mathematical Sciences (New York) 92, 3955–3959.

O. Yanushkevichiene (2006), *Optimal rates of convergence of second degree polynomials in several metrics*, Journal of Mathematical Sciences (New York) 138, 5472–5479.