Bounds for non-Gaussian approximations of U-statistics

Yuri V.Borovskikh* Neville C. Weber†

16 August, 2000

Running Head: Bounds for U-statistics.

MSC 2000 classification. Primary 60F05; secondary 62E20.

Mailing Address:
N. C. Weber
School of Mathematics and Statistics, F07
University of Sydney
N.S.W. 2006
AUSTRALIA
Telephone: +61 2 9351 4249.

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

†School of Mathematics and Statistics, F07, University of Sydney, N.S.W. 2006, Australia. Email: neville@maths.usyd.edu.au

Abstract

Let $X, \bar{X}, X_1, \ldots, X_n$ be i.i.d. random variables taking values in a measurable space $(\mathfrak{X}, \mathfrak{B})$. Consider U-statistics of degree two

$$U_n = n^{-1} \sum_{1 \le i < j \le n} \Phi(X_i, X_j),$$

with symmetric, degenerate kernel $\Phi : \mathfrak{X}^2 \mapsto \mathbb{R}$. Let

$$U_\infty = \frac{1}{2} \sum_{j=1}^{\infty} q_j (\tau_j^2 - 1),$$

where $\{q_j\}$ are eigenvalues of the Hilbert-Schmidt operator associated with the kernel Φ and $\{\tau_j\}$ are i.i.d. standard normal random variables. If $\mathbb{E}\Phi^2(X, \bar{X}) < \infty$ then

$$\Delta_n := \sup_x |\mathbb{P}\{U_n \le x\} - \mathbb{P}\{U_\infty \le x\}| \to 0 \quad \text{as} \quad n \to \infty.$$

Upper bounds for Δ_n are established under the moment condition $\mathbb{E}\Phi^2(X, \bar{X}) = 1$, provided that at least thirteen eigenvalues of the operator do not vanish. In Theorem 1.1 the bound is expressed via terms containing tail and truncated moments. The proof is based on the method developed by Bentkus and Götze (1999).

Key words and phrases. U-statistics, U-statistical sums, symmetric degenerate kernel, Gaussian random variables, tail moments.

1 Introduction

Let $X, \bar{X}, X_1, \dots, X_n$ be independent and identically distributed (i.i.d.) random variables taking values in an arbitrary measurable space $(\mathfrak{X}, \mathfrak{B})$. Let $\Phi : \mathfrak{X}^2 \mapsto \mathbb{R}$ denote a measurable symmetric function, that is, $\Phi(x,y) = \Phi(y,x)$, for all $x, y \in \mathfrak{X}$. Consider U-statistics of degree two

$$U_n = n^{-1} \sum_{1 \le i < j \le n} \Phi(X_i, X_j). \tag{1.1}$$

In what follows we assume $\sigma^2 = \mathbb{E}\Phi^2(X, \bar{X}) < \infty$ and that the kernel Φ is degenerate, that is, $\mathbb{E}\Phi(x, X) = 0$ for all $x \in \mathfrak{X}$. The variance of U_n is $(n-1)\sigma^2/(2n)$.

Let $\{q_j\}$ denote the eigenvalues of the Hilbert-Schmidt operator, $\mathbb{Q}$, associated with the kernel Φ,

$$\mathbb{Q}f(x) = \int_{\mathfrak{X}} \Phi(x,y) f(y) P(dy) = \mathbb{E}\Phi(x, X) f(X), \quad f \in L^2, \tag{1.2}$$

where $L^2 = L^2(\mathfrak{X}, \mathfrak{B}, P)$ denotes the Hilbert space of square integrable real functions defined on the measurable space $(\mathfrak{X}, \mathfrak{B}, P)$ with measure $P = \mathcal{L}(X)$. Further

$$\mathbb{E}\Phi^2(X, \bar{X}) = \sum_{j=1}^{\infty} q_j^2.$$

Without loss of generality we shall assume throughout that $|q_1| \ge |q_2| \ge \dots$. Let $\{e_j\}$ denote a corresponding orthogonal, complete system of eigenfunctions of the operator $\mathbb{Q}$. Then in $L^2(\mathfrak{X}^2, \mathfrak{B}^2, P \times P)$

$$\Phi(x,y) = \sum_{j=1}^{\infty} q_j e_j(x) e_j(y), \tag{1.3}$$

since the kernel Φ is degenerate.

Let $\{\tau_j\}$ be a sequence of i.i.d. standard normal random variables. The asymptotic distribution of U_n is non-Gaussian and is the same as the distribution of the random variable

$$U_\infty = \frac{1}{2} \sum_{j=1}^{\infty} q_j (\tau_j^2 - 1). \tag{1.4}$$

Introduce

$$\Delta_n = \sup_x |\mathbb{P}\{U_n \le x\} - \mathbb{P}\{U_\infty \le x\}|. \tag{1.5}$$

Our basic aim is to establish upper bounds for Δ_n under the minimal moment condition $\mathbb{E}\Phi^2(X, \bar{X}) < \infty$ provided at least $q_{13} \ne 0$. Without loss of generality we may assume that $\sigma = 1$.

To formulate our results we need the following definitions and notation. Define

$$\|x\| = \Big(\mathbb{E}\Phi^2(X, x)\Big)^{1/2}, \quad x \in \mathfrak{X} \tag{1.6}$$

and the degenerate, truncated kernel

$$h(x,y)=\int_{\mathfrak{X}}\int_{\mathfrak{X}}\Phi(z,\bar z)I(|\Phi(z,\bar z)|\le n)I(\|z\|\le\sqrt{n})I(\|\bar z\|\le\sqrt{n})$$
$$\times\Big(\delta_x(dz)-P(dz)\Big)\Big(\delta_y(d\bar z)-P(d\bar z)\Big),\quad x,y\in\mathfrak{X},\tag{1.7}$$

where $I(\cdot)$ denotes the indicator function and $\delta_x(\cdot)$ is Dirac's δ-measure. Introduce the tail and truncated moments

$$\alpha_n=\mathbb{E}\|X\|^2 I(\|X\|>\sqrt{n}),\quad \beta_n=\mathbb{E}\Phi^2(X,\bar X)I(|\Phi(X,\bar X)|>n),$$
$$\gamma_n=n^{-1}\mathbb{E}\|X\|^4 I(\|X\|\le\sqrt{n}),\quad \lambda_n=n^{-2}\mathbb{E}\Phi^4(X,\bar X)I(|\Phi(X,\bar X)|\le n),$$
$$\mu_n=n^{-1}|\mathbb{E}\Phi^3(X,\bar X)I(|\Phi(X,\bar X)|\le n)|$$

and let

$$L_n=\alpha_n+\beta_n+\gamma_n+\lambda_n,\quad n\ge 1.$$

Note that $L_n\to 0$ as $n\to\infty$ and since $\mathbb{E}\|X\|^2=1$

$$n^{-1}\le 2(n^{-1}\alpha_n^2+\gamma_n)\quad\text{and}\quad n^{-2}\le 8(n^{-2}\alpha_n^4+\gamma_n^2).\tag{1.8}$$

Further denote

$$p=\exp\{-(16s)^3/|q_s|\},\tag{1.9}$$

provided that $q_s\ne 0$, $s\ge 1$, and define an integer $m(s,q_s)\ge 1$ by the formula

$$m(s,q_s)=\min\{m\ge 1: L_m\le 2^{-15}s^{-3}p\}.$$

In what follows we use c and $c(\cdot)$, with or without indices, to denote absolute, generic, positive constants and constants depending on the parameters in the brackets. In particular, $c(q_s)=\exp\{c/|q_s|\}$ if $q_s\ne 0$.

Theorem 1.1. *If $q_{13}\ne 0$ then*

$$\Delta_n\le\exp\{c/|q_{13}|\}(L_n+\mu_n+n^{-1}m(13,q_{13})).\tag{1.10}$$

Throughout we shall assume that

$$L_n\le p/8.\tag{1.11}$$

Since $\Delta_n\le 1$, if $L_n\ge p/8$ then $\Delta_n\le 8p^{-1}L_n\le c(q_{13})L_n$. That is (1.10) holds. Similarly we may assume that $np>m(13,q_{13})$.

Under additional moment conditions the estimate (1.10) implies a bound of order $O(n^{-1})$.

Corollary 1.1. *Let $a:=\mathbb{E}\|X\|^4<\infty$ and assume*

$$b:=\sup_x\Big(x^3\mathbb{P}(|\Phi(X,\bar X)|>x)\Big)<\infty.$$

Then

$$\alpha_n\le n^{-1}a,\qquad \beta_n\le 3n^{-1}b,\qquad \gamma_n\le n^{-1}a,\qquad \lambda_n\le 4n^{-1}b$$

and

$$m(13, q_{13}) \leq 2^{15} 13^3 (2a + 7b) p^{-1} + 1.$$

If

$$|\mathbb{E}\Phi^3(X, \bar{X}) I(|\Phi(X, \bar{X})| \leq n)| \leq d < \infty \tag{1.12}$$

then

$$\Delta_n \leq c(q_{13})\,(a + b + d)\, n^{-1}. \tag{1.13}$$

Note that if $\mathbb{E}|\Phi|^3 < \infty$ then the condition (1.12) is satisfied with $d = \mathbb{E}|\Phi|^3$ and moreover $b \leq \mathbb{E}|\Phi|^3$. In this case the Corollary 1.1 is reduced to the result by Bentkus and Götze (1999)

$$\Delta_n \leq c(q_{13}) \Big(\mathbb{E}\|X\|^4 + \mathbb{E}|\Phi(X, \bar{X})|^3 \Big) n^{-1}.$$

However the conditions $b < \infty$ and $d < \infty$ can satisfied but nevertheless $\mathbb{E}|\Phi|^3 = \infty$, as shown in the following example.

EXAMPLE 1.1. Let X be a random variable with the uniform distribution on $\mathfrak{X} = [-1, 1]$. For $x, y \in \mathfrak{X}$ introduce the functions

$$\phi(x, y) = (x + y)|x - y|^{-1/3},$$

$$\phi_1(x) = \mathbb{E}\phi(x, X) = \frac{3}{10}(4x - 1)(1 + x)^{2/3} + \frac{3}{10}(4x + 1)(1 - x)^{2/3}.$$

It is obvious that

$$\mathbb{E}\phi(X, \bar{X}) = 0, \quad \phi(x, y) = \phi(y, x), \quad \phi(x, y) = -\phi(-x, -y), \quad x, y \in \mathfrak{X}.$$

Define the degenerate kernel $\Phi(x, y) = \phi(x, y) - \phi_1(x) - \phi_1(y)$. Then

$$\Phi(x, y) = \Phi(y, x), \quad \Phi(x, y) = -\Phi(-x, -y), \quad x, y \in \mathfrak{X},$$

and therefore for this kernel $\mathbb{E}\Phi^3(X, \bar{X}) I(|\Phi(X, \bar{X})| \leq n) = 0$. Simple calculations show that

$$\sup_x \Big(x^3 \mathbb{P}(|\Phi(X, \bar{X})| > x) \Big) \leq 2^{16},$$

$$\|x\| = \Big(\mathbb{E}\Phi^2(x, X) \Big)^{1/2} \leq 12, \quad x \in \mathfrak{X}$$

but

$$\mathbb{E}|\Phi(X, \bar{X})|^3 = \infty.$$

Consequently, in Corollary 1.1 we obtain the optimal bound of order $O(n^{-1})$ in (1.13) even though the third absolute moment of the kernel Φ is not finite. Thus $\mathbb{E}|\Phi|^3 < \infty$ is not necessary to achieve a bound of order $O(n^{-1})$ for Δ_n.

In the sequel let $\mathbb{R}^\infty$ be the space of all real sequences $x = (x_1, x_2, \ldots)$, $x_j \in \mathbb{R}$. Let $l_2 \subset \mathbb{R}^\infty$ be the Hilbert space consisting of $x \in \mathbb{R}^\infty$, such that

$$|x|^2 = (x, x), \quad |x| < \infty, \quad (x, y) = \sum_{j=1}^{\infty} x_j y_j.$$

Define $\mathbb{Q}x = (q_1x_1, q_2x_2, \ldots)$ with $x \in \mathbb{R}^\infty$. The random eigenfunctions vector $\mathbf{e}(X) = (e_1(X), e_2(X), \ldots)$ takes values in $\mathbb{R}^\infty$. This vector has mean zero and identity covariance matrix, since $\{e_j(X)\}$ is a system of uncorrelated random variables with unit variances. According to (1.3) we can write

$$\Phi(X, \bar{X}) = (\mathbb{Q}\mathbf{e}(X), \mathbf{e}(\bar{X})). \tag{1.14}$$

Equation (1.14) enables us to assume that the measurable space $\mathfrak{X}$ is $\mathbb{R}^\infty$ and the random variable X is a random vector taking values in $\mathbb{R}^\infty$, with mean zero and identity covariance. Also without loss of generality we can assume that the kernel $\Phi(x, y)$ is linear function in each of its arguments.

Introduce the Gaussian random vector $G = (\tau_1, \tau_2, \ldots)$ taking values in $\mathbb{R}^\infty$, where $\tau_1, \tau_2, \ldots$ denote i.i.d. standard normal random variables. In this case in (1.14) $\mathbf{e}(G) = G$ and $e_j(G) = \tau_j$, $j = 1, 2, \ldots$. The random vectors G and X have mean zero and equal covariances, that is

$$\mathbb{E}\Phi(X, x) = \mathbb{E}\Phi(G, x) = 0,$$

$$\mathbb{E}\Phi(X, x)\Phi(X, y) = \mathbb{E}\Phi(G, x)\Phi(G, y). \tag{1.15}$$

The paper is organized as follows. Lemma 2.1 in Section 2 provides a truncation inequality for U-statistical type sums which enables us to introduce simultaneously truncated random variables and truncated kernels. In Lemma 2.2 the eigenvalues associated with the original and truncated kernels are compared. Section 3 contains the non-degeneracy condition of Bentkus and Götze (1999) for the non-Gaussian random vector under the minimal moment condition $\sigma = 1$. The fourth section provides bounds for associated characteristic functions and in Section 5 we prove Theorem 1.1 by applying Lemma 4.2 of Bentkus and Götze (1999).

2 Double truncation

Introduce the truncated functions

$$f(x, y) = \Phi(x, y)I(\|x\| \le \sqrt{n})I(\|y\| \le \sqrt{n}), \quad g(x, y) = f(x, y)I(|\Phi(x, y)| \le n),$$

$x, y \in \mathfrak{X}$, and the truncated random variables $X_j^{tr} = X_j I(\|X_j\| \le \sqrt{n})$, $j = 1, \ldots, n$.

Let $s(\Phi) = s(\Phi; X_1, \ldots, X_n)$, where

$$s(\Phi; X_1, \ldots, X_n) = \sum_{1 \le i < j \le n} \Phi(X_i, X_j).$$

The following lemma contains the double truncation inequality, which enables us to introduce X_j^{tr} instead of X_j, $j = 1, \ldots, n$ and the random variables $g(X_i, X_j)$ instead of $\Phi(X_i, X_j)$, $1 \le i < j \le n$.

Lemma 2.1. *For all* $x \in \mathbb{R}$

$$|\mathbb{P}(s(\Phi) \le x) - \mathbb{P}(s(g) \le x)|$$

$$\le n\,\mathbb{P}(\|X\| > \sqrt{n}) + \frac{1}{2}n(n-1)\mathbb{P}(|\Phi(X, \bar{X})| > n). \tag{2.1}$$

PROOF. First we shall show that for $x \in \mathbb{R}$

$$|\mathbb{P}(s(\Phi) \le x) - \mathbb{P}(s(f) \le x)| \le n\,\mathbb{P}(\|X\| > \sqrt{n}). \tag{2.2}$$

Consider the events

$$A_j = \{\|X_j\| \le \sqrt{n}\}, \quad A_j^c = \{\|X_j\| > \sqrt{n}\}, \quad j = 1, \dots, n,$$

$$A = \bigcap_{j=1}^{n} A_j, \quad A^c = \bigcup_{j=1}^{n} A_j^c,$$

where A^c denotes the complement of A.

