Bahadur-Kiefer theory for sample quantiles of weakly dependent linear processes

Rafał Kulik
University of Sydney
School of Mathematics and Statistics
F07, University of Sydney
NSW 2006, Australia
rkuli@maths.usyd.edu.au

July 16, 2007

Abstract

In this paper we establish the Bahadur-Kiefer representation for sample quantiles for a class of weakly dependent linear processes. The rate of approximation is the same as for i.i.d. sequences and thus it is optimal.

Keywords: Bahadur representation, empirical processes, law of the iterated logarithms, linear processes, sample quantiles, strong approximation.
Running title: Sample quantiles and weak dependence

1 Introduction

Consider the class of stationary linear processes

$$X_i = \sum_{k=0}^{\infty} c_k \epsilon_{i-k} \ ,$$

where $\{\epsilon_i, i \in \mathbb{Z}\}$ is an i.i.d. sequence and $\sum_{k=0}^{\infty} |c_k| < \infty$. Assume that X_1 has continuous distribution function $F(x) = P(X_1 \le x)$ and let f and Q denote the associated density and quantile function. Given a sample $X_1, \ldots, X_n$, let $F_n(x) = \frac{1}{n}\sum_{i=1}^{n} 1_{\{X_i \le x\}}$ and let $Q_n(y)$ denote the corresponding empirical quantile function. Define

$$\beta_n(x) = n^{1/2}(F_n(x) - F(x)), \quad x \in \mathbb{R} \ ,$$

$$q_n(y) = n^{1/2}(Q(y) - Q_n(y)), \quad y \in (0,1) \ ,$$

the *general empirical* and the *general quantile* processes, respectively. With $U_i = F(X_i)$, $i \geq 1$, let $E_n(x) = \frac{1}{n}\sum_{i=1}^{n} 1_{\{U_i \leq x\}} = F_n(Q(x))$ and $U_n(y)$ be the uniform empirical distribution and uniform empirical quantile functions. Let

$$\alpha_n(x) = n^{1/2}(E_n(x) - x), \quad x \in (0,1) \ ,$$

$$u_n(y) = n^{1/2}(y - U_n(y)), \quad y \in (0,1) \ ,$$

be the corresponding *uniform empirical* and *uniform quantile* processes.

Assume initially that X_i, $i \geq 1$, are i.i.d. Fix $y \in (0,1)$. Let I_y be a neighborhood of $Q(y)$. Assuming that $\inf_{x \in I_y} f(x) > 0$ and $\sup_{x \in I_y} |f^{'}(x)| < \infty$, Bahadur, [2], obtained the following representation

$$f(Q(y))q_n(y) - \alpha_n(y) =: R_n(y), \tag{1}$$

where $R_n(y) = O_{a.s}(n^{-1/4}(\log n)^{1/2}(\log\log n)^{1/4})$. Later Kiefer, [11], proved that this can be strengthened to

$$R_n(y) = O_{a.s}(n^{-1/4}(\log\log n)^{3/4}), \tag{2}$$

which is the optimal rate. Continuing his study Kiefer, [12], established the uniform version of (1), referred later as the *Bahadur-Kiefer representation.* For $0 \leq a < b \leq 1$,

$$\sup_{y \in (a,b)} |f(Q(y))q_n(y) - \alpha_n(y)| =: R_n, \tag{3}$$

where

$$R_n = O_{a.s}(n^{-1/4}(\log n)^{1/2}(\log\log n)^{1/4}). \tag{4}$$

The above rate is also optimal. Kiefer obtained his result assuming

(K1) $\sup_{y \in (a,b)} |f^{'}(Q(y))| < \infty$,

(K2) $\inf_{y \in (a,b)} f(Q(y)) > 0$.

We shall refer to (K1)-(K2) as the *Kiefer conditions.* In particular, if $a = 0, b = 1$, (K1)-(K2) imply that f has a finite support.

Further, Csörgő and Révész, [5], relaxed the conditions for Kiefer's result (3) and introduced the *Csörgő-Révész conditions* (see Section 2; cf. also [4, Theorem 3.2.1]).

The main purpose of this paper is to obtain Bahadur-Kiefer type representations for sample quantiles of linear processes with the *optimal* rate. Some results are available for weakly dependent random variables. For ϕ-mixing sequences, under (K1)-(K2), the optimal rates have been obtained in [1]. These results were improved in [7] and [19] through less restrictive mixing rates and Csörgő-Révész-type conditions. The rate of approximation was $R_n = O_{a.s}((\log n)^{-\lambda})$ for some $\lambda > 0$.

Mixing is rather hard to verify and requires additional assumptions. In particular, to obtain a strong mixing for linear processes both regularity of the density of ϵ_1 and some constraints on c_k's are required (see e.g. [6] or [16]). Nevertheless, even if we are able to establish strong mixing, we do not attain the optimal rate in the Bahadur-Kiefer representation.

Another way of looking at linear processes is to approximate the sequence $\{X_i, i \geq 1\}$ by a sequence with finite memory and then use the classical Bernstein blocking technique. Hesse [8] obtained the Bahadur representation (1) with rate $R_n(y) = O_{a.s}(n^{-1/4+\lambda})$ for some $\lambda > 0$ via this technique. The method avoids some assumptions on the density of ϵ_1, but it leads to restrictive constraints on c_k's, and does not lead to the optimal rates.

The blocking technique or mixing require some strong assumptions. To overcome such restrictions, Ho, Hsing, Mielniczuk and Wu (see [9], [10], [17], [18]) developed a martingale based methods. In particular, for a class of linear processes Wu [17] obtained the exact rate (2) in the Bahadur representation (1). He also studied the Bahadur-Kiefer representation but was not able to attain the optimal rate (4) due to a lack of an appropriate version of the law of iterated logarithm for empirical processes.

In this paper we shall combine the Bernstein's blocking technique together with Wu's method to obtain the optimal rate in the Bahadur-Kiefer representation (3) under quite mild conditions on c_k's. The result is stated in Theorem 2.1. The methodology involves the recent strong approximation result of Berkes and Horváth [3]. Since part of our computations follows their proof, we include it in the Appendix. Further, we shall obtain the optimal rate under general conditions on F, on the whole interval $(0, 1)$, by considering an appropriately weighted process $f(Q(y))q_n(y) - \alpha_n(y)$ as in Theorem 2.2.

