U-processes, U-quantile processes and generalized linear statistics of dependent data

Martin Wendler[*]

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

Abstract

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 have only been studied under independence; in this paper, we develop an asymptotic theory for GL-statistics of sequences which are strongly mixing or L^1 near epoch dependent on an absolutely regular process. For this purpose, we prove an almost sure approximation of the empirical U-process by a Gaussian process. With the help of a generalized Bahadur representation, it follows that such a strong invariance principle also holds for the empirical U-quantile process and consequently for GL-statistics. We obtain central limit theorems and laws of the iterated logarithm for U-processes, U-quantile processes and GL-statistics as straightforward corollaries.

© 2011 Elsevier B.V. All rights reserved.

MSC: 62G30; 60G10; 60F17

Keywords: L-Statistic; U-statistics; Invariance principle; Bahadur representation; Mixing; Near epoch dependence

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

---

* Fax: +49 234 3214039.
  E-mail address: Martin.Wendler@rub.de.

following two estimators of scale are U-statistics:

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

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

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 central limit theorem (CLT) goes back to [22] and was extended to absolutely regular sequences by Yoshihara [37], to near epoch dependent sequences on absolutely regular processes by Denker and Keller [19] and to strongly mixing random variables by Dehling and Wendler [17]. The law of the iterated logarithm (LIL) under independence was proved by Serfling [32] and was extended to strongly mixing and near epoch dependent sequences by Dehling and Wendler [18].

Not only are U-statistics with fixed kernel g 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 [12].

The functional CLT for the empirical U-distribution function has been established by Arcones and Giné [5] for independent data, by Arcones and Yu for absolutely regular data [7], and by Borovkova et al. [12] 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 [2], Arcones and Giné [6] under independence. The strong invariance principle has been investigated by Dehling et al. [16]. We will show a strong invariance principle under dependence. As a corollary, we will obtain the LIL to sequences which are strongly mixing or L^1 near epoch dependent on an absolutely regular process and the CLT under conditions which are slightly different from the conditions in [12]. Let us now proceed with 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 $(\sqrt{n}(U_n(t) - U(t)))_{t \in \mathbb{R}}$.

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

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

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 non-decreasing in the third argument. The U-distribution function U is continuous. For all $x, y \in \mathbb{R} : \lim_{t \to \infty} h(x, y, t) = 1, \lim_{t \to -\infty} h(x, y, t) = 0$.

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 [19]):

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

$$E\left[\sup_{\|(x,y)-(X,Y)\| \leq \epsilon} |h(x, y, t) - h(X, Y, t)| \right] \leq L\epsilon,$$

where X, Y are independent with the same distribution as X_1 and $\|\cdot\|$ denotes the Euclidean 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 U at point p (in general, U does not have to be invertible, but this is guaranteed by our Assumption 3 at least in the interval I introduced in Theorem 2). A natural estimator of a U-quantile is the empirical U-quantile $U_n^{-1}(p)$, which is the generalized inverse of the empirical U-distribution function at point p:

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

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

is called the empirical U-quantile.

Empirical U-quantiles have applications in robust statistics.

Example 1.5. Let $h(x, y, t) := \mathbb{1}_{\{|x-y| \leq t\}}$. Then the 0.25-$U$-quantile is the Q_n estimator of scale proposed by Rousseeuw and Croux [31], 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 variation condition), if the U-distribution function is Lipschitz continuous. For every $\epsilon > 0$

$$E\left[\sup_{\|(x,y)-(X,Y)\| \leq \epsilon} \left| \mathbb{1}_{\{|x-y| \leq t\}} - \mathbb{1}_{\{|X-Y| \leq t\}} \right| \right] \leq P\left[t - \sqrt{2}\epsilon < |X - Y| \leq t + \sqrt{2}\epsilon \right]$$

$$\leq U(t + \sqrt{2}\epsilon) - U(t - \sqrt{2}\epsilon) \leq C\epsilon.$$

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

representation [10]:

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

where $u = U'$ is the derivative of the U-distribution function. For independent data and fixed p, Geertsema [20] established a generalized Bahadur representation with $R_n(p) = O\left(n^{-\frac{3}{4}}\log n\right)$ a.s. Dehling et al. [16] and Choudhury and Serfling [14] improved the rate to $R_n(p) = O\left(n^{-\frac{3}{4}}(\log n)^{\frac{3}{4}}\right)$. Arcones [4] 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 absolutely regular processes, we recently established rates of convergence for $R_n(p)$ which depend on the decrease of the mixing coefficients [34]. The CLT and the LIL for $U_n^{-1}(p)$ are straightforward 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 $\left(U_n^{-1}(p)\right)_{p \in I}$ 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. In order to do this, we will examine the rate of convergence of $\sup_{p \in I} R_n(p)$ and use the approximation of the empirical U-process by a Gaussian process. As we divide 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 $0 < \inf_{t \in [C_1, C_2]} u(t) \le \sup_{t \in [C_1, C_2]} u(t) < \infty$ $(u(t) = U'(t))$ and

$$\sup_{t, t' \in [C_1, C_2]: |t - t'| \le x} \left| U(t) - U(t') - u(t)(t - t') \right| = O\left(x^{\frac{5}{4}}\right).$$

The Bahadur representation for the sample quantile process goes back to [24] under independence. Babu and Singh [9] proved such a representation for mixing data and Kulik [26] and Wu [35] 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 $p_1, \ldots, p_d \in I, b_1, \ldots, b_d \in \mathbb{R}$ and let J be a bounded function, that is continuous a.e. and vanishes outside of I. We call a statistic of the form

$$T_n = T\left(U_n^{-1}\right) := \int_I J(p) U_n^{-1}(p) dp + \sum_{j=1}^{d} b_j U_n^{-1}(p_j)$$

$$= \sum_{i=1}^{\frac{n(n-1)}{2}} \int_{\frac{2(i-1)}{n(n-1)}}^{\frac{2i}{n(n-1)}} J(t) \, dt \cdot U_n^{-1}\left(\frac{2i}{n(n-1)}\right) + \sum_{j=1}^{d} b_j U_n^{-1}(p_j)$$

generalized linear statistic (GL-statistic).

This generalization of L-statistics was introduced by Serfling [33]. U-statistics, U-quantiles and L-statistics can be written as GL-statistics (though this might be somewhat artificially). For a U-statistics, just take $h(x, y, t) = \mathbb{1}_{\{g(x,y) \le t\}}$ 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 $h(x, y, t) := \frac{1}{2} \left(\mathbb{1}_{\{x \leq t\}} + \mathbb{1}_{\{y \leq t\}} \right)$, $p_1 = 0.25$, $p_2 = 0.75$, $b_1 = -1$, $b_2 = 1$, and $J = 0$. Then a short calculation shows that the related GL-statistic is

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

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 $h(x, y, t) := \mathbb{1}_{\left\{ \frac{1}{2}(x-y)^2 \leq t \right\}}$, $p_1 = 0.75$, $b_1 = 0.25$ 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 [33], there seem 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 [13]. 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, see [1].