Note that $\mathbb{P}(A^c) \le n\,\mathbb{P}(\|X\| > \sqrt{n})$. Taking account of the linearity of $\Phi(x, y)$ in each of its arguments we obtain

$$\mathbb{P}\Big(\{s(\Phi; X_1, \dots, X_n) \le x\} \cap A\Big) = \mathbb{P}\Big(\{s(\Phi; X_1^{tr}, \dots, X_n^{tr}) \le x\} \cap A\Big)$$

$$= \mathbb{P}\Big(\{s(f; X_1, \dots, X_n) \le x\} \cap A\Big).$$

Hence

$$\mathbb{P}\Big(s(\Phi) \le x\Big) = \mathbb{P}\Big(\{s(f) \le x\} \cap A\Big) + \mathbb{P}\Big(\{s(\Phi) \le x\} \cap A^c\Big)$$

$$\le \mathbb{P}\Big(s(f) \le x\Big) + \mathbb{P}(A^c).$$

Similarly

$$\mathbb{P}\Big(s(f) \le x\Big) \le \mathbb{P}\Big(s(\Phi) \le x\Big) + \mathbb{P}(A^c).$$

This proves (2.2). Next we shall show that for $x \in \mathbb{R}$

$$|\mathbb{P}\Big(s(f) \le x\Big) - \mathbb{P}\Big(s(g) \le x\Big)| \le \frac{1}{2} n(n-1)\mathbb{P}\Big(|\Phi(X, \bar{X})| > n\Big). \tag{2.3}$$

Consider the events

$$A_{ij} = \{|f(X_i, X_j)| \le n\}, \quad 1 \le i < j \le n, \quad A = \bigcap_{1 \le i < j \le n} A_{ij}, \quad A^c = \bigcup_{1 \le i < j \le n} A_{ij}^c.$$

Here $\mathbb{P}(A^c) \le 2^{-1} n(n-1)\mathbb{P}\Big(|\Phi(X, \bar{X})| > n\Big)$.

Further we write

$$\mathbb{P}\Big(s(f) \le x\Big) = \mathbb{P}\Big(\{s(g) \le x\} \cap A\Big) + \mathbb{P}\Big(\{s(f) \le x\} \cap A^c\Big)$$

$$\le \mathbb{P}\Big(s(g) \le x\Big) + \mathbb{P}(A^c).$$

Similarly

$$\mathbb{P}\Big(s(g) \le x\Big) \le \mathbb{P}\Big(s(f) \le x\Big) + \mathbb{P}(A^c).$$

This proves (2.3) and the Lemma follows.

To complete this section we shall compare the eigenvalues of the Hilbert-Schmidt operators associated with the initial kernel Φ and the truncated kernel h defined at (1.7). Introduce the Hilbert-Schmidt operator $\bar{\mathbb{Q}} : L^2 \mapsto L^2$, associated with the kernel h,

$$\bar{\mathbb{Q}}f(x) = \int_{\mathfrak{X}} h(x,y)f(y)P(dy) = \mathbb{E}h(x,X)f(X), \quad f \in L^2. \tag{2.4}$$

Let $\{\bar{q}_j\}$ denote the eigenvalues of $\bar{\mathbb{Q}}$. Without loss of generality assume that $|\bar{q}_1| \geq |\bar{q}_2| \geq \dots$. Further let $\{\bar{e}_j\}$ denote an orthonormal, complete system of eigenfunctions of $\bar{\mathbb{Q}}$ ordered by decreasing absolute values of the corresponding eigenvalues $\bar{q}_1, \bar{q}_2, \dots$. Since the kernel h is degenerate

$$h(x,y) = \sum_{j=1}^{\infty} \bar{q}_j \bar{e}_j(x) \bar{e}_j(y) \tag{2.5}$$

in $L^2(\mathfrak{X}^2, \mathfrak{B}^2, P \times P)$. The following lemma contains lower and upper bounds for $\bar{q}_j$ in terms of q_j.

Lemma 2.2. *If $q_s \neq 0$ for some $s \geq 1$ then*

$$\frac{1}{2}|q_j| \leq |\bar{q}_j| \leq 2|q_j|, \quad \text{for} \quad 1 \leq j \leq s \tag{2.6}$$

provided that $\alpha_n + \beta_n \leq p$.

Proof. Define the symmetric kernel

$$\Phi_1(x,y) = \int_{\mathfrak{X}} \Phi(x,z)\Phi(y,z)P(dz). \tag{2.7}$$

According to (1.2) for $f \in L^2$

$$\begin{aligned} \mathbb{Q}^2 f(x) = \mathbb{Q}(\mathbb{Q}f)(x) &= \mathbb{Q}\Big(\int_{\mathfrak{X}} \Phi(x,y)f(y)P(dy)\Big) \\ &= \int_{\mathfrak{X}} \Phi(x,z)\Big(\int_{\mathfrak{X}} \Phi(z,y)f(y)P(dy)\Big)P(dz) \\ &= \int_{\mathfrak{X}} \Phi_1(x,y)f(y)P(dy). \end{aligned}$$

Hence $\mathbb{Q}^2 : L^2 \mapsto L^2$ is the Hilbert-Schmidt operator, associated with the kernel Φ_1. Further for any $f \in L^2$

$$(\mathbb{Q}^2 f, f) = \int_{\mathfrak{X}} \Big\{ \int_{\mathfrak{X}} \Phi(x,z)f(x)P(dx) \Big\}^2 P(dz), \tag{2.8}$$

that is, the operator $\mathbb{Q}^2$ is positive definite. From (1.3)

$$\int_{\mathfrak{X}} \Phi_1(x,y)e_j(y)P(dy) = q_j^2 e_j(x), \quad j \geq 1 \tag{2.9}$$

so that $\{e_j\}$ and $\{q_j^2\}$ are the orthonormal complete system of eigenfunctions and corresponding eigenvalues of the operator $\mathbb{Q}^2$. Therefore by the theory of Schmidt-Courant (see, for example, Dunford and Schwartz (1961), Chapter 10)

$$q_1^2 = \sup\{(\mathbb{Q}^2 f, f) : |f| = 1\}, \tag{2.10}$$

$$q_j^2 = \inf\Big\{\sup\{(\mathbb{Q}^2 f, f) : |f| = 1, (f, f_i) = 0\} : f_i \in L^2, \quad i = 1, \ldots, j-1\Big\}, j \geq 2.$$

Here the supremum is taken for fixed f_i and then the infimum is taken over all possible $f_i \in L^2$, $i = 1 \ldots, j-1$.

Further define the symmetric kernel

$$h_1(x, y) = \int_{\mathfrak{X}} h(x, z)\, h(y, z)\, P(dz). \tag{2.11}$$

and the Hilbert-Schmidt operator $\bar{\mathbb{Q}}^2 : L^2 \mapsto L^2$, associated with the kernel h_1. The eigenfunctions and the eigenvalues of $\bar{\mathbb{Q}}^2$ are $\{\bar{e}_j\}$ and $\{\bar{q}_j^2\}$ respectively. A similar relation to (2.10) holds for the operator $\bar{\mathbb{Q}}^2$ and the eigenvalues $\{\bar{q}_j^2\}$. Hence, for all $j \geq 1$, we can write

$$|q_j^2 - \bar{q}_j^2| \leq \sup\Big\{\Big|\Big((\mathbb{Q}^2 - \bar{\mathbb{Q}}^2)f, f\Big)\Big| : |f| = 1\Big\}. \tag{2.12}$$

By the definition of the operators $\mathbb{Q}^2$ and $\bar{\mathbb{Q}}^2$

$$\Big((\mathbb{Q}^2 - \bar{\mathbb{Q}}^2)f, f\Big) = \int_{\mathfrak{X}} \Big(\Big\{\int_{\mathfrak{X}} \Phi(x, y) f(x) P(dx)\Big\}^2 - \Big\{\int_{\mathfrak{X}} h(x, y) f(x) P(dx)\Big\}^2\Big) P(dy). \tag{2.13}$$

From (1.7)

$$\begin{aligned} h(x, y) \;=\; & \Phi(x, y) I(|\Phi(x, y)| \leq n) I(\|x\| \leq \sqrt{n}) I(\|y\| \leq \sqrt{n}) \\ & - \mathbb{E}\Phi(x, \bar{X}) I(|\Phi(x, \bar{X})| \leq n) I(\|x\| \leq \sqrt{n}) I(\|\bar{X}\| \leq \sqrt{n}) \\ & - \mathbb{E}\Phi(X, y) I(|\Phi(X, y)| \leq n) I(\|X\| \leq \sqrt{n}) I(\|y\| \leq \sqrt{n}) \\ & + \mathbb{E}\Phi(X, \bar{X}) I(|\Phi(X, \bar{X})| \leq n) I(\|X\| \leq \sqrt{n}) I(\|\bar{X}\| \leq \sqrt{n}). \end{aligned} \tag{2.14}$$

Hence

$$h(x, y) = \Phi(x, y) + \phi(x, y) \tag{2.15}$$

with the representation

$$\phi(x, y) = \theta(x, y) - \mathbb{E}\theta(x, \bar{X}) - \mathbb{E}\theta(X, y) + \mathbb{E}\theta(X, \bar{X}), \tag{2.16}$$

where

$$\begin{aligned} \theta(x, y) \;=\; & -\Phi(x, y) I(\|x\| > \sqrt{n}) - \Phi(x, y) I(\|y\| > \sqrt{n}) \\ & + \Phi(x, y) I(\|x\| > \sqrt{n}) I(\|y\| > \sqrt{n}) \\ & - \Phi(x, y) I(|\Phi(x, y)| > n) I(\|x\| \leq \sqrt{n}) I(\|y\| \leq \sqrt{n}). \end{aligned} \tag{2.17}$$

Simple calculations show that

$$\mathbb{E}\phi^2(X, \bar{X}) = \mathbb{E}\theta^2(X, \bar{X}) + 3(\mathbb{E}\theta(X, \bar{X}))^2, \tag{2.18}$$

and,

$$\mathbb{E}\theta^2(X, \bar{X}) \leq 12\alpha_n + 4\beta_n \quad \text{and} \quad |\mathbb{E}\theta(X, \bar{X})| \leq \frac{1}{n}(\alpha_n + \beta_n). \tag{2.19}$$

Then

$$\mathbb{E}\phi^2(X, \bar{X}) \leq 18\alpha_n + 10\beta_n. \tag{2.20}$$

By substituting (2.15) in (2.13) and using the Hölder's inequality and (2.20) we obtain

$$\sup\Big\{\Big|\Big((\mathbb{Q}^2 - \bar{\mathbb{Q}}^2)f, f\Big)\Big| : |f| = 1\Big\} \leq 2(18\alpha_n + 10\beta_n)^{1/2} + 18\alpha_n + 10\beta_n. \tag{2.21}$$

The condition $\alpha_n + \beta_n \leq p$ of this lemma is valid by assumption (1.11), since $\alpha_n + \beta_n \leq L_n$. Therefore from (2.12) and (2.21) we have for all $j \geq 1$

$$|q_j^2 - \bar{q}_j^2| \leq 40\sqrt{p}. \tag{2.22}$$

If $j = 1, \ldots, s$ then $|q_j| \geq |q_s| > 0$ and

$$p = \exp\{-(16s)^3/|q_s|\} \leq \exp\{-(16s)^3/|q_j|\} \leq 16^{-4}q_j^4.$$

From here and (2.22) we get (2.6).

As in (2.15) we can define the degenerate kernel $\bar{h}(x, y)$ and write

$$\bar{h}(x, y) = \Phi(x, y) + \bar{\phi}(x, y) \tag{2.23}$$

truncating at m instead of n. Using (2.5) we can assume throughout that $h(x, y)$ and $\bar{h}(x, y)$ are linear in their arguments $x, y \in \mathbb{R}^\infty$.

3 Non-degeneracy condition

For a matrix $\mathbb{A} = (a_{ij})$, $1 \leq i, j \leq s$, we denote by $|\mathbb{A}|_\infty = \max\{|a_{ij}| : 1 \leq i, j \leq s\}$. Define the random matrix $\mathbb{A}(z, h) = (a_{ij}(h))$, $1 \leq i, j \leq s$, with entries $a_{ij}(h) = h(z_i, \bar{z}_j)$ where z_i and $\bar{z}_j$ are independent copies of z. Following Bentkus and Götze (1999) we say that z and h satisfy the non-degeneracy condition $N(p, \delta, s, z, h)$ if

$$\mathbb{P}\Big\{|\mathbb{A}(z, h) - \mathbb{I}|_\infty \leq \delta\Big\} \geq p,$$

where $\mathbb{I}$ denotes the identity matrix, $0 < p \leq 1$, $\delta \geq 0$ and $s \geq 1$. From their Lemma 5.3 for a Gaussian random vector G and $\bar{q}_s \neq 0$ for some $s \geq 1$ the condition $N(2p, \delta, s, \epsilon G, h)$ is fulfilled with $p = \exp\{-c_s/(\epsilon^2|\bar{q}_s|)\}$, provided that $\delta \leq 1$ and $4s\bar{\sigma}\epsilon^2 \leq \delta$, where $\bar{\sigma}^2 = \mathbb{E}h^2(X, \bar{X})$ and c_s is a positive constant. Their inequality (5.25) gives the constant

$$c_1(s) = 1/(2^{s/2}\Gamma(1 + s/2)) \geq (s + 2)^{-(s+1)/2}\exp(s/2),$$

and moreover $c_s = 8s$. Since $\sigma = 1$, using (2.14) we have $\bar\sigma \le 4$. Hence taking into account (2.6) we have

$$\mathbb{P}\Big\{|\mathbb{A}(\epsilon G, h) - \mathbb{I}|_\infty \le \delta\Big\} \ge 2p \tag{3.1}$$

with $p = \exp\{-16s/(\epsilon^2|q_s|)\}$ for $\delta \le 1$ and $16s\epsilon^2 \le \delta$. Set $\epsilon = \delta = 1/(16s)$. Then in (3.1) $p = \exp\{-(16s)^3/(|q_s|)\}$.

The following lemma provides conditions for the non-degeneracy of the non-Gaussian random vectors under the minimal moment condition $\mathbb{E}\Phi^2(X, \bar X) = 1$. Denote

$$Y = (2m)^{-1/2}(\tilde X_1 + \ldots + \tilde X_m), \quad m \ge 1 \tag{3.2}$$

and let $\bar Y = (2m)^{-1/2}(\tilde{\bar X}_1 + \ldots + \tilde{\bar X}_m)$ be a independent copy of Y. Here $\tilde X_k = X_k - X_k^1$, $\tilde{\bar X}_l = \bar X_l - \bar X_l^1$, $1 \le k, l \le m$, are independent symmetrizations and X_k, X_k^1, $\bar X_l$, $\bar X_l^1$, $1 \le k, l \le m$ are independent copies of X. Hence we can write

$$h(Y, \bar Y) = \frac{1}{2m} \sum_{k=1}^m \sum_{l=1}^m h(\tilde X_k, \tilde{\bar X}_l). \tag{3.3}$$

Further let $Y_1, \ldots, Y_s$ and $\bar Y_1, \ldots, \bar Y_s$ be independent copies of Y, that is,

$$Y_i = (2m)^{-1/2} \sum_{k=1}^m \tilde X_{ik}, \quad \bar Y_j = (2m)^{-1/2} \sum_{l=1}^m \tilde{\bar X}_{jl},$$

with the symmetrizations $\tilde X_{ik} = X_{ik} - X_{ik}^1$, $\tilde{\bar X}_{jl} = \bar X_{jl} - \bar X_{jl}^1$, where X_{ik}, $\bar X_{jl}$ and X_{ik}^1, $\bar X_{jl}^1$, $1 \le k, l \le m$, are independent copies of the sequence X_i, $\bar X_j$, $1 \le i, j \le s$. Therefore by (3.3) the entries $a_{ij}(h)$ of the random matrix $\mathbb{A}(\epsilon Y, h)$ can be written as

$$a_{ij}(h) = \frac{\epsilon^2}{2m} \sum_{k=1}^m \sum_{l=1}^m h(\tilde X_{ik}, \tilde{\bar X}_{jl}). \tag{3.4}$$

A similar expression holds for the entries $a_{ij}(\bar h)$ of the random matrix $\mathbb{A}(\epsilon Y, \bar h)$.

