U-Quantile Processes and Generalized Linear Statistics of Dependent Data

Martin Wendler*

Fakultät für Mathematik, Ruhr-Universität Bochum, 44780 Bochum, Germany

September 28, 2010

Keywords: L-Statistic; U-statistics; U-quantiles; Bahadur representation; mixing; near epoch dependence
62G30; 60G10; 60F17

Generalized linear statistics are a unifying class that contains U-statistics, U-quantiles, L-statistics as well as trimmed and winsorized U-statistics. For example, many commonly used estimators of scale fall into this class. GL-statistics only have been studied under independence; in this paper, we establish the central limit theorem (CLT) and the law of the iterated logarithm (LIL) for GL-statistics of sequences which are strongly mixing or L^1 near epoch dependent on an absolutely regular process. We first investigate the empirical U-process. With the help of a generalized Bahadur representation, the CLT and the LIL for the empirical U-quantile process follow. As GL-statistics are linear functionals of the U-quantile process, the CLT and the LIL for GL-statistics are straightforward corollaries.

1 Introduction

U-Statistics and the Empirical U-Process

In the whole paper, $(X_n)_{n\in\mathbb{N}}$ shall be a stationary, real valued sequence of random variables. A U-statistic $U_n(g)$ can be described as generalized mean, i.e. the mean of the values $g(X_i, X_j)$, $1 \leq i < j \leq n$, where g is a bivariate, symmetric and measurable kernel. The following to estimators of scale are U-statistics:

*Martin.Wendler@rub.de

Example 1.1. Consider $g(x,y) = \frac{1}{2}(x-y)^2$. A short calculation shows that the related U-statistic is the well known variance estimator

$$U_n(g) = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2.$$

Example 1.2. Let $g(x,y) = |x-y|$. Then the corresponding U-statistic is

$$U_n(g) = \frac{2}{n(n-1)}\sum_{1\leq i<j\leq n}|X_i - X_j|,$$

known as Gini's mean difference.

For U-statistics of independent random variables, the CLT dates back to Hoeffding [17] and was extended to absolutely regular sequences by Yoshihara [32], to near epoch dependent sequences on absolutely regular processes by Denker and Keller [15] and to strongly mixing random variables by Dehling and Wendler [13]. The LIL under independence was proved by Serfling [27] and by Dehling and Wendler [14] under strong mixing and near epoch dependence on absolutely regular processes.

Not only U-statistics with fixed kernel g are of interest, but also the empirical U-distribution function $(U_n(t))_{t\in\mathbb{R}}$, which is for fixed t a U-statistic with kernel $h(x,y,t) := \mathbb{1}_{\{g(x,y)\leq t\}}$. The Grassberger-Procaccia and the Takens estimator of the correlation dimension in a dynamical system are based on the empirical U-distribution function, see Borovkova, Burton, Dehling [9]. The functional CLT for the empirical U-distribution function has been established by Arcones and Giné [3] for independent data, by Arcones and Yu for absolutely regular data [5], and by Borovkova, Burton and Dehling [9] for data, which is near epoch dependent on absolutely regular processes. The functional LIL for the empirical U-distribution function has been proved by Arcones [1], Arcones and Giné [4] under independence. We will extend the LIL to sequences which are strongly mixing or L^1 near epoch dependent on an absolutely regular process and give a CLT under conditions which are slightly different from the conditions in Borovkova et al. [9]. Let us now give precise definitions:

Definition 1.3. *We call a measurable function $h: \mathbb{R}\times\mathbb{R}\times\mathbb{R}\to\mathbb{R}$, which is symmetric in the first two arguments a kernel function. For fixed $t\in\mathbb{R}$, we call*

$$U_n(t) := \frac{2}{n(n-1)}\sum_{1\leq i<j\leq n} h(X_i, X_j, t)$$

the U-statistic with kernel $h(\cdot,\cdot,t)$ and the process $(U_n(t))_{t\in\mathbb{R}}$ the empirical U-distribution function. We define the U-distribution function as $U(t) := E[h(X,Y,t)]$, where X, Y are independent with the same distribution as X_1, and the empirical U-process as $\left(\sqrt{n}\,(U_n(t) - U(t))\right)_{t\in\mathbb{R}}$.

The main tool for the investigation of U-statistics is the Hoeffding decomposition into a linear and a so-called degenerate part:

$$U_n(t) = U(t) + \frac{2}{n}\sum_{i=1}^{n} h_1(X_i, t) + \frac{2}{n(n-1)}\sum_{1\leq i<j\leq n} h_2(X_i, X_j, t)$$

where

$$h_1(x,t) := Eh(x,Y,t) - U(t)$$
$$h_2(x,y,t) := h(x,y,t) - h_1(x,t) - h_1(y,t) - U(t).$$

We need some technical assumptions to guarantee the convergence of the empirical U-process:

Assumption 1. *The kernel function h is bounded and nondecreasing in the third argument. The U-distribution function U is Lipschitz-continuous.*

Furthermore, we will consider dependent random variables, so we need an additional continuity property of the kernel function (which was introduced by Denker and Keller [15]):

Assumption 2. *h satisfies the uniform variation condition, that means there is a constant L, such that for all $t \in \mathbb{R}$, $\epsilon > 0$*

$$E\left[\sup_{\|(x,y)-(X,Y)\|\leq\epsilon,\ \|(x',y')-(X,Y)\|\leq\epsilon} |h(x,y,t) - h(x',y',t)|\right] \leq L\epsilon,$$

where X, Y are independent with the same distribution as X_1 and $\|\cdot\|$ denotes the Eucledean norm.

Empirical U-Quantiles and GL-Statistics

For $p \in (0,1)$, the $p-th$ U-quantile $t_p = U^{-1}(p)$ is the inverse of the U-distribution function at point p (where one needs additional conditions to ensure that t_p is uniquely determined). A natural estimator of a U-quantile is the empirical U-quantile $U_n^{-1}(p)$, which is the $p-th$ quantile of the empirical U-distribution function:

Definition 1.4. *Let be $p \in (0,1)$ and U_n the empirical U-distribution function.*

$$U_n^{-1}(p) := \inf\left\{t \middle| U_n(t) \geq p\right\}$$

is called the empirical U-quantile.

Empirical U-quantiles have application in robust statistics.

Example 1.5. Let be $h(x,y,t) := \mathbb{1}_{\{|x-y|\leq t\}}$. Then the 0.25-$U$-quantile is the Q_n estimator of scale proposed by Rousseeuw and Croux [26], which is highly robust, as its breakdown point is 50%.

The kernel function $h(x,y,t) := \mathbb{1}_{\{|x-y|\leq t\}}$ satisfies Assumption 2 (uniform varition condition), if Assumption 1 (Lipschitz continuity of U) holds. For every $\epsilon > 0$

$$E\left[\sup_{\|(x,y)-(X,Y)\|\leq\epsilon,\ \|(x',y')-(X,Y)\|\leq\epsilon} \left|\mathbb{1}_{\{|x-y|\leq t\}} - \mathbb{1}_{\{|x'-y'|\leq t\}}\right|\right]$$
$$\leq P\left[t - \sqrt{2}\epsilon < |X-Y| \leq t + \sqrt{2}\epsilon\right] \leq U(t+\sqrt{2}\epsilon) - U(t-\sqrt{2}\epsilon) \leq C\epsilon.$$

The empirical U-quantile and the empirical U-distribution function have a converse behaviour: $U_n^{-1}(p)$ is greater than t_p iff $U_n(t_p)$ is smaller than p. This motivates a generalized Bahadur representation [7]:

$$U_n^{-1}(p) = t_p + \frac{p - U_n(t_p)}{u(t_p)} + R_n(p). \tag{1}$$

where $u = U'$ is the derivative of the U-distribution function. For independent data and fixed p, Geertsema [16] established a generalized Bahadur representation with $R_n(p) = O\left(n^{-\frac{3}{4}} \log n\right)$ a.s.. Dehling, Denker, Philipp [12] and Choudhury and Serfling [11] improved the rate to $R_n(p) = O\left(n^{-\frac{3}{4}} (\log n)^{\frac{3}{4}}\right)$. Arcones [2] proved the exact order $R_n(p) = O\left(n^{-\frac{3}{4}} (\log\log n)^{\frac{3}{4}}\right)$ as for sample quantiles. Under strong mixing and near epoch dependence on an absolutely regular processes, we recently established rates of convergence for $R_n(p)$ which depend on the decrease of the mixing coefficients [30]. The CLT and the LIL for $U_n^{-1}(p)$ are straightfoward corollaries of the convergence of R_n and the corresponding theorems for $U_n(t_p)$.

