EDGEWORTH EXPANSIONS AND BOOTSTRAP FOR DEGENERATE VON MISES STATISTICS

F. Götze[1] R.Zitikis[2]

Fakultät für Mathematik
Universität Bielefeld

Institute of Mathematics and
Informatics, Vilnius

December 1993

Abstract. We prove Edgeworth expansions for degenerate von Mises statistics like the Beran, Watson, and Cramér-von Mises goodness-of-fit statistics. Furthermore, we show that the bootstrap approximation works up to an error of order $O(N^{-1/2})$ and that Bootstrap based confidence regions attain a prescribed confidence level up to the order $O(n^{-1})$.

1. Introduction, main results, and some examples

Let $\{\Omega,\ \mathcal{A},\ \mathbb{P}\}$ be a probability space, and let $\mathbf{X} : \{\Omega,\ \mathcal{A}\} \to \{\Lambda, \mathcal{F}\}$ be a random variable. The distribution function of $\mathbf{X}$ will be denoted by $\mathbb{F}$. Furthermore, we shall denote by $\mathbf{X}_1, \mathbf{X}_2$, ... independent copies of $\mathbf{X}$.

In this paper we consider various approximations for the (degenerate) von Mises statistic

$$V_N \triangleq \frac{1}{N}\sum_{j=1}^{N}\sum_{k=1}^{N} H(\mathbf{X}_j, \mathbf{X}_k),$$

[1]Research supported by the SFB 343 at Bielefeld.

[2]Research supported by the SFB 343 and Graduiertenkolleg at the University of Bielefeld.

1991 *Mathematics Subject Classification.* Primary 62E20; secondary 60F05.

Key words and phrases. Bootstrap, coverage probabilities, confidence intervals, Edgeworth expansions, von Mises statistics, U-statistics.

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

where $H : \Lambda \times \Lambda \to \mathbb{R}$ is a symmetric measurable function which is assumed to be degenerate with respect to $\mathbb{P}$, that is $\mathbb{E}\big(H(\mathbf{X}_1, \mathbf{X}_2)|\mathbf{X}_2\big) = 0\,\mathbb{P}$-a.s. Furthermore, we assume $\mathbb{E}\big|H(\mathbf{X}, \mathbf{X})\big| + \mathbb{E}H^2(\mathbf{X}_1, \mathbf{X}_2) < +\infty$. We use the notation $P_N(x) \triangleq \mathbb{P}\{V_N \le x\}$.

Our goal is to investigate the following three closely related problems for the statistic V_N:

1) A bootstrap approximation;

2) The bootstrap based coverage probabilities;

3) Edgeworth expansions.

The bootstrap version of V_N we use in this paper is defined (see Bretagnolle (1983)) as follows:

$$V_n^* \triangleq \frac{1}{n} \sum_{j=1}^{n} \sum_{k=1}^{n} H^*(\mathbf{X}_j^*, \mathbf{X}_k^*),$$

$$H^*(\mathbf{x}, \mathbf{y}) \triangleq H(\mathbf{x}, \mathbf{y}) - \mathbb{E}^* H(\mathbf{x}, \mathbf{X}_2^*) - \mathbb{E}^* H(\mathbf{X}_1^*, \mathbf{y}) - \mathbb{E}^* H(\mathbf{X}_1^*, \mathbf{X}_2^*),$$

where $\mathbb{E}^*$ denotes the expectation with respect to the empirical distribution function $\mathbb{F}^*$ given the sample $\mathbf{X}_1, ..., \mathbf{X}_N$.

We shall use the notation $P_n^*(x) \triangleq \mathbb{P}^*\{V_n^* \le x\}$.

The limit distribution of V_N, say P_∞, was described by von Mises (1947), and it is the distribution of the random variable

$$V_\infty \triangleq \mathbb{E}H(\mathbf{X}, \mathbf{X}) + \sum_{k=1}^{\infty} \lambda_k (G_k^2 - 1),$$

where $\lambda_1, \lambda_2, \ldots$, $|\lambda_1| \ge |\lambda_2| \ge \ldots$ are eigenvalues of the self–adjoint Hilbert–Schmidt operator $S : L_2(\Lambda, \mathbb{F}) \to L_2(\Lambda, \mathbb{F})$ defined by

$$Sf \triangleq \mathbb{E}H(\cdot, \mathbf{X}) f(\mathbf{X}) = \int H(\cdot, \mathbf{y}) f(\mathbf{y}) \mathbb{F}(d\mathbf{y}).$$

Here $G_1, G_2, \ldots$ are independent copies of a Gaussian random variable G with mean 0 and variance 1. See Gregory (1977), Rubin and Vitale (1980), Serfling (1980), Dynkin and Mandelbaum (1983) for more details on this representation.

In the same way we may describe the distribution function P_∞^* of the random variable V_∞^* ($\triangleq$ the weak limit, as $n \to +\infty$, of the statistic V_n^* when the random variables $\mathbf{X}_1, ..., \mathbf{X}_N$ are fixed).

If it is not stated otherwise, we shall always assume that the following holds:

The operator S has an infinite number of non-zero eigenvalues λ_j and the random variables $H(\mathbf{X}, \mathbf{X})$ and $H(\mathbf{X}_1, \mathbf{X}_2)$ have moments of all orders.

In the sequel we use $\|\cdot\|_\infty$ to denote the sup-norm.

Theorem 1.1. *We have*

$$\big\|P_N - P_N^*\big\|_\infty = O_{\mathbb{P}}\big(N^{-1/2}\big). \tag{1.1}$$

For the general theory of the bootstrap we refer to Efron (1979), Bickel and Freedman (1981), Efron (1982), Athreya (1983), Hinkley (1988), and the monograph of Hall (1992).

Previous results on the CLT for degenerate von Mises statistics (which will be discussed below) show that the rate of convergence in the CLT for these statistics might be of order $O(N^{-1+\varepsilon})$ for any $\varepsilon > 0$, and even of order $O(N^{-1})$ provided some smoothness conditions on the kernel H and the distribution function $\mathbb{F}$ are imposed. These facts are the main ingredients in the proof of the following theorem on the bootstrap based confidence regions; see Hall (1986), Beran (1987), Hall (1992) on this topic. In the theorem below we use $P_N^{*-1}(\alpha)$ to denote the α-th quantile of the function P_N^*.

Theorem 1.2. *For each $\alpha \in (0,1)$ and $\varepsilon > 0$*

$$\mathbb{P}\big\{V_N \geq P_N^{*-1}(\alpha)\big\} = 1 - \alpha + O\big(N^{-1+\varepsilon}\big), \quad N \to +\infty. \tag{1.2}$$

At this point a natural question arises: Does Theorem 1.2 hold with $\varepsilon = 0$? An inspectation of the proof of Theorem 1.2 shows that (1.2) holds with $\varepsilon = 0$ if the distribution functions P_N and P_N^* have Edgeworth expansions of order $O(N^{-1})$ and $O_{\mathbb{P}}(N^{-1-\delta})$ for some $\delta > 0$, respectively. In order to prove such a result we need to require more than moment conditions and assumptions on the limit distribution. As an example reflecting this fact we formulate the following result:

There exist infinitely differentiable and rapidly decreasing functions $a_1, \dots, a_k$ such that the asymptotic expansion

$$P_N(x) = P_\infty(x) + a_1(x)N^{-1} + \dots + a_k(x)N^{-k} + R(x) \tag{1.3}$$

holds with a remainder term R satisfying, for any $\varepsilon > 0$ and some constant $c \geq 0$,

$$\sup_{x\in\mathbb{R}} \big|x\big|^m \big|R(x)\big| \leq \frac{c}{N^{k+1}} + c \sup_{0\leq\mu\leq m} \int_{N^{1-\varepsilon}\leq|t|\leq N^{k+1}} \bigg| \Big(\frac{d}{dt}\Big)^\mu \mathbf{E}\exp\big\{itV_N\big\}\bigg| dt \tag{1.4}$$

for all $N \in \mathbb{N}$.

This result appeared in slightly different forms in Götze (1979, 1984, 1985); see also Bentkus, Götze and Zitikis (1993) and the survey paper by Bentkus Götze, Paulauskas and Račkauskas (1990).

Thus our further task is to investigate which assumptions on H and the distribution of the $\mathbf{X}$ yield bounds like

$$\sup_{|t|\geq \varrho(N)} \left|(d/dt)^m \mathbb{E}\exp\{itV_N\}\right| = O\big(N^{-L}\big), \quad N \to +\infty$$

with some function $\varrho : \mathbb{N} \to [0, +\infty)$ such that $\varrho(N) \to +\infty$ for $N \to +\infty$, and for every $L \geq 0$.

We assume that $\Lambda = \mathbb{R}^d$, and that the random variable $\mathbf{X}$ has a non-zero absolutely continuous component. This means that for some $\alpha \in (0, 1]$ the distribution function $\mathbb{F}$ of the $\mathbf{X}$ allows the representation

$$\mathbb{F} = \alpha\mathbb{F}_0 + (1-\alpha)\mathbb{F}_1, \tag{1.5}$$

where $\mathbb{F}_0$ and $\mathbb{F}_1$ are distribution functions, and $\mathbb{F}_0$ is absolutely continuous. We use $\mathbf{Y}, \mathbf{Y}_1, ..., \mathbf{Y}_N$ to denote i.i.d. $\mathbb{R}^d$-valued random variables having the distribution function $\mathbb{F}_0$.

We use $\mathcal{D}f(\mathbf{x})$ to denote the gradient of a function f at a point $\mathbf{x} \in \mathbb{R}^d$. Also, if ξ is a random variable, let $\tilde{\xi}$ denote a symmetrization of ξ.

Theorem 1.3. *Let $m \in \mathbb{N} \cup \{0\}$ and*

$$\mathbb{E}\big|H(\mathbf{X}_1, \mathbf{X}_1)\big|^m + \mathbb{E}\big|H(\mathbf{X}_1, \mathbf{X}_2)\big|^m < +\infty. \tag{1.6}$$

Let $\mathbf{B}$ and $\mathbf{C}$ be some cubes in the space $\mathbb{R}^d$ such that

$$\mathbb{P}\big\{\mathbf{Y} \in \mathbf{B}\big\} > 0, \quad \mathbb{P}\big\{\mathbf{Y} \in \mathbf{C}\big\} > 0. \tag{1.7}$$

Let $\mathbf{Z}_1, \mathbf{Z}_2, ...$ be independent $\mathbb{R}^d$-valued random variables having the distribution function $\mathbb{F}_0\{\cdot|\mathbf{B}\}$. Denote$S_N(\mathbf{x}) \triangleq N^{-1/2}\big\{\widetilde{H}(\mathbf{Z}_1, \mathbf{x}) + ... + \widetilde{H}(\mathbf{Z}_N, \mathbf{x})\big\}$. Furthermore, assume that for every $\varepsilon > 0$ there exist numbers $\lambda \in \mathbb{R}$, $\kappa > 0$, and $\nu > 0$ such that λ, κ, $\nu \leq \varepsilon$ and such that for every $D \geq 0$ the following three conditions hold:

$$\sup_{N\in\mathbb{N}} N^D \mathbb{P}\Big\{\sup_{\mathbf{x}\in\mathbf{C}} \big\|\mathcal{D}S_N(\mathbf{x})\big\| \geq N^\lambda\Big\} < +\infty; \tag{1.8}$$

$$\sup_{N\in\mathbb{N}} N^D \mathbb{P}\Big\{\sup_{\mathbf{x}\in\mathbf{C}} \big\|\mathcal{D}S_N(\mathbf{x})\big\| \leq 1/N^\kappa\Big\} < +\infty; \tag{1.9}$$

$$\sup_{N\in\mathbb{N}} N^D \mathbb{P}\Big\{\sup_j \sup_{\mathbf{x},\mathbf{y}\in\mathbf{C}_j} \big\|\mathcal{D}S_N(\mathbf{x}) - \mathcal{D}S_N(\mathbf{y})\big\| \geq 1/(2dN^\kappa)\Big\} < +\infty, \tag{1.10}$$

where $\mathbf{C}_j$, $j = 1, ..., [N^{\nu/d}]^d$ *are the subcubes of the cube* $\mathbf{C}$ *such that* $Vol(\mathbf{C}_1) = Vol(\mathbf{C}_j)$ *for all* $j = 1, ..., [N^{\nu/d}]^d$, *and* $\|\cdot\|$ *denotes the Euclidean norm in the space* $\mathbb{R}^d$. *Then, for every* $\varepsilon > 0$ *and* $L \geq 0$,

$$\sup_{t:|t|\geq N^\varepsilon} \big|(d/dt)^m \mathbb{E}\exp\{itV_N\}\big| = O\big(N^{-L}\big), \quad N \to +\infty. \tag{1.11}$$

Some simple corollaries of Theorem 1.3 are given in the paragraph 5 (see Corollaries 5.1–5.3 below). Let us note that in some special cases the validity of the bound (1.11) was investigated by Sadikova (1961), Yurinskii (1972), van Zwet (1977), Csörgő and Stachó (1980), Götze (1984), Zitikis (1988, 1990), Helmers (1991), Bentkus, Götze and Zitikis (1993), etc. See also the survey paper by Bentkus, Götze, Paulauskas and Račkauskas (1990).