Throughout the paper C will denote a generic constant which may be different at each appearance. Also, we write $a_n \sim b_n$ if $\lim_{n\to\infty} a_n/b_n = 1$. For any stationary sequence $\{Z_i, i \geq 1\}$ of random variables, Z will be a random variables with the same distribution as Z_1.

Some further notation. Let $b_n = n^{-1/4}(\log n)^{1/2}(\log\log n)^{1/4}$ and $\lambda_n = n^{-1/2}(2\log\log n)^{1/2}$. For any function $h(x)$ defined on $\mathbb{R}$ and $x < y$, we write $h(x,y) := h(y) - h(x)$.

2 Results

Assume that the following moment and dependence conditions hold. For $\alpha \geq 2$,

$$\mathrm{E}|\epsilon|^\alpha < \infty. \tag{5}$$

and for some $\rho \in (0, \frac{1}{2})$,

$$\sum_{k=i}^{\infty} c_k^2 = O(i^{-2/\rho}(\log i)^{-3}). \tag{6}$$

Theorem 2.1 *Assume (5), (6) and (K1)-(K2). Furthermore, assume that for f_ϵ, a density of ϵ, we have*

$$\sup_{x\in\mathbf{R}}(f_\epsilon(x) + |f'_\epsilon(x)| + |f''_\epsilon(x)|) < \infty. \tag{7}$$

Then,

$$\sup_{y\in(a,b)} |f(Q(y))q_n(y) - \alpha_n(y)| = O_{a.s}(n^{-1/4}(\log n)^{1/2}(\log\log n)^{1/4}).$$

To obtain the bound without the Kiefer conditions (K1)-(K2), we shall consider *Csörgő-Révész conditions*:

(CsR1) f' exists on (a,b), where $a = \sup\{x : F(x) = 0\}$, $b = \inf\{x : F(x) = 1\}$, $-\infty \leq a < b \leq \infty$,

(CsR2) $\inf_{x\in(a,b)} f(x) > 0$,

(CsR3(i)) $f(Q(y)) \sim y^{\gamma_1} L_1(y^{-1})$ as $y \downarrow 0$;

(CsR3(ii)) $f(Q(y)) \sim (1-y)^{\gamma_2} L_2((1-y)^{-1})$ as $y \uparrow 1$;

(CsR4) (i) $0 < A := \lim_{y\downarrow 0} f(Q(y)) < \infty$, $0 < B := \lim_{y\uparrow 1} f(Q(y)) < \infty$, or
(ii) if $A = 0$ (respectively $B = 0$) then f is nondecreasing (respectively nonincreasing) on an interval to the right of $Q(0+)$ (respectively to the left of $Q(1-)$).

Theorem 2.2 *Assume Csörgő-Révész conditions with* $\gamma := \min\{\gamma_1, \gamma_2\} \geq 1$. *As in Theorem 2.1, assume (5), (6) and (7). Then for arbitrary* $\nu > \max\{2\gamma, 3\gamma - 2\}$,

$$\sup_{y\in(0,1)} (y(1-y))^{\nu} \left|f(Q(y))q_n(y) - \alpha_n(y)\right| = O_{a.s}(n^{-1/4}(\log n)^{1/2}(\log\log n)^{1/4}).$$

3 Proof of Theorem 2.1

Denote $\mathcal{F}_i = \sigma(\epsilon_i, \epsilon_{i-1}, \ldots,)$. Note that $F(x) = \mathrm{E}F_\epsilon(x - X_{i,i-1}) = \mathrm{E}F_\epsilon(x - X_{1,0})$. Denote also $Y_i(x) = F_\epsilon(x - X_{i,i-1}) - F(x)$. Then we have

$$\begin{aligned} &\frac{1}{n}\sum_{i=1}^{n}(1_{\{X_i\leq x\}} - F(x)) = \\ &= \frac{1}{n}\sum_{i=1}^{n}\Big(1_{\{X_i\leq x\}} - \mathrm{E}(1_{\{X_i\leq x\}}|\mathcal{F}_{i-1})\Big) + \frac{1}{n}\sum_{i=1}^{n} Y_i(x) =: M_n(x) + N_n(x). \end{aligned}$$

Then $nM_n(x)$, $n \geq 1$, is a martingale and $N_n(x)$ is differentiable.

The plan of our proof is the following. First, we obtain a strong approximation of the differentiable part N_n by an appropriate Gaussian process. In order to do this, we will replace the original sequence $\{X_i, i \geq 1\}$ with a one with finite memory. From that approximation we will establish the uniform law of the iterated logarithm (ULIL) for the differentiable part N_n, which together with the ULIL for the martingale part M_n will imply the ULIL for the empirical process β_n (see Section 3.1).

Then using a modification of Lemma 13, [17], we will be able to control increments of the empirical processes β_n and α_n (Section 3.2), which together with the ULIL will imply the result (Section 3.3).

3.1 Approximation of the differentiable part by a Gaussian process and laws of the iterated logarithm

Proposition 3.1 *Assume (5), (6) and (7). Then there exists a centered Gaussian process* $K(x,t)$ *with* $\mathrm{E}K(x,t)K(y,t') = t \wedge t'\Gamma(x,y)$ *such that*

$$\sup_{0\le t\le 1}\sup_{x\in\mathbf{R}} |[nt]N_{[nt]}(x) - K(x,[nt])| = o(n^{1/2}(\log n)^{-\lambda}) \quad \text{almost surely}$$

with some $\lambda > 0$.

The law of iterated logarithm follows from Proposition 3.1 and the ULIL for K.

Corollary 3.2 *If (5), (6), (7) and Kiefer conditions are fulfilled, then*

$$\limsup_{n\to\infty} \frac{1}{(2\log\log n)^{1/2}} \sup_{y\in(0,1)} |q_n(y)| = C \quad \text{almost surely.} \tag{8}$$

Proof. Proposition 3.1, together with the ULIL for the martingale M_n (see e.g. [17, Lemma 7]) yields

$$\limsup_{n\to\infty} \frac{1}{(2\log\log n)^{1/2}} \sup_{-\infty<x<\infty} |\beta_n(x)| = C \quad \text{almost surely.}$$

Consequently,

$$\limsup_{n\to\infty} \frac{1}{(2\log\log n)^{1/2}} \sup_{y\in(0,1)} |u_n(y)| = \limsup_{n\to\infty} \frac{1}{(2\log\log n)^{1/2}} \sup_{x\in(0,1)} |\alpha_n(x)| = C, \tag{9}$$

almost surely.