In this paper, we will study not a single U-quantile, but the empirical U-quantile process $(U_n^{-1}(p))_{p\in I}$ under dependence, where the interval I is given by $I = [\tilde{C}_1, \tilde{C}_2]$ with $U(C_1) < \tilde{C}_1 < \tilde{C}_2 < U(C_2)$ and the constants C_1, C_2 from Assumption 3 below. In order to do this, we will examine the rate of convergence of $\sup_{p\in I} R_n(p)$ and use the CLT and the LIL for the empirical U-process. As we devide by u in the Bahadur representation, we have to assume that this derivative behaves nicely. Furthermore, we need U to be a bit more than differentiable (but twice differentiable is not needed).

Assumption 3. *U differentiable on an interval $[C_1, C_2]$ with $0 < \inf_{t\in[C_1,C_2]} u(t) \leq \sup_{t\in[C_1,C_2]} u(t) < \infty$ ($u(t) = U'(t)$) and*

$$\sup_{s,t\in[C_1,C_2]:\ |t-s|\leq x} |U(t) - U(s) - u(t)(t-s)| = O\left(x^{\frac{5}{4}}\right).$$

The Bahadur representation for sample quantile process dates back to Kiefer [20] under independence, Babu and Singh [6] proved such an representation for mixing data and Kulik [21] and Wu [31] for linear processes, but there seem to be no such results for the U-quantile process.

Furthermore, we are interested in linear functionals of the U-quantile process.

Definition 1.6. *Let be $p_1, \ldots, p_d \in I$ and J a bounded function, that is continuous a.e. and vanishes outside of I. We call a statisic of the form*

$$\begin{aligned} T_n = T\left(U_n^{-1}\right) &:= \int_I J(p)\, U_n^{-1}(p) dp + \sum_{j=1}^{d} b_j U_n^{-1}(p_j) \\ &= \sum_{i=1}^{\frac{n(n-1)}{2}} \int_{\frac{2(i-1)}{n(n-1)}}^{\frac{2i}{n(n-1)}} J(t)\, dt \cdot U_n^{-1}\left(\frac{2i}{n(n-1)}\right) + \sum_{j=1}^{d} b_j U_n^{-1}(p_j) \end{aligned}$$

generalized linear statistic (GL-statistic).

This generalization of L-statistics was introduced by Serfling [28]. U-statistics, U-quantiles and L-statistics can be written as GL-statistics (though this might be somewhat artifically). For a U-statistics, just take $h(x,y,t) = \mathbb{1}_{\{g(x,y)\leq t\}}$ and $J = 1$ (this only works if we can consider the interval $I = [0,1]$). The following example shows how to deal with an ordinary L-statistic.

Example 1.7. Let be $h(x,y,t) := \frac{1}{2}\left(\mathbb{1}_{\{x\leq t\}} + \mathbb{1}_{\{y\leq t\}}\right)$, $p_1 = 0.25$, $p_1 = 0.75$, $b_1 = -1$, $b_2 = 1$, and $J = 0$. Then a short calculation shows that the related GL-statistic is

$$T_n = F_n^{-1}(0.75) - F_n^{-1}(0.25), \tag{2}$$

where F_n^{-1} denotes the empirical sample quantile function. This is the well known inter quartile distance, a robust estimator of scale with 25% breakdown point.

Example 1.8. Let be $h(x,y,t) := \mathbb{1}_{\left\{\frac{1}{2}(x-y)^2\leq t\right\}}$, $p_1 = 0.75$, $b_1 = 0.25$ and $J(x) = \mathbb{1}_{\{x\in[0,0.75]\}}$. The related GL-statistic is called winsorized variance, a robust estimator of scale with 13% breakdown point.

The uniform variation condition also holds in this case, as $h(x,y,t) = \mathbb{1}_{\left\{\frac{1}{2}(x-y)^2\leq t\right\}} = \mathbb{1}_{\left\{|x-y|\leq\sqrt{2t}\right\}}$ and this is the kernel function of Example 1.5.

Dependent Sequences of Random Variables

While the theory of GL-statistics under independence has been studied by Serfling [28], there seems to be no results under dependence. But many dependent random sequences are very common in applications. Strong mixing and near epoch dependence are widely used concepts to describe short range dependence.

Definition 1.9. *Let $(X_n)_{n\in\mathbb{N}}$ be a stationary process. Then the strong mixing coefficient is given by*

$$\alpha(k) = \sup\left\{|P(A\cap B) - P(A)P(B)| : A \in \mathcal{F}_1^n, B \in \mathcal{F}_{n+k}^\infty, n \in \mathbb{N}\right\},$$

where $\mathcal{F}_a^l$ is the σ-field generated by random variables $X_a, \ldots, X_l$., and $(X_n)_{n\in\mathbb{N}}$ is called strongly mixing, if $\alpha(k) \to 0$ as $k \to \infty$.

Strong mixing in the sense of α-mixing is the weakest of the well known strong mixing conditions, see Bradley [10]. But this class of weak dependent processes is too strong for many applications, as it excludes examples like linear processes with innovations that do not have a density or data from dynamical systems.

Example 1.10. Let $(Z_n)_{n\in\mathbb{N}}$ be independent r.v.'s with $P[Z_n = 1] = P[Z_n = 0] = \frac{1}{2}$, X_0 uniformly distributed on $[0,1]$, independent of $(Z_n)_{n\in\mathbb{N}}$ and

$$X_{n+1} = \frac{1}{2}X_n + \frac{1}{2}Z_{n+1}.$$

Then the stationary autoregressive process $(X_n)_{n\in\mathbb{N}}$ is not strongly mixing, as

$$\Bigg| P\left[X_1 \in \left[0, \frac{1}{2}\right],\ X_k \in \bigcup_{i=1}^{2^{(k-1)}} \left[(2i-2)2^{-k}, (2i-1)2^{-k}\right]\right] \\ - P\left[X_1 \in \left[0, \frac{1}{2}\right]\right] P\left[X_k \in \bigcup_{i=1}^{2^{(k-1)}} \left[(2i-2)2^{-k}, (2i-1)2^{-k}\right]\right]\Bigg| = \frac{1}{2} - \frac{1}{2}\cdot\frac{1}{2} = \frac{1}{4}.$$

We will consider sequences which are near epoch dependent on absolutely regular processes, as this class covers the example above and data from dynamical systems, which are deterministic except for the initial value. Let $T : [0,1] \to [0,1]$ be a piecewise smooth and expanding map such that $\inf_{x\in[0,1]} |T'(x)| > 1$. Then there is a stationary process $(X_n)_{n\in\mathbb{N}}$ such that $X_{n+1} = T(X_n)$ which can be represented as a functional of an absolutely regular process, for details see Hofbauer, Keller [18]. Linear processes (as in the example above) and GARCH processes are also near epoch dependent, see Pötscher, Prucha [25]. Near epoch dependent random variables are also called approximating functionals (for example in Borovkova et al. [9])

Definition 1.11. *Let $(X_n)_{n\in\mathbb{N}}$ be a stationary process.*

1. *The absolute regularity coefficient is given by*

$$\beta(k) = \sup_{n\in\mathbb{N}} E \sup\{\left|P(A|\mathcal{F}_{-\infty}^n) - P(A)\right| : A \in \mathcal{F}_{n+k}^\infty\},$$

   and $(X_n)_{n\in\mathbb{N}}$ is called absolutely regular, if $\beta(k) \to 0$ as $k \to \infty$.