We shall now discuss applications of Theorems 1.1–1.3 to some goodness-of-fit statistics. Let $\Lambda = (0,1)$, and let $\mathbf{X} = U$ where U denotes a $(0,1)$-uniform random variable.

Beran's statistic B_N^2.

Let $b_0, b_1, \dots, \sum b_l^2 < +\infty$, denote the Fourier coefficients with respect to the basis $\{e^{i2\pi lx}, l \in \mathbb{Z}\}$ of a probability density f on the circle S^1 of unit circumference. If the function H is given by

$$H_1(x,y) \triangleq 2 \sum_{l=1}^{\infty} b_l^2 \cos 2\pi l(x-y), \tag{1.12}$$

then V_N is Beran's statistic B_N^2; see Beran (1968, 1969). Further investigations of the statistic B_N^2, and more general ones as well, are done by Mardia (1972), Giné (1976), Prentice (1978), and Baringhaus (1991).

Statement 1.1. *Assume that the number of non-zero coefficients* $b_0, b_1, \ldots$ *is infinite and* $\sum_{l=0}^\infty b_l^2 l^2 < +\infty$. *Then*

(i) *Theorems 1.1 and 1.2 hold for Beran's statistic* B_N^2;

(ii) *For every fixed* $m \in \mathbb{N} \cup \{0\}$, $L \geq 0$ *and* $\varepsilon > 0$ *the bound (1.11) holds for all* $t \in \mathbb{R}$ *such that* $|t| \geq N^\varepsilon$;

(iii) *There exist infinitely differentiable and rapidly decreasing functions* $a_1, a_2, \ldots$ *such that, for every fixed* k *and* $m \in \mathbb{N} \cup \{0\}$, *the asymptotic expansion (1.3) holds with the remainder term* R *satisfying*

$$\sup_{x\in\mathbb{R}} |x|^m \big|R(x)\big| = O\big(N^{-k-1}\big), \quad N \to +\infty. \tag{1.13}$$

In the range $N^{1/2+\varepsilon} \le |t| \le c_1 N$ for some c_1 and any $\varepsilon > 0$, the bound (1.11) is given in the monograph by Koroliuk and Borovskich (1984).

Thus, in view of Statement 1.1 (iii) it follows that for Beran's statistic B_N^2 the bound of Theorem 1.2 holds with $\varepsilon = 0$ as well.

Watson's statistic W_N^2.

Here the function H is given by

$$H_2(x, y) \triangleq \mathbb{E} I(x, U) I(y, U)$$

where $I(x, v) \triangleq \mathbb{I}\{x \le v\} - v - \mathbb{E}\big(\mathbb{I}\{x \le U\} - U\big)$. Thus V_N is Watson's statistic W_N^2; see Watson (1961).

Statement 1.2. (i) *Theorems 1.1 and 1.2 hold for Watson's statistic W_N^2;*

(ii) *For every fixed $m \in \mathbb{N} \cup \{0\}$, $L \ge 0$ and $\varepsilon > 0$ the bound (1.11) holds for all $t \in \mathbb{R}$ such that $|t| \ge N^{\varepsilon}$;*

(iii) *There exist infinitely differentiable and rapidly decreasing functions $a_1, a_2, \ldots$ such that, for every fixed k and $m \in \mathbb{N} \cup \{0\}$, the asymptotic expansion (1.3) holds with a remainder term R satisfying (1.13).*

In the case $m = 0$ the result 1.2 (iii) was proved in Götze (1979).

Thus, in view of 1.2 (iii) the bound given in Theorem 1.2 for Watson's statistic W_N^2 is valid with $\varepsilon = 0$ as well.

The goodness-of-fit statistic $\omega_N^2(q)$.

Let $q : (0, 1) \to [0, +\infty)$ be a measurable function such that $\int x(1-x)q(x)dx < +\infty$. Here the function H is given by

$$H_3(x, y) \triangleq \mathbb{E} J(x, U) J(y, U) q(U)$$

where $J(x, v) \triangleq \mathbb{I}\{x \le v\} - v$, and V_N is the ω^2-square statistic; see Martynov (1979) for an exhaustive review on ω^2-square statistics. Let us recall that in the case $q(x) = 1$ for all $x \in (0, 1)$ this is Cramér-von Mises' statistic, and in the case $q(x) = 1/x(1-x)$ for all $x \in (0, 1)$ it is Anderson-Darling's statistic.

Statement 1.3. (i) *Theorems 1.1 and 1.2 hold for the Cramér-von Mises and Anderson-Darling statistics;*

(ii) *Let $m \in \mathbb{N} \cup \{0\}$ and assume that*

$$\mu(m) \triangleq \int_0^1 \Big\{ \int_0^u sq(s)ds \Big\}^m du + \int_0^1 \Big\{ \int_u^1 sq(s)ds \Big\}^m du < +\infty. \tag{1.14}$$

Furthermore, assume that there exists a non-empty interval $(\gamma,\delta)\subset(0,1)$ *such that*

$$q(x)>0 \quad \text{for all} \quad x\in(\gamma,\delta), \tag{1.15}$$

and, for some numbers $\tau>0$ *and* $c\geq 0$,

$$\big|q(x)-q(y)\big|\leq c|x-y|^{\tau} \quad \text{for all} \quad x,y\in(\gamma,\delta). \tag{1.16}$$

Then, for every fixed $\varepsilon>0$ *and* $L\geq 0$ *bound (1.11) holds for all* $t\in\mathbb{R}$ *such that* $|t|\geq N^{\varepsilon}$;

(iii) *Fix* $k,m\in\mathbb{N}\cup\{0\}$ *and assume that* $\mu(M)<+\infty$ *for* $M=k+2$ *and* $M=m$. *If assumptions (1.15) and (1.16) are satisfied, then there exist infinitely differentiable and rapidly decreasing functions* $a_1,a_2,\ldots$ *such that the asymptotic expansion (1.3) holds with the remainder* R *satisfying (1.13).*

In the case $p=2$ Statement 1.3 (ii) improves Theorem 2.7 by Bentkus, Götze and Zitikis (1993) where the same bound (1.11) was proved under the condition $\sup\big\{|q'(x)|: x\in(\gamma,\delta)\big\}<+\infty$ (instead of (1.16)) and in the region $|t|\geq N^{1/2+\varepsilon}$ only. Let us note that if assumption (1.15) is valid, then there is an infinite number of non-zero eigenvalues λ_j; see Bentkus, Götze and Zitikis (1993) and Bentkus Götze, Paulauskas and Račkauskas (1990) for more details on this subject. Statement 1.3 (iii) improves the corresponding results for ω^2-square statistics given by Bentkus, Götze and Zitikis (1993).

In view of Statement 1.3 we claim that for the Cramér-von Mises and Anderson-Darling statistics the bound given in Theorem 1.2 is valid with $\varepsilon=0$ as well.

2. Proof of Theorem 1.1.

Let $S^*: L_2(\Lambda,\mathbb{F}^*)\to L_2(\Lambda,\mathbb{F}^*)$ be the operator defined by the formula

$$S^*f\triangleq\mathbb{E}^*H^*(\cdot,\mathbf{X}^*)f(\mathbf{X}^*)=\int H^*(\cdot,\mathbf{y})f(\mathbf{y})\mathbb{F}^*(d\mathbf{y}).$$

Denote the eigenvalues of S^* by $\lambda_1^*,\lambda_2^*,\ \ldots\ ,\ |\lambda_1^*|\geq|\lambda_2^*|\geq\ldots$. Given the random variables $\mathbf{X}_1,\ldots,\mathbf{X}_N$, N fixed, V_n^* as n tends to infinity converges in distribution to the random variable

$$V_\infty^*\triangleq\mathbb{E}^*H^*(\mathbf{X}^*,\mathbf{X}^*)+\sum_{k=1}^{\infty}\lambda_k^*(G_k^2-1).$$

Lemma 1.1. *For every* $K\in\mathbb{N}$ *and every* $A\geq 0$ *there exists a constant* c *such that*

$$\mathbb{P}\big\{|\lambda_K^*|\leq|\lambda_K|/\sqrt{2K}\big\}\leq cN^{-A} \tag{2.1}$$

for all $N \in \mathbb{N}$.

Proof. The idea of the proof is based on the proof of Lemma 4.1 by Bickel, Götze and van Zwet (1986). Let us first note that without loss of generality we may assume $\lambda_K \neq 0$ and $A \geq 1$. Furthermore, let $e_1, ..., e_K$ be the eigenfunctions of the operator S corresponding to the eigenvalues λ_1, ... , λ_K, and let $S_2 : L_2(\Lambda, \mathbb{F}) \to L_2(\Lambda, \mathbb{F})$ be the (self–adjoint and positive–definite) Hilbert–Schmidt operator corresponding to the kernel

$$(\mathbf{x}, \mathbf{y}) \mapsto \mathbb{E}H(\mathbf{x}, \mathbf{X})H(\mathbf{X}, \mathbf{y}).$$

Then λ_1^2, ... ,λ_K^2 are the eigenvalues of the operator S_2 corresponding to the eigenfunctions $e_1, ..., e_K$. Define

$$\mathcal{A} \triangleq \bigcup_{p=1}^{K} \big\{ \|e_p\|_{\mathbb{F}^*}^2 \geq 1/2 \big\},$$

where $\|\cdot\|_{\mathbb{F}^*}$ denotes the norm in the space $L_2(\Lambda, \mathbb{F}^*)$. Since $\mathbb{E}e_p^2(\mathbf{X}) = 1$, the quantity $1 - \mathbb{P}(\mathcal{A})$ does not exceed $\sum_{p=1}^{K} \mathbb{P}\big\{\mathbb{E}e_p^2(\mathbf{X}) - \|e_p\|_{\mathbb{F}^*}^2 \geq 1/2\big\}$ and therefore, $1 - \mathbb{P}(\mathcal{A}) \leq cN^{-A}$. Thus we get the bound

$$\mathbb{P}\big\{|\lambda_K^*| \leq |\lambda_K|/\sqrt{2K}\big\} \leq \mathbb{P}\Big(\big\{|\lambda_K^*| \leq |\lambda_K|/\sqrt{2K}\big\} \cap \mathcal{A}\Big) + cN^{-A}. \tag{2.2}$$

Let δ_{pq} denotes the Kronecker delta, and let

$$\mathcal{B} \triangleq \bigcap_{p=1}^{K} \bigcap_{q=1}^{K} \Big\{ \Big| \frac{1}{N} \sum_{j=1}^{N} e_p(\mathbf{X}_j) e_q(\mathbf{X}_j) - \delta_{pq} \Big| \leq 1/(2K) \Big\}.$$

We clearly have

$$1 - \mathbb{P}(\mathcal{B}) \leq \sum_{p=1}^{K} \sum_{q=1}^{K} \mathbb{P}\Big\{ \Big| \frac{1}{N} \sum_{j=1}^{N} e_p(\mathbf{X}_j) e_q(\mathbf{X}_j) - \delta_{pq} \Big| \geq 1/2 \Big\}.$$

and, therefore, $1 - \mathbb{P}(\mathcal{B}) \leq cN^{-A}$. This bound together with (2.2) implies

$$\mathbb{P}\big\{|\lambda_K^*| \leq |\lambda_K|/\sqrt{2K}\big\} \leq \mathbb{P}\Big(\big\{|\lambda_K^*| \leq |\lambda_K|/\sqrt{2K}\big\} \cap \mathcal{A} \cap \mathcal{B}\Big) + cN^{-A}, \tag{2.3}$$

and therefore our further task is to show that bound (2.1) holds for the quantity $\mathbb{P}\big(\big\{|\lambda_K^*| \leq |\lambda_K|/\sqrt{2K}\big\} \cap \mathcal{A} \cap \mathcal{B}\big)$ instead of $\mathbb{P}\big\{|\lambda_K^*| \leq |\lambda_K|/\sqrt{2K}\big\}$.