Lemma 3.1. *Let $q_s \ne 0$ for some $s \ge 1$ and*

$$\epsilon = \delta = \frac{1}{16s}, \quad p = \exp\{-16s/(\epsilon^2|q_s|)\}. \tag{3.5}$$

Then the condition $N(2p, \delta, s, \epsilon G, h)$ implies the condition $N(p, 8\delta, s, \epsilon Y, h)$ provided that $m \ge m(s, q_s)$, where

$$m(s, q_s) = \min\{m \ge 1 : L_m \le 2^{-15}s^{-3}p\} \tag{3.6}$$

and $L_m = \alpha_m + \beta_m + \gamma_m + \lambda_m$.

Proof. We shall follow Bentkus and Götze (1999). It suffices to show that

$$\mathbb{P}\Big\{|\mathbb{A}(\epsilon Y, h) - \mathbb{I}|_\infty \le 8\delta\Big\} \ge p \tag{3.7}$$

for $m \ge m(s, q_s)$. First we shall show that for $m \ge 1$

$$\begin{aligned} &\mathbb{P}\Big\{|\mathbb{A}(\epsilon Y, h) - \mathbb{I}|_\infty \le 8\delta\Big\} \\ \ge\ &\mathbb{P}\Big\{|\mathbb{A}(\epsilon Y, \bar h) - \mathbb{I}|_\infty \le 4\delta\Big\} - \frac{9}{16^2}(\alpha_m + \beta_m) - \frac{9}{16^2}(\alpha_n + \beta_n). \end{aligned} \tag{3.8}$$

Using (2.15), (2.23) and (3.4) we write $\mathbb{A}(\epsilon Y,h)=\mathbb{A}(\epsilon Y,\bar{h})+\mathbb{A}(\epsilon Y,\phi-\bar{\phi})$ and

$$|\mathbb{A}(\epsilon Y,h)-\mathbb{I}|_\infty \le |\mathbb{A}(\epsilon Y,\bar{h})-\mathbb{I}|_\infty+|\mathbb{A}(\epsilon Y,\phi-\bar{\phi})|_\infty. \tag{3.9}$$

The random variables $a_{ij}(\phi-\bar{\phi})$ are identically distributed and

$$\begin{aligned}
\mathbb{P}\Big\{|\mathbb{A}(\epsilon Y,\phi-\bar{\phi})|_\infty \ge 4\delta\Big\} &\le \frac{1}{16\delta^2}\mathbb{E}\max_{1\le i,j\le s} a_{ij}^2(\phi-\bar{\phi}) \\
&\le \frac{s^2}{16\delta^2}\mathbb{E}a_{11}^2(\phi-\bar{\phi}) \\
&\le \frac{s^2}{8\delta^2}\Big(\mathbb{E}a_{11}^2(\phi)+\mathbb{E}a_{11}^2(\bar{\phi})\Big) \\
&\le \frac{s^2}{8\delta^2}\epsilon^4\Big(\mathbb{E}\phi^2(X,\bar{X})+\mathbb{E}\bar{\phi}^2(X,\bar{X})\Big) \\
&\le \frac{1}{8\cdot 16^2}(18\alpha_n+10\beta_n+18\alpha_m+10\beta_m), \qquad \text{using (2.21)} \\
&\le \frac{9}{16^2}(\alpha_m+\beta_m)+\frac{9}{16^2}(\alpha_n+\beta_n).
\end{aligned} \tag{3.10}$$

Further applying these results we have

$$\begin{aligned}
&\mathbb{P}\Big\{|\mathbb{A}(\epsilon Y,h)-\mathbb{I}|_\infty \le 8\delta\Big\} \\
&\ge \mathbb{P}\Big\{|\mathbb{A}(\epsilon Y,\bar{h})-\mathbb{I}|_\infty+|\mathbb{A}(\epsilon Y,\phi-\bar{\phi})|_\infty \le 8\delta\Big\} \\
&\ge \mathbb{E}I\Big\{|\mathbb{A}(\epsilon Y,\bar{h})-\mathbb{I}|_\infty+|\mathbb{A}(\epsilon Y,\phi-\bar{\phi})|_\infty \le 8\delta\Big\}I\Big\{|\mathbb{A}(\epsilon Y,\phi-\bar{\phi})|_\infty \le 4\delta\Big\} \\
&\ge \mathbb{E}I\Big\{|\mathbb{A}(\epsilon Y,\bar{h})-\mathbb{I}|_\infty \le 4\delta\Big\}I\Big\{|\mathbb{A}(\epsilon Y,\phi-\bar{\phi})|_\infty \le 4\delta\Big\} \\
&\ge \mathbb{P}\Big\{|\mathbb{A}(\epsilon Y,\bar{h})-\mathbb{I}|_\infty \le 4\delta\Big\}-\mathbb{P}\Big\{|\mathbb{A}(\epsilon Y,\phi-\bar{\phi})|_\infty \ge 4\delta\Big\} \\
&\ge \mathbb{P}\Big\{|\mathbb{A}(\epsilon Y,\bar{h})-\mathbb{I}|_\infty \le 4\delta\Big\}-\frac{9}{16^2}(\alpha_m+\beta_m)-\frac{9}{16^2}(\alpha_n+\beta_n).
\end{aligned} \tag{3.11}$$

This proves (3.8).

Next we shall show that for $m\ge 1$

$$\mathbb{P}\Big\{|\mathbb{A}(\epsilon Y,\bar{h})-\mathbb{I}|_\infty \le 4\delta\Big\} \ge \mathbb{P}\Big\{|\mathbb{A}(\epsilon G,\bar{h})-\mathbb{I}|_\infty \le 2\delta\Big\}-3\cdot 4^3\cdot s^3\lambda_m-15\cdot 4^4\cdot s^3\gamma_m. \tag{3.12}$$

Introduce the four times continuously differentiable function $\varphi:\mathbb{R}\mapsto\mathbb{R}$ with bounded derivatives such that $\varphi(x)=1$, for $|x|\le 1$, $\varphi(x)=0$, for $|x|\ge 2$, and

$$\varphi(x)=\Big(1-(\frac{1}{3}x^2-\frac{1}{3})^5\Big)^5, \quad \text{for} \quad 1\le|x|\le 2.$$

It is obvious that $|\varphi|_\infty=1$, where $|\varphi|_\infty=\sup_x|\varphi(x)|$. The first four derivatives of $\varphi(x)$ for $1\le|x|\le 2$ have the form

$$\varphi^{(1)}(x) = -\frac{50}{3}x\Big\{(\frac{1}{3}x^2-\frac{1}{3})-(\frac{1}{3}x^2-\frac{1}{3})^6\Big\}^4,$$

$$
\begin{aligned}
\varphi^{(2)}(x) &= -\frac{50}{3}\Big\{(\frac{1}{3}x^2-\frac{1}{3})-(\frac{1}{3}x^2-\frac{1}{3})^6\Big\}^3 \\
&\quad\times\Big\{\frac{8}{3}+9(\frac{1}{3}x^2-\frac{1}{3})-16(\frac{1}{3}x^2-\frac{1}{3})^5-49(\frac{1}{3}x^2-\frac{1}{3})^6\Big\}, \\
\varphi^{(3)}(x) &= -\frac{50}{3}x^2\Big\{(\frac{1}{3}x^2-\frac{1}{3})-(\frac{1}{3}x^2-\frac{1}{3})^6\Big\}^2\varphi_1(x), \\
\varphi^{(4)}(x) &= -\frac{50}{3}\Big\{(\frac{1}{3}x^2-\frac{1}{3})-(\frac{1}{3}x^2-\frac{1}{3})^6\Big\} \\
&\quad\times\Big\{(\frac{2}{3}-4(\frac{1}{3}x^2-\frac{1}{3})^5)\varphi_1(x)+((\frac{1}{3}x^2-\frac{1}{3})-(\frac{1}{3}x^2-\frac{1}{3})^6)\varphi_2(x)\Big\},
\end{aligned}
$$

where

$$
\begin{aligned}
\varphi_1(x) &= \frac{16}{3}+18(\frac{1}{3}x^2-\frac{1}{3})-64(\frac{1}{3}x^2-\frac{1}{3})^5-206(\frac{1}{3}x^2-\frac{1}{3})^6 \\
&\quad+192(\frac{1}{3}x^2-\frac{1}{3})^{10}+588(\frac{1}{3}x^2-\frac{1}{3})^{11}, \\
\varphi_2(x) &= \frac{52}{3}+54(\frac{1}{3}x^2-\frac{1}{3})-\frac{640}{3}(\frac{1}{3}x^2-\frac{1}{3})^4-1528(\frac{1}{3}x^2-\frac{1}{3})^5 \\
&\quad-2678(\frac{1}{3}x^2-\frac{1}{3})^6+1280(\frac{1}{3}x^2-\frac{1}{3})^9+6424(\frac{1}{3}x^2-\frac{1}{3})^{10} \\
&\quad+13524(\frac{1}{3}x^2-\frac{1}{3})^{11}.
\end{aligned}
$$

Note that $\varphi^{(k)}(x)=0$, $k=1,\ldots,4$ for $|x|\le 1$ and $|x|\ge 2$. Thus

$$
|\varphi^{(1)}|_\infty \le 34, \quad |\varphi^{(2)}|_\infty \le 1309, \quad |\varphi^{(3)}|_\infty \le 7160, \quad |\varphi^{(4)}|_\infty \le 108719. \tag{3.13}
$$

From the definition of φ we can write

$$
\begin{aligned}
\mathbb{P}\Big\{|\mathbb{A}(\epsilon Y,\bar{h})-\mathbb{I}|_\infty \le 4\delta\Big\} &\ge \mathbb{E}\prod_{1\le i,j\le s}\varphi\Big(\frac{\epsilon^2\bar{h}(Y_i,\bar{Y}_j)-\delta_{ij}}{2\delta}\Big) \\
&= \mathbb{E}\prod_{1\le i,j\le s}\varphi\Big(\frac{\epsilon^2\bar{h}(G_i,\bar{G}_j)-\delta_{ij}}{2\delta}\Big)+\mathcal{R} \\
&\ge \mathbb{P}\Big\{|\mathbb{A}(\epsilon G,\bar{h})-\mathbb{I}|_\infty \le 2\delta\Big\}-|\mathcal{R}|,
\end{aligned}
\tag{3.14}
$$

where

$$
\mathcal{R}=\mathbb{E}\prod_{1\le i,j\le s}\varphi\Big(\frac{\epsilon^2\bar{h}(Y_i,\bar{Y}_j)-\delta_{ij}}{2\delta}\Big)-\mathbb{E}\prod_{1\le i,j\le s}\varphi\Big(\frac{\epsilon^2\bar{h}(G_i,\bar{G}_j)-\delta_{ij}}{2\delta}\Big),
$$

and δ_{ij} is the Kronecker delta. In what follows we shall show that

$$
|\mathcal{R}| \le 3\cdot 4^3 s^3\lambda_m + 15\cdot 4^4 s^3\gamma_m. \tag{3.15}
$$

In $\mathcal{R}$ we shall replace the random vectors Y_j (or $\bar{Y}_j$) by corresponding Gaussian random vectors G_j (respectively, $\bar{G}_j$) for $1\le j\le s$ and then show that after $2s$

replacements the general error is bounded by the right hand side of (3.15). First we write the formula containing all $2s$ replacements

$$\begin{aligned}\mathcal{R} \;=\; & \sum_{l=1}^{s}\mathbb{E}\Big\{\prod_{i=l+1}^{s}\mathbb{E}_Y\prod_{j=1}^{s}\varphi\Big(\frac{\epsilon^2\bar h(Y,\bar Y_j)-\delta_{ij}}{2\delta}\Big)\prod_{i=1}^{l-1}\mathbb{E}_G\prod_{j=1}^{s}\varphi\Big(\frac{\epsilon^2\bar h(G,\bar Y_j)-\delta_{ij}}{2\delta}\Big)\\
&\times\Big[\mathbb{E}_Y\prod_{j=1}^{s}\varphi\Big(\frac{\epsilon^2\bar h(Y,\bar Y_j)-\delta_{lj}}{2\delta}\Big)-\mathbb{E}_G\prod_{j=1}^{s}\varphi\Big(\frac{\epsilon^2\bar h(G,\bar Y_j)-\delta_{lj}}{2\delta}\Big)\Big]\Big\}\\
&+\sum_{l=1}^{s}\mathbb{E}\Big\{\prod_{j=l+1}^{s}\mathbb{E}_{\bar Y}\prod_{i=1}^{s}\varphi\Big(\frac{\epsilon^2\bar h(G_i,\bar Y)-\delta_{ij}}{2\delta}\Big)\prod_{j=1}^{l-1}\mathbb{E}_{\bar G}\prod_{i=1}^{s}\varphi\Big(\frac{\epsilon^2\bar h(G_i,\bar G)-\delta_{ij}}{2\delta}\Big)\\
&\times\Big[\mathbb{E}_{\bar Y}\prod_{i=1}^{s}\varphi\Big(\frac{\epsilon^2\bar h(G_i,\bar Y)-\delta_{il}}{2\delta}\Big)-\mathbb{E}_{\bar G}\prod_{i=1}^{s}\varphi\Big(\frac{\epsilon^2\bar h(G_i,\bar G)-\delta_{il}}{2\delta}\Big)\Big]\Big\}.\end{aligned}$$

Since $|\varphi|_\infty \le 1$ then from this formula we get

$$|\mathcal{R}| \le s\Big(\mathbb{E}J(\bar Y)+\mathbb{E}J(G)\Big), \tag{3.16}$$

where

$$J(\bar Y) = |\mathbb{E}_Y\prod_{j=1}^{s}\varphi\Big(\frac{\epsilon^2\bar h(Y,\bar Y_j)-\delta_{1j}}{2\delta}\Big)-\mathbb{E}_G\prod_{j=1}^{s}\varphi\Big(\frac{\epsilon^2\bar h(G,\bar Y_j)-\delta_{1j}}{2\delta}\Big)|,$$

$$J(G) = |\mathbb{E}_Y\prod_{j=1}^{s}\varphi\Big(\frac{\epsilon^2\bar h(Y,G_j)-\delta_{1j}}{2\delta}\Big)-\mathbb{E}_G\prod_{j=1}^{s}\varphi\Big(\frac{\epsilon^2\bar h(G,G_j)-\delta_{1j}}{2\delta}\Big)|.$$

Now we shall estimate $J(\bar Y)$. Fix $\bar y_j = \bar Y_j$ for $1 \le j \le s$. Introduce the random vectors

$$\mathbf{y} = \Big(\epsilon^2\bar h(Y,\bar y_1),\dots,\epsilon^2\bar h(Y,\bar y_s)\Big),\quad \mathbf{g} = \Big(\epsilon^2\bar h(G,\bar y_1),\dots,\epsilon^2\bar h(G,\bar y_s)\Big)$$

and the Frechet differentiable function

$$\psi(x) = \varphi\Big(\frac{x_1-1}{2\delta}\Big)\varphi\Big(\frac{x_2}{2\delta}\Big)\cdots\varphi\Big(\frac{x_s}{2\delta}\Big),\quad x=(x_1,\dots,x_s)\in\mathbb{R}^s.$$