2. *We say that $(X_n)_{n\in\mathbb{N}}$ is L^1 near epoch dependent on a process $(Z_n)_{n\in\mathbb{Z}}$ with approximation constants $(a_l)_{l\in\mathbb{N}}$, if*

$$E\left|X_1 - E(X_1|\mathcal{G}_{-l}^l)\right| \le a_l \qquad l = 0, 1, 2\ldots$$

   where $\lim_{l\to\infty} a_l = 0$ and $\mathcal{G}_{-l}^l$ is the σ-field generated by $Z_{-l}, \ldots, Z_l$.

In the literature one often finds L^2 near epoch dependence (where the L^1 norm in the second part of definition 1.11 is replaced by the L^2 norm), but this requires second moments and we are interested in robust estimation. So we want to allow heavier tails and consider L^1 near epoch dependence. Furthermore, we do not require that the underlying process is independent, it only has to be weakly dependent in the sense of absolute regularity.

Assumption 4. *Let one of the following two conditions hold:*

1. *$(X_n)_{n\in\mathbb{N}}$ is strongly mixing with mixing coefficients $\alpha(n) = O(n^{-\alpha})$ for $\alpha \ge 8$ and let be $E|X_i|^r < \infty$ for a $r > \frac{1}{5}$.*

2. *$(X_n)_{n\in\mathbb{N}}$ is near epoch dependent on an absolutely regular process with mixing coefficients $\beta(n) = O(n^{-\beta})$ for $\beta \ge 8$ with appoximation constants $a(n) = O(n^{-a})$ for $a = \max\{\beta + 3, 12\}$.*

2 Main Results

Empirical U-Process

The CLT and the LIL for the empirical U-process make use of the Hoeffding decomposition, recall that $h_1(x,t) := E\left[g(x,Y,t)\right] - U(t)$. Under Assumptions 1, 2 and 4, the following covariance function converges absolutely and is continuous (compare Theorem 5 of Borovkova et al. [9]):

$$\begin{aligned} K(s,t) &= 4\operatorname{Cov}\left[h_1\left((X_1),s\right), h_1\left((X_1),t\right)\right] \\ &\quad + 4\sum_{k=1}^{\infty}\operatorname{Cov}\left[h_1\left((X_1),s\right), h_1\left((X_{k+1}),t\right)\right] + 4\sum_{k=1}^{\infty}\operatorname{Cov}\left[h_1\left((X_{k+1}),s\right), h_1\left((X_1),t\right)\right]. \end{aligned}$$

We need the following assumption on K

Assumption 5. *Let K be positive definite on $\mathbb{R}$.*

Before we can give our results about the empirical U-process, we have to introduce the reproducing kernel Hilbert space:

Definition 2.1. *We define*

$$\mathcal{K}_m := \left\{ f : \mathbb{R} \to \mathbb{R} \middle| f(x) = \sum_{i=1}^{m} b_i K(x, y_i),\ b_1, \ldots, b_m, y_1, \ldots, y_m \in \mathbb{R} \right\}.$$

For $f(x) = \sum_{i=1}^{m_1} b_i K(x,y_i) \in \mathcal{K}_{m_1}$, $g(x) = \sum_{j=1}^{m_2} c_i K(x,z_i) \in \mathcal{K}_{m_2}$, the inner product of f and g is given by

$$(f,g) = \sum_{i=1}^{m_1}\sum_{j=1}^{m_2} b_i c_j K(y_i, z_j)$$

and $\sqrt{(f,f)}$ is a norm on every $\mathcal{K}_m$. We call $\mathcal{K} = \overline{\bigcup_{m=1}^{\infty}}\mathcal{K}_m$ (the completion of the union) reproducing kernel Hilbert space.

Theorem 1. *Under the assumptions 1, 2, 4 and 5 the empirical U-process*

$$(\sqrt{n}(U_n(t) - Ut(t))_{t\in\mathbb{R}}$$

converges weakly to a centered Gaussian Process $(W_t)_{t\in\mathbb{R}}$ with covariance function K.

$$\left((\sqrt{\frac{n}{2n\log\log n}} (U_n(t) - Ut(t))_{t\in\mathbb{R}} \right)_{n\in\mathbb{N}}$$

is almost surely compact in the space of bounded continuous functions $C(\mathbb{R})$ (equipped with the supremum norm) and the limit set is the unit ball U_K of the reproducing kernel Hilbert space $\mathcal{K}$ associated wtih the covariance function K.

The first part of this theorem is similar to Theorem 9 of Borovkova, Burton Dehling [9]. The main differences are that they use a continuity condition that is different from our Assumption 2 and that our theorem is not restricted to bounded random variables. Part 2 seems to be the first functional LIL for empirical U-processes under dependence.

Generalized Bahadur Representation

Recall that the remainder term in the Bahadur representation is defined as

$$R_n(p) = U_n^{-1}(p) - t_p - \frac{p - U_n(t_p)}{u(t_p)}$$

and that we write $t_p := U^{-1}(p)$

Theorem 2. *Under the Assumptions 1, 2, 3 and 4*

$$\sup_{p\in I} |R_n(p)| = \sup_{p\in I} \left| U_n^{-1}(p) - t_p - \frac{p - U_n(t_p)}{u(t_p)} \right| = o(n^{-\frac{1}{2}-\frac{\gamma}{8}} \log n)$$

almost surely with $I = [\tilde{C}_1, \tilde{C}_2]$ with $U(C_1) < \tilde{C}_1 < \tilde{C}_2 < U(C_2)$, $\gamma := \frac{\alpha-2}{\alpha}$ (if the first part of Assumption 4 holds) respectively $\gamma := \frac{\beta-3}{\beta+1}$ (if the second part of Assumption 4 holds).

Empirical U-Quantiles and GL-Statistics

Using the Bahadur representation, we can deduce the asymptotic bahaviour of the empirical U-quantile process from Theorem 1

Theorem 3. *Under the Assumptions 1, 2, 3, 4 and 5*

$$\left(\sqrt{n}\left(U_n^{-1}(p) - t_p\right)\right)_{p\in I} \xrightarrow{\mathcal{D}} \left(\frac{1}{u(t_p)} W(t_p)\right)_{p\in I},$$

where W is the Gaussian process introduced in Theorem 1 and I the interval introduced in Theorem 3. The sequence

$$\left(\left(\sqrt{\frac{n}{2\log\log n}}(U_n^{-1}(p) - t_p)\right)_{p\in(U(C_1),U(C_2))}\right)_{n\in\mathbb{N}}$$

is almost surely compact with limit set $\left\{ f \middle| f(p) = \frac{1}{u(t_p)} g(t_p),\ g \in U_K \right\}$.