Let us examine the event $\mathcal{B}$ more closely. Assume that $c_1 e_1 + ... + c_K e_K = 0$ in the space $L_2(\Lambda, \mathbb{F}^*)$. This means that $\sum_{p=1}^K c_p e_p(\mathbf{X}_j) = 0$ for all $j = 1, ..., N$. Thus

$$0 = \frac{1}{N}\sum_{j=1}^N \Big|\sum_{p=1}^K c_p e_p(\mathbf{X}_j)\Big|^2 = \sum_{p=1}^K c_p^2 + \sum_{p=1}^K\sum_{q=1}^K c_p c_q \Big\{\frac{1}{N}\sum_{j=1}^N e_p(\mathbf{X}_j)e_q(\mathbf{X}_j) - \delta_{pq}\Big\}$$
$$\geq \sum_{p=1}^K c_p^2 - \frac{1}{2K}\Big\{\sum_{p=1}^K |c_p|\Big\}^2$$
$$\geq \frac{1}{2}\sum_{p=1}^K c_p^2$$

Therefore $c_1 = ... = c_K = 0$. Thus the functions $e_1, ..., e_K$ are linearly independent in the space $L_2(\Lambda, \mathbb{F}^*)$.

Hence $e_p^\diamond \triangleq e_p/\|e_p\|_{F^*}$, $p = 1, ..., K$ are well-defined (on the set $\mathcal{A}$) and linearly independent in the space $L_2(\Lambda, \mathbb{F}^*)$ (on the set $\mathcal{B}$). Thus, $\mathcal{G} \triangleq$ span $\{e_1^\diamond, ..., e_K^\diamond\}$ is a K-dimensional subspace of the space $L_2(\Lambda, \mathbb{F}^*)$.

Furthermore, let $S_2^* : L_2(\Lambda, \mathbb{F}^*) \to L_2(\Lambda, \mathbb{F}^*)$ be the (self–adjoint and positive–definite) Hilbert–Schmidt operator corresponding to the kernel

$$(\mathbf{x}, \mathbf{y}) \mapsto \mathbb{E}^* H^*(\mathbf{x}, \mathbf{X}^*) H^*(\mathbf{X}^*, \mathbf{y}).$$

If $e_1^*, ..., e_K^*$ are the eigenfunctions of the operator S^* corresponding to the eigenvalues $\lambda_1^*, \dots, \lambda_K^*$, then $\lambda_1^{*2}, \dots, \lambda_K^{*2}$ are the eigenvalues of the operator S_2^* corresponding to the eigenfunctions $e_1^*, ..., e_K^*$. Since $\mathcal{G}$ is a K-dimensional subspace of the space $L_2(\Lambda, \mathbb{F}^*)$ the following estimates hold:

$$\lambda_K^{*2} = \sup_{\mathcal{H}} \inf_{f \in \mathcal{H}} \frac{\|S_2^*\|_{\mathbb{F}^*}}{\|f\|_{\mathbb{F}^*}} \geq \inf_{f \in \mathcal{G}} \frac{\|S_2^*\|_{\mathbb{F}^*}}{\|f\|_{\mathbb{F}^*}} \geq \inf_{f \in \mathcal{G}} \frac{\big\langle S_2^* f, f\big\rangle_{\mathbb{F}^*}}{\|f\|_{\mathbb{F}^*}^2}, \tag{2.4}$$

where $\mathcal{H}$ denotes an arbitrary K-dimensional subspace of $L_2(\Lambda, \mathbb{F}^*)$. Since $\delta_{pq}\lambda_p^2 = \big\langle S_2 e_p, e_q\big\rangle_{\mathbb{F}}$, we have

$$\big\langle S_2^* e_p^\diamond, e_q^\diamond\big\rangle_{\mathbb{F}^*} = \delta_{pq}\lambda_p^2 + \xi_{pq}, \tag{2.5}$$

where $\langle\cdot,\cdot\rangle_{\mathbb{F}}$ and $\langle\cdot,\cdot\rangle_{\mathbb{F}^*}$ denote the inner products in the spaces $L_2(\Lambda, \mathbb{F})$ and $L_2(\Lambda, \mathbb{F}^*)$ respectively, and

$$\xi_{pq} \triangleq \big\langle S_2 e_p, e_q\big\rangle_{\mathbb{F}} - \big\langle S_2^* e_p^\diamond, e_q^\diamond\big\rangle_{\mathbb{F}^*}.$$

Furthermore, since $f=\sum_{p=1}^{K} c_p e_p^\diamond$ for some $\mathbf{c} \triangleq (c_1,...,c_K)$ and $\|e_p^\diamond\|_{F^*}^2=1$, we obtain $\|f\|_{\mathbb{F}^*}^2 \le K\|\mathbf{c}\|^2$ where $\|\mathbf{c}\|^2 \triangleq c_1^2+...+c_K^2$ and, therefore,

$$\begin{aligned}\big\langle S_2^* f, f\big\rangle_{\mathbb{F}^*} &= \big\langle \big(\delta_{pq}\lambda_p^2\big)_{p,q=1,\ldots,K}\mathbf{c},\mathbf{c}\big\rangle + \big\langle \big(\xi_{pq}\big)_{p,q=1,\ldots,K}\mathbf{c},\mathbf{c}\big\rangle \\ &\ge \lambda_K^2\|\mathbf{c}\|^2 - \max_{p,q=1,\ldots,K}|\xi_{pq}|\|\mathbf{c}\|^2 \\ &\ge \big\{\lambda_K^2 - \max_{p,q=1,\ldots,K}|\xi_{pq}|\big\}\|f\|_{\mathbb{F}^*}^2/K\end{aligned}$$

where $\langle\cdot,\cdot\rangle$ denotes the inner product in the space $\mathbb{R}^K$. Applying the just obtained estimate in the right-hand side of (2.4), we get

$$\lambda_K^{*2} \ge \big\{\lambda_K^2 - \max_{p,q=1,\ldots,K}|\xi_{pq}|\big\}/K$$

which implies

$$\begin{aligned}\mathbb{P}\Big(\big\{|\lambda_K^*| \le |\lambda_K|/\sqrt{2K}\big\}\cap\mathcal{A}\cap\mathcal{B}\Big) &\le \mathbb{P}\big\{\lambda_K^2/(2K) \ge \big\{\lambda_K^2 - \max_{p,q=1,\ldots,K}|\xi_{pq}|\big\}/K.\big\} \\ &\le \mathbb{P}\big\{\lambda_K^2/2 \le \max_{p,q=1,\ldots,K}|\xi_{pq}|\big\} \qquad (2.6)\end{aligned}$$

Let us show that the right-hand side of (2.6) does not exceed cN^{-A}. This could be proved as follows: Write ξ_{pq} as the sum $\Delta_1+\Delta_2+\Delta_3$, where

$$\begin{aligned}\Delta_1 &\triangleq \big\langle S_2 e_p, e_q\big\rangle_{\mathbb{F}} - \big\langle S_2 e_p, e_q\big\rangle_{\mathbb{F}^*}, \\ \Delta_2 &\triangleq \big\langle S_2 e_p^\diamond, e_q^\diamond\big\rangle_{\mathbb{F}^*}\big\{\|e_p\|_{F^*}\|e_q\|_{\mathbb{F}^*}-1\big\}, \\ \Delta_3 &\triangleq \big\langle S_2 e_p^\diamond, e_q^\diamond\big\rangle_{\mathbb{F}^*} - \big\langle S_2^* e_p^\diamond, e_q^\diamond\big\rangle_{\mathbb{F}^*}.\end{aligned}$$

One could easily prove that for every $j=1,2,3$, every $A\ge 0$ and every positive constant c_1, the quantities $\mathbb{P}\big\{|\Delta_j|\ge c_1\big\}$ do not exceed cN^{-A}. Since $\lambda_K^2>0$, this completes the proof of the lemma. $\square$

Lemma 1.2. *For every $A\ge 0$ there exists a constant $c\ge 0$ such that*

$$\mathbb{P}\big\{\sqrt{N}\big\|P_\infty - P_\infty^*\big\|_\infty \ge a\big\} \le cN^{-A}+ca^{-A}$$

for all $N\in\mathbb{N}$ and $a\ge 0$.

Proof. Because of Lemma 1.1 we only need to show that

$$\mathbb{P}\big\{\sqrt{N}\big\|P_\infty - P_\infty^*\|_\infty \ge a, \quad |\lambda_k^*| \ge |\lambda_k|/\sqrt{2k}, \quad k=1,...,K\big\} \le cN^{-A}+ca^{-A} \qquad (2.7)$$

for a number K depending on A. In the following we have to estimate $\big\|P_\infty - P_\infty^*\big\|_\infty$ using Fourier's inversion formula. Let us first state some auxiliary results.

Let $L_2(\Lambda, \mathbb{F})$ denote the complex Hilbert space $L_2(\Lambda, \mathbb{C}, \mathbb{F})$ of all measurable functions $f : \Lambda \to \mathbb{C}$ such that $\|f\|_{\mathbb{F}}^2 \triangleq \mathbb{E} f(\mathbf{X})\,\overline{f}(\mathbf{X}) = \int_\Lambda f\,\overline{f} d\mathbb{F} < +\infty$. The inner product in this space will be denoted by $\langle\cdot,\cdot\rangle_{\mathbb{F}}$. Also, let I denote the identity operator, and $\mathcal{T}_{\mathbb{F}}$ denote the trace operator in the space $L_2(\Lambda, \mathbb{F})$. Then

$$\mathbb{E}\exp\{itV_\infty\} = \exp\{T_t\},$$

where

$$T_t \triangleq it\mathbb{E}H(\mathbf{X}, \mathbf{X}) - 2\int_0^t \mathcal{T}_{\mathbb{F}}\big\{S(I - 2ivS)^{-1}S\big\}v\,dv.$$

A similar representation holds for the quantity $\mathbb{E}^* \exp\{itV_\infty^*\}$.