We can rewrite $J(\bar y)$ as

$$J(\bar y) = |\mathbb{E}\psi(\mathbf{y})-\mathbb{E}\psi(\mathbf{g})|,\quad \bar y=(\bar y_1,\dots,\bar y_s)\in\mathbb{R}^s.$$

Recall that $Y \overset{d}{=} (\tilde X_1+\cdots+\tilde X_m)/\sqrt{2m}$ and we can write G, as
$G \overset{d}{=} (\tilde G_1+\cdots+\tilde G_m)/\sqrt{2m}$. Therefore

$$\mathbf{y} \overset{d}{=} \xi_1+\cdots+\xi_m,\quad \text{with}\quad \xi_j = \Big(\epsilon^2\bar h(\tilde X_j,\bar y_l)/\sqrt{2m}\Big)_{1\le l\le s},$$

$$\mathbf{g} \overset{d}{=} \zeta_1+\cdots+\zeta_m,\quad \text{with}\quad \zeta_j = \Big(\epsilon^2\bar h(\tilde G_j,\bar y_l)/\sqrt{2m}\Big)_{1\le l\le s},$$

in addition the random vectors $\xi_1, \dots, \xi_m$ (respectively, $\zeta_1, \dots, \zeta_m$) take values in $\mathbb{R}^s$ and are independent and have identical, symmetric distributions. The Gaussian random vectors ζ_j have mean zero and their covariances are equal to those of ξ_j. Further, denoting

$$\eta_j = |\mathbb{E}\psi(\mathbf{s}_j + \xi_j) - \mathbb{E}\psi(\mathbf{s}_j + \zeta_j)|, \quad \mathbf{s}_j = \zeta_1 + \dots + \zeta_{j-1} + \xi_{j+1} + \dots + \xi_m,$$

we get

$$J(\bar{y}) \le \eta_1 + \dots + \eta_m. \tag{3.17}$$

Expanding into a Taylor series in powers of ξ_j and ζ_j and using symmetry and the equality of the covariances of ξ_j and ζ_j we obtain

$$\eta_j \le \sum_{\nu_1 + \dots + \nu_s = 4} \frac{1}{\nu_1! \cdots \nu_s!} \Big(\mathbb{E}|\xi_j|^4 + \mathbb{E}|\zeta_j|^4\Big) \sup\Big\{\Big|\Big(D^\nu \psi\Big)(x)\Big| : x \in \mathbb{R}^s\Big\} \tag{3.18}$$

where

$$\mathbb{E}|\xi_j|^4 = \frac{1}{(2m)^2}\epsilon^8 \mathbb{E}\Big(\sum_{l=1}^s \bar{h}^2(\tilde{X}, \bar{y}_l)\Big)^2, \quad \mathbb{E}|\zeta_j|^4 = \frac{1}{(2m)^2}\epsilon^8 \mathbb{E}\Big(\sum_{l=1}^s \bar{h}^2(\tilde{G}, \bar{y}_l)\Big)^2,$$

and $D^\nu \psi$ is derivative of order $\nu = (\nu_1, \dots, \nu_s)$ of the function ψ, that is,

$$\Big(D^\nu \psi\Big)(x) = \frac{\partial^{\nu_1}}{\partial x_1^{\nu_1}} \varphi\Big(\frac{x_1 - 1}{2\delta}\Big) \frac{\partial^{\nu_2}}{\partial x_2^{\nu_2}} \varphi\Big(\frac{x_2}{2\delta}\Big) \cdots \frac{\partial^{\nu_s}}{\partial x_s^{\nu_s}} \varphi\Big(\frac{x_s}{2\delta}\Big).$$

For $\nu_1 + \dots + \nu_s = 4$, $\nu_j \ge 0$, $j = 1, \dots, s$, under the condition $|\varphi|_\infty \le 1$ we have

$$\sup\Big\{\Big|\Big(D^\nu \psi\Big)(x)\Big| : x \in \mathbb{R}^s\Big\}$$

$$\le (2\delta)^{-4} \max\Big\{|\varphi^{(4)}|_\infty, |\varphi^{(1)}|_\infty \cdot |\varphi^{(3)}|_\infty, |\varphi^{(1)}|_\infty^2 \cdot |\varphi^{(2)}|_\infty, |\varphi^{(2)}|_\infty^2, |\varphi^{(1)}|_\infty^4\Big\}.$$

According to (3.13)

$$\sup\Big\{\Big|\Big(D^\nu \psi\Big)(x)\Big| : x \in \mathbb{R}^s\Big\} \le 17^2 \cdot 77^2 \cdot (2\delta)^{-4}. \tag{3.19}$$

Further in (3.18)

$$\sum_{\nu_1 + \dots + \nu_s = 4} \frac{1}{\nu_1! \cdots \nu_s!} = \frac{1}{4!} s^4$$

and

$$\mathbb{E}|\xi_j|^4 \le \frac{4s\epsilon^8}{m^2} \sum_{l=1}^s \mathbb{E}\bar{h}^4(X, \bar{y}_l).$$

Fixing $x = X$ and taking into account (3.2) – (3.3) we can write

$$\begin{aligned} \mathbb{E}\bar{h}^4(x, \bar{Y}_l) &= \mathbb{E}\Big\{\frac{1}{\sqrt{2m}} \sum_{j=1}^m \bar{h}(x, \tilde{\bar{X}}_j)\Big\}^4 \\ &= \frac{1}{4m}\mathbb{E}\bar{h}^4(x, \tilde{\bar{X}}) + \frac{3(m-1)}{4m}\Big(\mathbb{E}\bar{h}^2(x, \tilde{\bar{X}})\Big)^2 \\ &\le \frac{2}{m}\mathbb{E}\bar{h}^4(x, \bar{X}) + 3\Big(\mathbb{E}\bar{h}^2(x, \bar{X})\Big)^2. \end{aligned}$$

Here

$$\mathbb{E}\bar{h}^4(X,\bar{X}) \le 4^3\mathbb{E}\Phi^4 I(|\Phi| \le m) + 3\cdot 4^3\mathbb{E}\|X\|^4 I(\|X\| \le \sqrt{m})$$

and

$$\mathbb{E}\Big(\mathbb{E}_{\bar{X}}\bar{h}^2(X,\bar{X})\Big)^2 \le 4^4\mathbb{E}\|X\|^4 I(\|X\| \le \sqrt{m}).$$

Since the random variable $\bar{h}(G,\bar{y}_l)$ is Gaussian then for $\mathbb{E}|\zeta_j|^4$ we have the bounds

$$\begin{aligned}\mathbb{E}|\zeta_j|^4 &\le \frac{4s\epsilon^8}{m^2}\sum_{l=1}^{s}\mathbb{E}\bar{h}^4(G,\bar{y}_l)\\ &= \frac{12s\epsilon^8}{m^2}\sum_{l=1}^{s}\Big(\mathbb{E}\bar{h}^2(X,\bar{y}_l)\Big)^2\\ &\le \frac{12s\epsilon^8}{m^2}\sum_{l=1}^{s}\mathbb{E}\bar{h}^4(X,y_l).\end{aligned}$$

Therefore from (3.17) – (3.19) we see that

$$\mathbb{E}J(\bar{Y}) \le 3\cdot 4^3 s^2\lambda_m + 3^2 4^3 5 s^2\gamma_m. \tag{3.20}$$

A similar bound is valid for $\mathbb{E}J(G)$, that is,

$$\mathbb{E}J(G) \le 15\cdot 4^3 s^2\gamma_m. \tag{3.21}$$

Combining (3.16), (3.20) and (3.21) we get the estimate (3.15).

Further we shall show that in (3.12)

$$\mathbb{P}\{|\mathbb{A}(\epsilon G,\bar{h}) - \mathbb{I}|_\infty \le 2\delta\} \ge \mathbb{P}\{|\mathbb{A}(\epsilon G,h) - \mathbb{I}|_\infty \le \delta\} - 5(L_m + L_n). \tag{3.22}$$

Arguing as in (3.9) – (3.11) with G replacing Y we have

$$\begin{aligned}&\mathbb{P}\{|\mathbb{A}(\epsilon G,\bar{h}) - \mathbb{I}|_\infty \le 2\delta\} \qquad (3.23)\\ &\ge \mathbb{P}\{|\mathbb{A}(\epsilon G,h) - \mathbb{I}|_\infty \le \delta\} - \frac{1}{128}(\mathbb{E}\phi^2(G,\bar{G}) + \mathbb{E}\bar{\phi}^2(G,\bar{G})).\end{aligned}$$

Simple calculations show that

$$\mathbb{E}\phi^2(G,\bar{G}) = \mathbb{E}\theta^2(G,\bar{G}) + 2\mathbb{E}(\mathbb{E}_X\theta(X,G))^2 + (\mathbb{E}\theta(X,\bar{X}))^2. \tag{3.24}$$

In (3.24), using (1.8)

$$\begin{aligned}\mathbb{E}\theta^2(G,\bar{G}) &\le 12\mathbb{E}\|G\|^2 I(\|G\| > \sqrt{n}) + 4\mathbb{E}\Phi^2(G,\bar{G})I(|\Phi(G,\bar{G})| > n) \qquad (3.25)\\ &\le \frac{12}{n}\mathbb{E}\|G\|^4 + \frac{4}{n}\mathbb{E}\Phi^4(G,\bar{G})\\ &= \frac{24}{n}\mathbb{E}(\mathbb{E}_X\Phi^2(X,G))^2\\ &= \frac{24}{n}\mathbb{E}\Big(\sum_{j=1}^{\infty}q_j^2\tau_j^2\Big)^2 \le \frac{72}{n} \le 144(\alpha_n + \gamma_n).\end{aligned}$$

Similarly

$$\begin{aligned}\mathbb{E}\theta^2(G,\bar G) &\le \frac{1}{n^4}\Big(12\mathbb{E}\|G\|^{10}+4\mathbb{E}\Phi^6(G,\bar G)\Big) \\ &\le \frac{612}{n^4}\le\frac{5000}{n^2}(\alpha_n^2+\gamma_n^2).\end{aligned}\tag{3.26}$$

Further in (3.24) we have

$$\begin{aligned}&\mathbb{E}\Big(\mathbb{E}_X\theta(G,X)\Big)^2 \\ &\le \frac{2}{n}\mathbb{E}\Big(\mathbb{E}_X|\Phi(G,X)|\|X\|\,I(\|X\|>\sqrt{n})\Big)^2 \\ &\quad+\frac{2}{n^6}\mathbb{E}\Big(\mathbb{E}_X\Phi^4(G,X)I(\|X\|>\sqrt{n})\Big)^2 \\ &= \frac{2}{n}\mathbb{E}\|X\|\,\|\bar X\|\,I(\|X\|>\sqrt{n})I(\|\bar X\|>\sqrt{n})\mathbb{E}_G|\Phi(G,X)|\,|\Phi(G,\bar X)| \\ &\quad+\frac{2}{n^6}\mathbb{E}I(\|X\|\le\sqrt{n})I(\|\bar X\|\le\sqrt{n})\mathbb{E}_G\Phi^4(G,X)\Phi^4(G,\bar X) \\ &\le \frac{2}{n}\mathbb{E}\|X\|\,\|\bar X\|\,I(\|X\|>\sqrt{n})I(\|\bar X\|>\sqrt{n})\Big(\mathbb{E}_G\Phi^2(G,X)\Big)^{1/2}\Big(\mathbb{E}_G\Phi^2(G,\bar X)\Big)^{1/2} \\ &\quad+\frac{2}{n^6}\mathbb{E}I(\|X\|\le\sqrt{n})I(\|\bar X\|\le\sqrt{n})\Big(\mathbb{E}_G\Phi^8(G,X)\Big)^{1/2}\Big(\mathbb{E}_G\Phi^8(G,\bar X)\Big)^{1/2} \\ &= \frac{2}{n}\alpha_n^2+\frac{210}{n}\gamma_n^2,\end{aligned}\tag{3.27}$$

since the random variable $\Phi(G,x)$ is Gaussian. Combining (2.19) and (3.24) – (3.26) we obtain the inequalities

$$\mathbb{E}\phi^2(G,\bar G)\le 600(\alpha_n+\beta_n+\gamma_n),\quad \mathbb{E}\phi^2(G,\bar G)\le\frac{6000}{n^2}(\alpha_n^2+\beta_n^2+\gamma_n^2).\tag{3.28}$$

The bound (3.22) follows from (3.23) and the first inequality (3.28). From (3.8), (3.12) and (3.22) we see that for all $m\ge 1$

$$\mathbb{P}\{|\mathbb{A}(\epsilon Y,h)-\mathbb{I}|_\infty\le 8\delta\}\ge\mathbb{P}\{|\mathbb{A}(\epsilon G,h)-\mathbb{I}|_\infty\le\delta\}-6L_n-2^{12}s^3L_m.\tag{3.29}$$

By the assumption (1.11), definition (3.6) and (3.29), if $m\ge m(s,q_s)$ then (3.7) holds. This proves Lemma 3.1.

4 Estimates of characteristic functions

In what follows we shall use the abridged notation $U_n(\Phi)$ or $U(\Phi;X_1,\dots,X_n)$ for U-statistics of the type (1.1) to show the dependence on Φ or $X_1,\dots,X_n$. For example, using (2.14)

$$U_n(g)=U_n(h)+S+s,\tag{4.1}$$

where

$$S = \frac{n-1}{n}\sum_{j=1}^{n}(\mathbb{E}_{\bar{X}}g(X_j,\bar{X}) - \mathbb{E}g(X,\bar{X})), \quad s = \frac{1}{2}(n-1)\mathbb{E}g(X,\bar{X})$$

and

$$g(x,y) = \Phi(x,y)I(|\Phi(x,y)| \le n)I(\|x\| \le \sqrt{n})I(\|y\| \le \sqrt{n}), \quad x,y \in \mathfrak{X}.$$

Define the random variable

$$\Phi_\infty(G,G) = \sum_{j=1}^{\infty} q_j(\tau_j^2 - 1). \tag{4.2}$$

Here the series converges in L^2 and moreover $\mathbb{E}\Phi_\infty^2(G,G) = 2\mathbb{E}\Phi^2(X,\bar{X})$. Substituting $G \overset{d}{=} (G_1 + \cdots + G_n)/\sqrt{n}$ and using (4.2) we have

$$\Phi_\infty(G,G) \overset{d}{=} \frac{2}{n}\sum_{1\le i<j\le n}\Phi(G_i,G_j) + \frac{1}{n}\sum_{j=1}^{n}\Phi_\infty(G_j,G_j)$$

with $\Phi(G,\bar{G}) = (\mathbb{Q}G,\bar{G})$. Then for the random variable $U_\infty = U_\infty(\Phi)$ defined by (1.4) we can write

$$U_\infty(\Phi) \overset{d}{=} U(\Phi; G_1,\ldots,G_n) + \frac{1}{2n}\sum_{j=1}^{n}\Phi_\infty(G_j,G_j). \tag{4.3}$$

Note that the representation of U_∞ equivalent to (4.3) can be given in the form of stochastic integrals (see, for example, Borovskikh(1996)).