We will consider sequences which are near epoch dependent on absolutely regular processes, as this class covers linear processes 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 [23]. Linear processes (even with discrete innovations) and GARCH processes are also near epoch dependent, see [21]. Near epoch dependent random variables are also called approximating functionals (for example in [12])

Definition 1.10. Let $(X_n)_{n \in \mathbb{N}}$ be a stationary process.

1. The absolute regularity coefficient is given by

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

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

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

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 1.10 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 $\alpha \geq 8$ and $E|X_i|^r < \infty$ 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 $\beta \geq 8$ with approximation constants $a(n) = O(n^{-a})$ for $a = \max \{\beta + 3, 12\}$.

Kiefer–Müller processes

For uniformly on [0, 1] distributed and independent random variables $(X_n)_{n \in \mathbb{N}}$, Müller [29] determined the limit distribution of the empirical process

$$\left(\frac{1}{\sqrt{n}} \sum_{1 \leq i \leq sn} (\mathbb{1}_{\{X_i \leq t\}} - t) \right)_{t,s \in [0,1]}.$$

It converges weakly towards a Gaussian process $(K(t, s))_{s,t \in [0,1]}$ with covariance function $E K(t, s) K(t', s') = \min\{s, s'\}(\min\{t, t'\} - tt')$. Kiefer [25] proved an almost sure invariance principle: After enlarging the probability space, there exists a copy of the Kiefer–Müller process K such that the empirical process and K are close together with respect to the supremum norm. Berkes and Philipp [11] extended this to dependent random variables. For sample quantiles, Csörgő and Révész [15] established a strong invariance principle, but only under independence. We will extend this to dependent data and to U-quantiles.

A strong invariance principle is a very interesting asymptotic theorem, as the limit behavior of Gaussian processes is well understood and it is then possible to conclude that the approximated process has the same asymptotic properties. Note that a Kiefer–Müller processes can be described as a functional Brownian motion, as its increments in s direction are independent Brownian Bridges. We have the following scaling behavior: $\left(\frac{1}{\sqrt{n}} K(t, ns) \right)_{s,t \in [0,1]}$ has the same distribution as $(K(t, s))_{s,t \in [0,1]}$.

Furthermore, a functional LIL holds: The sequence

$$\left(\left(\frac{1}{\sqrt{2n \log \log n}} K(t, ns) \right)_{s,t \in [0,1]} \right)_{n \in \mathbb{N}}$$

is almost surely relatively compact (with respect to the supremum norm). The limit set is the unit ball of the reproducing kernel Hilbert space associated with the covariance function of the process $(K(t, s))_{s,t \in [0,1]}$. For details about the reproducing kernel Hilbert space, see [8] or [27].

2. Main results

Empirical U-process

The asymptotic theory for the empirical U-process makes use of the Hoeffding decomposition. Recall that $h_1(x, t) := E[h(x, Y, t)] - U(t)$. Under Assumptions 1, 2 and 4, the following covariance function converges absolutely and is continuous (compare to Theorem 5 of [12]):

$$\Gamma(t, t') = 4\text{Cov}\left[h_1(X_1, t), h_1(X_1, t')\right] + 4\sum_{k=1}^{\infty}\text{Cov}\left[h_1(X_1, t), h_1(X_{k+1}, t')\right]$$

$$+ 4\sum_{k=1}^{\infty}\text{Cov}\left[h_1(X_{k+1}, t), h_1(X_1, t')\right].$$

Theorem 1. *Under the Assumptions 1, 2 and 4 there exists a centered Gaussian process $(K(t, s))_{t,s\in\mathbb{R}}$ (after enlarging the probability space if necessary) with covariance function*

$$E K(t, s) K(t', s') = \min\{s, s'\} \Gamma(t, t')$$

such that almost surely

$$\sup_{\substack{t\in\mathbb{R} \\ s\in[0,1]}} \frac{1}{\sqrt{n}}\left|\lfloor ns\rfloor(U_{\lfloor ns\rfloor}(t) - U(t)) - K(t, ns)\right| = O\left(\log^{-\frac{1}{3840}} n\right).$$

The rate of convergence to zero in this theorem is very slow, but the same as in [11], as we strongly use their method of proof. By the scaling property of the process K, we obtain the asymptotic distribution of $U_{\lfloor ns\rfloor}(t)$, and by Theorem 2.3 of [3] a functional LIL:

Corollary 1. *Under the Assumptions 1, 2 and 4 the empirical U-process*

$$\left(\frac{\lfloor ns\rfloor}{\sqrt{n}}(U_{\lfloor ns\rfloor}(t) - U(t))\right)_{t\in\mathbb{R}, s\in[0,1]}$$

converges weakly in the space $D(\mathbb{R} \times [0, 1])$ (equipped with the supremum norm) to a centered Gaussian Process $(K(t, s))_{t,s\in\mathbb{R}}$ introduced in Theorem 1. The sequence

$$\left(\left(\frac{\lfloor ns\rfloor}{\sqrt{2n\log\log n}}(U_{\lfloor ns\rfloor}(t) - U(t))\right)_{t\in\mathbb{R}, s\in[0,1]}\right)_{n\in\mathbb{N}}$$

is almost surely relatively compact in the space $D(\mathbb{R} \times [0, 1])$ (equipped with the supremum norm) and the limit set is the unit ball of the reproducing kernel Hilbert space associated with the covariance function of the process K.

The first part of this corollary is very similar to Theorem 9 of [12] (they use a continuity condition that is different from our Assumption 2). Up to our knowledge, Part 2 is the first functional LIL for empirical U-processes under dependence.

Generalized Bahadur representation

Recall that the remainder term in the generalized 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 $t_p := U^{-1}(p)$. We set $U_0^{-1}(p) := 0$ as it is not possible to find a generalized inverse of $U_0 = 0$.

Theorem 2. *Under the Assumptions 1–4*

$$\sup_{\substack{p \in I \\ s \in [0,1]}} \frac{\lfloor ns \rfloor}{\sqrt{n}} |R_{\lfloor ns \rfloor}(p)| = o(n^{-\frac{\gamma}{8}} \log n)$$

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

Note that for a fast decay of the mixing coefficients, the rate becomes close to $n^{-\frac{1}{8}}$, while the optimal rate for sample quantiles of independent data is $n^{-\frac{1}{4}} (\log n)^{\frac{1}{2}} (\log \log n)^{\frac{1}{4}}$.