Let us prove that there exists a constant $c \geq 0$ which might depend on K and λ_K only, and such that, for all $w \in [0, 1]$,

$$\Psi \triangleq \big|\exp\{wT_t^* + (1 - w)T_t\}\big| \leq c\big(1 + |t|\big)^{-K/2} \tag{2.8}$$

Indeed, the representation

$$\mathcal{T}_{\mathbb{F}}\big\{S(I - 2ivS)^{-1}S\big\} = \mathcal{T}_{\mathbb{F}}\big\{S(I + 4v^2S^2)^{-1}S\big\} + 2iv\mathcal{T}_{\mathbb{F}}\big\{S(I + 4v^2S^2)^{-1}S^2\big\}$$

shows that

$$\mathrm{Re}\ \mathcal{T}_{\mathbb{F}}\{S(I - 2ivS)^{-1}S\} = \sum_{k=1}^{\infty} \frac{\lambda_k^2}{1 + 4v^2\lambda_k^2}. \tag{2.9}$$

A similar representation holds for the quantity $Re\mathcal{T}_{\mathbb{F}^*}\big\{S^*(I - 2ivS^*)^{-1}S^*\big\}$ as well. Therefore,

$$\Psi \leq \exp\bigg\{-2w\int_0^t \sum_{k=1}^{K} \frac{\lambda_k^{*2}}{1 + 4v^2\lambda_k^{*2}}v\,dv - 2(1 - w)\int_0^t \sum_{k=1}^{K} \frac{\lambda_k^2}{1 + 4v^2\lambda_k^2}v\,dv\bigg\}. \tag{2.10}$$

Since $\lambda_k^{*2} \geq \lambda_k^2/(2k)$ for all $k = 1, ..., K$ the bound (2.10) implies $\Psi \leq c\big(1+|t|\big)^{-K/2}$ which completes the proof of (2.8).

Starting from (2.8) we may use Fourier's inversion formula to bound the quantity $\big\|P_\infty - P_\infty^*\big\|_\infty$. We get

$$\big\|P_\infty - P_\infty^*\big\|_\infty \leq \Phi_1 \triangleq \int_{\mathbb{R}} \frac{1}{|t|}\Big|\mathbb{E}\exp\{itV_\infty\} - \mathbb{E}^*\exp\{itV_\infty^*\}\Big|dt. \tag{2.11}$$

Furthermore, the bound (2.8) implies

$$\Phi_1 = \int_{\mathbb{R}} \int_0^1 \Big| \exp\Big\{ wT_t^* + (1-w)T_t \Big\} \Big| \Big| t^{-1}\{T_t^* - T_t\} \Big| dwdt \le c\Phi_2, \tag{2.12}$$

where

$$\Phi_2 \triangleq \int_{\mathbb{R}} \big(1+|t|\big)^{-K/2} \Big| t^{-1}\big\{T_t^* - T_t\big\} \Big| dt.$$

Let us prove the estimate

$$\mathbb{P}\big\{\sqrt{N}\Phi_2 \ge a\big\} \le cN^{-A} + ca^{-A} \tag{2.13}$$

which clearly completes the proof of the lemma.

Let us introduce some additional notation. Denote the resolvent operator of S by $R(\cdot, S)$, that is
$R(z,S) \triangleq (zI - S)^{-1}$ for any complex number z. Unless stated otherwise, we shall always assume $z = 1/(2iv)$. With these notations we may rewrite T_t as follows:

$$T_t = it\mathbb{E}H(\mathbf{X},\mathbf{X}) + \int_0^t \mathbb{E}\Big\langle iR(z,S)H(\mathbf{X},\cdot), H(\mathbf{X},\cdot)\Big\rangle_{\mathbb{F}} dv.$$

Also, a similar representation holds for the quantity T_t^* as well. Therefore,

$$t^{-1}\{T_t^* - T_t\} = ih_1 + t^{-1}\int_0^t h_2(v)dv,$$

where

$$h_1 \triangleq \mathbb{E}^* H^*(\mathbf{X}^*,\mathbf{X}^*) - \mathbb{E}H(\mathbf{X},\mathbf{X}) \text{ and}$$

$$h_2(v) \triangleq \mathbb{E}^* \Big\langle iR(z,S^*)H^*(\mathbf{X}^*,\cdot), H^*(\mathbf{X}^*,\cdot)\Big\rangle_{\mathbb{F}^*} - \mathbb{E}\Big\langle iR(z,S)H(\mathbf{X},\cdot), H(\mathbf{X},\cdot)\Big\rangle_{\mathbb{F}}.$$

Consequently,

$$\Phi_2 \le c|h_1| + \Delta[h_2],$$

where

$$\Delta[h_2] \triangleq \int_{\mathbb{R}} \big(1+|t|\big)^{-K/2} |t|^{-1} \int_0^t \big|h_2(v)\big| dvdt.$$

It is easy to see that $\mathbb{P}\big\{\sqrt{N}|h_1| \ge a\big\} \le cN^{-A} + ca^{-A}$. Therefore, in order to prove (2.13) we only need to show

$$\mathbb{P}\big\{\sqrt{N}\Delta[h_2] \ge a\big\} \le cN^{-A} + ca^{-A} \tag{2.14}$$

Here we use the estimate $\Delta[h_2] \le \Delta[h_3] + \Delta[h_4] + \Delta[h_5]$, where $h_3(v)$, $h_4(v)$, and $h_5(v)$ are the following three quantities, respectively:

$$\mathbb{E}^* \Big\langle R(z,S^*)H^*(\mathbf{X}^*,\cdot), H^*(\mathbf{X}^*,\cdot)\Big\rangle_{\mathbb{F}^*} - \mathbb{E}^* \Big\langle R(z,S^*)H(\mathbf{X}^*,\cdot), H(\mathbf{X}^*,\cdot)\Big\rangle_{\mathbb{F}^*},$$

$$\mathbb{E}^* \Big\langle R(z,S^*)H(\mathbf{X}^*,\cdot), H(\mathbf{X}^*,\cdot)\Big\rangle_{\mathbb{F}^*} - \mathbb{E}^* \Big\langle R(z,S)H(\mathbf{X}^*,\cdot), H(\mathbf{X}^*,\cdot)\Big\rangle_{\mathbb{F}^*},$$

$$\mathbb{E}^* \Big\langle R(z,S)H(\mathbf{X}^*,\cdot), H(\mathbf{X}^*,\cdot)\Big\rangle_{\mathbb{F}^*} - \mathbb{E}\Big\langle R(z,S)H(\mathbf{X},\cdot), H(\mathbf{X},\cdot)\Big\rangle_{\mathbb{F}}.$$

Some straightforward calculations show that

$$\mathbb{P}\big\{\sqrt{N}\Delta[h_j] \ge a\big\} \le cN^{-A} + ca^{-A}$$

for $j = 3,4,5$ which proves (2.14) and the lemma. □

We will now prove Theorem 1.1. By Theorem (2.3) in Götze (1979) we have

$$\big\|P_N - P_\infty\big\|_\infty \le c\beta_3^3\lambda_{31}^{-9}/\sqrt{N}, \tag{2.15}$$

where $\beta_3 \triangleq \mathbb{E}\big|H(\mathbf{X},\mathbf{X})\big|^3 + \mathbb{E}\big|H(\mathbf{X}_1,\mathbf{X}_2)\big|^3$. Let us note that the cited result also yields the bound

$$\big\|P_N^* - P_\infty^*\big\|_\infty \le c\big(\beta_3^*\big)^3\big(\lambda_{31}^*\big)^{-9}/\sqrt{N}, \tag{2.16}$$

where β_3^* is similar to β_3 with distribution function $\mathbb{F}$ replaced by the empirical distribution function $\mathbb{F}^*$. Thus, the bounds (2.15) and (2.16) and Lemmas 1.1 and 1.2 together complete the proof of the theorem. □

3. Proof of Theorem 1.2.

For the proof of Theorem 1.2 we need the following lemma (compare Lemma 1.1 by Beran (1969)) several times:

Lemma 3.1. *For each $\alpha \in (0,1)$ we have*

$$P_\infty'\big(P_\infty^{-1}(\alpha)\big) > 0.$$

Proof. Without loss of generality we may assume that $\lambda_1 > 0$. It is clear that the lemma follows from the following result:

There exists a point $x_0 \in [-\infty,+\infty)$ such that $P_\infty'(x) = 0$ for all $x \le x_0$ and $P_\infty'(x) > 0$ for all $x > x_0$.

Let p and q denote, respectively, the densities of the random variables $\lambda_1(G_1^2-1)$ and $\sum_{k=2}^{\infty}\lambda_k(G_k^2-1)$. The density p vanishes for all $x\le -\lambda_1$ and is positive for all $x>-\lambda_1$. Since the function P_∞ is continuous, the density q is not degenerated. Our claim follows from the representation $P'_\infty(x)=\int p(x-z)q(z)dz$ such that $P'_\infty(x)=0$ for $x\le x_0$ and $P'_\infty(x)>0$ for all $x_0<x<z$ for some $z\in\mathbb{R}$. The assumption $P'_\infty(z)=0$ now leads to a contradiction with $P'_\infty(x)>0$ for $x_0<x<z$. $\square$

Without loss of generality we assume $\varepsilon>0$ in the proof of Theorem 1.2 to be small (say $\varepsilon<10^{-10}$) and $N\ge N_0$ where N_0 is a large fixed constant (which might depend on $\varepsilon>0$ and $\alpha\in(0,1)$).

We have

$$\begin{aligned}\mathbb{P}\big\{V_N\ge P_N^{*-1}(\alpha)\big\}&=\mathbb{P}\big\{P_N^*(V_N)\ge\alpha\big\}\\&=\mathbb{P}\big\{(P_N^*-P_\infty^*)(V_N)+P_\infty^*(V_N)\ge\alpha\big\}.\end{aligned}\tag{3.1}$$

Let $\beta_0=N^{-1+\varepsilon}$.

Using the bound (2.5) of Götze (1979) (in the case $s=4$) and Lemma 1.1 we get that, for every $A\ge 0$,

$$\mathbb{P}\big\{\|P_N^*-P_\infty^*\|_\infty\ge\beta_0\big\}=O(N^{-A}).$$

Therefore, using a Slutzky argument (see Lemma 3 on p.16 in Petrov (1975)) in the right-hand side of (3.1), we get the bound (1.2) provided that

$$\mathbb{P}\big\{V_N\ge P_\infty^{*-1}(\gamma)\big\}=1-\gamma+O\big(N^{-1+\varepsilon}\big)\tag{3.2}$$

with $\gamma=\alpha\pm\beta_0$.

Furthermore, since $P_\infty\big(P_\infty^{-1}(\gamma)\big)=\gamma$ we obtain

$$\begin{aligned}\mathbb{P}\big\{V_N\ge P_\infty^{*-1}(\gamma)\big\}&=1-\gamma\\&\quad+\mathbb{P}\big\{V_N-P_\infty^{*-1}(\gamma)+P_\infty^{-1}(\gamma)\ge P_\infty^{-1}(\gamma)\big\}-\mathbb{P}\big\{V_\infty\ge P_\infty^{-1}(\gamma)\big\}.\end{aligned}$$

Thus in order to show (3.2) it is enough to prove

$$\Delta\triangleq\big\|\mathbb{P}\big\{V_N-P_\infty^{*-1}(\gamma)+P_\infty^{-1}(\gamma)\ge\cdot\big\}-\mathbb{P}\big\{V_\infty\ge\cdot\big\}\big\|_\infty=O\big(N^{-1+\varepsilon}\big).\tag{3.3}$$

Denote $q(f)\triangleq q_\gamma(f)\triangleq f^{-1}(\gamma)$, and let $h\triangleq P_\infty^*-P_\infty$.

Since $P'_\infty\big(P_\infty^{-1}(\gamma)\big)\ge const(P_\infty,\alpha)>0$ (recall that $N\ge c_0$ where c_0 is a large fixed constant, and note that $P'_\infty\big(P_\infty^{-1}(\alpha)\big)>0$), we obtain that the quantity

$$q'(P_\infty)(h)\triangleq -h\big(P_\infty^{-1}(\gamma)\big)/P'_\infty\big(P_\infty^{-1}(\gamma)\big)$$

is well defined.

(We shall show at the end of the proof that on a set of probability close to 1 the quantity $q'(P_\infty)(h)$ is actually the directional derivative of the function q at the point P_∞ in the direction h.)