With $\Delta_{n,y} = Q_n(y) - Q(y)$, we have

$$\begin{aligned}
&\limsup_{n\to\infty} \sup_{y\in(a,b)} \frac{n^{1/2}\Delta_{n,y}}{(\log\log n)^{1/2}} \\
\le\ & \limsup_{n\to\infty} \sup_{1\le k\le n} \frac{n^{1/2}|Q(F(X_{k:n})) - Q(\frac{k}{n})|}{(\log\log n)^{1/2}} + \limsup_{n\to\infty} \sup_{1\le k\le n} \sup_{y\in(\frac{k-1}{n},\frac{k}{n}]} \frac{n^{1/2}|Q(\frac{k}{n}) - Q(y)|}{(\log\log n)^{1/2}}.
\end{aligned}$$

Now, if the Kiefer conditions hold, then $\sup_{y\in(a,b)} |Q''(y)| < \infty$. Using Taylor's expansion one obtains $Q(\frac{k}{n}) = Q(y) + Q'(y)(\frac{k}{n} - y) + O\left(\frac{1}{n}\right)$ and

$$Q(F(X_{k:n})) = Q(U_{k:n}) = Q\left(\frac{k}{n}\right) + Q'\left(\frac{k}{n}\right)\left(U_{k:n} - \frac{k}{n}\right) + O_{a.s}\left(\left(U_{k:n} - \frac{k}{n}\right)^2\right).$$

Thus, by (9),

$$\limsup_{n\to\infty} \sup_{y\in(a,b)} \frac{n^{1/2}\Delta_{n,y}}{(\log\log n)^{1/2}} \le \sup_{y\in(a,b)} |Q'(y)| \limsup_{n\to\infty} \sup_{1\le k\le n} \frac{n^{1/2}|U_{k:n}-\frac{k}{n}|}{(\log\log n)^{1/2}} \le C.$$

□

Before proving Proposition 3.1, we need some notation. With ρ from (6), define

$$\hat{X}_i := \hat{X}_i(\rho) = \sum_{k=0}^{i^\rho-1} c_k\epsilon_{i-k}.$$

Without loss of generality we may assume that $c_0 = 1$. Let $X_{i,i-1} = \sum_{k=1}^{\infty} c_k\epsilon_{i-k}$ and define its truncated version $\hat{X}_{i,i-1} = \sum_{k=1}^{i^\rho-1} c_k\epsilon_{i-k}$.

By Rosenthal's inequality, for any $\alpha \ge 2$,

$$\mathrm{E}|X_i - \hat{X}_i|^\alpha = \mathrm{E}\left|\sum_{k=i^\rho}^{\infty} c_k\epsilon_{i-k}\right|^\alpha \le C\left(\sum_{k=i^\rho}^{\infty} c_k^2\right)^{\alpha/2} + C\sum_{k=i^\rho}^{\infty} |c_k|^\alpha. \tag{10}$$

This estimate is true for $\mathrm{E}|X_{i,i-1} - \hat{X}_{i,i-1}|^\alpha$ as well. Thus, replacing i with i^ρ in (6),

$$\left\|\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} \hat{X}_i\right\|_\alpha \le \sum_{i=1}^{\infty} O(i^{-1}(\log i)^{-3/2}) < \infty \tag{11}$$

and the same estimate is valid for $\left\|\sum_{i=1}^{n} X_{i,i-1} - \sum_{i=1}^{n} \hat{X}_{i,i-1}\right\|_\alpha$.

Define

$$F_n^*(x) = \frac{1}{n}\sum_{i=1}^{n} F_\epsilon(x - X_{i,i-1}),$$

a conditional empirical distribution function, and

$$\hat{F}_n^*(x) = \frac{1}{n}\sum_{i=1}^{n} F_\epsilon(x - \hat{X}_{i,i-1}),$$

its corresponding version based on the truncated random variables.

Let $\hat{F}_i(x) := P(\hat{X}_i \le x) = \mathrm{E}F_\epsilon(x - \hat{X}_{i,i-1})$, $i \ge 1$. Since f_ϵ exists, we may define $f_n^*(x) = dF_n^*(x)/dx$, $\hat{f}_n^*(x) = d\hat{F}_n^*(x)/dx$, $\hat{f}_n(x) = \frac{1}{n}\sum_{i=1}^{n} d\hat{F}_i(x)/dx$.

Further, denote $\hat{Y}_i(x) = F_\epsilon(x - \hat{X}_{i,i-1}) - \hat{F}_i(x)$ and

$$\hat{N}_n(x) = \frac{1}{n}\sum_{i=1}^{n}(F_\epsilon(x - \hat{X}_{i,i-1}) - \hat{F}_i(x)) = \frac{1}{n}\sum_{i=1}^{n}\hat{Y}_i(x).$$

Lemma 3.3 *Assume (5), (6) and (7). Then with some $D_0 > 0$,*

$$||n\hat{N}_n(x,y)||_2^2 \le D_0 n(y-x)^2. \tag{12}$$

Proof. We have for any $-\infty < x < y < \infty$,

$$\begin{aligned} ||nN_n(x,y)||_2^2 &= n^2\mathrm{E}\left|\int_x^y (f_n^*(u) - f(u))du\right|^2 \\ &\le \quad n^2(y-x)^2\mathrm{E}\sup_{x\le u\le y}|f_n^*(u) - f(u)|^2 \le Cn(y-x)^2 \end{aligned} \tag{13}$$

by Lemma 9 in [17]. Next,

$$\begin{aligned} \mathrm{E}\sup_{x\le u\le y}|\hat{f}_n^*(u) - \hat{f}_n(u)|^2 \le 3\Bigg(&\mathrm{E}\sup_{x\le u\le y}|\hat{f}_n^*(u) - f_n^*(u)|^2 \\ &+ \mathrm{E}\sup_{x\le u\le y}|f_n^*(u) - f(u)|^2 + \sup_{x\le u\le y}|f(u) - \hat{f}_n(u)|^2\Bigg). \end{aligned} \tag{14}$$