Empirical U-quantiles and GL-statistics

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

Theorem 3. *Under the Assumptions 1–4, there exists a centered Gaussian process $(K'(p, s))_{p \in I, s \in \mathbb{R}}$ (after enlarging the probability space if necessary), where I is the interval introduced in Theorem 2, with covariance function*

$$E K'(p, s) K'(p', s') = \min \{s, s'\} \frac{1}{u(t_p) u(t_{p'})} \Gamma(t_p, t_{p'})$$

such that

$$\sup_{\substack{p \in I \\ s \in [0,1]}} \frac{1}{\sqrt{n}} \left| \lfloor ns \rfloor (U_{\lfloor ns \rfloor}^{-1}(p) - t_p) - K'(p, ns) \right| = O \left(\log^{-\frac{1}{3840}} n \right).$$

K' is a Gaussian process with independent increments in s direction, so we have the following consequences:

Corollary 2. *Under the Assumptions 1–4*

$$\left(\frac{\lfloor ns \rfloor}{\sqrt{n}} (U_{\lfloor ns \rfloor}^{-1}(p) - t_p) \right)_{p \in I, s \in [0,1]}$$

converges weakly in the space $D(I \times [0, 1])$ (equipped with the supremum norm) to the centered Gaussian Process $(K'(p, s))_{p \in I, s \in \mathbb{R}}$ introduced in Theorem 3. The sequence

$$\left(\left(\frac{\lfloor ns \rfloor}{\sqrt{2n \log \log n}} (U_{\lfloor ns \rfloor}^{-1}(p) - t_p) \right)_{p \in I, s \in [0,1]} \right)_{n \in \mathbb{N}}$$

is almost surely relatively compact in the space $D(I \times [0, 1])$ (equipped with the supremum norm) and the limit set is the unit ball of the reproducing kernel Hilbert space associated with the covariance function of the process K'.

As GL-statistics are linear functionals of the empirical U-quantile process, we get an approximation for T_n:

Theorem 4. *Let* $p_1, \ldots, p_d \in I$ *and let* J *be a bounded function, that is continuous a.e. and vanishes outside of* I. *Under the [Assumptions](#) 1–4, there exists (after enlarging the probability space if necessary) a Brownian motion* B, *such that for* T_n *defined in [Definition](#) 1.6 and*

$$\sigma^2 = \int_{\tilde{C}_1}^{\tilde{C}_2} \int_{\tilde{C}_1}^{\tilde{C}_2} \frac{\Gamma(t_p, t_q)}{u(t_p)u(t_q)} J(p)J(q)dpdq$$

$$+ 2 \sum_{j=1}^{d} b_j \int_{\tilde{C}_1}^{\tilde{C}_2} \frac{\Gamma(t_{p_j}, t_p)}{u(t_{p_j})u(t_p)} J(p)dp + \sum_{i,j=1}^{d} b_i b_j \frac{\Gamma(t_{p_i}, t_{p_j})}{u(t_{p_i})u(t_{p_j})}$$

we have that

$$\sup_{s \in [0,1]} \frac{1}{\sqrt{n}} \left| \lfloor ns \rfloor (T_{\lfloor ns \rfloor} - T(U^{-1})) - \sigma B(ns) \right| = O\left(\log^{-\frac{1}{3840}} n \right)$$

almost surely.

By the well-known properties of Brownian motions, we have:

Corollary 3. *Let* $p_1, \ldots, p_d \in I$ *and let* J *be a bounded function. Under the [Assumptions](#) 1–4 for* T_n *defined in [Definition](#) 1.6:*

$$\frac{\lfloor ns \rfloor}{\sqrt{n}} (T_{\lfloor ns \rfloor} - T(U^{-1}))$$

converges weakly to the Brownian motion $\sigma B(s)$ *with* σ^2 *as in [Theorem](#) 4. Furthermore, we have that the sequence*

$$\left(\frac{\lfloor ns \rfloor}{\sqrt{2n \log \log n}} (T_{\lfloor ns \rfloor} - T(U^{-1}))_{s \in [0,1]} \right)_{n \in \mathbb{N}}$$

is almost surely relatively compact in the space of bounded continuous functions $C[0,1]$ *(equipped with the supremum norm) and the limit set is*

$$\left\{ f : [0,1] \to \mathbb{R} \,|\, f(0) = 0, \int_0^1 f'^2(s)ds \leq \sigma^2 \right\}.$$

3. Preliminary results

Proposition 3.1. *Under the [Assumptions](#) 1, 2 and 4 there exists a centered Gaussian process* $(K(t,s))_{t,s \in \mathbb{R}}$ *(after enlarging the probability space if necessary) with covariance function*

$$EK(t,s)K(t',s') = \min\left\{s, s'\right\} \Gamma(t,t')$$

such that almost surely

$$\sup_{\substack{t \in \mathbb{R} \\ s \in [0,1]}} \frac{1}{\sqrt{n}} \left| \left(2 \sum_{1 \leq i \leq ns} h_1(X_i, t) - K(t, ns) \right) \right| = O\left(\log^{-\frac{1}{3840}} n \right).$$

Proof. This proposition is basically Theorem 1 of [11], which we have to generalize in three aspects:

1. Berkes and Philipp assume that the covariance kernel Γ is positive definite, we want to avoid this condition here.

2. Berkes and Philipp consider indicator functions $\mathbb{1}_{\{x \leq t\}}$, while in this version of the proposition, we deal with more general functions $Eh(x, Y, t)$.

3. Theorem 1 of Berkes and Philipp is restricted to the distribution function $F(t) = E\mathbb{1}_{\{X_i \leq t\}} = t$, we will extend this to a function U according to our Assumption 1.

The mixing condition of Berkes and Philipp is the same as our Assumption 4.

1. In the proof of their Theorem 1, Berkes and Philipp use the fact that Γ is positive definite only for two steps. Their Proposition 4.1 (page 124) also holds if this is not the case. It is easy to see that the characteristic functions of the finite dimensional distributions then might converge to 1 at some points, but with the required rate. Furthermore, we have to show (page 135) that for all $t_1, \ldots, t_{d_k} \in [0, 1]$, $P[\|(K(t_1, 1), \ldots, K(t_{d_k}, 1))\| \geq \frac{1}{4} T_k] \leq \delta_k$, where T_k and δ_k are defined in their article. Let $\Gamma_{d_k} = \left(\Gamma(t_i, t_j)\right)_{1 \leq i, j \leq d_k}$ be the covariance matrix of $K(t_1, 1), \ldots, K(t_{d_k}, 1)$ and ρ its biggest eigenvalue. We first consider the case that $\rho > 0$. As Γ_{d_k} is symmetric and positive semidefinite, there exists a matrix $\Gamma_{d_k}^{\frac{1}{2}}$ such that $\left(\Gamma_{d_k}^{\frac{1}{2}}\right)^t \Gamma_{d_k}^{\frac{1}{2}} = \Gamma_{d_k}$ and the vector $K(t_1, 1), \ldots, K(t_{d_k}, 1)$ has the same distribution as $\Gamma_{d_k}^{\frac{1}{2}}(W_1, \ldots, W_{d_k})^t$, where $W_1, \ldots, W_{d_k}$ are independent standard normal random variables. So it follows that

$$
\begin{aligned}
P\left[\|(K(t_1, 1), \ldots, K(t_{d_k}))\| \geq \frac{1}{4} T_k\right] &= P\left[\|\Gamma_{d_k}^{\frac{1}{2}}(W_1, \ldots, W_{d_k})\| \geq \frac{1}{4} T_k\right] \\
&\leq P\left[\sqrt{\rho}\|(W_1, \ldots, W_{d_k})^t\| \geq \frac{1}{4} T_k\right] \\
&= \frac{1}{(2\pi)^{\frac{1}{2} d_k}} \int_{\|(x_1, \ldots, x_{d_k})\| \geq \frac{1}{4\sqrt{\rho}} T_k} \exp\left(-\frac{1}{2}(x_1^2 + \cdots + x_{d_k}^2)\right) dx_1 \ldots dx_{d_k}.
\end{aligned}
$$

The rest of the proof is then exactly the same as in [11]. In the case $\rho = 0$, we have that $\Gamma = 0$, so trivially $P[\|(K(t_1, 1), \ldots, K(t_{d_k}))\| \geq \frac{1}{4} T_k] = 0 \leq \delta_k$.