An application of Slutzky arguments shows that the quantity Δ does not exceed $\Delta_1 + \Delta_2 + \Delta_3$, where

$$\begin{aligned}
\Delta_1 &\triangleq \Big\| \mathbb{P}\big\{V_N - q'(P_\infty)(h) \geq \cdot\big\} - \mathbb{P}\big\{V_\infty \geq \cdot\big\}\Big\|_\infty, \\
\Delta_2 &\triangleq \mathbb{P}\big\{|q(P_\infty^*) - q(P_\infty) - q'(P_\infty)(h)| \geq \beta_0\big\}, \\
\Delta_3 &\triangleq \sup_{x\in\mathbb{R}} \mathbb{P}\big\{x - \beta_0 \leq V_\infty \leq x + \beta_0\big\}.
\end{aligned}$$

Using the fact that the Fourier transformation of P_∞ decreases rapidly, we easily arrive at the bound $\Delta_3 \leq c\beta_0 = O\big(N^{-1+\varepsilon}\big)$. To prove the same bound for the quantity Δ_1 note first (see some details below) that,for any fixed real x,

$$h(x) = \frac{1}{N}\sum_{j=1}^{N} f(X_j, x) + r_N, \tag{3.4}$$

where f is a measurable function, the i.i.d random variables $f(X_j, x)$ are centered and has finite moments of all orders. Furthermore, the remainder term r_N is such that for every $A \geq 0$ there exists a constant $c \geq 0$ such that

$$\mathbb{P}\big\{N|r_N| \geq a\big\} \leq cN^{-A} + ca^{-A} \tag{3.5}$$

for all $N \in \mathbb{N}$ and $a \geq 0$.

The asymptotic expansion (3.4) with the remainder term r_N as in (3.5) could be proved using Fourier transformation and (with slight modifications) following the lines of the proof of Lemma 1.2; we omit the details of the proof. Using Slutzky arguments together with the bounds (3.5) and $\Delta_3 \leq c\beta_0$ we obtain, for every $A \geq 0$,

$$\Delta_1 \leq \Big\|\mathbb{P}\Big\{V_N + \frac{1}{N}\sum_{j=1}^{N} f(X_j, x_0)/P_\infty'(x_0) \geq \cdot\Big\} - \mathbb{P}\Big\{V_\infty \geq \cdot\Big\}\Big\|_\infty + O\big(N^{-1+\varepsilon}\big), \tag{3.6}$$

where $x_0 \triangleq P_\infty^{-1}(\gamma)$. The first summand in the right-hand side of (3.6) is of the order $O\big(N^{-1+\varepsilon}\big)$; this fact follows from Corollary (3.20) by Götze (1985) (see Example 3.7 there as well). Thus we have proved $\Delta_1 = O\big(N^{-1+\varepsilon}\big)$, and still have to show

$$\Delta_2 = O\big(N^{-1+\varepsilon}\big). \tag{3.7}$$

Let $\beta_1 \triangleq N^{-1/2+\varepsilon}$ and denote $\mathcal{A} \triangleq \mathcal{A}_1 \cap ... \cap \mathcal{A}_4$, where

$$\begin{aligned} \mathcal{A}_1 &\triangleq \{|\lambda_K^*| \geq |\lambda_K|/\sqrt{2K}\}, \\ \mathcal{A}_{2+\kappa} &\triangleq \{\|h^{(\kappa)}\|_\infty \leq \beta_1\}, \quad \kappa = 0,1,2. \end{aligned}$$

($h^{(\kappa)}$ denotes the κ-th derivative of the h; $h^{(0)} \triangleq h$.) We know already that $1 - \mathbb{P}(\mathcal{A}_1) = O(N^{-A})$ (Lemma 1.1) and $1 - \mathbb{P}(\mathcal{A}_2) = O(N^{-A})$ (Lemma 1.2). Some slight modifications of the proof of Lemma 1.2 leads to the bound (valid for every fixed $\kappa = 0, 1, 2, ...$)

$$\mathbb{P}\{\sqrt{N}\|h^{(\kappa)}\|_\infty \geq a\} \leq cN^{-A} + ca^{-A} \tag{3.8}$$

which shows that $1 - \mathbb{P}(\mathcal{A}_{2+\kappa}) = O(N^{-A})$ for $\kappa = 1, 2$ (we omit the details since they are straightforward).

Therefore we may restrict our analysis to the set $\mathcal{A}$ only. On this set we have that, for some constants c_1 and c_2 which do not depend upon N and $\tau \in [0, 1]$, the following two bounds hold: $H'_\tau(H_\tau^{-1}(\gamma)) \geq c_1 > 0$ and $|H''_\tau(H_\tau^{-1}(\gamma))| \leq c_2$, where $H_\tau \triangleq P_\infty + \tau h$. This leads to the following Taylor expansion (with all the quantities well-defined) for $q(P^*_\infty)$:

$$q(P^*_\infty) = q(P_\infty) + q'(P_\infty)(h) + \int_0^1 (1-\tau) q''(H_\tau)(h)^2 d\tau,$$

where

$$\begin{aligned} q''(H_\tau)(h)^2 &= 2h\big(H_\tau^{-1}(\gamma)\big)h'\big(H_\tau^{-1}(\gamma)\big)/H'_\tau\big(H_\tau^{-1}(\gamma)\big)^2 \\ &\quad - h\big(H_\tau^{-1}(\gamma)\big)^2 H''_\tau\big(H_\tau^{-1}(\gamma)\big)/H'_\tau\big(H_\tau^{-1}(\gamma)\big)^3. \end{aligned}$$

After some tedious but elementary calculations, we obtain

$$\Delta_2 \leq \mathbb{P}\{\|h\|_\infty \|h'\|_\infty + \|h\|^2_\infty \geq \beta_0\} + O(N^{-A})$$

which together with bound (3.8) completes the proof of (3.7), and of the theorem as well. $\square$

4. Proof of Theorem 1.3.

In the proof we shall use the following lemma.

Lemma 4.1. *Let $0 \leq a < b \leq 1$ be any numbers, and let g be a differentiable on (a, b) function such that $\operatorname{sgn} g'(x) = const$ for all $x \in (a, b)$, and, for some numbers $\Phi > 0$ and Ψ,*

$$\Phi \leq |g'(x)| \leq \Psi$$

for all $x \in (a, b)$.

Furthermore, let h *be an integrable on* (a, b) *function such that, for some* $\Upsilon > 0$,

$$h(x) \geq \Upsilon$$

for all $x \in (a, b)$. *Then, for all* $\tau_0 \geq 0$ *and all* τ *such that* $|\tau| \geq \tau_0$,

$$\begin{aligned} I \triangleq & \Big| \int_a^b \exp\{i\tau g(x)\} h(x) dx \Big| \\ & \leq \int_a^b h(x) dx - \frac{\Upsilon \min\{\Phi^3, 1\}}{48\, \Psi} \min\{\tau_0^2, 1\}(b-a)^3. \end{aligned} \tag{4.1}$$

Proof. With some slight modifications we shall follow the lines of the proof of Lemma 2.1 by Götze and Hipp (1992). Without loss of generality we assume that $sgn\, g'(x) = 1$. Because of $h(g^{-1}(x)) \geq \Upsilon$ and $g'(g^{-1}(x)) \leq \Psi$ for all $x \in (g(a), g(b))$, we get that

$$\begin{aligned} I \leq & \Big| \int_{g(a)}^{g(b)} \exp\{i\tau x\} \Big\{ \frac{h(g^{-1}(x))}{g'(g^{-1}(x))} - \frac{\Upsilon}{\Psi} \Big\} dx \Big| + \frac{\Upsilon}{\Psi} \Big| \int_{g(a)}^{g(b)} \exp\{i\tau x\} dx \Big| \\ \leq & \int_a^b h(x) dx - \frac{\Upsilon}{\Psi}\{g(b) - g(a)\} + \frac{\Upsilon}{\Psi} \Big| \int_{g(a)}^{g(b)} \exp\{i\tau x\} dx \Big|. \end{aligned} \tag{4.2}$$

Furthermore,

$$\Big| \int_{g(a)}^{g(b)} \exp\{i\tau x\} dx \Big| = \{g(b) - g(a)\} \Big| \frac{\sin v}{v} \Big|, \tag{4.3}$$

where $v \triangleq \tau\{g(b) - g(a)\}/2$. To estimate the quantity $|v^{-1} \sin v|$ in the right-hand side of (4.3) we use the bound $|x^{-1} \sin x| \leq 1 - \min\{d_0^2, 1\}/12$ which is true for all $d_0 \geq 0$ and all $|x| \geq d_0$. Let us find a number d_0 such that $|v| \geq d_0$. It is clear that $|g(b) - g(a)| \geq \Phi(b - a)$. Therefore, $d_0 = \tau_0 \Phi(b - a)/2$ and so we have

$$\Big| \frac{\sin v}{v} \Big| \leq 1 - \frac{\min\{\Phi^2, 1\}}{48} \min\{\tau_0^2, 1\}(b-a)^2 \tag{4.4}$$

Now taking bounds (4.2)–(4.4) together, and using bound $|g(b) - g(a)| \geq \Phi(b - a)$ once again, we get the lemma proved. $\square$

Let us note at the beginning of the proof of Theorem 1.3 that we do not specify the constants c in the text below; we only want to emphasize is that all of them are non-random, do not depend on both N and t, and are non-negative.

We subdivide the proof into several steps.

Step 1. (*Reduction of the general case to the case* $m = 0$.) Because of the moment condition (1.6), we interchange the signs of differentiation and integration in the quantity $(d/dt)^m \mathbb{E}\exp\{itV_N\}$. Then we write V_N^m as a multiply sum and use the fact that $\mathbf{X}_1, ..., \mathbf{X}_N$ are identically distributed. We get

$$\big|(d/dt)^m \mathbb{E}\exp\{itV_N\}\big| \le cN^{2m}\mathbb{E}\big|\mathbb{E}^{\ominus}\exp\{itV_N\}\big|,$$

where $\mathbb{E}^{\ominus}$ denotes the conditional expectation with respect to $\mathbf{X}_1, ..., \mathbf{X}_{N-2m}$ when all the other random vectors are fixed. Thus the theorem follows if

$$\mathbb{E}\big|\mathbb{E}^{\ominus}\exp\{itV_N\}\big| \le cN^{-L}. \tag{4.5}$$

Step 2. (*Using the absolutely continuous part.*) Here we employ arguments used by Bikelis (1973). Without loss of generality we assume that $N \ge c_1$ where c_1 is a large constant (which might depend only on m and α). Denote $N_1 \triangleq N - 2m$ and fix $\mathbf{X}_{N-2m+1}, ..., \mathbf{X}_N$. Because of decomposition (1.5), we have

$$\begin{aligned}\mathbb{E}^{\ominus}\exp\{itV_N\} &= \int \exp\{itV_N\}\big(d\mathbb{F}\big)^{N_1} \\ &= \sum_{j=0}^{N_1}\binom{N_1}{j}\alpha^j(1-\alpha)^{N_1-j}\int\int \exp\{itV_N\}\big(d\mathbb{F}_0\big)^j\big(d\mathbb{F}_1\big)^{N_1-j}.\end{aligned} \tag{4.6}$$

Now split the summation in (4.6) into two parts: $\sum' \triangleq$ the sum taken over all $j = 0, ..., N_1$ such that $|j - \alpha N_1| < \sqrt{N_1}\log N_1$, and $\sum'' \triangleq$ the sum taken over all $j = 0, ..., N_1$ such that $|j - \alpha N_1| \ge \sqrt{N_1}\log N_1$. A theorem by Bernstein (1924)yields the bound

$$\sum_{(j-\alpha N_1)\ge 2x\sqrt{N_1\alpha(1-\alpha)}}\binom{N_1}{j}\alpha^j(1-\alpha)^{N_1-j} \le 2\exp\{-x^2\} \tag{4.7}$$

for every $x \le \sqrt{N_1\alpha(1-\alpha)/4}$. If we take $x = x_0$, where x_0 is the solution of the equation $2x\sqrt{\alpha(1-\alpha)} = \log N_1$, the bound (4.7) implies that the quantity $\sum''\binom{N_1}{j}\alpha^j(1-\alpha)^{N_1-j}$ does not exceed cN^{-L}. This bound and representation (4.6) together yield

$$\begin{aligned}\big|\mathbb{E}^{\ominus}\exp\{itV_N\}\big| &\le cN^{-L} \\ &+ {\sum}'\binom{N_1}{j}\alpha^j(1-\alpha)^{N_1-j}\int\Big|\int \exp\{itV_N\}\big(d\mathbb{F}_0\big)^j\Big|\big(d\mathbb{F}_1\big)^{N_1-j}.\end{aligned} \tag{4.8}$$