The second term is $O(n^{-1})$ as in (13). As for the first term we have

$$\begin{aligned} \mathrm{E}\sup_{x\le u\le y}|\hat{f}_n^*(u) - f_n^*(u)|^2 &= \frac{1}{n^2}\mathrm{E}\sup_{x\le u\le y}\left|\sum_{i=1}^{n}(f_\epsilon(u - X_{i,i-1}) - f_\epsilon(u - \hat{X}_{i,i-1}))\right|^2 \\ = \quad &\frac{1}{n^2}\mathrm{E}\sup_{x\le u\le y}\left|\sum_{i=1}^{n}\int_{u-X_{i,i-1}}^{u-\hat{X}_{i,i-1}} f_\epsilon'(v)dv\right|^2 \le \frac{1}{n^2}\sup_{x\in\mathbf{R}}|f_\epsilon'(x)|^2\mathrm{E}\left(\sum_{i=1}^{n}|X_{i,i-1} - \hat{X}_{i,i-1}|\right)^2 \\ \le \quad &\frac{1}{n}\sup_{x\in\mathbf{R}}|f_\epsilon'(x)|^2\left(\sum_{i=1}^{n}||X_{i,i-1} - \hat{X}_{i,i-1}||_\alpha\right)^2 \le C\frac{1}{n} \end{aligned} \tag{15}$$

as $\alpha \ge 2$ and the comment following (11). Also,

$$\sup_{x\le u\le y}|f(u) - \hat{f}_n(u)|^2 \le \sup_{x\le u\le y}\mathrm{E}|\hat{f}_n^*(u) - f_n^*(u)|^2 \le C\frac{1}{n}. \tag{16}$$

Putting together (14), (15), (16) we obtain (12).

□

Next, we derive an exponential inequality for $nN_n(x,y)$.

Lemma 3.4 *Assume (5), (6) and (7). Then for any $z > 0$*

$$P\left(\left|\sum_{i=1}^{n} Y_i(x,y)\right| > z\right) \le C_1 z^{-\alpha} + C_2 \exp(-C_3 z^2/(n(y-x)^2)) + C_4 \exp(-C_5 z/n^{\rho}),$$

where C_1, C_2, C_3, C_4, C_5 are positive constants.

Proof. With the help of differentiability and the comment following (10),

$$||F_\epsilon(x - X_{i,i-1}) - F_\epsilon(x - \hat{X}_{i,i-1})||_\alpha = O\left(\left(\left(\sum_{k=i^\rho}^{\infty} c_k^2\right)^{\alpha/2} + \sum_{k=i^\rho}^{\infty} |c_k|^\alpha\right)^{1/\alpha}\right).$$

Further,

$$|F(x) - \hat{F}_i(x)|^\alpha \le \mathrm{E}|F_\epsilon(x - X_{i,i-1}) - F_\epsilon(x - \hat{X}_{i,i-1})|^\alpha.$$

Consequently,

$$||Y_i(x) - \hat{Y}_i(x)||_\alpha = O\left(\left(\left(\sum_{k=i^\rho}^{\infty} c_k^2\right)^{\alpha/2} + \sum_{k=i^\rho}^{\infty} |c_k|^\alpha\right)^{1/\alpha}\right) \tag{17}$$

and

$$||nN_n(x) - n\hat{N}_n(x)||_\alpha \le C_0 := \sum_{i=1}^{\infty} i^{-1}(\log i)^{-3/2}.$$

From this and the Markov inequality we get

$$P(|nN_n(x,y)| > z) \le C_0 z^{-\alpha} + P(|n\hat{N}_n(x,y)| > z/2).$$

To obtain the bound for the second part, divide $[1,n]$ into blocks I_1, J_1, I_2, J_2,..., I_M, J_M with the same length n^ρ, where ρ is defined in (6). Thus, $M \sim n^{1-\rho}$. Let $\hat{U}_k = \sum_{i \in I_k} \hat{Y}_i(x,y)$, $\hat{V}_k = \sum_{i \in J_k} \hat{Y}_i(x,y)$ and $n\hat{N}_n^{(1)} = \sum_{k=1}^{M} \hat{U}_k$, $n\hat{N}_n^{(2)} = \sum_{k=1}^{M} \hat{V}_k$. Both $(\hat{U}_1, \ldots, \hat{U}_M)$ and $(\hat{V}_1, \ldots, \hat{V}_M)$ are vectors of independent random variables. Also, $\max_{k=1,\ldots,M} \hat{U}_k \le [I_k] \le Cn^\rho$. The equation (12) yields $||U_k||_2^2 \le D_0 n^\rho (y-x)^2$. Recall that $\rho \in (0, \frac{1}{2})$. Applying the result in [14, p. 293] to the centered sequence $\hat{U}_1, \ldots, \hat{U}_M$ with $M_n = n^\rho$, $B_n = D_0 n(y-x)^2$ we obtain

$$P(|n\hat{N}_n^{(1)}| > z) \le \exp(-z^2/(4D_0 n(y-x)^2)) + \exp(-z/(4n^\rho)).$$

The same applies to $P(|n\hat{N}_n^{(2)}| > z)$ and hence the result follows.

□

To state an approximation result, note that $Y_0(x)$ and $\hat{Y}_i(x)$, $i \geq 1$ are independent. Thus, we obtain

$$\begin{aligned} |\mathrm{E}Y_0(x)Y_i(y)| &\leq |\mathrm{E}Y_0(x)Y_i(y) - Y_0(x)\hat{Y}_i(y)| + |\mathrm{E}Y_0(x)\hat{Y}_i(y)| \\ &\leq ||Y_i(y) - \hat{Y}_i(y)||_\alpha + 0. \end{aligned}$$

Then, in view of (6) and (17),

$$\Gamma(x,y) = \mathrm{E}Y_0(x)Y_0(y) + \sum_{i=1}^{\infty} \left(\mathrm{E}Y_0(x)Y_i(y) + \mathrm{E}Y_0(y)Y_i(x)\right) \tag{18}$$

is absolutely convergent for all $x, y \in \mathbb{R}$.

Having Lemma 3.4, (18) and $\alpha \geq 2$ we may proceed in the very same way as in [3] to obtain the approximation result of Proposition 3.1 (see the Appendix for details).

3.2 Controlling increments

Let $l_q(n) = (\log n)^{1/q}(\log\log n)^{2/q}$, $q > 2$, and recall that $\sum_{k=0}^{\infty} |c_k| < \infty$.