2. The proof uses different properties of the indicator functions. If the process $(X_n)_{n\in\mathbb{N}}$ is near epoch dependent with constants $(a_n)_{n\in\mathbb{N}}$, then as a consequence of Lemma 3.2.1 of [30] the process $\left(\mathbb{1}_{\{X_n \leq t\}}\right)_{n\in\mathbb{N}}$ is near epoch dependent with constants $(\sqrt{a_n})_{n\in\mathbb{N}}$. The same holds for the sequence $(h_1(X_n, t))_{n\in\mathbb{N}}$ by Assumption 2, Lemmas 3.5 and 3.10 of [34].

Furthermore, h and U are non-decreasing in t. Berkes and Philipp used different moment properties, which we also assume: $h_1(X_n, t)$ is bounded by 1 and $E|h_1(X_n, t) - h_1(X_n, t')| \leq C|t - t'|$ for $t, t' \in \mathbb{R}$, so consequently for $m \geq 1$ $\|h_1(X_n, t)\|_m \leq 1$ and $\left\|h_1(X_n, t) - h_1(X_n, t')\right\|_m \leq |t - t'|^{\frac{1}{m}}$. So this more general version can be proved along the lines of the proof in [11].

3. If $U(t) = t$ does not hold, note that $Eh_1(X_i, t_p) = U(t_p) = p$ with $t_p = U^{-1}(p) := \inf\{t \in \mathbb{R} | U(t) \geq p\}$, because U is continuous. Clearly, Assumptions 1 and 2 hold for $h(x, y, U^{-1}(p))$. Furthermore, notice that if $U(t) = U(s)$, we have that $h_1(X_i, t) = h_1(X_i, s)$ 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. From the first two parts of the proof, we know that there is a centered Gaussian process $K^\star$ with covariance function

$$E[K^\star(p, s)K^\star(p', s')] = \min\left\{s, s'\right\}\Gamma(t_p, t_{p'})$$

with

$$\sup_{\substack{p\in[0,1]\\s\in[0,1]}}\frac{1}{\sqrt{n}}\left|\left(2\sum_{1\le i\le ns}h_1(X_i, t_p) - K^\star(p, ns)\right)\right| = O\left(\log^{-\frac{1}{3840}}n\right)$$

almost surely. The Gaussian process K with $K(t, s) = K^\star(U(t), s)$ has the required covariance function and

$$\sup_{\substack{t\in\mathbb{R}\\s\in[0,1]}}\frac{1}{\sqrt{n}}\left|\left(2\sum_{1\le i\le ns}h_1(X_i, t) - K(t, ns)\right)\right|$$

$$= \sup_{\substack{t\in\mathbb{R}\\s\in[0,1]}}\frac{1}{\sqrt{n}}\left|\left(2\sum_{1\le i\le ns}h_1(X_i, t_{U(t)}) - K^\star(U(t), ns)\right)\right| = O\left(\log^{-\frac{1}{3840}}n\right). \quad\square$$

Lemma 3.2. *Let C_3, C_4, L be positive constants. Under Assumption 4, part 1 (strong mixing), there exists a constant C, such that for all measurable, non-negative functions $g : \mathbb{R} \to \mathbb{R}$ that are bounded by C_3 with $E|g(X_1) - Eg(X_1)| \ge C_4 n^{-\frac{\alpha}{\alpha+1}}$ and satisfy the variation condition with constant L, and all $n \in \mathbb{N}$ we have*

$$E\left(\sum_{i=1}^n g(X_i) - E[g(X_1)]\right)^4 \le Cn^2(\log n)^2(E|g(X_1)|)^{1+\gamma},$$

where γ is defined in Theorem 2. The same statement holds under Assumption 4, Part 2 (near epoch dependence on absolutely regular sequence) for functions $g : \mathbb{R} \to \mathbb{R}$ with $E|g(X_1) - Eg(X_1)| \ge C_4 n^{-\frac{\beta}{\beta+1}}$.

This is Lemma 3.4 respectively 3.6 of [34].

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

$$\sum_{i_1, j_1, i_2, j_2=1}^n\left|E\left[h_2(X_{i_1}, X_{j_1}, t)h_2(X_{i_2}, X_{j_2}, t)\right]\right| \le Cn^2.$$

This is Lemma 4.4 of [18].

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

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

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. Without loss of generality, we can assume that $U(t) = t$, otherwise we use the same transformation as in the proof of Proposition 3.1 and study the kernel function $h(x, y, U^{-1}(p))$. We define $Q_n(t) := \sum_{1 \leq i < j \leq n} h_2(X_i, X_j, t)$. For $l \in \mathbb{N}$, let $k = k_l = 2^{\lceil \frac{5}{8}l \rceil}$ and $t_{r,l} = \frac{r}{k_l}$ for $r = 0, \ldots, k_l$, so that $Cn^{\frac{5}{8}} \leq t_{r,l} - t_{r-1,l} = \frac{1}{k_l} \leq C'n^{\frac{5}{8}}$ for all $n \in \mathbb{N}$ with $2^{l-1} \leq n < 2^l$ and some constants C, C'. By Assumption 1, h and U are non-decreasing in t, so we have for any $t \in [t_{r-1,l}, t_{r,l}], n < 2^l$

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

So we have that

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

We will treat these three summands separately. By the choice of $t_1, \ldots, t_{k-1}$, we have $|U(t_{r,l}) - U(t_{r-1,l})| = t_{r,l} - t_{r-1,l} = \frac{1}{k_l}$, so for the last summand and $2^{l-1} \leq n < 2^l$ we know that $\max_{r=0,\ldots,k} n(n-1)|U(t_{r,l}) - U(t_{r-1,l})| \leq Cn^{2-\frac{5}{8}} = o\left(n^{\frac{3}{2}-\frac{\gamma}{8}}\right)$. For the first summand, we obtain by similar arguments as the ones used by Wu [36] to prove his inequality (6) of his Proposition 1 or by Dehling and Wendler [18] to prove their line (5)

$$
\begin{aligned}
E &\left[\max_{n=2^{l-1},\ldots,2^l-1} \max_{r=0,\ldots,k} |Q_n(t_{r,l})|^2 \right] \\
&\leq \sum_{r=0}^{k} E \left[\left(\sum_{d=1}^{l} \max_{i=1,\ldots,2^{l-d}} |Q_{i2^{d-1}}(t_{r,l}) - Q_{(i-1)2^{d-1}}(t_{r,l})| \right)^2 \right]
\end{aligned}
$$