To estimate the second summand in the right-hand side of (4.8) we use the fact that $|j-\alpha N_1|<\sqrt{N_1}\log N_1$ (which implies that $j\ge[N_1\alpha/2]$). This fact and the estimate (4.8) show that the theorem will follow if we show

$$\mathbb{E}\Big|\int\exp\{itV_N\}\big(d\mathbb{F}_0\big)^{[N_1\alpha/2]}\Big|\le cN^{-L}. \tag{4.9}$$

Step 3. (*Reduction to the set* $\mathbf{B}\times\mathbf{C}$.) The main idea of this step is based on arguments by Bikelis (1973), Götze (1979), and Bickel, Götze, van Zwet (1986). Let $\beta\triangleq\mathbb{P}\{\mathbf{Y}\in\mathbf{B}\}$ and $\gamma\triangleq\mathbb{P}\{\mathbf{Y}\in\mathbf{C}\}$, and let

$$\mathbb{F}_2(\cdot)\triangleq\beta^{-1}\mathbb{P}\{\mathbf{Y}\in\cdot\cap\mathbf{B}\},$$

$$\mathbb{F}_3(\cdot)\triangleq\gamma^{-1}\mathbb{P}\{\mathbf{Y}\in\cdot\cap\mathbf{C}\}.$$

Furthermore, denote $N_2\triangleq[N_1\alpha/4]$ and write

$$\int\exp\{itV_N\}\big(d\mathbb{F}_0\big)^{[N_1\alpha/2]}=\int\exp\{itV_N\}\big(d\mathbb{F}_0\big)^{N_2}\big(d\mathbb{F}_0\big)^{N_2}\big(d\mathbb{F}_0\big)^{[N_1\alpha/2]-2N_2} \tag{4.6}$$

In the first group of $d\mathbb{F}_0$ in the right-hand side of (4.10) we decompose each $\mathbb{F}_0$ as follows: $\mathbb{F}_0=\beta\mathbb{F}_2+(1-\beta)$(“some distribution function”), and in the second group of $d\mathbb{F}_0$ in the right-hand side of (4.10) we decompose each $\mathbb{F}_0$ as follows: $\mathbb{F}_0=\gamma\mathbb{F}_3+(1-\gamma)$(“some distribution function”). Now going along the lines of Step 2, we get from (4.10) that in order to prove (4.9) it is enough to prove the bound

$$I\triangleq\mathbb{E}\Big|\int\exp\{itV_N\}\big(d\mathbb{F}_2\big)^{[N_2\beta/2]}\big(d\mathbb{F}_3\big)^{[N_2\gamma/2]}\Big|\le cN^{-L}.$$

Let $\mathbf{Z},\mathbf{Z}_1,\mathbf{Z}_2,...$ and $\mathbf{W},\mathbf{W}_1,\mathbf{W}_2,...$ be independent random vectors having the distribution functions $\mathbb{F}_2$ and $\mathbb{F}_3$, respectively. Using the symmetrization technique of the proof of Part 1 of Lemma (3.37) by Götze (1979), we get

$$I\le J\triangleq\mathbb{E}\Big|\mathbb{E}^{\oplus}\exp\Big\{i\frac{2t}{N}\sum_{j=1}^{[N_2\beta/2]}\sum_{k=1}^{[N_2\gamma/2]}\big\{H(\mathbf{Z}_j^*,\mathbf{W}_k)-H(\mathbf{Z}_j^{**},\mathbf{W}_k)\big\}\Big\}\Big|,$$

where $\mathbf{Z}_j^*$ and $\mathbf{Z}_j^{**}$ are independent and have the same distribution function $\mathbb{F}_2$. The $\mathbb{E}^{\oplus}$ denotes the conditional expectation with respect to the random vectors $\mathbf{W}_k$, $k=1,..,[N_2\gamma/2]$. Therefore,

$$J=\mathbb{E}\Big|\mathbb{E}^{\oplus}\exp\Big\{i\frac{2t}{N}\sum_{j=1}^{[N_2\beta/2]}\big\{H(\mathbf{Z}_j^*,\mathbf{W})-H(\mathbf{Z}_j^{**},\mathbf{W})\big\}\Big\}\Big|^{[N_2\gamma/2]},$$

where $\mathbb{E}^{\oplus}$ *this time* denotes the conditional expectation with respect to the random vector $\mathbf{W}$. Thus in order to complete the proof of the theorem we need to show

$$J \le cN^{-L}. \tag{4.11}$$

Step 4. (*Proof of the bound* (4.11).) Let us rewrite the quantity J in a more convenient way for further calculations. Denote $M \triangleq [N_2\beta/2]$, $\tau \triangleq 2t\sqrt{M}/N$, and let

$$\Delta_1 \triangleq \frac{1}{\gamma}\Big|\int_{\mathbf{C}} \exp\{i\tau S_M(\mathbf{x})\}p_0(\mathbf{x})d\mathbf{x}\Big|,$$

where p_0 denotes the density of $\mathbb{F}_0$. Then $J = \mathbb{E}\Delta_1^{[N_2\gamma/2]}$.

Fix an elementary event ω. Because of (1.8)–(1.10), we may assume without loss of generality that there exists a direction $k \in \{1,...,d\}$ and a point $\mathbf{x}_0$ such that

$$\sup_{\mathbf{x}\in\mathbf{C}}\Big|\big\langle \mathcal{D}S_M(\mathbf{x}), \mathbf{e}_k\big\rangle\Big| \le M^{\lambda}, \tag{4.12}$$

$$\Big|\big\langle \mathcal{D}S_M(\mathbf{x}_0), \mathbf{e}_k\big\rangle\Big| \ge 1/(dM^{\kappa}), \tag{4.13}$$

$$\sup_{\mathbf{x},\mathbf{y}\in\mathbf{C}_r}\Big|\big\langle \mathcal{D}S_M(\mathbf{x}) - \mathcal{D}S_M(\mathbf{y}), \mathbf{e}_k\big\rangle\Big| \le 1/(2dM^{\kappa}), \tag{4.14}$$

where $\mathbf{C}_r \in \{\mathbf{C}_1, ..., \mathbf{C}_{[M^{\nu}]}\}$ is such that $\mathbf{x} \in \mathbf{C}_r$. Rewrite $\mathbf{C}_r$ as the Cartesian product $I_1 \times \cdots \times I_d$ of intervals. Then we have with dx_k^{ν} denoting integration with respect to the variables different form x_k

$$\begin{aligned}\Delta_1 &\le \frac{1}{\gamma}\int_{\mathbf{C}\setminus\mathbf{C}_r} p_0(\mathbf{x})d\mathbf{x} + \frac{1}{\gamma}\Big|\int_{\mathbf{C}_r} \exp\{i\tau S_M(\mathbf{x})\}p_0(\mathbf{x})d\mathbf{x}\Big| \\ &\le \frac{1}{\gamma}\int_{\mathbf{C}\setminus\mathbf{C}_r} p_0(\mathbf{x})d\mathbf{x} + \frac{1}{\gamma}\int\Big|\int_{I_k} \exp\{i\tau S_M(\mathbf{x})\}p_0(\mathbf{x})dx_k\Big|d\mathbf{x}_k^{\nu}.\end{aligned} \tag{4.15}$$

To estimate the quantity $|\int_{I_k} ...dx_k|$ we are going to use Lemma 4.1. For this, fix $x_1,..., x_{k-1}, x_{k+1},..., x_d$, and let $g(x_k) \triangleq S_M(\mathbf{x})$, $h(x_k) \triangleq p_0(\mathbf{x})$. Also, let a and b be numbers such that $I_k = (a,b)$. For every number $\mathbf{x} \in \mathbf{C}_r$ we have $g'(x_k) = \langle \mathcal{D}S_M(\mathbf{x}), \mathbf{e}_k\rangle$ which implies

$$\big|g'(x_k)\big| \ge \Big|\big\langle \mathcal{D}S_M(\mathbf{x}_0), \mathbf{e}_k\big\rangle\Big| - \Big|\big\langle \mathcal{D}S_M(\mathbf{x}) - \mathcal{D}S_M(\mathbf{x}_0), \mathbf{e}_k\big\rangle\Big|. \tag{4.16}$$

Using (4.13) and (4.14) in the right-hand side of (4.16), we get that $|g'(x_k)| \ge 1/(2dM^{\kappa})$. Therefore, we may choose

$$\Phi = 1/(2dM^{\kappa}).$$

The same proof shows that for all $\mathbf{x} \in \mathbf{C}_r$ the sign of $g'(x_k)$ is the same (and equals to $sgn\langle \mathcal{D}S_M(\mathbf{x}_0), \mathbf{e}_k\rangle)$. For a number Ψ, as it is easy to see from (4.12), we may use

$$\Psi = M^{\lambda}.$$

Since $\mathbb{P}\{\mathbf{Y} \in \mathbf{C}\} > 0$, then there exists a cub $\mathbf{C}_1 \subset \mathbf{C}$ such that for some $c_1 > 0$ we have $p_0(\mathbf{x}) > c_1$ for all $\mathbf{x} \in \mathbf{C}_1$. Thus (if necessary, replace the set $\mathbf{C}$ by $\mathbf{C}_1$) we get that for a number Υ we may use

$$\Upsilon = c_1.$$

Thus, using Lemma 4.1 we get that

$$\left| \int_{I_k} \exp\big\{i\tau S_M(\mathbf{x})\big\} p_0(\mathbf{x}) dx_k \right| \leq \int_{I_k} p_0(\mathbf{x}) dx_k - cM^{-\lambda-3\kappa} \min\{\tau_0^2, 1\}(b-a)^3.$$

Hence because of the bound (4.15), we have

$$\Delta_1 \leq 1 - cVol(\mathbf{C}_r) M^{-\lambda-3\kappa} \min\{\tau_0^2, 1\}(b-a)^2 \tag{4.17}$$

If we now use the bound $1 - x \leq \exp\{-x\}$ in the right-hand side of (4.17), then we obtain

$$J \leq \exp\big\{cM^{1-10\varepsilon} \min\{\tau_0^2, 1\}\big\}.$$

Take now $\tau_0^2 = M^{-1+10\varepsilon}(\log M)^2$ and note that the number $\varepsilon > 0$ may be taken as small as we want, say $\varepsilon \leq 10^{-10}$. Then the bound (4.11) follows immediately. This completes the proof. $\square$

5. Proof of the statements

Let us first prove Statement 1.3. To do that we give a special case of Theorem 1.3.