First, we generalize [17, Lemma 13] to the real line.

Lemma 3.5 *Assume (5), (7) and (K1)-(K2). Let* $-\infty \leq a < b \leq \infty$. *Then, under Kiefer conditions, for any positive bounded sequence* d_n *such that* $\log n = o(nd_n)$ *we have*

$$\sup_{|x-y|\leq d_n, x,y\in[a,b]} |F_n(x) - F(x) - (F_n(y) - F(y))| = O_{a.s}\left(\frac{\sqrt{d_n \log n}}{\sqrt{n}} + \frac{d_n l_q(n)}{\sqrt{n}}\right). \tag{19}$$

Proof. Define $U_i = F(X_i)$. Clearly, in order to show (19) it suffices to prove

$$\sup_{|u-v|\leq d_n, u,v\in(a_1,b_1)} |E_n(v) - v - (E_n(u) - u)| = O_{a.s}\left(\frac{\sqrt{d_n \log n}}{\sqrt{n}} + \frac{d_n l_q(n)}{\sqrt{n}}\right). \tag{20}$$

Let $y \in (a_1, b_1)$. Decompose

$$\frac{1}{n}\sum_{i=1}^{n}(1_{\{U_i \leq y\}} - y) =$$

$$
\begin{aligned}
&= \frac{1}{n}\sum_{i=1}^{n}\Big(1_{\{U_i\le y\}} - \mathrm{E}(1_{\{U_i\le y\}}|\mathcal{F}_{i-1})\Big) + \frac{1}{n}\sum_{i=1}^{n}\Big(\mathrm{E}(1_{\{U_i\le y\}}|\mathcal{F}_{i-1}) - y\Big) \\
&=: \tilde{M}_n(x) + \tilde{N}_n(x),
\end{aligned}
$$

where

$$
\tilde{N}_n(y) = \frac{1}{n}\sum_{i=1}^{n}\Big(\mathrm{E}(1_{\{U_i\le y\}}|\mathcal{F}_{i-1}) - y\Big) = \frac{1}{n}\sum_{i=1}^{n}\left(F_\epsilon(Q(y) - X_{i,i-1}) - y\right)
$$

so that

$$
\tilde{N}_n'(y) = \frac{1}{n}\sum_{i=1}^{n}\Big(Q^{'}(y) f_\epsilon(Q(y) - X_{i,i-1}) - 1\Big) \tag{21}
$$

and

$$
\tilde{N}_n^{''}(y) = \frac{1}{n}\sum_{i=1}^{n}\Big(Q^{''}(y) f_\epsilon(Q(y) - X_{i,i-1})\Big) + (Q^{'}(y))^2 \frac{1}{n}\sum_{i=1}^{n} f_\epsilon^{'}(Q(y) - X_{i,i-1}). \tag{22}
$$

Define the projection operator $\mathcal{P}_k\xi = \mathrm{E}(\xi|\mathcal{F}_k) - \mathrm{E}(\xi|\mathcal{F}_{k-1})$. Let $g_y(\mathcal{F}_i) = f_\epsilon(Q(y) - X_{i,i-1})$. As in the proof of Lemma 3 in [17], with $\alpha = q$,

$$
||\mathcal{P}_0 g_y(\mathcal{F}_i)||_\alpha = O(|c_i|)
$$

uniformly in $y \in (a_1, b_1)$. The same holds for $g_y(\mathcal{F}_i) = f_\epsilon^{'}(Q(y) - X_{i,i-1})$. Further, under Kiefer conditions, $\sup_{y\in(a_1,b_1)}(|Q^{'}(y)| + |Q^{''}(y)|) < \infty$. Thus, applying Proposition 1 in [17], $\max_{y\in(a_1,b_1)}(||\tilde{N}_n^{'}(y)||_\alpha + ||\tilde{N}_n^{''}(y)||_\alpha) = O(n^{-1/2})$. Consequently,

$$
\mathrm{E}\left[\max_{y\in(a_1,b_1)} |\tilde{N}_n(y)|^\alpha\right] \le \mathrm{E}\left[\int_0^1 |\tilde{N}_n'(y)| dy\right]^\alpha = O(n^{-\alpha/2}).
$$

Similarly, $\mathrm{E}\left[\max_{y\in(a_1,b_1)} |\tilde{N}_n^{'}(y)|^\alpha\right] = O(n^{-\alpha/2})$. Thus, by the same argument as in [17, Lemma 9] we have $\sup_{y\in(a_1,b_1)} |\tilde{N}_n^{'}(y)| = o_{a.s}(l_q(n)/\sqrt{n})$. Further,

$$
\sup_{|u-v|\le d_n, u,v\in(a_1,b_1)} |\tilde{N}_n(v) - \tilde{N}_n(u)| \le d_n \sup_{y\in(a_1,b_1)} |\tilde{N}_n'(y)| = o_{a.s}(d_n l_q(n)/\sqrt{n}).
$$

This, together with appropriate estimates for the martingale part yields (20).

□

3.3 Conclusion of the proof of Theorem 2.1

We apply Lemma 3.5 with $d_n = \lambda_n$. Then the second part in (19) is negligible. Thus, we have

$$\sup_{x \in \mathbf{R}} \sup_{|x-y| \le \lambda_n} |\beta_n(x) - \beta_n(y)| = O_{a.s}(b_n).$$

On account of (8) it yields

$$\sup_{y \in (0,1)} |\beta_n(Q_n(y)) - \beta_n(Q(y))| = O_{a.s}(b_n).$$

Equivalently,

$$n^{1/2} \sup_{y \in (0,1)} |F_n(Q_n(y)) - F(Q_n(y)) - (F_n(Q(y)) - F(Q(y)))| = O_{a.s}(b_n).$$

Since $|F_n(Q_n(y)) - F(Q(y))| \le 1/n$ one obtains

$$n^{1/2} \sup_{y \in (0,1)} |F(Q_n(y)) - F(Q(y)) - (F(Q(y)) - F_n(Q(y)))| = O_{a.s}(b_n).$$

Set $\Delta_{n,y} = Q_n(y) - Q(y)$. Using the Taylor's expansion $F(Q_n(y)) = F(Q(y)) + f(Q(y))\Delta_{n,y} + O_{a.s}(\Delta_{n,y}^2)$ we finish the proof of Theorem 2.1. $\square$