$$\leq \sum_{r=0}^{k} l \sum_{d=1}^{l} \sum_{i=1}^{2^{l-d}} E\left[\left(Q_{i2^{d-1}}(t_{r,l}) - Q_{(i-1)2^{d-1}}(t_{r,l})\right)^2\right]$$

$$\leq \sum_{r=0}^{k} l \sum_{d=1}^{l} \sum_{i_1,j_1,i_2,j_2=1}^{2^l} \left| E\left[h_2(X_{i_1}, X_{j_1}, t) h_2(X_{i_2}, X_{j_2}, t)\right]\right|$$

$$\leq C k l^2 2^{2(l+1)} \leq C l^2 2^{\left(2+\frac{5}{8}\right)l},$$

where we used Lemma 3.3 in the last line. With the Chebyshev inequality, it follows for every $\epsilon > 0$

$$\sum_{l=1}^{\infty} P\left[\max_{n=2^{l-1},\dots,2^l-1} \max_{r=0,\dots,k} \left|Q_n(t_{r,l})\right| > \epsilon 2^{l\left(\frac{3}{2}-\frac{\gamma}{8}\right)}\right]$$

$$\leq \sum_{l=1}^{\infty} \frac{1}{\epsilon^2 2^{l\left(3-\frac{\gamma}{4}\right)}} E\left[\max_{n=1,\dots,2^l} \max_{r=0,\dots,k} \left|Q_n(t_{r,l})\right|^2\right]$$

$$\leq \sum_{l=1}^{\infty} \frac{1}{\epsilon^2 2^{l\left(3-\frac{\gamma}{4}\right)}} l^2 2^{\left(2+\frac{5}{8}\right)l} < \infty,$$

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

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

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

$$E\left(\max_{n=2^{l-1},\dots,2^l-1} \max_{r=1,\dots,k} (n-1) \left|\sum_{i=1}^{n}(h_1(X_i, t_{r,l}) - h_1(X_i, t_{r-1,l}))\right|\right)^4$$

$$\leq 2^{4(l+1)} \sum_{r=1}^{k} E\left(\max_{n=2^{l-1},\dots,2^l-1} \left|\sum_{i=1}^{n}(h_1(X_i, t_{r,l}) - h_1(X_i, t_{r-1,l}))\right|\right)^4$$

$$\leq C 2^{6l} l^2 k \left(\max_{r=1,\dots,k} |t_{r,l} - t_{r-1,l}|\right)^{1+\gamma} \leq C l^2 2^{\left(6-\frac{5}{8}\gamma\right)l},$$

where we used Corollary 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 $|t_{r,l} - t_{r-1,l}| \geq \frac{1}{2^{\frac{5}{8}l}}$. We conclude that

$$\sum_{l=0}^{\infty} P\left[\max_{n=2^{l-1},\dots,2^l-1} \max_{r=1,\dots,k} (n-1) \left|\sum_{i=1}^{n}(h_1(X_i, t_{r,l}) - h_1(X_i, t_{r-1,l}))\right| > \epsilon 2^{\left(\frac{3}{2}-\frac{\gamma}{8}\right)l}\right]$$

$$\leq \sum_{l=0}^{\infty} \frac{C}{\epsilon^4 2^{l\left(6-\frac{\gamma}{2}\right)}} E\left(\max_{n=2^{l-1},\dots,2^l-1} \max_{r=1,\dots,k} (n-1) \left|\sum_{i=1}^{n}(h_1(X_i, t_{r,l}) - h_1(X_i, t_{r-1,l}))\right|\right)^4$$

$$\leq \sum_{l=0}^{\infty} \frac{C}{\epsilon^4 2^{l\left(6-\frac{\gamma}{2}\right)}} l^2 2^{\left(6-\frac{5}{8}\gamma\right)l} = \sum_{l=0}^{\infty} \frac{C l^2}{\epsilon^4 2^{\frac{\gamma}{8}l}} < \infty.$$

The Borel–Cantelli lemma completes the proof. $\quad\square$

Lemma 3.5. *Let F be a non-decreasing function, $c, l > 0$ constants and $[C_1, C_2] \subset \mathbb{R}$. If for all $t, t' \in [C_1, C_2]$ with $|t - t'| \leq l + 2c$*

$$|F(t) - F(t') - (t - t')| \leq c,$$

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

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

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

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

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

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

$$F^{-1}(p') = \inf\{t | F(t) \geq p'\} \leq F^{-1}(p) + (p' - p) + c + \epsilon.$$

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

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

On the other hand

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

So we have that

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

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

Lemma 3.6. *Under the Assumptions 1–4 for any constant $C > 0$*

$$\sup_{\substack{t,t' \in [C_1, C_2]: \\ |t-t'| \leq C\sqrt{\frac{\log\log n}{n}}}} \left|U_n(t) - U_n(t') - u(t)(t - t')\right| = o(n^{-\frac{1}{2} - \frac{\gamma}{8}} \log n).$$

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

$$\sup_{\substack{t,t' \in [C_1, C_2]: \\ |t-t'| \leq C\sqrt{\frac{\log\log n}{n}}}} \left|U(t) - U(t') - u(t)(t - t')\right| = o(n^{-\frac{1}{2} - \frac{\gamma}{8}} \log n),$$

so it suffices to show that

$$\sup_{\substack{t,t' \in [C_1, C_2]: \\ |t-t'| \leq C\sqrt{\frac{\log\log n}{n}}}} \left|U_n(t) - U_n(t') - (U(t) - U(t'))\right| = o(n^{-\frac{1}{2} - \frac{\gamma}{8}} \log n).$$

Without loss of generality, we can assume that $U(t) = t$, otherwise we use the same transformation as in the proof of Proposition 3.1 and study the kernel function $h(x, y, U^{-1}(p))$. Note that in this case, we can consider the supremum over $[0, 1]$. Furthermore, we will consider

only the case $C = 1$, we will prove

$$K_n := \sup_{\substack{t,t' \in [0,1]: \\ |t-t'| \leq \sqrt{\frac{\log \log n}{n}}}} \left| U_n(t) - U_n(t') - (t - t') \right| = o(n^{-\frac{1}{2} - \frac{\gamma}{8}} \log n).$$

For $l \in \mathbb{N}$, let $k = k_l = C2^{\lfloor \frac{1}{2}(l - \log \log l) \rfloor}$, so that for all $n = 2^{l-1}, \ldots, 2^l - 1$, we have that $\sqrt{\frac{\log \log n}{n}} \leq \frac{1}{k_l} \leq C\sqrt{\frac{\log \log n}{n}}$. We define for $r = 0, \ldots, k_l$ the real numbers $t_{r,l} := \frac{r}{k_l}$. Clearly

$$K_n \leq 2 \max_{r=1,\ldots,k} \sup_{t,t' \in [t_{r-1,l}, t_{r,l}]} \left| U_n(t) - U_n(t') - (t - t') \right|$$

$$\leq 4 \max_{r=1,\ldots,k} \sup_{t \in [t_{r-1,l}, t_{r,l}]} \left| U_n(t) - U_n(t_{r-1,l}) - (t - t_{r-1,l}) \right|.$$