According to (2.15) and (4.3)

$$U_\infty(\Phi) = U_\infty(h) - U_\infty(\phi). \tag{4.4}$$

Now we shall estimate the difference of the characteristic functions

$$\mathbb{E}e\{tU_n(g)\} - \mathbb{E}e\{tU_\infty(\Phi)\}, \quad t \in \mathbb{R}, \tag{4.5}$$

where $e\{x\} = \exp\{ix\}$, $x \in \mathbb{R}$. With the help of (4.1) and (4.4) write

$$\begin{aligned}\mathbb{E}e\{tU_n(g)\} &= \mathbb{E}e\{tU_n(h)\} + \varrho_1(t),\\ \mathbb{E}e\{tU_\infty(\Phi)\} &= \mathbb{E}e\{tU_\infty(h)\} + \varrho_2(t),\end{aligned} \tag{4.6}$$

where

$$\begin{aligned}\varrho_1(t) &= \mathbb{E}e\{tU_n(h)\}(e\{t(S+s)\} - 1),\\ \varrho_2(t) &= \mathbb{E}e\{tU_\infty(h)\}(e\{-tU_\infty(\phi)\} - 1).\end{aligned}$$

Further we shall use the abbreviations

$$f = th/n, \quad f_\infty = th_\infty/(2n) \tag{4.7}$$

and the following bound $\kappa(t) = \kappa_1(t) + \kappa_2(t)$ of characteristic functions, where

$$\kappa_1(t) = \sup_L \Big| \mathbb{E} e\Big\{ \sum_{1 \le i < j \le k} f(X_i, X_j) + L(X_1, \dots, X_k) \Big\} \Big|,$$

$$\kappa_2(t) = \sup_L \Big| \mathbb{E} e\Big\{ \sum_{1 \le i < j \le k} f(G_i, G_j) + L(G_1, \dots, G_k) \Big\} \Big|,$$

and the supremum is taken over all linear statistics L, of the form $L(x_1, \dots, x_k) = l_1(x_1) + \dots + l_k(x_k)$ for some functions $l_1, \dots, l_k$, and $k = [(n-2)/20]$ is the integer part of $(n-2)/20$.

Below we write $A \ll B$ if $A \le cB$ and we shall use Taylor's expansion for the function $e\{x\}$ in the form

$$e\{x\} = \sum_{s=0}^{r-1} (ix)^s / s! + \mathbb{E}(1-v)^{r-1} e\{vx\} (ix)^r / (r-1)!\,, \quad x \in \mathbb{R} \tag{4.8}$$

where v is a random variable uniformly distributed on $[0,1]$ and independent of all other random variables.

LEMMA 4.1. *For* $t \in \mathbb{R}$

$$|\varrho_1(t)| \ll (|t| + t^2)(\alpha_n + \beta_n + \gamma_n)\kappa(t). \tag{4.9}$$

PROOF. Using the Taylor expansion (4.8) with $r = 2$

$$\begin{aligned} \varrho_1(t) \;\; &= \;\; it\,\mathbb{E} S e\{tU_n(h)\} + its\mathbb{E} e\{tU_n(h)\} \\ &\quad -t^2 \mathbb{E}(1-v)(S+s)^2 e\{tU_n(h) + vt(S+s)\}. \end{aligned} \tag{4.10}$$

Here by the i.i.d. assumption

$$\begin{aligned} &\mathbb{E} S e\{tU_n(h)\} \\ &= (n-1)\mathbb{E}(\mathbb{E}_{\bar X} g(X_1, \bar X) - \mathbb{E} g(X, \bar X)) e\{\frac{t}{n} \sum_{j=2}^n h(X_1, X_j) + \bar U(X_2, \dots, X_n)\}, \end{aligned}$$

where

$$\bar U(X_2, \dots, X_n) = \sum_{2 \le i < j \le n} f(X_i, X_j).$$

Again by (4.8) with $r = 1$ we obtain

$$\begin{aligned} \mathbb{E}\, S e\{tU_n(h)\} \;\; &= \;\; \frac{n-1}{n} it \mathbb{E}\Big(\mathbb{E}_{\bar X} g(X_1, \bar X) - \mathbb{E} g(X, \bar X) \Big) \\ &\quad \times \Big(\sum_{j=2}^n h(X_1, X_j) \Big) e\{\frac{vt}{n} \sum_{j=2}^n h(X_1, X_j) + \bar U(X_2, \dots, X_n)\}. \end{aligned} \tag{4.11}$$

Let

$$S_1 = \sum_{2 \le j \le n/2} h(X_1, X_j), \quad S_2 = \sum_{n/2 < j \le n} h(X_1, X_j).$$

Then from (4.11) we get

$$\Big|\mathbb{E}\,Se\{tU_n(h)\}\Big| \le |t|\kappa(t)\Big\{\mathbb{E}\Big|\mathbb{E}_{\bar X}g(X_1,\bar X) - \mathbb{E}g(X,\bar X)\Big|(|S_1|+|S_2|)\Big\}. \tag{4.12}$$

For fixed $X_1 = x$ we write taking into account the i.i.d. assumption, the degeneracy of h and Hölder's inequality

$$\mathbb{E}(|S_1|+|S_2|) = \mathbb{E}\Big|\sum_{2\le j\le n/2} h(x,X_j)\Big| + \mathbb{E}\Big|\sum_{n/2<j\le n} h(x,X_j)\Big| \le 2\sqrt{n}\Big(\mathbb{E}_X h^2(x,X)\Big)^{1/2}.$$

By virtue of definition of the function $g(x,y)$ in Lemma 2.1

$$\begin{aligned}\mathbb{E}_{\bar X}g(x,\bar X) &= -\mathbb{E}_{\bar X}\Phi(x,\bar X)I(\|\bar X\|>\sqrt{n})I(\|x\|\le\sqrt{n}) \\ &\quad -\mathbb{E}_{\bar X}\Phi(x,\bar X)I(|\Phi(x,\bar X)|>n)I(\|\bar X\|\le\sqrt{n})I(\|x\|\le\sqrt{n})\end{aligned} \tag{4.13}$$

and

$$\begin{aligned}|\mathbb{E}_{\bar X}g(x,\bar X)| &\le \frac{1}{\sqrt{n}}\mathbb{E}_{\bar X}|\Phi(x,\bar X)|\|\bar X\|I(\|\bar X\|>\sqrt{n})I(\|x\|\le\sqrt{n}) \\ &\quad +\frac{1}{n}\mathbb{E}_{\bar X}\Phi^2(x,\bar X)I(|\Phi(x,\bar X)|>n)I(\|x\|\le\sqrt{n}),\end{aligned} \tag{4.14}$$

$$|\mathbb{E}g(X,\bar X)| \le \frac{1}{n}(\alpha_n+\beta_n), \tag{4.15}$$

According to (2.14)

$$\begin{aligned}\mathbb{E}_X h^2(x,X) &\le 8\|x\|^2 I(\|x\|\le\sqrt{n}) + 4\mathbb{E}_X(\mathbb{E}_{\bar X}g(X,\bar X))^2 \\ &\le 8\Big(\|x\|^2 I(\|x\|\le\sqrt{n}) + \alpha_n+\beta_n\Big), \quad x\in\mathfrak{X}.\end{aligned} \tag{4.16}$$

Hence in (4.12)

$$\begin{aligned}&\mathbb{E}\Big|\mathbb{E}_{\bar X}g(X_1,\bar X) - \mathbb{E}g(X,\bar X)\Big|(|S_1|+|S_2|) \\ &\ll \mathbb{E}_X\Big\{(\|X\|^2+\alpha_n+\beta_n)^{1/2}\Big(\mathbb{E}_{\bar X}|\Phi(X,\bar X)|\,\|\bar X\|\,I(\|\bar X\|>\sqrt{n}) \\ &\quad +\frac{1}{\sqrt{n}}\mathbb{E}_{\bar X}\Phi^2(X,\bar X)I(|\Phi(X,\bar X)|>n)I(\|X\|\le\sqrt{n}) + \alpha_n+\beta_n\Big)\Big\}.\end{aligned} \tag{4.17}$$

For fixed $\bar X = x$ we have by Hölder's inequality

$$\mathbb{E}(\|X\|^2+\alpha_n+\beta_n)^{1/2}|\Phi(X,x)| \le \|x\|,$$

that is in (4.17)

$$\mathbb{E}_X\Big\{(\|X\|^2+\alpha_n+\beta_n)^{1/2}\mathbb{E}_{\bar X}|\Phi(X,\bar X)|\,\|\bar X\|\,I(\|\bar X\|>\sqrt{n}) \le 2\alpha_n.$$

Using the inequality $ab \le a^2 + b^2$ we see that in (4.17)

$$\begin{aligned}
&\frac{1}{\sqrt{n}}\mathbb{E}_X(\|X\|^2 + \alpha_n + \beta_n)^{1/2}\mathbb{E}_{\bar{X}}\Phi^2(X,\bar{X})I(|\Phi(X,\bar{X})| > n)I(\|X\| \le \sqrt{n}) \\
&= \mathbb{E}\Big\{\frac{1}{\sqrt{n}}(\|X\|^2 + \alpha_n + \beta_n)^{1/2}|\Phi(X,\bar{X})|I(\|X\| \le \sqrt{n})\Big\} \\
&\quad \times\Big\{|\Phi(X,\bar{X})|\, I(|\Phi(X,\bar{X})| > n)\Big\} \\
&\le \frac{1}{n}(\mathbb{E}\|X\|^4 I(\|X\| \le \sqrt{n}) + \alpha_n + \beta_n) + \beta_n.
\end{aligned}$$

Thus in (4.12)

$$\Big|\mathbb{E}\, Se\{tU_n(h)\}\Big| \ll |t|\kappa(t)(\alpha_n + \beta_n + \gamma_n). \tag{4.18}$$

The second expectation in (4.10) is bounded using (4.14), (4.15) and $\kappa(t)$.

Consider the third expectation in (4.10). Split $S + s = s_1 + s_2 + s_3$, where

$$s_1 = s + \sum_{1 \le j \le n/3} \varphi(X_j), \quad s_2 = \sum_{n/3 < j \le 2n/3} \varphi(X_j),$$

and $s_3 = S + s - s_1 - s_2$, and $\varphi(x) = (\mathbb{E}_{\bar{X}} g(x,\bar{X}) - \mathbb{E}g(X,\bar{X}))(n-1)/n$. Then in (4.10)

$$\Big|\mathbb{E}^v(S+s)^2 e\{tU_n(h) + vt(S+s)\}\Big| \le \sum_{1 \le j,k \le 3}\Big|\mathbb{E}^v s_j s_k e\{tU_n(h) + vt(S+s)\}\Big|,$$

where $\mathbb{E}^v$ denotes the conditional expectation given v. For example, consider $s_1 s_2$. Conditioning on s_1 and s_2, using the i.i.d. assumption and the definition of $\kappa(t)$ we obtain

$$\begin{aligned}
&\Big|\mathbb{E}^v s_1 s_2 e\{tU_n(h) + vt(S+s)\}\Big| \\
&\le \mathbb{E}^v|s_1|\,|s_2|\,\Big|\mathbb{E}_{s_3} e\{tU_n(h) + vts_3\}\Big| \\
&\le (\mathbb{E}s_1^2 + \mathbb{E}s_2^2)\kappa(t) \ll \Big(n\mathbb{E}_X(\mathbb{E}_{\bar{X}} g(X,\bar{X}))^2 + n^2(\mathbb{E}g(X,\bar{X}))^2\Big)\kappa(t).
\end{aligned} \tag{4.19}$$

Here by (4.13)

$$\begin{aligned}
&\mathbb{E}_X(\mathbb{E}_{\bar{X}} g(X,\bar{X}))^2 \\
&\le 2\mathbb{E}_X\Big(\mathbb{E}_{\bar{X}}\Phi(X,\bar{X})I(\|\bar{X}\| > \sqrt{n})\Big)^2 \\
&\quad + 2\mathbb{E}_X\Big(\mathbb{E}_{\bar{X}}\Phi(X,\bar{X})I(|\Phi(X,\bar{X})| > n)I(\|\bar{X}\| \le \sqrt{n})I(\|X\| \le \sqrt{n})\Big)^2 \\
&\le \frac{2}{n}\Big\{\mathbb{E}_X\Big(\mathbb{E}_{\bar{X}}|\Phi(X,\bar{X})|\,\|\bar{X}\|\, I(\|\bar{X}\| > \sqrt{n})\Big)^2 \\
&\quad + \frac{1}{n}\mathbb{E}_X(\mathbb{E}_{\bar{X}}\Phi^2(X,\bar{X}))^2 I(\|X\| \le \sqrt{n})\Big\} \ll \frac{1}{n}(\alpha_n + \gamma_n),
\end{aligned} \tag{4.20}$$

where

$$\begin{aligned}
&\mathbb{E}_X\Big(\mathbb{E}_{\bar X}|\Phi(X,\bar X)|\,\|\bar X\|\,I(\|\bar X\|>\sqrt{n})\Big)^2\\
&\quad= \mathbb{E}_{\bar X_1}\mathbb{E}_{\bar X_2}\|\bar X_1\|\,\|\bar X_2\|I(\|\bar X_1\|>\sqrt{n})I(\|\bar X_2\|>\sqrt{n})\mathbb{E}_X|\Phi(X,\bar X_1)|\,|\Phi(X,\bar X_2)|\\
&\quad\le \alpha_n^2\le\alpha_n.
\end{aligned}$$

Combining (4.10), (4.18) – (4.20) we obtain (4.9).

LEMMA 4.2. *For* $t\in\mathbb{R}$

$$|\varrho_2(t)|\ll(\alpha_n+\beta_n+\gamma_n)\Big((|t|+t^2)\kappa(t)+|t|n^{-3}\Big). \tag{4.21}$$

PROOF. According to (2.15) and (4.3) we can write the representation

$$U_\infty(\phi)=U(\theta;G_1,\dots,G_n)+S(\theta)+S_0,$$

where

$$S(\theta)=\sum_{j=1}^n\theta_j,\quad \theta_j=-\frac{n-1}{n}\mathbb{E}_X\theta(G_j,X)+\frac{1}{2n}\phi_\infty(G_j,G_j),$$

$$S_0=\frac{1}{2}(n-1)\mathbb{E}\theta(X,\bar X).$$

Applying (4.8) three times and taking account of this representation we obtain

$$\varrho_2(t)=-it\mathbb{E}S(\theta)e\{tU_\infty(h)\}-t^2\varrho_{21}(t)-it\varrho_{22}(t)-it\varrho_{23}(t), \tag{4.22}$$

where

$$\begin{aligned}
\varrho_{21}(t)&=\mathbb{E}(1-v)S^2(\theta)e\{tU_\infty(h)-vtS(\theta)\},\\
\varrho_{22}(t)&=\mathbb{E}U(\theta;G_1,\dots,G_n)e\{tU_\infty(h)-tS(\theta)-tS_0-vtU(\theta;G_1,\dots,G_n)\},\ \text{and}\\
\varrho_{23}(t)&=S_0\mathbb{E}e\{tU_\infty(h)-tS(\theta)-vtS_0\}.
\end{aligned}$$

First estimate $\varrho_{21}(t)$. Split the sum $S(\theta)$ into the three parts

$$S(\theta)=s_1(\theta)+s_2(\theta)+s_3(\theta),$$

where

$$s_1(\theta)=\sum_{1\le j\le n/3}\theta_j,\quad s_2(\theta)=\sum_{n/3<j\le 2n/3}\theta_j,\quad s_3(\theta)=S(\theta)-s_1(\theta)-s_2(\theta).$$

Just as in the proof of (4.19) we have

$$|\varrho_{21}(t)|\ll\sum_{j=1}^3\mathbb{E}s_j^2(\theta)\kappa(t), \tag{4.23}$$

where

$$\mathbb{E}s_j^2(\theta)\ll n\mathbb{E}(\mathbb{E}_X\theta(G,X))^2+\frac{1}{n}\mathbb{E}\phi_\infty^2(G,G).$$

Here according to (1.3), (2.5), (2.15) and (4.2)

$$\begin{aligned}
\mathbb{E}\phi_\infty^2(G,G) &= \mathbb{E}(\Phi_\infty(G,G)-h_\infty(G,G))^2 \\
&= \mathbb{E}\Big(\sum_{j=1}^\infty (q_j-\bar q_j)(\tau_j^2-1)\Big)^2 \\
&= 2\sum_{j=1}^\infty (q_j-\bar q_j)^2 \\
&= 2\mathbb{E}\Big(\sum_{j=1}^\infty (q_j-\bar q_j)\tau_j\bar\tau_j\Big)^2 \\
&= 2\mathbb{E}\Big(\Phi(G,\bar G)-h(G,\bar G)\Big)^2 = 2\mathbb{E}\phi^2(G,\bar G).
\end{aligned}$$