4 Proof of Theorem 2.2

Note that $Q^{'}(y) = \frac{1}{f(Q(y))}$ and $Q^{''}(y) = -\frac{f'(Q(y))}{f^3(Q(y))}$. Let $h(y) = (y(1-y))^\nu$. Moreover,

$$(h(y)\tilde{N}_n(y))' = h'(y)\tilde{N}_n(y) + h(y)\tilde{N}_n'(y).$$

Thus, in view of (21), we need $h'(y)$ and $h(y)Q'(y)$ to be uniformly bounded, which is achieved by $\nu > 1$ and $\nu > \gamma$, respectively. Moreover,

$$(h(y)\tilde{N}_n(y))^{''} = h^{''}(y)\tilde{N}_n(y) + 2h'(y)\tilde{N}_n'(y) + h(y)\tilde{N}_n^{''}(y).$$

Thus, in view of (21), (22) we need $h^{''}(y)$, $h'(y)Q'(y)$, $h(y)(Q'(y))^2$ and $h(y)Q^{''}(y)$ to be uniformly bounded. The first claim is achieved by $\nu > 2$, the second by $\nu - 1 > \gamma$, the third by $\nu > 2\gamma$, respectively. As for the fourth one we have

$$h(y)|Q^{''}(y)| = \left(\frac{|f'(Q(y))|}{f^2(Q(y))}(y(1-y))\right)\left(\frac{(y(1-y))^{\nu-1}}{f(Q(y))}\right).$$

Since (CsR3(i)), (CsR3(ii)) are fulfilled, the first part is $O(1)$ uniformly in $y \in (0,1)$. The second part is $O(1)$ since $\nu - 1 > 2\gamma - 1 > \gamma$ (recall $\gamma > 1$).

Let $g(y) = h(y)\tilde{N}_n(y)$. Thus, proceeding exactly as in Lemma 3.5 we obtain $\mathrm{E}\left[\max_{y\in(0,1)} |g'(y)|^\alpha\right] = O(n^{-\alpha/2})$. Thus, $\sup_{y\in(0,1)} |g'(y)| = o_{a.s}(l_q(n)/\sqrt{n})$ and

$$\sup_{|u-v|\le d_n, u,v\in(0,1)} |g(v) - g(u)| \le d_n \sup_{y\in(0,1)} |g'(y)| = o_{a.s}(l_q(n)/\sqrt{n}).$$

Further, by differentiability of h,

$$\sup_{|u-v|\le d_n, u,v\in(0,1)} |h(u) - h(v)||\tilde{N}_n(u)| \le d_n \sup_{y\in(0,1)} |\tilde{N}_n(y)| = o_{a.s}(l_q(n)/\sqrt{n}).$$

Consequently,

$$\sup_{|u-v|\le d_n, u,v\in(0,1)} h(v)|\tilde{N}_n(v) - \tilde{N}_n(u)| = o_{a.s}(l_q(n)/\sqrt{n}).$$

Therefore, taking into account the martingale part and $d_n = \lambda_n$

$$\sup_{|u-v|\le \lambda_n, u,v\in(0,1)} h(v)|(E_n(v) - v) - (E_n(u) - u)| = O_{a.s}\left(\frac{\sqrt{d_n \log n}}{\sqrt{n}}\right). \quad (23)$$

Note that

$$u_n(y) \quad = \quad \sqrt{n}(E_n(U_n(y)) - U_n(y)) + O_{a.s}(n^{-1/2}).$$

Consequently, by (23) and (9)

$$\sup_{y\in(0,1)} h(y)|u_n(y) - \alpha_n(y)| = \sup_{y\in(0,1)} h(y)|\alpha_n(U_n(y)) - \alpha_n(y)| + O_{a.s}(n^{-1/2}) = O_{a.s}(b_n).$$

Let $(k-1)/n < y \le k/n$. Further, let $\delta_n = 2C^* n^{-1/2}(\log\log n)^{1/2}$, $C^* = C$, C from (9). As in [5], with $\theta = \theta_n(y)$ such that $|\theta - y| \le n^{-1/2}u_n(y) = O_{a.s}(\delta_n)$,

$$\begin{aligned}
&\sup_{y\in(\delta_n, 1-\delta_n)} h(y)|f(Q(y))q_n(y) - u_n(y)| \le \\
&\quad = \quad n^{-1/2}u_n^2(y)\left(\frac{f'(Q(\theta))}{f^2(Q(\theta))}\theta(1-\theta)\right)\frac{f(Q(y))}{f(Q(\theta))}\frac{(y(1-y))^\nu}{\theta(1-\theta)}.
\end{aligned}$$

In view of [5, Lemma 1] we have

$$\frac{f(Q(y))}{f(Q(\theta))} \le \left\{\frac{y\vee\theta}{y\wedge\theta}\frac{1-y\wedge\theta}{1-y\vee\theta}\right\}^\gamma. \quad (24)$$

Further on, if $y \geq 2\delta_n$,

$$\frac{y}{\theta} = \frac{y-\theta}{\theta} + 1 \leq \frac{\delta_n}{y - n^{-1/2}u_n(y)} + 1 \leq 2. \tag{25}$$

Consequently, by (CsR3(i)), (CsR3(ii)), (24), (25) we have

$$\sup_{y \in (\delta_n, 1-\delta_n)} h(y)|f(Q(y))q_n(y) - u_n(y)| = O_{a.s}(n^{-1/2} \log\log n).$$

If $1 \geq U_{k:n} \geq y$ then

$$\begin{aligned}
&\sup_{y \in (0,\delta_n)} h(y)|f(Q(y))q_n(y)| = \sup_{y \in (0,\delta_n)} n^{1/2} h(y) \int_y^{U_{k:n}} \frac{f(Q(y))}{f(Q(u))} du \\
&\leq \; C \sup_{y \in (0,\delta_n)} n^{1/2} h(y) \int_y^{U_{k:n}} \left(\frac{u}{y}\right)^{\gamma_1} du \\
&\leq \; C \sup_{y \in (0,\delta_n)} n^{1/2} U_{k:n}^{\gamma_1+1} h(y) y^{-\gamma_1} = O_{a.s}(n^{-1/2} \log\log n).
\end{aligned}$$