Now chose $m = m_l \in \mathbb{N}$ such that $m_l k_l \approx 2^{\left(\frac{1}{2} + \frac{\gamma}{8}\right)l}$. So for all $n = 2^{l-1}, \ldots, 2^l - 1$ and some constants C, C', we have that $Cn^{-\frac{1}{2} - \frac{\gamma}{8}} \leq \frac{1}{k_l m_l} \leq C'n^{-\frac{1}{2} - \frac{\gamma}{8}}$. We define for $r = 1, \ldots, k_l$ and $r^\star = 0, \ldots, m_l$ the real numbers $t^\star_{r^\star,r,l} = t_{r,l} + \frac{r^\star}{k_l m_l}$. As U_n and U are non-decreasing, we have for $t \in (t^\star_{r^\star-1,r,l}, t^\star_{r^\star,r,l})$

$$\left| U_n(t) - U_n(t_{r-1,l}) - (t - t_{r-1,l}) \right| \leq \max \left\{ \left| U_n(t^\star_{r^\star,r,l}) - U_n(t_{r-1,l}) - (t - t_{r-1,l}) \right|, \right.$$
$$\left. \left| U_n(t^\star_{r^\star-1,r,l}) - U_n(t_{r-1,l}) - (t - t_{r-1,l}) \right| \right\}$$
$$\leq \max \left\{ \left| U_n(t^\star_{r^\star,r,l}) - U_n(t_{r-1,l}) - (t^\star_{r^\star,r,l} - t_{r-1,l}) \right|, \right.$$
$$\left. \left| U_n(t^\star_{r^\star-1,r,l}) - U_n(t_{r-1,l}) - (t^\star_{r^\star-1,r,l} - t_{r-1,l}) \right| \right\}$$
$$+ |t^\star_{r^\star,r,l} - t^\star_{r^\star-1,r,l}|,$$

and consequently

$$K_n \leq 4 \max_{r=1,\ldots,k} \max_{r^\star=1,\ldots,m} \left| U_n(t^\star_{r^\star,r,l}) - U_n(t_{r-1,l}) - (t^\star_{r^\star,r,l} - t_{r-1,l}) \right|$$

$$+ 4 \max_{r=1,\ldots,k} \max_{r^\star=1,\ldots,m} |t^\star_{r^\star,r,l} - t^\star_{r^\star-1,r,l}|$$

$$\leq 8 \max_{r=1,\ldots,k} \max_{r^\star=1,\ldots,m} \left| \frac{1}{n} \sum_{1 \leq i \leq n} h_1(X_i, t^\star_{r^\star,r,l}) - \frac{1}{n} \sum_{1 \leq i \leq n} h_1(X_i, t_{r-1,l}) \right|$$

$$+ 4 \max_{r=1,\ldots,k} \max_{r^\star=1,\ldots,m} \left| \frac{2}{n(n-1)} \left(\sum_{1 \leq i < j \leq n} h_2(X_i, X_j, t^\star_{r^\star,r,l}) \right. \right.$$

$$\left. \left. - \sum_{1 \leq i < j \leq n} h_2(X_i, X_j, t_{r-1,l}) \right) \right|$$

$$+ 4 \max_{r=1,\ldots,k} \max_{r^\star=1,\ldots,m} |t^\star_{r^\star,r,l} - t^\star_{r^\star-1,r,l}|.$$

By our construction of the numbers $t^\star_{r^\star,r,l}$, we have that $t^\star_{r^\star,r,l} - t^\star_{r^\star-1,r,l} = \frac{1}{k_l m_l}$ and obtain for all $n = 2^{l-1}, \ldots, 2^l - 1$

$$\max_{r=1,\ldots,k} \max_{r^\star=1,\ldots,m} |t^\star_{r^\star,r,l} - t^\star_{r^\star-1,r,l}| \leq \sup_{t \in [C_1,C_2]} u(t) 2^{-\left(\frac{1}{2} - \frac{\gamma}{4}\right)l}$$

$$\leq Cn^{-\frac{1}{2} - \frac{\gamma}{8}} = o(n^{-\frac{1}{2} - \frac{\gamma}{8}} \log n).$$

With the help of Lemma 3.4, it follows that

$$\max_{r=1,\ldots,k} \max_{r^\star=1,\ldots,m} \left| \frac{2}{n(n-1)} \left(\sum_{1 \le i < j \le n} h_2(X_i, X_j, t^\star_{r^\star,r,l}) - \sum_{1 \le i < j \le n} h_2(X_i, X_j, t_{r-1,l}) \right) \right|$$

$$\le \frac{4}{n(n-1)} \sup_{t \in \mathbb{R}} \left| \sum_{1 \le i < j \le n} h_2(X_i, X_j, t) \right| = o\left(n^{-\frac{1}{2} - \frac{\gamma}{8}} \right).$$

Furthermore, we have for the linear part by Lemma 3.2 and Corollary 1 of [28] (which gives moment bounds for the maximum other multidimensional partial sums)

$$E \left[\left(\max_{n=2^{l-1},\ldots,2^l-1} \max_{r=1,\ldots,k} \max_{r^\star=1,\ldots,m} \left| \sum_{i=1}^{n} h_1(X_i, t^\star_{r^\star-1,r,l}) - \sum_{i=1}^{n} h_1(X_i, t_{r-1,l}) \right| \right)^4 \right]$$

$$\le \sum_{r=1}^{k} E \left[\left(\max_{n=2^{l-1},\ldots,2^l-1} \max_{m_1=1,\ldots,m} \left| \sum_{i=1}^{n} \sum_{r^\star=1}^{m_1} \left(h_1(X_i, t^\star_{r^\star,r,l}) - h_1(X_i, t^\star_{r^\star-1,r,l}) \right) \right| \right)^4 \right]$$

$$\le C k 2^{2l} l^2 \left(\sqrt{\frac{\log l}{2^l}} \right)^{1+\gamma} = C l^2 (\log l)^{\frac{\gamma}{2}} 2^{(2-\frac{\gamma}{2})l},$$

as $E|h_1(X_i, t) - h_1(X_i, t')| \le |t - t'|$ and by our construction $t^\star_{m,r,l} - t^\star_{0,r,l} = t_{r+1,l} - t_{r,l} = \frac{1}{k_l} \le C\sqrt{\frac{\log l}{2^l}}$. So we can conclude that for any $\epsilon > 0$

$$\sum_{l=1}^{\infty} P \left[\max_{n=2^{l-1},\ldots,2^l-1} \max_{r \le k} \max_{r^\star \le m} \left| \sum_{i=1}^{n} \left(h_1(X_i, t^\star_{r^\star-1,r,l}) - h_1(X_i, t_{r-1,l}) \right) \right| \ge \epsilon 2^{\frac{1}{2} - \frac{\gamma}{8} l} l \right]$$

$$\le C \sum_{l=1}^{\infty} \frac{2^{(2-\frac{\gamma}{2})l} l^2 (\log l)^{\frac{\gamma}{2}}}{\epsilon^4 l^4 2^{(2-\frac{\gamma}{2})l}} = C \sum_{l=1}^{\infty} \frac{(\log l)^{\frac{\gamma}{2}}}{l^2} < \infty.$$