With the help of the second inequality (3.28) we obtain

$$\mathbb{E}\phi_\infty^2(G,G) \le \frac{12000}{n^2}(\alpha_n^2+\beta_n^2+\gamma_n^2). \tag{4.24}$$

Thus by (3.27) and (4.24) in (4.23)

$$|\varrho_{21}(t)| \ll (\alpha_n+\beta_n+\gamma_n)\kappa(t). \tag{4.25}$$

To estimate $\varrho_{22}(t)$ we write

$$\begin{aligned}
|\varrho_{22}(t)| &\le \mathbb{E}|U(\theta;G_1,\dots,G_n)| \le n\mathbb{E}|\theta(G,\bar G)| \qquad (4.26)\\
&\le 3n\,\mathbb{E}\|G\|\,I(\|G\|>\sqrt{n}) + n\,\mathbb{E}|\Phi(G,\bar G)|\,I(|\Phi(G,\bar G)|>\sqrt{n}) \\
&\le \frac{3}{n^4}\Big(\mathbb{E}\|G\|^{12}+\mathbb{E}\Phi^6(G,\bar G)\Big) \ll \frac{1}{n^4}.
\end{aligned}$$

Further we have taking into account (2.19)

$$|\varrho_{23}(t)| \le n\,|\mathbb{E}\theta(X,\bar X)|\kappa(t) \le (\alpha_n+\beta_n)\kappa(t). \tag{4.27}$$

Next we bound the expectation in (4.22). By the i.i.d. assumption and (4.8) we have

$$\begin{aligned}
&\mathbb{E}S(\theta)e\{tU_\infty(h)\} \qquad (4.28)\\
&= -(n-1)\mathbb{E}\,\mathbb{E}_X\theta(G,X)e\Big\{\frac{t}{n}\sum_{j=2}^n h(G,G_j)+\bar U(G_2,\dots,G_n)\Big\} \\
&\quad +\frac{1}{2}\mathbb{E}\,\phi_\infty(G,G)e\Big\{\frac{t}{n}\sum_{j=2}^n h(G,G_j)+\bar U(G_2,\dots,G_n)\Big\} \\
&= -(n-1)\mathbb{E}\theta(G,X)\,\mathbb{E}e\{\bar U(G_2,\dots,G_n)\} - \frac{n-1}{n}it\varrho_3(t)+\frac{it}{2n}\varrho_4(t),
\end{aligned}$$

where by (2.17)

$$\begin{aligned}
\mathbb{E}\theta(G,X) &= \mathbb{E}\Phi(G,X)I(\|G\|>\sqrt{n})I(\|X\|>\sqrt{n}) \\
&\quad -\mathbb{E}\Phi(G,X)I(|\Phi(G,X)|>n)I(\|G\|\le\sqrt{n})I(\|X\|\le\sqrt{n})
\end{aligned}$$

and

$$\begin{aligned}
\varrho_3(t) &= \mathbb{E}\,\mathbb{E}_X\theta(G,X)\sum_{j=2}^n h(G,G_j)e\Big\{\frac{vt}{n}\sum_{j=2}^n h(G,G_j)+\bar U(G_2,\dots,G_n)\Big\},\\
\varrho_4(t) &= \mathbb{E}\,\phi_\infty(G,G)\sum_{j=2}^n h(G,G_j)e\Big\{\frac{vt}{n}\sum_{j=2}^n h(G,G_j)+\bar U(G_2,\dots,G_n)\Big\}.
\end{aligned}$$

Just as in (4.11) – (4.12) since $\mathbb{E}\Phi^2(X,\bar X)=1$, $\mathbb{E}h^2(X,\bar X)\le 16$ and (3.27) we find

$$|\varrho_3(t)|\ll \sqrt{n}\Big(\mathbb{E}(\mathbb{E}_X\theta(G,X))^2\Big)^{1/2}\kappa(t)\ll(\alpha_n+\gamma_n)\kappa(t) \tag{4.29}$$

Similarly for $\varrho_4(t)$ with using (4.24) we find

$$|\varrho_4(t)|\ll \sqrt{n}\Big(\mathbb{E}\phi_\infty^2(G,G)\Big)^{1/2}\kappa(t)\ll(\alpha_n+\beta_n+\gamma_n)\kappa(t). \tag{4.30}$$

From (4.22) – (4.30) we obtain (4.21).

In what follows we shall estimate the difference of the characteristic functions

$$\mathbb{E}e\{tU_n(h)\}-\mathbb{E}\{tU_\infty(h)\},\quad t\in\mathbb{R}$$

when the kernel is degenerate. Just as in Bentkus and Götze (1999) we shall apply the Bergström type identity

$$\begin{aligned}
&\mathbb{E}e\{tU_n(h)\}-\mathbb{E}\{tU_\infty(h)\} \qquad\qquad (4.31)\\
&= \mathbb{E}e\{\mathfrak{R}_{[1,n]}tU_n(h)\}-\mathbb{E}e\{tU_\infty(h)\}+n\Big(\mathbb{E}e\{\mathfrak{R}_{[2,n]}tU_n(h)\}-\mathbb{E}e\{\mathfrak{R}_{[1,n]}tU_n(h)\}\Big)\\
&\quad+\sum_{k=2}^n(k-1)\Big(\mathbb{E}e\{\mathfrak{R}_{[3,k]}tU_n(h)\}-\mathbb{E}e\{\mathfrak{R}_{[2,k]}tU_n(h)\}\\
&\quad-\mathbb{E}e\{\mathfrak{R}_{[1,k]\setminus\{2\}}tU_n(h)\}+\mathbb{E}e\{\mathfrak{R}_{[1,k]}tU_n(h)\}\Big),
\end{aligned}$$

where the operator $\mathfrak{R}_A$ replaces the random variables X_j, $j\in A$, in $U(h;X_1,\dots,X_n)$ by G_j, $j\in A$. As a result of these replacements in (4.31)

$$\begin{aligned}
\mathfrak{R}_{[1,k]}tU_n(h) &= f(G_1,S_k)+f(G_2,S_k)+f(G_1,G_2)+V_k,\\
\mathfrak{R}_{[2,k]}tU_n(h) &= f(X_1,S_k)+f(G_2,S_k)+f(X_1,G_2)+V_k,\\
\mathfrak{R}_{[3,k]}tU_n(h) &= f(X_1,S_k)+f(X_2,S_k)+f(X_1,X_2)+V_k,\\
\mathfrak{R}_{[1,k]\setminus\{2\}}tU_n(h) &= f(G_1,S_k)+f(X_2,S_k)+f(G_1,X_2)+V_k,
\end{aligned}$$

where

$$S_k=\sum_{j=k+1}^n X_j+\sum_{j=3}^k G_j,\quad \text{and}$$

$$V_k=\sum_{i=3}^{k-1}\sum_{j=i+1}^k f(G_i,G_j)+\sum_{i=3}^k\sum_{j=k+1}^n f(G_i,X_j)+\sum_{i=k+1}^{n-1}\sum_{j=k+1}^n f(X_i,X_j).$$

Following Bentkus and Götze (1999) we shall now estimate the difference of the expectations on the right in (4.31). Under the condition $\mathbb{E}\Phi^2(X,\bar{X})=1$ the bounds depend on the moments of the truncated kernel h.

LEMMA 4.3. *For* $t \in \mathbb{R}$

$$\Big|\mathbb{E}e\{\mathfrak{R}_{[1,n]}tU_n(h)\} - \mathbb{E}e\{tU_\infty(h)\}\Big| \ll \frac{1}{n}(t^2+|t|^3)\kappa(t), \tag{4.32}$$

and

$$n\Big|\mathbb{E}e\{\mathfrak{R}_{[2,n]}tU_n(h)\} - \mathbb{E}e\{\mathfrak{R}_{[1,n]}tU_n(h)\}\Big| \ll \gamma_n t^4\kappa(t). \tag{4.33}$$

PROOF. According to (4.3), (4.4)

$$tU_\infty(h) = t\mathfrak{R}_{[1,n]}U_n(h) + \mathfrak{U}, \quad \text{where} \quad \mathfrak{U} = \sum_{j=1}^n f_\infty(G_j, G_j).$$

Using (4.8) with $r=2$ we write

$$\begin{aligned}&\mathbb{E}e\{\mathfrak{R}_{[1,n]}tU_n(h)\} - \mathbb{E}e\{tU_\infty(h)\}\\ &= i\mathbb{E}\,\mathfrak{U}e\{\mathfrak{R}_{[1,n]}tU_n(h)\} - \mathbb{E}\mathfrak{U}^2(1-v)e\{\mathfrak{R}_{[1,n]}tU_n(h)+v\mathfrak{U}\}.\end{aligned} \tag{4.34}$$

By the i.i.d. assumption

$$\mathbb{E}\,\mathfrak{U}e\{\mathfrak{R}_{[1,n]}tU_n(h)\} = n\mathbb{E}f_\infty(G,G)e\{S+\bar{U}(G_2,\dots,G_n)\},$$

where

$$S = f_\infty(G,G_2)+\dots+f_\infty(G,G_n).$$

Further taking into account $\mathbb{E}f_\infty(G,G)=0$ and $\mathbb{E}_G f_\infty(G,G)f_\infty(G,G_j)=0$, $2\le j\le n$, we get by (4.8)

$$\Big|\mathbb{E}\,\mathfrak{U}e\{\mathfrak{R}_{[1,n]}tU_n(h)\}\Big| \le n\mathbb{E}|\mathbb{E}^v f_\infty(G,G)S^2 e\{vS+\bar{U}(G_2,\dots,G_n)\}|.$$

To estimate the expectation on the right we repeat the reasoning used for the estimation of the expectation in (4.10). As a result we find

$$\Big|\mathbb{E}\,\mathfrak{U}e\{\mathfrak{R}_{[1,n]}tU_n(h)\}\Big| \ll \frac{1}{n}|t|^3\kappa(t). \tag{4.35}$$

Similarly spliting the sum $\mathfrak{U}$ in (4.34) into a sum of three approximately equal parts we obtain the bound

$$\Big|\mathbb{E}\,\mathfrak{U}^2(1-v)e\{\mathfrak{R}_{[1,n]}tU_n(h)+v\mathfrak{U}\}\Big| \ll \frac{1}{n}|t|^2\kappa(t). \tag{4.36}$$

The bound (4.32) follows from (4.34) – (4.36). The bound (4.33) follows from the inequality (8.61) in Bentkus and Götze (1999).

$$n\Big|\mathbb{E}e\{\mathfrak{R}_{[2,n]}tU_n(h)\} - \mathbb{E}e\{\mathfrak{R}_{[1,n]}tU_n(h)\}\Big| \ll \frac{1}{n}\mathbb{E}\Big(\mathbb{E}_X h^2(X,\bar{X})\Big)^2 t^4\kappa(t) \ll \gamma_n t^4\kappa(t),$$

since $\bar\beta_4 = 0$ and $\hat F_1(t) = 0$ due to the symmetry of G_j and the linearity of $h(x,y)$. This proves Lemma 4.3.

It remains to estimate the sum over $2 \le k \le n$ on the right in (4.31). Denote

$$\begin{aligned}
J_1(k) &= \mathbb{E}e\{f(X_1,S_k) + f(X_2,S_k) + f(X_1,X_2) + V_k\},\\
J_2(k) &= \mathbb{E}e\{f(G_1,S_k) + f(X_2,S_k) + f(G_1,X_2) + V_k\},\\
J_3(k) &= \mathbb{E}e\{f(X_1,S_k) + f(G_2,S_k) + f(X_1,G_2) + V_k\},\\
J_4(k) &= \mathbb{E}e\{f(G_1,S_k) + f(G_2,S_k) + f(G_1,G_2) + V_k\}.
\end{aligned}$$

Note that by the i.i.d. assumption $J_2(k) = J_3(k)$.

Lemma 4.4. *For $t \in \mathbb{R}$*

$$\sum_{k=2}^{n} k\Big|J_1(k) - J_2(k) - J_3(k) + J_4(k)\Big| \ll (|t| + t^6)\kappa(t)(L_n + \mu_n), \tag{4.37}$$

where $\mu_n = n^{-1}|\mathbb{E}\Phi^3(X,\bar X)I(|\Phi(X,\bar X)| \le n)|$.

Proof. To prove (4.37) it suffices to verify that

$$\Big|J_1(k) - J_2(k) - J_3(k) + J_4(k)\Big| \ll \frac{1}{n^2}(L_n + \mu_n)(|t| + t^6)\kappa(t). \tag{4.38}$$

To simplify the notation we shall provide the arguments for $k = 2$. The case $k > 2$ is similar since here in S_k and V_k the random vectors $X_3, \ldots, X_k$ are replaced by the Gaussian random vectors $G_3, \ldots, G_k$. For the case $k = 2$ introduce the notation

$$\hat x = \sum_{1 \le j \le n-2} f(x, X_j), \quad x \in \mathfrak{X}$$

and let $V = V_2$, $J_s = J_s(2)$, $1 \le s \le 4$. Further replacing $X_1, X_2, G_1, G_2, X_3, \ldots, X_n, G_3, \ldots, G_n$ by X, $\bar X$, G, $\bar G$, $X_1, \ldots, X_{n-2}$, $G_1, \ldots, G_{n-2}$ respectively, we can rewrite J_s, $1 \le s \le 4$, as in Bentkus and Götze (1999), in the form

$$\begin{aligned}
J_1 &= \mathbb{E}e\{V + \hat X + \hat{\bar X} + f(X,\bar X)\},\\
J_2 &= \mathbb{E}e\{V + \hat G + \hat{\bar X} + f(G,\bar X)\},\\
J_3 &= \mathbb{E}e\{V + \hat X + \hat{\bar G} + f(X,\bar G)\},\\
J_4 &= \mathbb{E}e\{V + \hat G + \hat{\bar G} + f(G,\bar G)\},
\end{aligned}$$

where $J_2 = J_3$. Expanding via Taylor series in powers of $f(\cdot,\cdot)$ up to terms of order $O(f^4)$ and conditioning, we get

$$\Big|J_s - J_{s0} - J_{s1} - \frac{1}{2}J_{s2} - \frac{1}{6}J_{s3}\Big| \ll \mathbb{E}f^4(X,\bar X)\kappa(t), \quad 1 \le s \le 4, \tag{4.39}$$

where we use the fact that

$$\mathbb{E}f^4(G,\bar G) \le 3\mathbb{E}f^4(X,G) \le 9\mathbb{E}f^4(X,\bar X).$$

Denoting $I_r = J_{1r} - J_{2r} - J_{3r} + J_{4r}$, $0 \le r \le 3$, and using (4.39), we obtain

$$|J_1 - J_2 - J_3 + J_4| \ll |I_0| + |I_1| + |I_2| + |I_3| + \mathbb{E}f^4(X,\bar X)\kappa(t).$$

Hence, instead of (4.38) it suffices to verify that

$$|I_r| \ll (|t|+t^6)\kappa(t)\frac{1}{n^2}(L_n+\mu_n), \quad \text{for} \quad 0 \le r \le 3. \tag{4.40}$$

We shall apply the following Rosenthal type inequality: for independent random variables $\xi_1,\dots,\xi_n$ in $\mathbb{R}$ with mean zero

$$\mathbb{E}|\sum_{j=1}^n \xi_j|^q \le c(q)\Big(\sum_{j=1}^n \mathbb{E}|\xi_j|^q + (\sum_{j=1}^n \mathbb{E}\xi_j^2)^{q/2}\Big), \quad 2 \le q < \infty. \tag{4.41}$$