Further, if $U_{k:n} \leq y$ then for $\gamma_1 > 1$

$$\sup_{y \in (0,\delta_n)} h(y)|f(Q(y))q_n(y)| \leq C n^{1/2} \delta_n^{\nu+\gamma_1} U_{k:n}^{-(\gamma_1 - 1)} \tag{26}$$

and for $\gamma_1 = 1$,

$$\sup_{y \in (0,\delta_n)} h(y)|f(Q(y))q_n(y)| \leq C n^{1/2} \delta_n^{\nu+1} \log(\delta_n / U_{k:n}). \tag{27}$$

Now,

$$P(U_{1:n} \leq n^{-2}(\log n)^{-3/2}) \leq \sum_{i=1}^{n} P(U_i \leq n^{-2}(\log n)^{-3/2}) \leq n^{-1}(\log n)^{-3/2}.$$

Consequently, via the Borel-Cantelli lemma, $U_{k:n} = o_{a.s}(n^{-2}(\log n)^{-3/2})$. Therefore, the bound in (27) is $O_{a.s}(n^{-1/2} \log\log n)$ and the bound in (26) is of the same order provided $\nu > 3\gamma_1 - 2$. The upper tail is treated in the similar way. Consequently, the result follows.

□

5 Remarks

Remark 5.1 Assume that $c_k \sim k^{-\tau}(\log k)^{-3/2}$. Then

$$\sum_{k=i}^{\infty} c_k^2 = \sum_{k=i}^{\infty} k^{-2\tau}(\log k)^{-3} = O(i^{-2\tau+1}(\log i)^{-3}),$$

and (6) is fulfilled if $\tau > 5/2$.

If X_1 has all moments finite then in [17], the summability assumption $\sum_{k=0}^{\infty} |c_k| < \infty$ provides the bound $O(n^{-3/4}(\log n)^{\frac{1}{2}+\eta})$, $\eta > 0$.

Remark 5.2 It should be pointed out that the process $K(x,t)$ in Proposition 3.1 approximates N_n, not β_n. We use this as a tool to achieve ULIL only. Using the blocking technique applied to β_n directly, we may establish Proposition 3.1 with $N_{[nt]}$ replaced by $([nt])^{1/2}\beta_{[nt]}$ and a different Gaussian process. However, it requires stronger assumptions than (6) on the coefficients c_k. In particular, $\tau > \frac{9}{2} + \frac{4}{\alpha}$. This is clear. Without exploiting differentiability, we have to compare indicators applied to the original sequences X_i with the indicators applied to the truncated one, see e.g. [3].

Remark 5.3 The main part of this paper is devoted to the law of the iterated logarithm for the empirical process. One may ask, whether imposing (6), we could obtain an uniform law of the iterated logarithm for empirical processes via strong mixing. From [16] one knows that if $\mathrm{E}|\epsilon_1|^\alpha < \infty$, $c_k = O(k^{-\tau})$ and $\tau > 1 + \frac{1}{\alpha} + \max\{1, \alpha^{-1}\}$, then $\alpha(n) = O(n^{-\lambda})$, where $\lambda = (\tau\alpha - \max\{\alpha, 1\})/(1+\alpha) - 1 > 0$ and $\alpha(n)$ is the strong mixing coefficient. In particular, if $\alpha \geq 2$, then this condition yields $\tau > 2 + \frac{1}{\alpha}$ and $\alpha(n) = O(n^{-\frac{\tau\alpha - 1 - 2\alpha}{1+\alpha}})$. In view of Rio [15] to obtain the LIL for partial sums of bounded strongly mixing random variables one needs $\alpha(n)$ to be summable, which would require $\tau > 3 + \frac{2}{\alpha}$. Thus, we need stronger conditions than (6). Also, Rio's result would provide the LIL for $\beta_n(x)$ at *fixed* x.

Appendix

Here, we re-prove the strong approximation of Proposition 3.1. The proof follows the lines of [3], thus we present the major steps only. For full details we refer to [13].

We are the keeping notation from Section 3.1. Recall that $||X||_\alpha = (\mathrm{E}|X|^\alpha)^{1/\alpha}$. For $\mathbf{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$, let $||\mathbf{x}||$ be the Euclidean norm and define random vectors in $\mathbb{R}^d$:

$$\xi_i = (Y_i(x_1), \ldots, Y_i(x_d)), \quad \hat{\xi}_i = (\hat{Y}_i(x_1), \ldots, \hat{Y}_i(x_d)).$$

Denote the covariance matrix of ξ_i by $\Gamma_d = \Gamma_d(\mathbf{x}) = (\Gamma(x_i, x_j), 1 \le i, j \le d)$. We shall use the following implication which is valid for any vectors ξ_i, η_i in $\mathbb{R}^d$:

$$\mathrm{E}||\xi_i - \eta_i|| \le A \;\text{ implies }\; |\mathrm{E}\exp(i < \mathbf{u}, \xi_i >) - \mathrm{E}\exp(i < \mathbf{u}, \eta_i >)| \le ||\mathbf{u}||A\,. \tag{28}$$

Also, we will use the following bound: for $\mathbf{x}_i \in \mathbb{R}^d$, $i = 1, \ldots, n$, we have $||\sum_{i=1}^n \mathbf{x}_i|| \le \sum_{i=1}^n ||\mathbf{x}_i||$.

Lemma 5.4 *Under the conditions of Proposition 3.1, for all* $\mathbf{u} \in \mathbb{R}^d$

$$\left|\mathrm{E}\exp(i < \mathbf{u}, n^{-1/2}\sum_{i=1}^n \xi_i >) - \mathrm{E}\exp(i < \mathbf{u}, n^{-1/2}\sum_{i=1}^n \hat{\xi}_i >)\right| \le ||\mathbf{u}|| d^{1/2} n^{-1/2}\,.$$

Proof. From (6) and (17) we have $||Y_i(x) - \hat{Y}_i(x)||_2 \le C i^{-1}(\log i)^{-3/2}$. Thus

$$\mathrm{E}||\xi_i - \hat{\xi}_i||^2 \le C d i^{-2}(\log i)^{-3} \tag{29}$$

and (29) implies

$$\mathrm{E}||\xi_i - \hat{\xi}_i|| \le C d^{1/2} i^{-1}(\log i)^{-3/2}.$$

Therefore,

$$\mathrm{E}\left\|\sum_{i=1}^n (\xi_i - \hat{\xi}_i)\right\| \le C d^{1/2}\,.$$

This, together with (28) implies the result.