With the Borel–Cantelli lemma, it follows that

$$\max_{r=1,\ldots,k} \max_{r^\star=1,\ldots,m} \left| \sum_{1 \le i \le n} h_1(X_i, t^\star_{r^\star,r,l}) - \sum_{1 \le i \le n} h_1(X_i, t_{r-1,l}) \right| = o(n^{\frac{1}{2} - \frac{\gamma}{8}} \log n)$$

almost surely and finally

$$\max_{r=1,\ldots,k} \max_{r^\star=1,\ldots,m} \left| \frac{1}{n} \sum_{1 \le i \le n} h_1(X_i, t^\star_{r^\star,r,l}) - \frac{1}{n} \sum_{1 \le i \le n} h_1(X_i, t_{r-1,l}) \right| = o(n^{-\frac{1}{2} - \frac{\gamma}{8}} \log n). \quad \square$$

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

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

Let K be a Gaussian process as in Proposition 3.1. Then

$$\sup_{\substack{t \in \mathbb{R} \\ s \in [0,1]}} \frac{1}{\sqrt{n}} \left| \lfloor ns \rfloor (U_{\lfloor ns \rfloor}(t) - U(t)) - K(t, ns) \right|$$

$$\leq \sup_{\substack{t \in \mathbb{R} \\ s \in [0,1]}} \frac{1}{\sqrt{ns}} \left| \left(2 \sum_{1 \leq i \leq ns} h_1(X_i, t) - K(t, ns) \right) \right| + \sup_{\substack{t \in \mathbb{R} \\ s \in [0,1]}} \frac{1}{n^{\frac{3}{2}} s} \left| \sum_{1 \leq i < j \leq ns} h_2\left(X_i, X_j, t \right) \right|$$

$$= O\left(\log^{-\frac{1}{3840}} n \right),$$

as by Lemma 3.4, we have

$$\sup_{\substack{t \in \mathbb{R} \\ s \in [0,1]}} \frac{1}{n^{\frac{3}{2}} s} \left| \sum_{1 \leq i < j \leq ns} h_2\left(X_i, X_j, t \right) \right| \leq n^{-\frac{\gamma}{8}} \sup_{\substack{t \in \mathbb{R} \\ n'=1,\ldots,n}} \frac{1}{(n')^{\frac{3}{2} - \frac{\gamma}{8}}} \left| \sum_{1 \leq i < j \leq n'} h_2\left(X_i, X_j, t \right) \right|$$

$$= O(n^{-\frac{\gamma}{8}}). \quad \square$$

Proof of Theorem 2. To simplify the notation, we will without loss of generality assume that $U(p) = p = t_p$ on the interval I. In the general case, one has to change the function $h(x, y, t)$ to $h(x, y, U^{-1}(t))$, as $Eh(X, Y, U^{-1}(p)) = U(U^{-1}(p)) = p$. The related empirical U-process $U_n \circ U^{-1}$, we have

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

$$= \frac{1}{u(t_p)} \left((U_n \circ U^{-1})^{-1}(p) - p - (p - U_n \circ U^{-1}(p)) \right)$$

$$+ o((U_n^{-1}(p) - U^{-1}(p))^{\frac{5}{4}}),$$

so Assumption 3 guarantees that $R_n(p)$ is only blown up by a constant because of this transformation. If $U(p) = p = t_p$, then we can write $R_n(p)$ as

$$R_n(p) = U_n^{-1}(p) - t_p + U_n(t_p) - p$$

$$= \left(U_n^{-1}(p) - U_n^{-1}(U_n(t_p)) + U_n(t_p) - p \right) + \left(U_n^{-1}(U_n(t_p)) - t_p \right).$$

Applying Lemmas 3.6 and 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

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

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

$$\sup_{p \in I} \left| U_n^{-1}(p) - U_n^{-1}(U_n(t_p)) + U_n(t_p) - p \right|$$

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

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 $U_n^{-1}(U_n(t_p)) - t_p \leq -\epsilon n^{-\frac{1}{2}-\frac{\gamma}{8}} \log n$ only if $U_n(t_p - \epsilon n^{-\frac{1}{2}-\frac{\gamma}{8}} \log n) \geq U_n(t_p)$. So we can conclude that

$$
P\left[\sup_{p \in I} |U_n^{-1}(U_n(t_p)) - t_p| > \epsilon n^{-\frac{1}{2}-\frac{\gamma}{8}} \log n \text{ i.o.} \right]
$$

$$
\leq P\left[\sup_{t \in [C_1, C_2 - \epsilon n^{-\frac{1}{2}-\frac{\gamma}{8}} \log n]} U_n(t + \epsilon n^{-\frac{1}{2}-\frac{\gamma}{8}} \log n) - U_n(t) \leq 0 \text{ i.o.} \right]
$$

$$
\leq P\left[\sup_{\substack{t,t' \in [C_1,C_2] \\ |t-t'|=\epsilon n^{-\frac{1}{2}-\frac{\gamma}{8}} \log n}} \left| U_n(t) - U_n(t') - (U(t) - U(t')) \right| \geq |U(t) - U(t')| \text{ i.o.} \right]
$$

$$
\leq P\left[\sup_{\substack{t,t' \in [C_1,C_2] \\ |t-t'| \leq \epsilon n^{-\frac{1}{2}-\frac{\gamma}{8}} \log n}} \left| U_n(t) - U_n(t') - (U(t) - U(t')) \right| \geq \frac{\epsilon \log n}{n^{\frac{1}{2}+\frac{\gamma}{8}} \inf_{t \in [C_1,C_2]} u(t)} \text{ i.o.} \right]
$$

$$
= 0,
$$

where the last line is a consequence of Lemma 3.6. We have proved that $\sup_{p \in I} |R_n(p)| = o(n^{-\frac{1}{2}-\frac{\gamma}{8}} \log n)$, and can finally conclude that

$$
\frac{n^{\frac{\gamma}{8}}}{\log n} \sup_{\substack{p \in I \\ s \in [0,1]}} \frac{\lfloor ns \rfloor}{\sqrt{n}} |R_{\lfloor ns \rfloor}(p)|
$$

$$
\leq \sup_{n' \leq \sqrt{n}} \left(\frac{n'}{n} \right)^{\frac{1}{2}-\frac{\gamma}{8}} \frac{\log n'}{\log n} \frac{n'^{\frac{1}{2}+\frac{\gamma}{8}}}{\log n'} \sup_{p \in I} |R_{n'}(p)| + \sup_{\sqrt{n} \leq n' \leq n} \frac{n'^{\frac{1}{2}+\frac{\gamma}{8}}}{\log n'} \sup_{p \in I} |R_{n'}(p)|
$$

$$
\leq C n^{-\frac{1}{4}+\frac{\gamma}{16}} \sup_{n' \in \mathbb{N}} \sup_{p \in I} |R_{n'}(p)| + \sup_{n' \geq \sqrt{n}} \frac{n'^{\frac{1}{2}+\frac{\gamma}{8}}}{\log n'} \sup_{p \in I} |R_{n'}(p)| \to 0. \quad \square
$$