Using the definition of I_r we write, for $0 \le r \le 3$,

$$|I_r| \ll \Big|\mathbb{E}e\{V\}\Big(\varphi_r(X,\bar X) - \varphi_r(G,\bar X) - \varphi_r(X,\bar G) + \varphi_r(G,\bar G)\Big)\Big|, \tag{4.42}$$

where

$$\varphi_r(x,y) = f^r(x,y)e\{\hat x + \hat y\}, \quad x,y \in \mathfrak{X}.$$

Verification of (4.40) for $r=0$. In (4.42) with the help of (4.8) and (1.15) we have

$$\begin{aligned}
&\mathbb{E}e\{V\}\Big(\varphi_0(X,\bar X) - \varphi_0(G,\bar X) - \varphi_0(X,\bar G) + \varphi_0(G,\bar G)\Big)\\
&\quad= \mathbb{E}e\{V\}(\mathbb{E}_X e\{\hat X\} - \mathbb{E}_G e\{\hat G\})^2\\
&\quad= -\frac{1}{4}\mathbb{E}e\{V\}[\mathbb{E}_v(1-v)^2(\mathbb{E}_X \hat X^3 e\{v\hat X\} - \mathbb{E}_G \hat G^3 e\{v\hat G\}]^2\\
&\quad= -\frac{1}{4}\mathbb{E}(1-v)^2(1-\bar v)^2[\mathbb{E}_1(\hat X\hat{\bar X})^3 e\{V + v\hat X + \bar v\hat{\bar X}\}\\
&\qquad+2\mathbb{E}_1(\hat X\hat G)^3 e\{V + v\hat X + \bar v\hat G\} - \mathbb{E}_1(\hat G\hat{\bar G})^3 e\{V + v\hat G + \bar v\hat{\bar G}\}],
\end{aligned}$$

where $\mathbb{E}_1$ denotes the expectation over $X_1,\dots,X_{n-2}$. From here and (4.42) we obtain for $r=0$

$$\begin{aligned}
|I_0| \ll\ & \mathbb{E}|\mathbb{E}_1(\hat X\hat{\bar X})^3 e\{V + v\hat X + \bar v\hat{\bar X}\}| \\
&+\mathbb{E}|\mathbb{E}_1(\hat X\hat G)^3 e\{V + v\hat X + \bar v\hat G\}| + \mathbb{E}|\mathbb{E}_1(\hat G\hat{\bar G})^3 e\{V + v\hat G + \bar v\hat{\bar G}\}|.
\end{aligned} \tag{4.43}$$

We will now show that in (4.43)

$$\mathbb{E}|\hat X\hat{\bar X}|^3 \ll t^6\frac{1}{n^2}L_n. \tag{4.44}$$

Similar inequalities hold for the expectations $\mathbb{E}|\hat X\hat G|^3$ and $\mathbb{E}|\hat G\hat{\bar G}|^3$. To prove (4.44) we write

$$\mathbb{E}|\hat X\hat{\bar X}|^3 = t^6\frac{1}{n^6}\mathbb{E}\Big|\Big(\sum_{i=1}^{n-2} h(X,X_i)\Big)\Big(\sum_{j=1}^{n-2} h(\bar X,X_j)\Big)\Big|^3. \tag{4.45}$$

Here

$$\Big(\sum_{i=1}^{n-2} h(X,X_i)\Big)\Big(\sum_{j=1}^{n-2} h(\bar X,X_j)\Big) = \sum_{j=1}^{n-2}\Big(h(X,X_j)h(\bar X,X_j) - \mathbb{E}_Z h(X,Z)h(\bar X,Z)\Big)$$

$$+\sum_{1\le i\ne j\le n-2} h(X,X_i)h(\bar X,X_j)+(n-2)\mathbb{E}_Z h(X,Z)h(\bar X,Z). \tag{4.46}$$

Due to (4.46) and since $|a+b+c|^3\le 9(|a|^3+|b|^3+|c|^3)$ we have

$$\begin{aligned}
&\mathbb{E}\Big|\Big(\sum_{i=1}^{n-2} h(X,X_i)\Big)\Big(\sum_{j=1}^{n-2} h(\bar X,X_j)\Big)\Big|^3 \qquad (4.47)\\
&\ll\ \mathbb{E}\Big|\sum_{j=1}^{n-2}\Big(h(X,X_j)h(\bar X,X_j)-\mathbb{E}_Z h(X,Z)h(\bar X,Z)\Big)\Big|^3\\
&\qquad +\mathbb{E}\Big|\sum_{1\le i\ne j\le n-2} h(X,X_i)h(\bar X,X_j)\Big|^3+n^3\mathbb{E}|\mathbb{E}_Z h(X,Z)h(\bar X,Z)|^3.
\end{aligned}$$

By (4.41) and conditioning

$$\begin{aligned}
&\mathbb{E}\Big|\sum_{j=1}^{n-2}\Big(h(X,X_j)h(\bar X,X_j)-\mathbb{E}_Z h(X,Z)h(\bar X,Z)\Big)\Big|^3\\
&\quad\ll\ n\mathbb{E}\Big|h(X,Z)h(\bar X,Z)\Big|^3+n^{3/2}\mathbb{E}(\mathbb{E}_Z h^2(X,Z)h^2(\bar X,Z))^{3/2}.
\end{aligned}$$

Here by (2.14) and (4.16)

$$\begin{aligned}
\mathbb{E}\Big|h(X,Z)h(\bar X,Z)\Big|^3 &\ll n^2\mathbb{E}_Z\Big(\mathbb{E}_X h^2(X,Z)\Big)^2\\
&\ll n^2\mathbb{E}\|X\|^4 I(\|X\|\le\sqrt n)+n^2(\alpha_n+\beta_n)^2
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}\Big(\mathbb{E}_Z h^2(X,Z)h^2(\bar X,Z)\Big)^{3/2} &\ll n\mathbb{E}\Big(\mathbb{E}_Z h^2(X,Z)h^2(\bar X,Z)\Big)\Big(\mathbb{E}_Z h^2(X,Z)\Big)^{1/2}\\
&\ll n^{3/2}\mathbb{E}_Z\Big(\mathbb{E}_X h^2(X,Z)\Big)^2\\
&\ll n^{3/2}\Big(\mathbb{E}\|X\|^4 I(\|X\|\le\sqrt n)+\alpha_n+\beta_n\Big).
\end{aligned}$$

Further in (4.47) by the decoupling inequality for U-statistics (see, for example, de la Peña (1992))

$$\mathbb{E}\Big|\sum_{1\le i\ne j\le n-2} h(X,X_i)h(\bar X,X_j)\Big|^3\ll\mathbb{E}\Big|\sum_{1\le i\ne j\le n-2} h(X,X_i)h(\bar X,X_j^1)\Big|^3, \tag{4.48}$$

where X_j^1 is a independent copy of X_j. Hence in (4.48)

$$\sum_{1\le i\ne j\le n-2} h(X,X_i)h(\bar X,X_j^1)=\sum_{i=1}^{n-2}\xi_i, \tag{4.49}$$

where

$$\xi_i = h(X, X_i) \sum_{\substack{j=1 \\ j \neq i}}^{n-2} h(\bar{X}, X_j).$$

Again by (4.41) and conditioning we obtain

$$\mathbb{E}\Big| \sum_{1 \le i \neq j \le n-2} h(X, X_i) h(\bar{X}, X_j^1)\Big|^3 \ll \sum_{i=1}^{n-2} \mathbb{E}|\xi_i|^3 + \mathbb{E}\Big(\sum_{i=1}^{n-2} \mathbb{E}_{X_i} \xi_i^2\Big)^{3/2}, \tag{4.50}$$

where

$$\begin{aligned} \sum_{i=1}^{n-2} \mathbb{E}|\xi_i|^3 &= (n-2)\mathbb{E}|h(X, Z)|^3 \mathbb{E}|\sum_{j=1}^{n-3} h(\bar{X}, X_j^1)|^3 \quad \text{and} \\ \mathbb{E}\Big(\sum_{i=1}^{n-2} \mathbb{E}_{X_i} \xi_i^2\Big)^{3/2} &= (n-2)^{3/2} \mathbb{E}(\mathbb{E}_Z h^2(X, Z))^{3/2} \mathbb{E}\| \sum_{j=1}^{n-3} h(\bar{X}, X_j)\|^3. \end{aligned}$$

Once more by (4.41) and conditioning

$$\mathbb{E}| \sum_{j=1}^{n-3} h(\bar{X}, X_j^1)|^3 \ll (n-3)\mathbb{E}|h(X, \bar{X})|^3 + (n-3)^{3/2} \mathbb{E}(\mathbb{E}_Z h^2(\bar{X}, Z))^{3/2}. \tag{4.51}$$

From (4.48) – (4.51) we obtain

$$\begin{aligned} &\mathbb{E}\Big| \sum_{1 \le i \neq j \le n-2} h(X, X_i) h(\bar{X}, X_j)\Big|^3 \qquad (4.52) \\ &\ll n^2 (\mathbb{E}|h(X, \bar{X})|^3)^2 + n^{5/2} \mathbb{E}|h(X, \bar{X})|^3 \mathbb{E}(\mathbb{E}_X h^2(X, \bar{X}))^{3/2} \\ &\quad + n^3 \Big(\mathbb{E}(\mathbb{E}_X h^2(X, \bar{X}))^{3/2}\Big)^2. \end{aligned}$$

In (4.52) due to (2.14), (4.16) and $ab \le a^2 + b^2$

$$\Big(\mathbb{E}|h(X, \bar{X})|^3\Big)^2 \le \mathbb{E}h^2(X, \bar{X}) \mathbb{E}h^4(X, \bar{X}) \ll n^2 \lambda_n, \tag{4.53}$$

and

$$\begin{aligned} \frac{\mathbb{E}|h(X, \bar{X})|^3}{n} \frac{\mathbb{E}(\mathbb{E}_X h^2(X, \bar{X}))^{3/2}}{\sqrt{n}} &\le \frac{\Big(\mathbb{E}|h(X, \bar{X})|^3\Big)^2}{n^2} + \frac{\Big(\mathbb{E}(\mathbb{E}_X h^2(X, \bar{X}))^{3/2}\Big)^2}{n} \\ &\ll \lambda_n + \gamma_n, \end{aligned}$$

since

$$\begin{aligned} \Big(\mathbb{E}(\mathbb{E}_X h^2(X, \bar{X}))^{3/2}\Big)^2 &\ll \Big(\mathbb{E}\|X\|^3 I(\|X\| \le \sqrt{n})\Big)^2 + \alpha_n + \beta_n \\ &\ll \mathbb{E}\|X\|^4 I(\|X\| \le \sqrt{n}) + \alpha_n + \beta_n. \end{aligned}$$

Further in (4.47) by Hölder inequality

$$\mathbb{E}|\mathbb{E}_Z h(X,Z)h(\bar{X},Z)|^3 \le \Big(\mathbb{E}(\mathbb{E}_Z h^2(X,Z))^{3/2}\Big)^2. \tag{4.54}$$

From (4.45) – (4.54) we obtain (4.44).

Finally, to get (4.40) for $r = 0$ from (4.43), it sufficies to repeat the arguments leading to (4.19): partition the sums of the type $\hat{X}$ in (4.43) into a sum of independent parts with approximately equal numbers of summands; condition on the part which is absent in the product to obtain the bound $\kappa(t)$; then estimate the expectations by the analogy with (4.44). More specifically, partition $[1, n-2] = N_1 \cup \cdots \cup N_l$ into an union of l disjoint subsets with approximately equal cardinalities. Further we write

$$\hat{X} = \sum_{i=1}^{l} a_i, \quad a_i = \sum_{j \in N_i} f(X, X_j), \quad \text{and}$$

$$\hat{\bar{X}} = \sum_{i=1}^{l} b_i, \quad b_i = \sum_{j \in N_i} f(\bar{X}, X_j).$$

It is well known that for $l \ge r$

$$(x_1 + \ldots + x_l)^r = \sum_{k=1}^{r} \sum_{1 \le i_1, < \cdots < i_k \le l} \sum_{\substack{\nu_1 + \cdots + \nu_k = r \\ \nu_j \ge 1}} \frac{r!}{\nu_1! \cdots \nu_k!} x_{i_1}^{\nu_1} \cdots x_{i_k}^{\nu_k}, \quad x_i \in \mathbb{R}.$$

Apply this formula for $x_i = a_i$ and $x_i = b_i$ provided $2r < l < 20$. By this formula the product

$$(a_1 + \cdots + a_l)^r (b_1 + \cdots + b_l)^r$$

is decomposed into a linear combination of terms of the form

$$a_{i_1}^{\nu_1} \cdots a_{i_k}^{\nu_k} \cdot b_{j_1}^{\mu_1} \cdots b_{j_s}^{\mu_s} \tag{4.55}$$

with

$$\nu_1 + \cdots + \nu_k = \mu_1 + \cdots + \mu_s = r, \quad \nu_i \ge 1, \quad \mu_i \ge 1,$$

and

$$1 \le i_1 < \cdots < i_k \le l, \quad 1 \le j_1 < \cdots < j_s \le l, \quad 1 \le k, s \le r.$$

Since $l > 2r$ then in the product (4.55) at least one pair (a_q, b_q) is absent. By conditioning on this pair (a_q, b_q) we obtain for $2r < l < 20$

$$\Big|\mathbb{E}_1(\hat{X}\hat{\bar{X}})^r e\{V + v\hat{X} + \bar{v}\hat{\bar{X}}\}\Big| \ll \Big(\sum_{i=1}^{l} \mathbb{E}_1|a_i b_i|^r + \sum_{1 \le i \ne j \le l} \mathbb{E}_1|a_i|^r \mathbb{E}_1|b_j|^r\Big)\kappa(t). \tag{4.56}$$

Applying (4.56) with $r = 3$ and $l = 7$

$$\begin{aligned} \mathbb{E}_1|a_i|^3 \le \mathbb{E}_1|\hat{X}|^3 &= \frac{t^3}{n^3}\mathbb{E}_1|\sum_{j=1}^{n-2} h(X, X_j)|^3 \\ &\ll \frac{t^3}{n^3}\Big(n\,\mathbb{E}_Z|h(X,Z)|^3 + n^{3/2}\Big(\mathbb{E}_Z h^2(X,Z)\Big)^{3/2}\Big). \end{aligned}$$

Similar estimate is valid for $\mathbb{E}_1|b_j|^3$. Hence for $1 \le i \ne j \le l$

$$\begin{aligned}\mathbb{E}|a_i|^3\mathbb{E}|b_j|^3 &\ll \frac{t^6}{n^2}\Big\{\frac{1}{n^2}\Big(\mathbb{E}|h(X,Z)|^3\Big)^2 + \frac{1}{n}\Big(\mathbb{E}\Big(\mathbb{E}_Z h^2(X,Z)\Big)^{3/2}\Big)^2\Big\} \\ &\ll \frac{t^6}{n^2}L_n.\end{aligned}$$

The expectation $\mathbb{E}|a_i b_i|^3$ is estimated with the help of (4.44). As a result we find from (4.56) that in (4.43)

$$\mathbb{E}\Big|\mathbb{E}_1(\hat{X}\hat{\bar{X}})^3 e\{V + v\hat{X} + \bar{v}\hat{\bar{X}}\}\Big| \ll t^6\kappa(t)\frac{1}{n^2}L_n.$$

This estimate is also valid for other similar expectations in (4.43). That is, we obtain (4.40) for $r = 0$.