As GL-statistics are linear functionals of the empirical U-quantile process, we can conclude that the CLT and the LIL hold also for T_n:

Theorem 4. *Let be $p_1, \ldots, p_d \in I$ and J a bounded function. Under the assumptions 1, 2, 3, 4 and 5 for T_n defined in Definition 1.6:*

$$\sqrt{n}\left(T_n - ET_n\right) \xrightarrow{\mathcal{D}} N\left(0, \sigma^2\right)$$

with

$$\begin{aligned}\sigma^2 = &\int_{\tilde{C}_1}^{\tilde{C}_2}\int_{\tilde{C}_1}^{\tilde{C}_2} \frac{\operatorname{Cov}\left[W(t_p), W(t_q)\right]}{u(t_p)u(t_q)} J(p)J(q)dpdq \\ &+ 2\sum_{j=1}^{d} b_j \int_{\tilde{C}_1}^{\tilde{C}_2} \frac{\operatorname{Cov}\left[W(t_{p_j}), W(t_p)\right]}{u(t_{p_j})u(t_p)} J(p)dp + 2\sum_{i,j=1}^{d} b_j b_j \frac{\operatorname{Cov}\left[W(t_{p_i}), W(t_{p_j})\right]}{u(t_{p_i})u(t_{p_j})}\end{aligned}$$

Furthermore, we have that

$$\limsup_{n\to\infty} \pm\sqrt{\frac{n}{2\sigma^2 \log\log n}}\,(T_n - ET_n) = 1$$

almost surely.

3 Peliminary Results

Proposition 3.1. *Under the assumptions 1, 2, 4 and 5*

$$\left(\frac{1}{\sqrt{n}}\left(\sum_{i=1}^{n} h_1(X_i,t)\right)\right)_{t\in\mathbb{R}} \xrightarrow{\mathcal{D}} \left(\frac{1}{2}W(t)\right)_{t\in\mathbb{R}}$$

where W is the Gaussian process introduced in Theorem 1 and the sequence

$$\left(\left(\frac{1}{\sqrt{2n\log\log n}}\left(\sum_{i=1}^{n} h_1(X_i)\right)\right)_{t\in\mathbb{R}}\right)_{n\in\mathbb{N}}$$

is almost surely compact, where the limit set is the unit ball $U_{\frac{1}{4}K}$ of the reproducing kernel Hilbert space associated with the covariance function of $\frac{1}{2}W$.

Proof. Without loss of generality, we may assume that $\lim_{t\to-\infty} U(t) = 0$ and $\lim_{t\to\infty} U(t) = 1$. We first study the case that $U(t) = t$ for $t \in [0,1]$. Then our proposition reduces to Theorem A and Theorem B of Berkes and Philipp [8], where the indicator function $\mathbb{1}_{x\le t} - t$ is replaced by $h_1(x,t)$. By Assumption 1, $Eh(x,Y,t)$ is nondecreasing in t and $|Eh_1(X_i,s) - Eh(h_1(X_i,t))| = |s-t|$. Furthermore, by Assumption 2, Lemma 3.5 and 3.10 of Wendler [30], $(h_1(X_n,t))_{n\in\mathbb{N}}$ is a near epoch dependent functional with approximations constants $C\sqrt{a_k} = O(k^{-6})$, so all properties needed in the proof of Theorem A and Theorem B of Berkes and Philipp [8] hold (see also the proofs in Philipp [24]).

To study the general case, note that $Eh_1(X_i,t_p) = U(t_p) = p$ with $t_p = U^{-1}(t_p)$, because U is continuous. So the functional CLT stated in our proposition holds for the process $\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} h_1(X_i,t_p)\right)_{p\in[0,1]}$. Furthermore, notice that if $U(t) = U(s)$, when $h_1(X_i,t) = h_1(X_i,s)$ almost surely by monotonicity of h, so

$$\sum_{i=1}^{n} h_1(X_i,t) = \sum_{i=1}^{n} h_1(X_i,t_{U(t)})$$

almost surely. The finite dimensional weak convergence of $\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} h_1(X_i,t)\right)_{t\in\mathbb{R}}$ follows directly, the tightness can also be deduced from the tightness of the transformed process $\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} h_1(X_i,t_p)\right)_{p\in[0,1]}$, we just need the fact that by Assumption 1 $|U(s) - U(t)| \le C|t-s|$, so the functional CLT follows. To prove the LIL in the general case, use the same transformation. $\square$

Lemma 3.2. *Let be $C_3 \in \mathbb{R}$. Under Assumptions 1, 2 and 4, there exists a constant C, such that for all $s, t \in \mathbb{R}$ with $|s-t| \geq C_3 n^{-\frac{\beta}{\beta+1}}$ and all $n \in \mathbb{N}$*

$$E\left(\sum_{i=1}^{n}\left(h_1\left(X_i, s\right)-h_1\left(X_1, t\right)\right)\right)^4 \leq C n^2 (\log n)^2 |s-t|^{1+\gamma},$$

where γ is defined in Theorem 2.

This lemma is a direct consequence of Lemma 3.4 and Lemma 3.6 of Wendler [30].

Lemma 3.3. *Under Assumptions 1, 2 and 4, there exists a constant C, such that for all $t \in \mathbb{R}$ and all $n \in \mathbb{N}$*

$$\sum_{i_1, i_2, i_3, i_4=1}^{n} E\left|h_2(X_{i_1}, X_{i_2}, t), h_2(X_{i_3}, X_{i_4}, t)\right| \leq C n^2.$$

This is Lemma 4.4 of Dehling, Wendler [12].

Lemma 3.4. *Under the Assumptions 1, 2 and 4*

$$\sup_{t \in \mathbb{R}}\left|\sum_{1 \leq i<j \leq n} h_2\left(X_i, X_j, t\right)\right|=o\left(n^{\frac{3}{2}-\frac{\gamma}{8}}\right)$$

almost surely with γ as in Theorem 2.

In all our proofs, C denotes a constant and may have different values from line to line.

Proof. We define $Q_n(t) := \sum_{1 \leq i<j \leq n} h_2\left(X_i, X_j, t\right)$. For $l \in \mathbb{N}$ chose $t_{1,l}, \ldots, t_{k-1,l}$ with $k = k_l = O\left(2^{\frac{5}{8}l}\right)$, such that

$$-\infty = t_{0,l} < t_{1,l} < \ldots < t_{k-1,l} < t_{k,l} = \infty$$

and $|t_{i,l} - t_{i-1,l}| \leq \frac{1}{2^{\frac{5}{8}l}}$, so that we have

$$|U(t_{i,l}) - U(t_{i-1,l})| \leq C \frac{1}{2^{\frac{5}{8}l}}.$$

By Assumption 1, h and U are nondecreasing in t, so we have for any $t \in [t_{i-1,l}, t_{i,l}]$

$$\begin{aligned}
|Q_n(t)| &= \left| \sum_{1 \le i < j \le n} \left(h\left(X_i, X_j, t\right) - h_1(X_i, t) - h_1(X_j, t)\right) - U(t)) \right| \\
&\le \max \left\{ \left| \sum_{1 \le i < j \le n} \left(h\left(X_i, X_j, t_{i,l}\right) - h_1(X_i, t) - h_1(X_j, t) - U(t)\right) \right|, \right. \\
&\left. \left| \sum_{1 \le i < j \le n} \left(h\left(X_i, X_j, t_{i-1,l}\right) - h_1(X_i, t) - h_1(X_j, t) - U(t)\right) \right| \right\} \\
&\le \max \left\{ |Q_n(t_{i,l})|, |Q_n(t_{i-1,l})| \right\} \\
&+ (n-1) \max \left\{ \left| \sum_{i=1}^{n} (h_1(X_i, t_{i,l}) - h_1(X_i, t))) \right|, \left| \sum_{i=1}^{n} (h_1(X_i, t) - h_1(X_i, t_{i-1,l}))) \right| \right\} \\
&+ \frac{n(n-1)}{2} |U(t_{i,l}) - U(t_{i-1,l})| \\
&\le \max \left\{ |Q_n(t_{i,l})|, |Q_n(t_{i-1,l})| \right\} \\
&+ (n-1) \left| \sum_{i=1}^{n} (h_1(X_i, t_{i,l}) - h_1(X_i, t_{i-1,l}))) \right| + 2 \frac{n(n-1)}{2} |U(t_{i,l}) - U(t_{i-1,l})|.
\end{aligned}$$