Corollary 5.1. *Let $m \in \mathbb{N} \cup \{0\}$ and*

$$\mathbb{E}\big|H(U_1, U_1)\big|^m + \mathbb{E}\big|H(U_1, U_2)\big|^m < +\infty. \tag{5.1}$$

Assume that there exists a non-empty rectangle $(\alpha, \beta) \times (\gamma, \delta) \subset (0,1) \times (0,1)$ such that for all $x \in (\alpha, \beta)$ the function $y \mapsto H(x, y)$ is differentiable on the interval (γ, δ). Denote $S_N \triangleq (T_1 + ... + T_N)/\sqrt{N}$, where $T_1, ..., T_N$ are independent copies of the random function $T \triangleq \widetilde{H}(\alpha + (\beta - \alpha)U, \cdot)$. Furthermore, assume that for *every $\varepsilon > 0$ one may find numbers $\lambda \in \mathbb{R}$, $\kappa > 0$, and $\nu > 0$ such that $\lambda,\ \kappa,\ \nu \leq \varepsilon$ and such that for every $D \geq 0$ the following three conditions hold:*

$$\sup_{N\in\mathbb{N}} N^D \mathbb{P}\big\{ \sup_{x\in(\gamma,\delta)} |S_N'(x)| \geq N^{\lambda} \big\} < +\infty; \tag{5.2}$$

$$\sup_{N\in\mathbb{N}} N^D \mathbb{P}\big\{ \| S_N' \|_{L_2(\gamma,\delta)} \leq 1/N^{\kappa} \big\} < +\infty; \tag{5.3}$$

$$\sup_{N\in\mathbb{N}} N^D \mathbb{P}\big\{ \sup_{j} \sup_{x,y\in I_j} |S_N'(x) - S_N'(y)| \geq 1/(2N^{\kappa}) \big\} < +\infty, \tag{5.4}$$

where I_j, $j = 1, ..., [N^{\nu}]$, denotes the partition of the interval (γ,δ) into subintervals $I_j \subset (\gamma,\delta)$ such that $Vol(I_j) = (\delta-\gamma)/[N^{\nu}]$. Then, for every $\varepsilon > 0$ and $L \geq 0$,

$$\sup_{t:|t|\geq N^{\varepsilon}} \big|(d/dt)^m \mathbb{E}\exp\{itV_N\}\big| = O\big(N^{-L}\big), \quad N \to +\infty.$$

Proof. The corollary is an easy consequence of Theorem 1.3. Let us note only that condition (1.9) follows from condition (5.3) using the inequality

$$\| S_N' \|_{L_2(\gamma,\delta)} \leq \| S_N' \|_{L_\infty(\gamma,\delta)} \sqrt{\delta-\gamma}. \tag{5.5}$$

□

Also, to prove Statement 1.3 we need the following lemma which is a simple consequence of Corollary 1.2 by Bentkus (1986).

Lemma 5.1. *Let $\mathcal{G}$ be a centered Gaussian $L_2(\gamma,\delta)$-valued random variable which covariance is the Hilbert-Schmidt operator $L_2(\gamma,\delta) \to L_2(\gamma,\delta)$ corresponding to the kernel*

$$(x,y) \mapsto \mathbb{E}\Big\{ \frac{d}{dx} H(\alpha + (\beta-\gamma)U, x) \frac{d}{dy} H(\alpha + (\beta-\gamma)U, y) \Big\} - \mathbb{E}\Big\{ \frac{d}{dx} H(\alpha + (\beta-\gamma)U, x) \Big\} \mathbb{E}\Big\{ \frac{d}{dy} H(\alpha + (\beta-\gamma)U, y) \Big\}.$$

If, for some $l > 2$,

$$\mathbb{E}\Big\{ \int_{\gamma}^{\delta} \Big| \frac{d}{dx} H\big(\alpha + (\beta-\alpha)U, x\big) \Big|^2 dx \Big\}^{l/2} < +\infty, \tag{5.6}$$

and the random variable $\mathcal{G}$ is not concentrated in any finite dimensional subspace of $L_2(\gamma,\delta)$, then condition (5.3) is satisfied.

Proof of Statement 1.3. i) This is an immediate consequence of Theorem 1.1.

ii) Let us verify the conditions of Corollary 5.1 with the function H_3 when $d = 1$, $\alpha = 0$, and $\beta = 1$. The equivalence of moment conditions (5.1) and (1.14) is easy to prove. Thus let us verify (5.2)-(5.4). Note first that

$$S'_N(x) = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} \frac{d}{dx} \widetilde{H}_3(U_j, x) = -\widetilde{\mathcal{E}}_N(x) q(x) \tag{5.7}$$

where $\mathcal{E}_N$ denotes the uniform empirical process. Let us look at the condition (5.2). It is clear that without loss of generality we may assume

$$\sup\{q(x) : x \in (\gamma, \delta)\} < +\infty. \tag{5.8}$$

Thus, because of (5.8) and Lemma 2.3 by Stute (1982), we obtain

$$\mathbb{P}\big\{ \sup_{x \in (\gamma,\delta)} |\widetilde{\mathcal{E}}_N(x) q(x)| \geq N^{\lambda} \big\} \leq \mathbb{P}\big\{ c \sup_{x \in (\gamma,\delta)} |\widetilde{\mathcal{E}}_N(x)| \geq N^{\lambda} \big\} \leq c N^{-D} \tag{5.9}$$

which completes the verification of (5.2).

Condition (5.3) is satisfied because of Lemma 2.1 and (5.1); compare the discussion concerning the infinite-dimensionality of the weighted Brownian bridge given just after Theorem 1.4 in the paper by Bentkus, Götze, Paulauskas and Račkauskas (1990).

Let us now show that (1.16) implies (5.4). Denote

$$\Delta \triangleq \mathbb{P}\big\{ \sup_j \sup_{x,y \in I_j} \big| \widetilde{\mathcal{E}}_N(x) q(x) - \widetilde{\mathcal{E}}_N(y) q(y) \big| \geq 1/(2N^{\kappa}) \big\}$$

Assumptions (1.16) and (5.8) imply that, for some constant $c_2 > 0$,

$$\Delta \leq \mathbb{P}\big\{ \sup_j \sup_{x,y \in I_j} |\widetilde{\mathcal{E}}_N(x) - \widetilde{\mathcal{E}}_N(y)| + \sup_j \sup_{x,y \in I_j} |\widetilde{\mathcal{E}}_N(x)| N^{-\nu\tau} \geq c_2 N^{-\kappa} \big\}. \tag{5.10}$$

Using bound (5.9) with $q(x) = 1$ in the right-hand side of (5.10), we get that, for a constant $c > 0$,

$$\begin{aligned} \Delta &\leq \mathbb{P}\big\{ \sup_j \sup_{x,y \in I_j} |\widetilde{\mathcal{E}}_N(x) - \widetilde{\mathcal{E}}_N(y)| + N^{\lambda - \nu\tau} \geq c N^{-\kappa} \big\} + c N^{-D} \\ &\leq \mathbb{P}\big\{ \sup_j \sup_{x,y \in I_j} |\widetilde{\mathcal{E}}(x) - \widetilde{\mathcal{E}}(y)| \geq c N^{-\kappa} \big\} + c N^{-D} \end{aligned} \tag{5.11}$$

if $\kappa < \nu\tau - \lambda$. Using Lemma 2.4 by Stute (1982), or Inequality 3.2 by Shorack and Wellner (1982), it follows that for small numbers $\nu > 0$ and $\kappa > 0$ the right-hand side of (5.11) does not exceed cN^{-D}. Let us note also that the numbers $\lambda > 0$,

$\nu > 0$ and $\kappa > 0$ could be chosen arbitrary small (but satisfying the condition $\kappa < \nu\tau - \lambda$).

iii) The result is a direct consequence of part ii) of this statement; see Section 3 (or Section 2) in Bentkus, Götze and Zitikis (1993) for more details. $\square$

Let us now prove Statement 1.2. For this we formulate another special case of Theorem 1.3.

Corollary 5.2. *Assume that*

$$\sup_{x,y\in(0,1)} \left| \frac{d}{dy} H(x,y) \right| < +\infty. \tag{5.12}$$

Denote $S_N \triangleq (T_1 + ... + T_N)/\sqrt{N}$, *where* $T_1, ..., T_N$ *are independent copies of the random function* $T \triangleq \widetilde{H}(U, \cdot\,)$. *Furthermore, assume that for every* $\varepsilon > 0$ *one may find numbers* $\kappa > 0$ *and* $\nu > 0$ *such that* $\kappa, \nu \le \varepsilon$ *and such that for every* $D \ge 0$ *the following condition holds:*

$$\sup_{N\in\mathbb{N}} N^D \mathbb{P}\big\{ \sup_j \sup_{x,y\in I_j} |S_N'(x) - S_N'(y)| \ge 1/(2N^\kappa) \big\} < +\infty, \tag{5.13}$$

where I_j, $j = 1, ..., [N^\nu]$, *denotes the partition of the interval* $(0,1)$ *into subintervals* $I_j \subset (0,1)$ *such that* $Vol(I_j) = 1/[N^\nu]$. *Furthermore, let* $\mathcal{G}$ *be a centered Gaussian* $L_2(0,1)$*-valued random variable which covariance is the Hilbert-Schmidt operator* $L_2(0,1) \to L_2(0,1)$ *corresponding to the kernel*

$$(x,y) \mapsto \mathbb{E}\Big\{ \frac{d}{dx} H(U,x) \frac{d}{dy} H(U,y) \Big\} - \mathbb{E}\Big\{ \frac{d}{dx} H(U,x) \Big\} \mathbb{E}\Big\{ \frac{d}{dy} H(U,y) \Big\}.$$

Assume that the random element $\mathcal{G}$ *is not concentrated in any finite dimensional subspace of* $L_2(0,1)$. *Then for every* $m \in \mathbb{N} \cup \{0\}$, $\varepsilon > 0$ *and* $L \ge 0$,

$$\sup_{t:|t|\ge N^\varepsilon} \big| (d/dt)^m \mathbb{E}\exp\{itV_N\} \big| = O\big(N^{-L}\big), \quad N \to +\infty.$$

Proof. We shall show that Corollary 5.1 implies the result. Take $\alpha = \gamma = 0$ and $\beta = \delta = 1$. Then, for every $j = 1, ..., [N^\nu]$, choose a non-random point $y_j \in I_j$. Using Slutzky's arguments, we get

$$\begin{aligned} \Delta \triangleq \mathbb{P}\big\{ \sup_{x\in(0,1)} \big|S_N'(x)\big| \ge N^\lambda \big\} &= \mathbb{P}\big\{ \sup_j \sup_{x\in I_j} \big|S_N'(x)\big| \ge N^\lambda \big\} \\ &\le \mathbb{P}\big\{ \sup_j \sup_{x,y\in I_j} \big|S_N'(x) - S_N'(y)\big| \ge 1/(2N^\kappa) \big\} \\ &\qquad + \mathbb{P}\big\{ \sup_j \big|S_N'(y_j)\big| + 1/(2N^\kappa) \ge N^\lambda \big\}. \end{aligned} \tag{5.14}$$

Because of (5.13), the first summand in the right-hand side of inequality (5.14) does not exceed cN^{-D}. Therefore, $\Delta \le cN^{-D}$ follows if

$$\mathbb{P}\big\{ \sup_j \big| S_N'(y_j) \big| \ge N^\lambda/2 \big\} \le cN^{-D}. \tag{5.15}$$

But the bound (5.15) (and (5.2) as well) are consequences of Markov's inequality and the bound $\mathbb{E}|S_N(y_k)|^D \le c$ which holds, because of the assumption (5.13), for all $k = 1, ..., [N^\nu]$ and all $D \ge 0$, where the constant $c \ge 0$ does not depend on N. Furthermore, because of Lemma 5.1 the assumption (5.3) is satisfied as well. This remark completes the proof of the theorem. $\square$

Proof of Statement 1.2. i) This is an easy consequence of Theorem 1.1.

ii) The proof is almost the same as that of Statement 1.3; use Corollary 5.2 instead of Corollary 5.1.

iii) This a consequence of part ii); use results from Section 3 by Bentkus, Götze and Zitikis (1993). $\square$

Finally as a simple consequence of Theorem 1.3 we have

Corollary 5.3. *Assume that*

$$\sup_{x,y\in(0,1)} \Big| \Big(\frac{d}{dy} \Big)^2 H(x,y) \Big| < +\infty. \tag{5.16}$$

Furthermore, let $\mathcal{G}$ be a centered Gaussian $L_2(0,1)$-valued random variable with covariance being the Hilbert-Schmidt operator $L_2(0,1) \to L_2(0,1)$ corresponding to the kernel

$$(x,y) \mapsto \mathbb{E}\Big\{ \frac{d}{dx} H(U,x) \frac{d}{dy} H(U,y) \Big\} - \mathbb{E}\Big\{ \frac{d}{dx} H(U,x) \Big\} \mathbb{E}\Big\{ \frac{d}{dy} H(U,y) \Big\}. \tag{5.17}$$