Verification of (4.40) for $r = 1$. Denoting

$$\alpha(x) = e\{x\} - ix - 1 = -x^2\mathbb{E}(1-v)e\{vx\}, \quad x \in \mathbb{R}$$

and using (1.15) we write in (4.42)

$$\begin{aligned}
&\mathbb{E}e\{V\}\Big(\varphi_1(X,\bar{X}) - \varphi_1(G,\bar{X}) - \varphi_1(X,\bar{G}) + \varphi_1(G,\bar{G})\Big) \qquad (4.57)\\
&= \mathbb{E}e\{V\}\Big[f(X,\bar{X})(e\{\hat{X}\}-1)(e\{\hat{\bar{X}}\}-1) - f(G,\bar{X})(e\{\hat{G}\}-1)(e\{\hat{\bar{X}}\}-1) \\
&\quad -f(X,\bar{G})(e\{\hat{X}\}-1)(e\{\hat{\bar{G}}\}-1) + f(G,\bar{G})(e\{\hat{G}\}-1)(e\{\hat{\bar{G}}\}-1)\Big] \\
&= \mathbb{E}e\{V\}\Big[f(X,\bar{X})(\alpha(\hat{X})+i\hat{X})(\alpha(\hat{\bar{X}})+i\hat{\bar{X}}) - f(G,\bar{X})(\alpha(\hat{G})+i\hat{G})(\alpha(\hat{\bar{X}})+i\hat{\bar{X}}) \\
&\quad -f(X,\bar{G})(\alpha(\hat{X})+i\hat{X})(\alpha(\hat{\bar{G}})+i\hat{\bar{G}}) + f(G,\bar{G})(\alpha(\hat{G})+i\hat{G})(\alpha(\hat{\bar{G}})+i\hat{\bar{G}})\Big] \\
&= \mathbb{E}e\{V\}\Big[f(X,\bar{X})\Big(\alpha(\hat{X})\alpha(\hat{\bar{X}}) + i\alpha(\hat{\bar{X}})\hat{X} + i\alpha(\hat{X})\hat{\bar{X}} - \hat{X}\hat{\bar{X}}\Big) \\
&\quad -f(G,\bar{X})\Big(\alpha(\hat{G})\alpha(\hat{\bar{X}}) + i\alpha(\hat{\bar{X}})\hat{G} + i\alpha(\hat{G})\hat{\bar{X}} - \hat{G}\hat{\bar{X}}\Big) \\
&\quad -f(X,\bar{G})\Big(\alpha(\hat{X})\alpha(\hat{\bar{G}}) + i\alpha(\hat{\bar{G}})\hat{X} + i\alpha(\hat{X})\hat{\bar{G}} - \hat{X}\hat{\bar{G}}\Big) \\
&\quad +f(G,\bar{G})\Big(\alpha(\hat{G})\alpha(\hat{\bar{G}}) + i\alpha(\hat{\bar{G}})\hat{G} + i\alpha(\hat{G})\hat{\bar{G}} - \hat{G}\hat{\bar{G}}\Big)\Big] \\
&= \mathbb{E}e\{V\}\Big[f(X,\bar{X})\alpha(\hat{X})\alpha(\hat{\bar{X}}) - f(G,\bar{X})\alpha(\hat{G})\alpha(\hat{\bar{X}}) \\
&\quad -f(X,\bar{G})\alpha(\hat{X})\alpha(\hat{\bar{G}}) + f(G,\bar{G})\alpha(\hat{G})\alpha(\hat{\bar{G}})\Big] \\
&= \mathbb{E}(1-v)(1-\bar{v})\Big[\mathbb{E}_1 f(X,\bar{X})(\hat{X}\hat{\bar{X}})^2 e\{V + v\hat{X} + \bar{v}\hat{\bar{X}}\} \\
&\quad -2\mathbb{E}_1 f(G,\bar{X})(\hat{G}\hat{\bar{X}})^2 e\{V + v\hat{G} + \bar{v}\hat{\bar{X}}\} + \mathbb{E}_1 f(G,\bar{G})(\hat{G}\hat{\bar{G}})^2 e\{V + v\hat{G} + \bar{v}\hat{\bar{G}}\}\Big].
\end{aligned}$$

From (4.42) and (4.57) we get for $r = 1$

$$|I_1| \ll \mathbb{E}|\mathbb{E}_1 f(X,\bar{X})(\hat{X}\hat{\bar{X}})^2 e\{V + v\hat{X} + \bar{v}\hat{\bar{X}}\}| \tag{4.58}$$
$$+\mathbb{E}|\mathbb{E}_1 f(G,\bar{X})(\hat{G}\hat{\bar{X}})^2 e\{V + v\hat{G} + \bar{v}\hat{\bar{X}}\}|$$
$$+\mathbb{E}|\mathbb{E}_1 f(G,\bar{G})(\hat{G}\hat{\bar{G}})^2 e\{V + v\hat{G} + \bar{v}\hat{\bar{G}}\}|.$$

We shall prove that in (4.58)

$$\mathbb{E}|f(X,\bar{X})|(\hat{X}\hat{\bar{X}})^2 \ll |t|^5 \frac{1}{n^2} L_n. \tag{4.59}$$

This inequality is also valid for the expectations

$$\mathbb{E}|f(G,\bar{X})|(\hat{G}\hat{\bar{X}})^2 \quad \text{and} \quad \mathbb{E}|f(G,\bar{G})|(\hat{G}\hat{\bar{G}})^2.$$

Due to (4.46) and $(a+b+c)^2 \le 3(a^2+b^2+c^2)$ we write

$$\begin{aligned}
&\mathbb{E}|f(X,\bar{X})|(\hat{X}\hat{\bar{X}})^2 \qquad\qquad (4.60)\\
&\ll \frac{|t|^5}{n^5}\Big[\mathbb{E}|h(X,\bar{X})|\mathbb{E}_1\Big(\sum_{j=1}^{n-2}(h(X,X_j)h(\bar{X},X_j) - \mathbb{E}_Z h(X,Z)h(\bar{X},Z))\Big)^2\\
&\quad +\mathbb{E}|h(X,\bar{X})|\mathbb{E}_1\Big(\sum_{1\le i\ne j\le n-2}(h(X,X_i)h(\bar{X},X_j)\Big)^2\\
&\quad +n^2\mathbb{E}|h(X,\bar{X})|\Big(\mathbb{E}_Z h(X,Z)h(\bar{X},Z)\Big)^2\Big]\\
&\ll \frac{|t|^5}{n^5}[n\mathbb{E}|h(X,\bar{X})|\mathbb{E}_Z h^2(X,Z)h^2(\bar{X},Z) + n^2\mathbb{E}|h(X,\bar{X})|\mathbb{E}_Z h^2(X,Z)\mathbb{E}_{\bar{Z}} h^2(\bar{X},\bar{Z})].
\end{aligned}$$

In (4.60), by (2.14) and (4.16),

$$\begin{aligned}
\mathbb{E}|h(X,\bar{X})|\mathbb{E}_Z h^2(X,Z)h^2(\bar{X},Z) &\le 4n\mathbb{E}_Z(\mathbb{E}_X h^2(X,Z))^2\\
&\ll n\mathbb{E}\|X\|^4 I(\|X\| \le \sqrt{n}) + n(\alpha_n+\beta_n)^2
\end{aligned}$$

and by Hölder's inequality

$$\begin{aligned}
&\mathbb{E}|h(X,\bar{X})|\mathbb{E}_Z h^2(X,Z)\mathbb{E}_{\bar{Z}} h^2(\bar{X},\bar{Z}) \qquad\qquad (4.61)\\
&\ll \mathbb{E}|h(X,\bar{X})|\Big(\|X\|^2 I(\|X\| \le \sqrt{n}) + \alpha_n + \beta_n\Big)\Big(\|\bar{X}\|^2 I(\|\bar{X}\| \le \sqrt{n}) + \alpha_n + \beta_n\Big)\\
&\ll \mathbb{E}\|X\|^4 I(\|X\| \le \sqrt{n}) + (\alpha_n+\beta_n)^2.
\end{aligned}$$

From (4.60) – (4.61) we get (4.59). Thus (4.40) for $r = 1$ follows from (4.56), (4.58), (4.59) and the partitioning argument for the sums $\hat{X}$ in (4.58).

Verification of (4.40) for $r = 2$. In (4.42) by Taylor's formula (4.8)

$$\begin{aligned}
&\mathbb{E}\exp\{V\}\Big(\varphi_2(X,\bar X) - 2\varphi_2(X,G) + \varphi_2(G,\bar G)\Big) \\
&\quad = \; -\mathbb{E}(1-v)\Big[f^2(X,\bar X)(\hat X + \hat{\bar X})^2 e\{V + v(\hat X + \hat{\bar X})\} \\
&\qquad -2f^2(X,G)(\hat X + \hat G)^2 e\{V + v(\hat X + \hat G)\} + f^2(G,\bar G)(\hat G + \hat{\bar G})^2 e\{V + v(\hat G + \hat{\bar G})\}\Big].
\end{aligned}$$

Hence in (4.42)

$$\begin{aligned}
|I_2| \;\ll\; & \mathbb{E}f^2(X,\bar X)\mathbb{E}_1|(\hat X + \hat{\bar X})^2 e\{V + v(\hat X + \hat{\bar X})\}| \\
&+\mathbb{E}f^2(X,G)\mathbb{E}_1|(\hat X + \hat G)^2 e\{V + v(\hat X + \hat G)\}| \\
&+\mathbb{E}f^2(G,\bar G)\mathbb{E}_1|(\hat G + \hat{\bar G})^2 e\{V + v(\hat G + \hat{\bar G})\}|.
\end{aligned} \tag{4.62}$$

We will prove that

$$\mathbb{E}f^2(X,\bar X)\hat{\bar X}^2 \ll t^4 \frac{1}{n^2} L_n.$$

Write, taking into account (4.16)

$$\begin{aligned}
\mathbb{E}f^2(X,\bar X)\hat{\bar X}^2 \;&=\; \mathbb{E}f^2(X,\bar X)\mathbb{E}_1\hat{\bar X}^2 \\
&=\; (n-2)\frac{t^4}{n^4}\mathbb{E}\Big(\mathbb{E}_{\bar X} h^2(X,\bar X)\mathbb{E}_Z h^2(X,Z)\Big) \\
&\ll\; \frac{t^4}{n^3}\mathbb{E}\Big(\|X\|^2\, I(\|X\| \le \sqrt{n}) + \alpha_n + \beta_n\Big)^2 \ll t^4 \frac{1}{n^2} L_n.
\end{aligned}$$

To get (4.40) from (4.62) it sufficies to partition the sums of type $\hat X + \hat{\bar X}$ into a sum of three parts with approximately equal numbers of summands, then conditioning to use the bound $\kappa(t)$ and approximations of the type (4.59) and (4.61) for the expectations.

Verification of (4.40) for $r = 3$. In (4.42) by (4.8), the symmetry of G and the i.i.d. assumption

$$\begin{aligned}
&\mathbb{E}\{V\}\Big(\varphi_3(X,\bar X) - 2\varphi_3(X,G) + \varphi_3(G,\bar G)\Big) \\
&\quad = \; \mathbb{E}f^3(X,\bar X)\mathbb{E}e\{V\} + 2i\mathbb{E}f^3(X,\bar X)\hat X\, e\{V + v(\hat X + \hat{\bar X})\} \\
&\qquad -2i\mathbb{E}f^3(X,G)(\hat X + \hat G)\, e\{V + v(\hat X + \hat G)\} \\
&\qquad +2i\mathbb{E}f^3(G,\bar G)\hat G\, e\{V + v(\hat G + \hat{\bar G})\}.
\end{aligned}$$

Therefore in (4.42)

$$\begin{aligned}
|I_3| \;\ll\; & |\mathbb{E}f^3(X,\bar X)|\kappa(t) + \mathbb{E}|f(X,\bar X)|^3|\mathbb{E}_1\hat X e\{V + v(\hat X + \hat{\bar X})\}| \\
&+|f(X,G)|^3|\mathbb{E}_1(\hat X + \hat G)e\{V + v(\hat X + \hat G)\}| \\
&+\mathbb{E}|f(G,\bar G)|^3|\mathbb{E}_1\hat G e\{V + v(\hat G + \hat{\bar G})\}|.
\end{aligned} \tag{4.63}$$

Here by (2.14)

$$|\mathbb{E}f^3(X,\bar{X})| \ll |t|^3 \frac{1}{n^2}(L_n+\mu_n).$$

Partitioning the sum $\hat{X}$ into a sum of two parts with approximately equal numbers of summands we find

$$\begin{aligned}
&\mathbb{E}|f(X,\bar{X})|^3|\mathbb{E}_1\hat{X}e\{V+v(\hat{X}+\hat{\hat{X}})\}| \qquad (4.64)\\
&\ll \sqrt{n}\kappa(t)\mathbb{E}|f(X,\bar{X})|^3\Big(\mathbb{E}_Z f^2(X,Z)\Big)^{1/2}\\
&= t^4\kappa(t)\frac{1}{n^2}n^{-3/2}\mathbb{E}|h(X,\bar{X})|^3\Big(\mathbb{E}_Z h^2(X,Z)\Big)^{1/2}.
\end{aligned}$$

In (4.64) by the inequality $ab \le a^2+b^2$, (2.16) and (4.16)

$$\begin{aligned}
&n^{-3/2}\mathbb{E}|h(X,\bar{X})|^3\Big(\mathbb{E}_Z h^2(X,Z)\Big)^{1/2} \qquad (4.65)\\
&= \mathbb{E}\Big\{(\frac{1}{n}\mathbb{E}_{\bar{X}}|h(X,\bar{X})|^3)(\frac{1}{n}\mathbb{E}_Z h^2(X,Z))^{1/2}\Big\}\\
&\ll n^{-2}\mathbb{E}h^4(X,\bar{X})+n^{-1}\mathbb{E}h^2(X,\bar{X}) \ll L_n.
\end{aligned}$$

The other expectations in (4.63) are estimated similarly. From (4.63) – (4.65) we have (4.40) for $r=3$. This completes the proof of (4.37).

Lemma 4.5. *For* $t \in \mathbb{R}$

$$|\mathbb{E}e\{tU_n(g)\} - \mathbb{E}e\{tU_\infty(\Phi)\}| \ll (|t|+t^6)\kappa(t)(L_n+\mu_n) + |t|n^{-3}L_n.$$

Proof. The result follows from (4.5) – (4.6), (4.9), (4.21), (4.31) – (4.33) and (4.37).

5 Proof of Theorem 1.1.

By Lemma 2.1.

$$\Delta_n \le \sup_x |\mathbb{P}\{U_n(\epsilon^2 g) \le x\} - \mathbb{P}\{U_\infty(\epsilon^2\Phi) \le x\}| + \alpha_n + \beta_n \qquad (5.1)$$

with $\epsilon = (16s)^{-1}$.

Choose $s=13$. Due to Lemma 3.1 and the representations (4.1) and (4.3) we can apply Lemma 4.2 by Bentkus and Götze (1999) with $m = m(13, q_{13})$. This gives

$$\sup_x |\mathbb{P}\{U_n(\epsilon^2 g) \le x\} - \mathbb{P}\{U_\infty(\epsilon^2\Phi) \le x\}| \ll J + R, \qquad (5.2)$$

where

$$J = \int_{-nt_*}^{nt_*} \Big|\mathbb{E}e\{t\epsilon^2 U_n(g)\} - \mathbb{E}e\{t\epsilon^2 U_\infty(\Phi)\}\Big| \, |t|^{-1}\, dt$$

with $t_* = m^{-1}(pn/m)^{-1/2}$ and $R \ll (pn)^{-1}m$. By our Lemma 4.5 and the estimate (4.20) in Bentkus and Götze (1999) we get

$$J \ll p^{-6}(L_n + \mu_n). \tag{5.3}$$

Collecting the bounds (5.1) – (5.3) we obtain (1.10). This proves Theorem 1.1.

REFERENCES

1. Bentkus, V. and Götze, F. (1999). Optimal bounds in non-Gaussian limit theorems for U-statistics. *Ann. Probab.* **27**, 454-521.

2. Borovskikh, Yu.V. (1996). *U-statistics in Banach Spaces,* VSP, Utrecht, The Netherlands.

3. Dunford, N. and Schwartz, J. (1961). *Linear Operators. Part II: Spectral Theory*, Interscience Publishers, New York.

4. de la Peña, V.H. (1992). Decoupling and Khintchine's inequalities for U-statistics. *Ann. Probab.* **20**, 1877-1892.