So we have that

$$\begin{aligned}
&\sup_{t \in \mathbb{R}} |Q_n(t)| \\
&\quad \le \max_{i=0,\ldots,k} |Q_n(t_i, l)| + \max_{i=1,\ldots,k} (n-1) \left| \sum_{i=1}^{n} (h_1(X_i, t_{i,l}) - h_1(X_i, t_{i-1,l}))) \right| \\
&\qquad + \max_{i=0,\ldots,k} n(n-1) |U(t_{i,l}) - U(t_{i-1,l})|.
\end{aligned}$$

We will treat these three summands separately. For $2^l \le n < 2^{l+1}$, we have for the last summand that $\max_{i=0,\ldots,k} n(n-1)|U(t_{i,l}) - U(t_{i-1,l})| \le C n^{2-\frac{5}{8}} = o\left(n^{\frac{3}{2}-\frac{7}{8}}\right)$ by the choice of $t_1,\dots,t_{k-1}$. For the first summand, we obtain

$$
\begin{aligned}
&E[\max_{n=2^l,\dots,2^{l+1}-1}\max_{i=0,\dots,k}|Q_n(t_{i,l})|^2]\\
&\qquad\le \sum_{i=0}^{k} E\left[\left(\sum_{d=1}^{l}\max_{i=1\dots 2^{l-d}}\left|Q_{2^{l-1}+i2^{d-1}}(t_{i,l})-Q_{2^{l-1}+(i-1)2^{d-1}}(t_{i,l})\right|\right)^2\right]\\
&\qquad\le \sum_{i=0}^{k} l\sum_{d=1}^{l}\sum_{i=1}^{2^{l-d}} E\left[\left(Q_{2^{l-1}+i2^{d-1}}(t_{i,l})-Q_{2^{l-1}+(i-1)2^{d-1}}(t_{i,l})\right)^2\right]\\
&\qquad\le \sum_{i=0}^{k} l\sum_{d=1}^{l}\sum_{i_1,i_2,i_3,i_4=1}^{2^{l+1}} E\left|h_2(X_{i_1},X_{i_2},t_{i,l}),h_2(X_{i_3},X_{i_4},t_{i,l})\right|\\
&\qquad\le Ckl^2 2^{2(l+1)}\le Cl^2 2^{(2+\frac{5}{8})l},
\end{aligned}
$$

where we used Lemma 3.2 in the last line. With the Chebyshev inequality, it follows for every $\epsilon>0$

$$
\begin{aligned}
&\sum_{l=1}^{\infty} P\left[\max_{n=2^l,\dots,2^{l+1}-1}\max_{i=0,\dots,k}|Q_n(t_{i,l})|>\epsilon 2^{l(\frac{3}{2}-\frac{\gamma}{8})}\right]\\
&\qquad\le \sum_{l=1}^{\infty}\frac{1}{\epsilon^2 2^{l(3-\frac{\gamma}{4})}}E[\max_{n=2^l,\dots,2^{l+1}-1}\max_{i=0,\dots,k}|Q_n(t_{i,l})|^2]\le\sum_{l=1}^{\infty}\frac{1}{\epsilon^2 2^{l(3-\frac{\gamma}{4})}}l^2 2^{(2+\frac{5}{8})l}<\infty,
\end{aligned}
$$

as $\gamma\le 1$, so by the Borel Cantelli lemma

$$
\begin{aligned}
&P\left[\max_{i=0,\dots,k}|Q_n(t_{i,l})|>\epsilon n^{\frac{3}{2}-\frac{\gamma}{8}}\text{ i.o.}\right]\\
&\qquad= P\left[\max_{n=2^l,\dots,2^{l+1}-1}\max_{i=0,\dots,k}|Q_n(t_{i,l})|>\epsilon 2^{l(\frac{3}{2}-\frac{\gamma}{8})}\text{ i.o.}\right]=0
\end{aligned}
$$

(the meaning of the abbreviation i.o. is "infinitely often"). It remains to show the convergence of the second summand:

$$
\begin{aligned}
&E\left(\max_{n=2^l,\dots,2^{l+1}-1}\max_{i=1,\dots,k}(n-1)\left|\sum_{i=1}^{n}(h_1(X_i,t_{i,l})-h_1(X_i,t_{i-1,l})))\right|\right)^4\\
&\qquad\le 2^{4(l+1)}\sum_{i=1}^{k}E\left(\max_{n=2^l,\dots,2^{l+1}-1}\left|\sum_{i=1}^{n}(h_1(X_i,t_{i,l})-h_1(X_i,t_{i-1,l})))\right|\right)^4\\
&\qquad\le C2^{6l}l^2k(\max_{i=1,\dots,k}|t_{i,l}-t_{i-1,l}|)^{1+\gamma}\le Cl^2 2^{(6-\frac{5}{8}\gamma)l},
\end{aligned}
$$

where we used Corolarry 1 of Móricz and Lemma 3.2 to obtain the last line. Remember that $k = k_l = O\left(2^{\frac{5}{8}l}\right)$ and that $|t_{i,l} - t_{i-1,l}| \leq \frac{1}{2^{\frac{5}{8}l}}$. We conclude that

$$\sum_{l=0}^{\infty} P\left[\max_{n=2^l,\dots,2^{l+1}-1}\max_{i=1,\dots,k}(n-1)\left|\sum_{i=1}^{n}(h_1(X_i,t_{i,l}) - h_1(X_i,t_{i-1,l})))\right| > \epsilon 2^{(\frac{3}{2}-\frac{\gamma}{8})l}\right]$$
$$\leq \sum_{l=0}^{\infty}\frac{C}{\epsilon^4 2^{l(6-\frac{\gamma}{2})}} E\left(\max_{n=2^l,\dots,2^{l+1}-1}\max_{i=1,\dots,k}(n-1)\left|\sum_{i=1}^{n}(h_1(X_i,t_{i,l}) - h_1(X_i,t_{i-1,l})))\right|\right)^4$$
$$\leq \sum_{l=0}^{\infty}\frac{C}{\epsilon^4 2^{l(6-\frac{\gamma}{2})}} l^2 2^{(6-\frac{5}{8}\gamma)l} = \sum_{l=0}^{\infty}\frac{Cl^2}{\epsilon^4 2^{\frac{\gamma}{8}l}} < \infty.$$

The Borel Cantelli lemma completes the proof. ☐

Lemma 3.5. *Let be F a nondecreasing function, $c, l > 0$ constants and $[C_1, C_2] \subset \mathbb{R}$. If for all $s, t \in [C_1, C_2]$ with $|t-s| \leq l + 2c$*

$$|F(t) - F(s) - (t-s)| \leq c,$$

then for all $p, q \in \mathbb{R}$ with $|p-q| \leq l$ and $F^{-1}(p), F^{-1}(q) \in (C_1 + 2c + l, C_2 - 2c - l)$

$$|F^{-1}(p) - F^{-1}(q) - (p-q)| \leq c$$

where $F^{-1}(p) := \inf\left\{t \middle| F(t) \geq p\right\}$ is the generalized inverse.