Assume that the the random element $\mathcal{G}$ is not concentrated in any finite dimensional subspace of $L_2(0,1)$. Then for every $m \in \mathbb{N} \cup \{0\}$, $\varepsilon > 0$ and $L \ge 0$,

$$\sup_{t:|t|\ge N^\varepsilon} \big| (d/dt)^m \mathbb{E}\exp\{itV_N\} \big| = O\big(N^{-L}\big), \quad N \to +\infty.$$

Proof. This corollary follows from Corollary 5.2. Note that assumption (5.12) holds because of (5.16). Furthermore, using the bound

$$\sup_{N\in\mathbb{N}} N^D \sup_{x,y\in I_j} \big| S_N'(x) - S_N'(y) \big| \le \|S_N''(\cdot)\|_{L_2(I_j)} / \sqrt{[N^\nu]},$$

the left-hand side of (5.13) does not exceed

$$\mathbb{P}\big\{ \|S_N''(\cdot)\|_{L_2(I_j)} \ge \sqrt{[N^\nu]}/(2N^\kappa) \big\}. \tag{5.18}$$

If $\nu > 2\kappa$, then for every $D \geq 0$ the quantity (5.18) does not exceed cN^{-D} because of the Markov inequality and (5.16). This completes the proof of the theorem. □

Proof of Statement 1.1. i) This is an easy consequence of Theorem 1.1.

ii) Because of $\sum_{l=0}^{\infty} b_l^2 l^2 < +\infty$ we have

$$\frac{d^2}{dx\,dy} H_1(x,y) = -2(2\pi)^2 \sum_{l=1}^{\infty} b_l^2 l^2 \cos 2\pi l(x-y) \tag{5.19}$$

for all $x, y \in (0,1)$, which shows that condition (5.16) holds. The assumption on infinite-dimensionality of the corresponding Gaussian random variable is satisfied because there is an infinite number of non-zero coefficients $b_0, b_1, \dots$. Thus, Corollary 5.3 implies the desired result.

iii) Because of (5.19) we have

$$B_N^2 \triangleq \frac{1}{N} \sum_{j=1}^{N} \sum_{k=1}^{N} H(U_j, U_k) = \int_0^1 \int_0^1 \mathcal{E}_N(x)\mathcal{E}_N(y) \frac{d^2}{dx\,dy} H(x,y)\,dx\,dy,$$

which shows that $B_N^2 = \pi_2(\mathcal{E}_N, \mathcal{E}_N)$, where π_2 is a polynomial of degree 2 in the Hilbert space $L_2(0,1)$. Therefore, results by Bentkus, Götze and Zitikis (1993) might be used to get the theorem proved. (In the case $m = 0$ one may use general results for von Mises statistics by Götze (1979, 1984).) □

Acknowledgement. The authors would like to thank V.Bentkus for interesting discussions and for several suggestions which led to simplifications of some of the proofs.

References

Anderson T.W. and Darling D.A., *Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes*, Ann. Math. Statistics **23** (1952), 193–212.

Athreya, K. B., *Strong law for the bootstrap*, Statist. Probab. Lett. **1** (1983), 147-150.

Athreya, K. B., Ghosh, M., Low, L. Y. and Sen, P. K., *Laws of large numbers for bootstrapped U-statistics*, J. Statist. Plann. Inference **9** (1984), 185–194.

Baringhaus, L., *Testing for spherical symmetry of a multivariate distribution*, Ann. Statist. **19** (1991), 899–917.

Bentkus, V., *On concentration functions of sums of independent random elements with values in a Banach space*, Litovsk. Mat. Sb **14** (1986), 32–39.

Bentkus V., Götze F., Paulauskas V. and Račkauskas A., *The accuracy of Gaussian approximation in Banach spaces*, SFB 343 "Diskrete Structuren in der Mathematik," Preprint 90 -100, Universität Bielefeld to appear in Encyclopedia of Mathematical Sciences, Springer), 1990.

Bentkus, V., Götze, F. and Zitikis, R., *Asymptotic expansions in the integral and local limit theorems in Banach spaces with applications to ω-statistics*, to appear in J. Theor. Probab. (1993).

Beran, R. J., *Testing for uniformity on a compact homogeneous space*, J. Appl. Prob. **5** (1968), 177–195.

Beran, R. J., *Asymptotic theory of a class of tests for uniformity of a circular distribution*, Ann. Math. Statist. **40** (1969), 1196–1206.

Beran, R., *Prepivoting to reduce level error of confidence sets*, Biometrika **74** (1987), 457–468.

Bernštein, S. N., *On modification of Čebyšev's inequality and an error in Laplace's formula*, Učen. Zap. Naučno–Issled. Kafedr. Ukrain. Otd. Math. **1** (1924), 38–49 (reprinted in Collected works Vol. IV, Nauka, Moscow (1964) p. 71–79; in Russian).

Bhattacharya R.N. and Ranga Rao R., *Normal Approximation and Asymptotic Expansions*, Wiley, New York, 1976.

Bickel, P.J. and Freedman, D., *Some asymptotic theory for the bootstrap*, Ann. Statist. **9** (1981), 1196–1217.

Bickel, P.J., Götze, F. and van Zwet, W.R., *The Edgeworth expansion for U–statistics of degree two*, Ann. Statist. **14** (1986), 1463–1484.

Bikelis A., *Remainder terms in asymptotic expansions for characteristic functions and their derivatives*, Selected Transl. Math. Stat. Prob. **11** (1973), 149–162.

Bikelis A., *Asymptotic expansions for the densities and distributions of sums of independent identically distributed random vectors*, Selected Transl. Math. Stat. Prob. **13** (1973), 213–234.

Bretagnolle, J., *Lois limites du Bootstrap de certaines functionnelles*, Ann. Inst. Henri Poincaré, Section B, Calcul de Probabilités et Statistique **XIX** (1983), 281–296.

Csörgő S. and Stachó L., *A step toward asymptotic expansions for the Cramér–von Mises statistic*, Analytic function methods in probability theory. North–Holland (1980), 45–67.

Dynkin, E. B. and Mandelbaum A., *Symmetric staistics, Poisson point processes, and multiple Wiener integrals*, Ann. Statist. **31** (1983), 739–745.

Efron, B., *Bootstrap methods: Another look at the jackknife*, Ann. Statist. **7** (1979), 1–26.

Efron, B., *The Jackknife, the Bootstrap and Resampling Plans*, SIAM, Philandelphia, 1982.

Filippova A.A., *Mises theorem of the asymptotic behavior of functionals of empirical distribution function and its statistical applications*, Theor. Prob. and Appl. **7** (1962), 26–60.

Giné, E. M., *Invariant tests for uniformity on compact Riemannian manifolds based on Sobolev norms*, Ann. Statist. **3** (1975), 1243–1266.

Götze, F., *Asymptotic expansions for bivariate von Mises functionals*, Z. Wahrscheinlichkeitstheorie verw. Gebiete **50** (1979), 333–355.

Götze F., *On Edgeworth expansions in Banach spaces*, Ann. Prob. **9** (1981), 852–859.

Götze, F., *Expansions for von Mises functionals*, Z. Wahrscheinlichkeitstheorie verw. Gebiete **65** (1984), 599–625.

Götze F., *Asymptotic expansion in functional limit theorem*, J. Multivar. Anal. **16** (1985), 1–20.

Götze F., *On the rate of convergence in the central limit theorem in Banach spaces*, Ann. Prob. **14** (1986), 922–942.

Götze F., *Edgeworth expansions in functional limit theorems*, Ann. Prob. **17** (1989), 1602–1634.

Götze F. and Hipp, C., *On the validity of Edgeworth–expansions*, SFB 343 "Diskrete Strukturen in der Mathematik," Preprint 92–045, Universität Bielefeld, 1992 (to appearAnnals of Statistics).

Gregory, G. G., *Large sample theory for U-statistics and tests of fit*, Ann. Statist. **5** (1977), 110–123.

Inglot, T., Kallenberg, W.C.M. and Ledwina, T., *Asymptotic behaviour of some bilinear functionals of the empirical process*, Memorandum of Universiteit Twente **1031** (1992), 1–27.

Hall, P., *On the bootstrap and confidence intervals*, Ann. Statist. **14** (1986), 1431–1452.

Hall, P., *The Bootstrap and Edgeworth Expansions*, Springer, New York, 1992.

Helmers R., *Edgeworth expansions for linear combinations of order statistics*, Math. Center Tracts vol. 103, Amsterdam, 1982.

Helmers R., *A local limit theorem for L-statistics*, CWI Preprint (1991).

Helmers, R., *On the Edgeworth expansion and the bootstrap approximation for a studentized U-statistic*, Ann. Statist. **19** (1991), 470–484.

Helmers, R., Janssen P. and Serfling R., *Berry–Esseen and bootstrap results for generalized L-statistics*, Scand. J. Statist. **17** (1990), 65–77.

Hinkley, D. V., *Bootstrap methods. (With discussion)*, J. Roy. Statist. Soc. Ser B. **50** (1988), 321–337.

Koroliuk V.S. and Borovskih Yu.V., *Asymptotic Analysis of Distributions of Statistics*, Naukova Dumka, Kiev, 1984 (in Russian).

Mardia, K. V., *Statistics of Directional Data*, Academic, New York, 1972.

Martynov G.V., *Omega–square Criterion*, Nauka, Moscow, 1979 (in Russian).

von Mises, R., *On the asymptotic distribution of differentiable statistical functions*, Ann. Math. Statist. **18** (1947), 309–348.

Petrov V.V., *Sums of Independent Random Variables*, Springer–Verlag, Berlin Heidelberg New York, 1975.

Pettitt A.N. and Stephens M.A., *Modified Cramér–von Mises statistics for censored data*, Biometrika **63** (1976), 291–298.

Prentice, M. J., *On invariant tests of uniformity for directions and orientations*, Ann. Statist. **6** (1978), 169–176.

Rubin, H. and Vitale, R. A., *Asymptotic distribution of symmetric statistics*, Ann. Statist. **8** (1980), 165–170.

Sadikova, S. M., *Some inequalities for characteristic functions*, Prob. Theory and Appl. **11** (1966), 441–447.

Serfling R.J., *Approximation Theorems of Mathematical Statistics*, Wiley, New York, 1980.

Shorack, G.R. and Wellner J.A., *Limit theorems and inequalities for the uniform empirical process indexed by intervals*, Ann. Probab. **10** (1982), 639–652.

Shorack G.R. and Wellner J.A., *Empirical Processes with Applications to Statistics*, Wiley, New York, 1986.

Smirnov N.V., *Probability theory and mathematical statistic. Collected papers (In Russian)*, Nauka, Moscow, 1970.

Stute, W., *The oscillation behavior of empirical processes*, Ann. Probab. **10** (1982), 86–107.

Watson G.S., *Goodness–of–fit tests on a circle*, Biometrika **48** (1961), 109–114.

Yurinskii V.V. (Jurinskii V.V.), *Bounds for characteristic functions of certain degenerate multidimensional distributions*, Theory Prob. and Appl. **12** (1972), 101–113.

Zitikis R., *Asymptotic expansions in the local limit theorem for the generalized Cramér–von Mises–Smirnov statistics*, Ph. D. Thesis. Vilnius University, Vilnius, 1988.

Zitikis R., *Smoothness of distribution function of $\mathcal{FL}$–statistic. I*, Lithuanian Math. J. **30** (1990), 231–240.

van Zwet W. R., *Asymptotic expansions for the distribution functions of linear combinations of order statistics*, Statistical decision theory and related topics. II. Ed. by S.S. Guppta and D.S. Moore. Academic Press. New York, San Francisco, London (1977), 421–437.

Friedrich Götze
Fakultät für Mathematik
Universität Bielefeld
Postfach 8640
4800 Bielefeld 1
Germany
E-mail address: goetze@mathematik.uni-bielefeld.de

R.Zitikis
Institute of Mathematics and Informatics
Akademijos 4
Vilnius 2600
Lithuania
E-mail address: zitikis@mathematik.uni-bielefeld.de