Proof of Theorem 3. Define $K'(p,s) := -\frac{1}{u(t_p)} K(t_p, s)$, there K is the Gaussian process introduced in Theorem 1. K' is then a Gaussian process with covariance function

$$
E K'(p,s) K'(p',s') = \min\{s,s'\} \frac{1}{u(t_p)u(t_{p'})} \Gamma(t_p, t_{p'})
$$

and by Theorems 1 and 2

$$
\sup_{\substack{p \in I \\ s \in [0,1]}} \frac{1}{\sqrt{n}} \left| \lfloor ns \rfloor (U_{\lfloor ns \rfloor}^{-1}(p) - t_p) - K'(p, ns) \right|
$$

$$
\leq \sup_{\substack{p \in I \\ s \in [0,1]}} \frac{1}{\sqrt{n}} \left| \lfloor ns \rfloor \left(U_{\lfloor ns \rfloor}^{-1}(p) - t_p - \frac{p - U_n(t_p)}{u(t_p)} \right) \right|
$$

$$+ \sup_{\substack{p \in I \\ s \in [0,1]}} \frac{1}{\sqrt{n}} \frac{1}{u(t_p)} \left| \lfloor ns \rfloor (U_{\lfloor ns \rfloor}(t_p) - p) - K(t_p, ns) \right|$$

$$\leq \sup_{\substack{p \in I \\ s \in [0,1]}} \frac{\lfloor ns \rfloor}{\sqrt{n}} |R_{\lfloor ns \rfloor}(p)| + \frac{1}{\inf_{p \in I} u(t_p)} \sup_{\substack{p \in I \\ s \in [0,1]}} \frac{1}{\sqrt{n}} \left| \lfloor ns \rfloor (U_{\lfloor ns \rfloor}(t_p) - p) - K(t_p, ns) \right|$$

$$= O \left(\log^{-\frac{1}{3840}} n \right)$$

almost surely. $\square$

Proof of Theorem 4. If $\sigma^2 > 0$, set

$$B(s) = \frac{1}{\sigma} T(K'(\cdot, s)) = \int_I J(p) K'(p, s) dp + \sum_{j=1}^d b_j U_n(p_j).$$

In the case $\sigma^2 = 0$, B may be an arbitrary Brownian motion. As J is a bounded function, T is a linear and Lipschitz continuous functional (with respect to the supremum norm), so

$$\sup_{s \in [0,1]} \frac{1}{\sqrt{n}} \left| \lfloor ns \rfloor (T(U_{\lfloor ns \rfloor}^{-1}) - T(U^{-1})) - \sigma B(ns) \right|$$

$$= \sup_{s \in [0,1]} \frac{1}{\sqrt{n}} \left| T \left(\lfloor ns \rfloor (U_{\lfloor ns \rfloor}^{-1} - U^{-1}) - K'(\cdot, ns) \right) \right|$$

$$\leq C \sup_{\substack{p \in I \\ s \in [0,1]}} \frac{1}{\sqrt{n}} \left| \lfloor ns \rfloor (U_{\lfloor ns \rfloor}^{-1}(p) - t_p) - K'(p, ns) \right|$$

$$= O \left(\log^{-\frac{1}{3840}} n \right).$$

It remains to show that B is a Brownian motion. Clearly, $E B(s) = 0$ for every $s \geq 0$. By the linearity of T, B is a Gaussian process with stationary independent increments. Furthermore

$$E[B^2(s)] = \frac{1}{\sigma^2} \int_{\tilde{C}_1}^{\tilde{C}_2} \int_{\tilde{C}_1}^{\tilde{C}_2} \frac{E[K(t_p, s) K(t_q, s)]}{u(t_p) u(t_q)} J(p) J(q) dp dq$$

$$+ \frac{1}{\sigma^2} 2 \sum_{j=1}^d b_j \int_{\tilde{C}_1}^{\tilde{C}_2} \frac{E[K(t_{p_j}, s) K(t_q, s)]}{u(t_{p_j}) u(t_p)} J(p) dp$$

$$+ \frac{1}{\sigma^2} \sum_{i,j=1}^d b_i b_j \frac{E[K(t_{p_i}, s) K(t_{p_j}, s)]}{u(t_{p_i}) u(t_{p_j})}$$

$$= s. \quad \square$$

Acknowledgments

References

[1] D.W.K. Andrews, Non-strong mixing autoregressive processes, J. Appl. Probab. 21 (1984) 930–934.

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

[3] M.A. Arcones, On the law of the iterated logarithm for Gaussian processes, J. Theoret. Probab. 8 (1995) 877–903.

[4] M.A. Arcones, The Bahadur–Kiefer representation for U-quantiles, Ann. Statist. 24 (1996) 1400–1422.

[5] M.A. Arcones, E. Giné, Limit theorems for U-processes, Ann. Probab. 21 (1993) 1494–1542.

[6] 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.

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

[8] N. Aronszajn, Theory of reproducing kernels, Trans. Amer. Math. Soc. 3 (1950) 337–404.

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

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

[11] 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.

[12] 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.

[13] R.C. Bradley, Introduction to Strong Mixing Conditions, vols. 1–3, Kendrick Press, Heber City, 2007.

[14] 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.

[15] M. Cörgő, P. Révész, Strong approximations of the quantile process, Ann. Statist. 4 (1978) 882–894.

[16] H. Dehling, M. Denker, W. Philipp, The almost sure invariance principle for the empirical process of U-statistic structure, Ann. Inst. H. Poincaré 23 (1987) 121–134.

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

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

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

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

[21] B.E. Hansen, GARCH(1,1) processes are near epoch dependent, Econom. Lett. 36 (1991) 181–186.

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

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

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

[25] J. Kiefer, Skorohod embedding of multivariate RV's, and the sample DF, Probab. Theory Related Fields 24 (1972) 1–35.

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

[27] T.L. Lai, Reproducing kernel Hilbert spaces and the law of the iterated logarithm for Gaussian processes, Probab. Theory Related Fields 29 (1974) 7–19.

[28] F. Móricz, A general moment inequality for the maximum of the rectangular partial sums of multiple series, Acta Math. Hungar. 43 (1983) 337–346.

[29] D.W. Müller, On Glivenko–Cantelli convergence, Probab. Theory Related Fields 16 (1970) 195–210.

[30] W. Philipp, A functional law of the iterated logarithm for empirical functions of weakly dependent random variables, Ann. Probab. 5 (1977) 319–350.

[31] P.J. Rousseeuw, C. Croux, Alternatives to the median absolute deviation, J. Amer. Statist. Soc. 88 (1993) 1273–1283.

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

[33] R.J. Serfling, Generalized L-, M-, and R-statistics, Ann. Probab. 12 (1984) 76–86.

[34] M. Wendler, Bahadur representation for U-quantiles of dependent data, J. Multivariate Anal. 102 (2011) 1064–1079.

[35] W.B. Wu, On the Bahadur representation of sample quantiles for dependent sequences, Ann. Statist. 33 (2005) 1934–1963.

[36] W.B. Wu, Strong invariance principles for dependent random variables, Ann. Probab. 35 (2007) 2294–2320.

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