Proof. Without loss of generality we assume that $p < q$. Let be $\epsilon \in (0, c)$. By our assumptions

$$F\left(F^{-1}(p) + (q-p) + c + \epsilon\right) \geq F\left(F^{-1}(p) + \epsilon\right) + (q-p) + c - c$$
$$\geq p + (q-p) = q.$$

By the definition of F^{-1}, it follows that

$$F^{-1}(q) = \inf\left\{t \middle| F(t) \geq q\right\} \leq F^{-1}(p) + (q-p) + c + \epsilon.$$

So taking the limit $\epsilon \to 0$, we obtain

$$F^{-1}(q) \leq F^{-1}(p) + (q-p) + c.$$

On the other hand

$$F\left(F^{-1}(p) + (q-p) - c - \epsilon\right) \geq F\left(F^{-1}(p) - \epsilon\right) + (q-p) - c + c$$
$$\geq p + (q-p) = q.$$

So we have that

$$F^{-1}(q) \geq F^{-1}(p) + (q-p) - c - \epsilon,$$

and hence $F^{-1}(q) \geq F^{-1}(p) + (q-p) - c$. Combining the upper and lower inequality for $F^{-1}(q)$, we conclude that $|F^{-1}(p) - F^{-1}(q) - (p-q)| \leq c$. ☐

Lemma 3.6. *Under the Assumptions 1, 2, 3 and 4 for any constand $C > 0$*

$$\sup_{\substack{s,t\in[C_1,C_2]:\\ |s-t|\le C\sqrt{\frac{\log\log n}{n}}}} |U_n(t) - U_n(s) - u(s)(t-s)| = o(n^{-\frac{1}{2}-\frac{\gamma}{8}}\log n).$$

Proof. As a consequence of Assumption 3 and $\gamma < 1$

$$\sup_{\substack{s,t\in[C_1,C_2]:\\ |s-t|\le C\sqrt{\frac{\log\log n}{n}}}} |U(t) - U(s) - u(s)(t-s)| = o(n^{-\frac{1}{2}-\frac{\gamma}{8}}\log n),$$

so it suffices to show that

$$K_n = \sup_{\substack{s,t\in[C_1,C_2]:\\ |s-t|\le C\sqrt{\frac{\log\log n}{n}}}} |U_n(t) - U_n(s) - (U(t) - U(s))| = o(n^{-\frac{1}{2}-\frac{\gamma}{8}}\log n).$$

For $l \in \mathbb{N}$ chose $t_{1,l}, \ldots, t_{k-1,l}$ with $k = k_l = O\left(\sqrt{\frac{2^l}{\log l}}\right)$, $C_1 = t_{0,l} < t_{1,l} < \ldots < t_{k-1,l} < t_{k,l} = C_2$ and $t_{i,l} - t_{i-1,l} \le \sqrt{\frac{\log l}{2^l}}$. Clearly

$$\begin{aligned} K_n &\le 2 \max_{i=1,\ldots,k} \sup_{s,t\in[t_{i-1,l},t_{i,l}]} |U_n(t) - U_n(s) - (U(t) - U(s))| \\ &\qquad \le 4 \max_{i=1,\ldots,k} \sup_{t\in[t_{i-1,l},t_{i,l}]} \left|U_n(t) - U_n(t_{i-1,l}) - (U(t) - U(t_{i-1,l}))\right|. \end{aligned}$$

Now chose for $i = 1, \ldots, k$ and $j = 1, \ldots, m-1$ real numbers $s_{j,i,l}$, such that $t_{i-1,l} = s_{0,i,l} < s_{1,i,l} < \ldots < s_{m-1,i,l} < s_{m,i,l} = t_{i,l}$ and $s_{j,i,l} - s_{j-1,i,l} \le 2^{-(\frac{1}{2}-\frac{\gamma}{4})l}$. As U_n and U are nondecreasing, we have for $t \in (s_{j-1,i,l}, s_{j,i,l})$

$$\begin{aligned} &|U_n(t) - U_n(t_{i-1,l}) - (U(t) - U(t_{i-1,l}))| \\ &\qquad \le \max\{|U_n(s_{j,i,l}) - U_n(t_{i-1,l}) - (U(t) - U(t_{i-1,l}))|, \\ &\qquad\qquad |U_n(s_{j-1,i,l}) - U_n(t_{i-1,l}) - (U(t) - U(t_{i-1,l}))|\} \\ &\qquad \le \max\{|U_n(s_{j,i,l}) - U_n(t_{i-1,l}) - (U(s_{j,i,l}) - U(t_{i-1,l}))|, \\ &\qquad |U_n(s_{j-1,i,l}) - U_n(t_{i-1,l}) - (U(s_{j-1,i,l}) - U(t_{i-1,l}))|\} + |U(s_{j,i,l}) - U(s_{j-1,i,l})|, \end{aligned}$$

and consequently for $2^l \le n < 2^{l+1}$

$$\begin{aligned} K_n &\le 4 \max_{i=1,\ldots,k} \max_{j=1,\ldots,m} |U_n(s_{j,i,l}) - U_n(t_{i-1,l}) - (U(s_{j,i,l}) - U(t_{i-1,l}))| \\ &\qquad + 4 \max_{i=1,\ldots,k} \max_{j=1,\ldots,m} |U(s_{j,i,l}) - U(s_{j-1,i,l})| \\ &\le 8 \max_{i=1,\ldots,k} \max_{j=1,\ldots,m} \left| \frac{1}{n}\sum_{i_1=1}^n h_1(X_{i_1}, s_{j,i,l}) - \frac{1}{n}\sum_{i_1=1}^n h_1(X_{i_1}, t_{i-1,l}) \right| \\ &+ 4 \max_{i=1,\ldots,k} \max_{j=1,\ldots,m} \left| \frac{2}{n(n-1)} \left(\sum_{1\le i_1<i_2\le n} h_2(X_{i_1}, X_{i_2}, s_{j,i,l}) - \sum_{1\le i_1,i_2} h_2(X_{i_1}, X_{i_2}, t_{i-1,l}) \right) \right| \\ &\qquad + 4 \max_{i=1,\ldots,k} \max_{j=1,\ldots,m} |U(s_{j,i,l}) - U(s_{j-1,i,l})|. \end{aligned}$$

From Assumption 3, we obtain

$$\max_{i=1,\dots,k}\max_{j=1,\dots,m}|U(s_{j,i,l})-U(s_{j-1,i,l})|\le\sup_{t\in[C_1,C_2]}u(t)2^{-(\frac{1}{2}-\frac{\gamma}{4})l}=o(n^{-\frac{1}{2}-\frac{\gamma}{8}}\log n).$$

With the help of Lemma 3.4, it follows that

$$\max_{i=1,\dots,k}\max_{j=1,\dots,m}\left|\frac{2}{n(n-1)}\left(\sum_{1\le i_1<i_2\le n}h_2(X_{i_1},X_{i_2},s_{j,i,l})-\sum_{1\le i_1,i_2}h_2(X_{i_1},X_{i_2},t_{i-1,l})\right)\right|$$
$$\le\frac{4}{n(n-1)}\sup_{t\in\mathbb{R}}\left|\sum_{1\le i<j\le n}h_2\left(X_i,X_j,t\right)\right|=o\left(n^{-\frac{1}{2}-\frac{\gamma}{8}}\right).$$

Furthermore, we have by Lemma 3.2 and Corollary 1 of Móricz [23]

$$E\left[\left(\max_{n=2^l,\dots,2^{l+1}-1}\max_{i=1,\dots,k}\max_{j=1,\dots,m}\left|\frac{1}{n}\sum_{i_1=1}^{n}h_1(X_{i_1},s_{j-1,i,l})-\frac{1}{n}\sum_{i_1=1}^{n}h_1(X_{i_1},t_{i-1,l})\right|\right)^4\right]$$
$$\le\frac{1}{2^{4l}}\sum_{i=1}^{k}E\left[\left(\max_{n=0,\dots,2^{l+1}-1}\max_{m_1=1,\dots,m}\left|\sum_{i_1=1}^{n}\sum_{j=1}^{m_1}(h_1(X_{i_1},s_{j,i,l})-h_1(X_{i_1},s_{j-1,i,l}))\right|\right)^4\right]$$
$$\le Ck\frac{1}{2^{4l}}2^{2l}l^2\left(\sqrt{\frac{\log l}{2^l}}\right)^{1+\gamma}=C\frac{l^2(\log l)^{\frac{\gamma}{2}}}{2^{(2+\frac{\gamma}{2})l}},$$

as $k\approx\sqrt{\frac{2^l}{\log l}}$ So we can conclude that for any $\epsilon>0$

$$\sum_{l=1}^{\infty}P\left[\max_{n=2^l,\dots,2^{l+1}-1}\max_{i\le k}\max_{j\le m}\left|\frac{1}{n}\sum_{i_1=1}^{n}(h_1(X_{i_1},s_{j-1,i,l})-h_1(X_{i_1},t_{i-1,l}))\right|\ge\epsilon 2^{-(\frac{1}{2}+\frac{\gamma}{8})l}l\right]$$
$$\le C\sum_{l=1}^{\infty}\frac{2^{(2+\frac{\gamma}{2})l}}{\epsilon^4l^4}\frac{l^2(\log l)^{\frac{\gamma}{2}}}{2^{(2+\frac{\gamma}{2})l}}=C\sum_{l=1}^{\infty}\frac{(\log l)^{\frac{\gamma}{2}}}{l^2}<\infty.$$

The Borel Cantelli lemma completes the proof. □

4 Proof of Main Results

In all our proofs, C denotes a constant and may have different values from line to line.

Proof of Theorem 1. We use the Hoeffding decomposition

$$U_n(t)=U(t)+\frac{2}{n}\sum_{i=1}^{n}h_1\left(X_i,t\right)+\frac{2}{n\left(n-1\right)}\sum_{1\le i<j\le n}h_2\left(X_i,X_j,t\right).$$

By Theorem 3.1 the CLT and the LIL hold for the linear part $\frac{2}{n}\sum_{i=1}^{n}h_1\left(X_i,t\right)$. The faster convergence of the degenerate part $\frac{2}{n(n-1)}\sum_{1\le i<j\le n}h_2\left(X_i,X_j,t\right)$ stated in Lemma 3.4 completes the proof. □

Proof of Theorem 2. By changing the random variables from X_n to $U(X_n)$, we can without loss of generality assume that $U(p) = p = t_p$ on the interval I (Assumption 3 guarantees that $R_n(p)$ is only blown up by a constant because of this transformation). Then we can write $R_n(p)$ as

$$\begin{aligned}R_n(p) &= U_n^{-1}(p) - t_p + U_n(t_p) - p \\ &= \big(U_n^{-1}(p) - U_n^{-1}(U_n(t_p)) + U_n(t_p) - p\big) + \big(U_n^{-1}(U_n(t_p)) - t_p\big)\end{aligned}$$

Applying Lemma 3.6 and Lemma 3.5 with $F = U_n$, $c = n^{-\frac{1}{2}-\frac{\gamma}{8}}\log n$ and $l = C\sqrt{\frac{\log\log n}{n}}$, we obtain

$$\sup_{\substack{p,q\in I:\\ |p-q|\le C\sqrt{\frac{\log\log n}{n}}}} \left|U_n^{-1}(p) - U_n^{-1}(q) - (p-q)\right| = o(n^{-\frac{1}{2}-\frac{\gamma}{8}}\log n).$$

almost surely. By Theorem 1 we have that $\sup_{t\in[C_1,C_2]}(U_n(t_p) - p) \le C\sqrt{\frac{\log\log n}{n}}$ almost surely, it follows that

$$\begin{aligned}&\sup_{p\in I}\left|U_n^{-1}(p) - U_n^{-1}(U_n(t_p)) + U_n(t_p) - p\right| \\ &\qquad\le \sup_{\substack{p,q\in I:\\ |p-q|\le C\sqrt{\frac{\log\log n}{n}}}} \left|U_n^{-1}(p) - U_n^{-1}(q) - (p-q)\right| = o(n^{-\frac{1}{2}-\frac{\gamma}{8}}\log n)\end{aligned}$$

almost surely. It remains to show the convergence of $U_n^{-1}(U_n(t_p)) - t_p$. For every $\epsilon > 0$ by the definition of the generalized inverse, $U_n^{-1}(U_n(t_p)) - t_p > \epsilon n^{-\frac{1}{2}-\frac{\gamma}{8}}\log n$ only if $U_n(t_p + \epsilon n^{-\frac{1}{2}-\frac{\gamma}{8}}\log n) < U_n(t_p)$ and $U_n^{-1}(U_n(t_p)) - t_p \le -\epsilon n^{-\frac{1}{2}-\frac{\gamma}{8}}\log n$ only if $U_n(t_p - \epsilon n^{-\frac{1}{2}-\frac{\gamma}{8}}\log n) \ge U_n(t_p)$. So we can conclude that

$$\begin{aligned}&P\left[\sup_{p\in I}|U_n^{-1}(U_n(t_p)) - t_p| > \epsilon n^{-\frac{1}{2}-\frac{\gamma}{8}}\log n \text{ i.o.}\right] \\ &\le P\left[\sup_{t\in[C_1,C_2-\epsilon n^{-\frac{1}{2}-\frac{\gamma}{8}}\log n]} U_n(t+\epsilon n^{-\frac{1}{2}-\frac{\gamma}{8}}\log n) - U_n(t) \le 0 \text{ i.o.}\right] \\ &\le P\left[\sup_{\substack{s,t\in[C_1,C_2]\\ |s-t|=\epsilon n^{-\frac{1}{2}-\frac{\gamma}{8}}\log n}} |U_n(t) - U_n(s) + (U(t) - U(s))| \ge |U(t) - U(s)| \text{ i.o.}\right] \\ &\le P\left[\sup_{\substack{s,t\in[C_1,C_2]\\ |s-t|\le\epsilon n^{-\frac{1}{2}-\frac{\gamma}{8}}\log n}} |U_n(t) - U_n(s) + (U(t) - U(s))| \ge \frac{\epsilon\log n}{n^{\frac{1}{2}+\frac{\gamma}{8}}\inf_{t\in[C_1,C_2]}u(t)} \text{ i.o.}\right] \\ &= 0,\end{aligned}$$

where the last line is a consequence of Lemma 3.6. $\square$

Proof of Theorem 3. We make use of the Bahadur representation

$$\sqrt{n}\left(U_n^{-1}(p) - t_p\right) = -\frac{1}{u(t_p)}\sqrt{n}\left(U_n(t_p) - U(t_p)\right) + \sqrt{n}R_n(p).$$

From Theorem 1, we have that the finite dimensional distribution of the rescaled empirical U-process $\left(-\frac{1}{u(t_p)}\sqrt{n}\left(U_n(t_p) - U(t_p)\right)\right)_{p\in I}$ converge to the finite dimensional distributions of the centered Gaussian process $\left(\frac{1}{u(t_p)}W(t_p)\right)_{p\in I}$. The tightness is inherited of the process $(\sqrt{n}\left(U_n(t) - U(t)\right))_{t\in[C_1,C_2]}$, as $|t_p - t_q| \leq \frac{|p-q|}{\inf_{t\in[C_1,C_2]} u(t)}$. The faster convergence of $(R_n(p))_{p\in I}$ (Theorem 2) completes the proof of the first half of this Theorem, the proof of the second half works in a similar way. □

Proof of Theorem 4. T defined in Definition 1.6 is a linear and continuous functional, so we have that

$$\sqrt{n}\left(T_n - ET_n\right) = T\left(\sqrt{n}\left(U_n^{-1} - U^{-1}\right)\right)$$

converges weakly to $T\left(\left(\frac{1}{u(t_p)}W(t_p)\right)_{p\in I}\right)$, which is a normal distributed random variable with variance σ^2. Similarly, we have that $\sqrt{\frac{n}{2\sigma^2\log\log n}}\left(T_n - ET_n\right)$ has almost surely the compact limit set $\left\{\frac{1}{\sqrt{2\sigma^2}}T(f)\middle|f(p) = \frac{1}{u(t_p)}g(t_p),\ g\in U_K\right\}$.

It remains to prove that this limit set is $[-1,1]$. This can be easily seen by the following argument: Let $(W_n)_{n\in\mathbb{N}}$ be a sequence of independent copies of the Gaussian process W introduced in Theorem 1. Then the limit set of $\tilde{W}_n := \sqrt{\frac{1}{2n\log\log n}}\sum_{i=1}^n W_i$ is U_K by the LIL for Hilbert space valued random variables, see Ledoux, Talagrand [22]. So the limit set of $\frac{1}{\sqrt{2\sigma^2}}T\left(\frac{1}{u(t_p)}\tilde{W}_n(t_p)\right)$ is $\left\{\frac{1}{\sqrt{2\sigma^2}}T(f)\middle|f(p) = \frac{1}{u(t_p)}g(t_p),\ g\in U_K\right\}$. On the other hand, by the linearity of T we have that

$$\frac{1}{\sqrt{2\sigma^2}}T\left(\left(\frac{1}{u(t_p)}\tilde{W}_n(t_p)\right)_{p\in I}\right) = \sqrt{\frac{1}{2\sigma^2 n\log\log n}}\sum_{i=1}^n T\left(\left(\frac{1}{u(t_p)}W_i(t_p)\right)_{p\in I}\right)$$

and has limit set $[0,1]$, as $T(\frac{1}{u(t_p)}W_i(t_p))$ is normal distributed with expectation 0 and vararaince σ^2.

□

Acknowledgement

The Research was supported by the German Academic Foundation (Studienstiftung des deutschen Volkes) and the Collaborative Research Center *Statistik nichtlinearer dynamischer Prozesse* (SFB 823) of the German Research Foundation (DFG).

References

[1] M.A. Arcones, The law of the iterated logarithm for U-processes, *J. Multivariate Anal.* **47** (1993) 139-151.

[2] M.A. Arcones, The Bahadur-Kiefer representation for U-quantiles, *Ann. Stat.* **24** (1996) 1400-1422.

[3] M.A. Arcones, E. Giné, Limit Theorems for U-processes, *Ann. Prob.* **21** (1993) 1494-1542.

[4] M.A. Arcones, E. Giné, On the law of the iterated logarithm for canonical U-statistics and processes, *Stochastic Process. Appl.* **58** (1995) 217-245.

[5] M.A. Arcones, B. Yu, Central limit theorem for empirical and U-processes of stationary mixing sequences, *J. Theoret. Probab.* **7** (1997) .47-53.

[6] G.J. Babu, K. Singh, On deviations between empirical and quantile processes for mixing random variables, *J. Multivariate Anal.*, **8** (1978) 532-549.

[7] R.R. Bahadur, A note on quantiles in large samples, *Ann. Math. Stat.* **37** (1966) 577-580.

[8] I. Berkes, W. Philipp An almost sure invariance principle for the empirical distribution function of mixing random variables, *Probab. Theory Related Fields* **41** (1977) 115-137.

[9] S. Borovkova, R. Burton, H. Dehling, Limit theorems for functionals of mixing processes with applications to U-statistics and dimension estimation, *Trans. Amer. Math. Soc.* **353** (2001) 4261–4318.

[10] R.C. Bradley, *Introduction to strong mixing conditions*, volume 1-3, Kendrick Press, Heber City (2007).

[11] J. Choudhury, R.J. Serfling, Generalized order statistics, Bahadur representations, and sequential nonparametric fixed-width confidence intervals, *J. Statist. Plann. Inference* **19** (1988) 269-282.

[12] H. Dehling, M. Denker, W. Philipp, The almost sure invariance principle for the empirical process of U-statistic structure, *Annales de l'I.H.P.* **23** (1987) 121-134.

[13] H. Dehling, M. Wendler, Central limit theorem and the bootstrap for U-statistics of strongly mixing data, *J. Multivariate Anal.*, **101** (2010) 126-137.

[14] H. Dehling, M. Wendler, Law of the iterated logarithm for U-statistics of weakly dependent observations, To appear in: Berkes, Bradley, Dehling, Peligrad, Tichy (Eds): *Dependence in Probability, Analysis and Number Theory*, Kendrick Press, Heber City (2010).

[15] M. Denker, G. Keller, Rigorous statistical procedures for data from dynamical systems, *J. Stat. Phys.* **44** (1986) 67-93.

[16] J.C. Geertsema, Sequential confidence intervals based on rank test, *Ann. Math. Stat.* **41** (1970) 1016-1026.

[17] W. Hoeffding, A class of statistics with asymptotically normal distribution, *Ann. Math. Stat.* **19** (1948) 293-325.

[18] F. Hofbauer, G. Keller, Ergodic properties of invariant measures for piecewise monotonic transformations, *Math. Z.* **180** (1982) 119-142.

[19] T. Hsing, W.B. Wu, On weighted U-statistics for stationary processes, *Ann. Prob.* **32** (2004) 1600-1631.

[20] J. Kiefer, Deviations between the sample quantile process and the sample df, in: M.L. Puri (Ed): *Nonparametric Techniques in Statistical Inference* (1970).

[21] R. Kulik, Bahadur-Kiefer theory for sample quantiles of weakly dependent linear processes, *Bernoulli* **13** (2007) 1071-1090.

[22] M. Ledoux, M. Talagrand, *Probabiliy in Banach Spaces*, Springer, New York (2002).

[23] F. Móricz, A general moment inequality for the maximum of the rectangular partial sums of multiple series, *Acta Math. Hung.* **43** (1983) 337-346.

[24] W. Philipp, A functional law of the iterated logarithm for empirical functions of weakly dependent random variables, *Ann. Prob.* **5** (1977) 319-350.

[25] B.M. Pötscher, I.R. Prucha, Basic structure of the asymptotic theory in dynamic nonlinear econometric models, part I, *Econometric Reviews* **10** (1991) 125-216.

[26] P.J. Rousseeuw, C. Croux, Alternatives to the median absolute deviation, *J. Amer. Stat. Soc.* **88** (1993) 1273-1283.

[27] R.J. Serfling, The law of the iterated logarithm for U-statistics and related von Mises statistics, *Ann. Math. Statist.* **42** (1971) 1794.

[28] R.J. Serfling, Generalized L-, M-, and R-statistics, *Ann. Prob.* **12** (1984) 76-86.

[29] W. Vervaat, Functional central limit theorems for processes with positive drift and their inverses, *Probab. Theory Related Fields* **23** (1972) 245-253.

[30] M. Wendler, Bahadur representation for U-quantiles of dependent data *preprint* arXiv:1004.2581 (2010).

[31] *W.B. Wu*, On the Bahadur representation of sample quantiles for dependent sequences, *Ann. Stat.* **33** (2005) 1934-1963.

[32] K. Yoshihara, Limiting behavior of U-statistics for stationary, absolutely regular processes, *Probab. Theory Related Fields* **35** (1976) 237-252.