□

Proposition 5.5 *Under the conditions of Theorem 3.1, for all* $\mathbf{u} \in \mathbb{R}^d$

$$\begin{aligned}&\left|\mathrm{E}\exp(i < \mathbf{u}, n^{-1/2}\sum_{i=1}^n \xi_i >) - \exp(-\frac{1}{2} < \mathbf{u}, \Gamma_d \mathbf{u} >)\right| \\ &\le \; C d ||\mathbf{u}|| n^{-\delta_1} + C||\mathbf{u}||^2 (n^{-1/4}(\log n)^{d/2}\exp(Cd) + d^{d/2}\exp(-C n^{1/2}))\end{aligned}$$

with some $\delta_1 > 0$.

Proof. Divide the interval $[1,n]$ into consecutive long and short blocks $I_1, J_1, I_2, J_2, \ldots$ with length $card(I_k) = [n^{\rho^*}]$, $card(J_k) = [n^{\rho}]$, $1 \le k \le M$, $\rho < \rho^* < \frac{1}{2}$. Then $M = M_n \sim n^{1-\rho^*}$ (the last block is possibly incomplete). Define $U_k = \sum_{i \in I_k} \xi_i$, $\hat{U}_k = \sum_{i \in I_k} \hat{\xi}_i$, $\hat{V}_k = \sum_{i \in J_k} \hat{\xi}_i$. By (29) we have

$$\mathrm{E}||U_k - \hat{U}_k|| \le \sum_{i \in I_k} \mathrm{E}||\xi_i - \hat{\xi}_i|| \le C d^{1/2} \sum_{i \in I_k} i^{-1} (\log i)^{-3/2} . \tag{30}$$

Note that $\hat{U}_k$, $k = 1, \ldots, M$ are independent. Thus, we can construct independent random vectors $\tilde{U}_1, \ldots, \tilde{U}_M$ such that $(\tilde{U}_k, \hat{U}_k) \stackrel{\mathrm{d}}{=} (U_k, \hat{U}_k)$, $k = 1, \ldots, M$. By (30) one obtains

$$\mathrm{E}\left\| \sum_{k=1}^{M} (\tilde{U}_k - \hat{U}_k) \right\| \le \sum_{k=1}^{M} \mathrm{E}||U_k - \hat{U}_k|| \le C d^{1/2} . \tag{31}$$

Further, $\mathrm{E}||\hat{V}_k|| \le d n^{\rho/2}$ which implies by independence

$$\mathrm{E}\left\| \sum_{k=1}^{M} \hat{V}_k \right\|^2 \le \sum_{k=1}^{M} \mathrm{E}\left\| \hat{V}_k \right\|^2 \le C d M n^{\rho} = d n^{\rho + 1 - \rho^*} . \tag{32}$$

We have

$$\sum_{i=1}^{n} \hat{\xi}_i = \sum_{k=1}^{M} (\hat{U}_k - \tilde{U}_k) + \sum_{k=1}^{M} \tilde{U}_k + \sum_{k=1}^{M} \hat{V}_k .$$

Thus, by (31), (32)

$$\mathrm{E}\left\| \sum_{i=1}^{n} \xi_i - \sum_{k=1}^{M} \tilde{U}_k \right\| \le C (d^{1/2} + d n^{(\rho + 1 - \rho *)/2}) . \tag{33}$$

Also, by [3, Lemma 2.9] and the same argument as at the end of Lemma 2.10 in the latter paper one obtains

$$\begin{aligned} &\left| \mathrm{E} \exp(i < \mathbf{u}, n^{-1/2} \sum_{k=1}^{M} \tilde{U}_k >) - \exp(-\frac{1}{2} < \mathbf{u}, \Gamma_d \mathbf{u} >) \right| \\ &\quad \le \quad C ||\mathbf{u}||^2 (n^{-(1-\rho^*)/2} (\log n)^{d/2} \exp(Cd) + d^{d/2} \exp(-C n^{1-\rho^*})) . \end{aligned} \tag{34}$$

Consequently, by (33), (28) and (34) the result follows.

□

5.1 Approximation

Let $\varepsilon \in (0, 1/4)$. Denote $t_k = \exp(k^{1-\varepsilon})$, $p_k = 2[t_k^{\rho}]$, $d_k = k^{1/2}$ and $x_i = (i-1)/d_k$, $i \in \mathbb{Z}$, $M_k = t_{k+1} - t_k - p_k$ so that $M_k \sim Ck^{-\varepsilon}\exp(k^{1-\varepsilon})$ and $\log M_k \sim Ck^{1-\varepsilon}$. Define random vectors in $\mathbb{R}^{d_k}$, $\eta_k = (\eta_{k1}, \ldots, \eta_{kd_k})$, where

$$\eta_{ki} = R(x_i, t_{k+1}) - R(x_i, t_k + p_k) = \sum_{j=t_k+p_k+1}^{t_{k+1}} Y_j(x_i) \; .$$

Also, $\hat{\eta}_{ki} = \sum_{j=t_k+p_k+1}^{t_{k+1}} \hat{Y}_j(x_i)$.

By Proposition 5.5

$$\begin{aligned}
&\left| \mathrm{E}\exp(i < \mathbf{u}, M_k^{-1/2}\eta_k >) - \exp(-\frac{1}{2} < \mathbf{u}, \Gamma_{d_k}\mathbf{u} >) \right| \\
&\leq \; C||\mathbf{u}||^2 \left(d_k M_k^{-\rho_1} + M_k^{-(1-\rho^*)/2}(\log M_k)^{d_k/2}\exp(Cd_k) + d_k^{d_k/2}\exp(-CM_k^{1-\rho^*}) \right) \\
&\leq \; C\exp(-Ck^{1-\varepsilon})||\mathbf{u}||^2.
\end{aligned}$$

Let ψ_{PL} be the Prokhorov-Levy distance. Choose $T := \exp(k^{\varepsilon})$; for sufficiently large k, $T > 10^8 d_k$. Then

$$\begin{aligned}
&\psi_{PL}(M_k^{-1/2}\eta_k, N(\mathbf{0}, \Gamma_{d_k})) \leq \frac{16d_k}{T}\log T + P(N(\mathbf{0}, \Gamma_{d_k}) > T/2) \\
&\qquad + T^{d_k}\int_{||\mathbf{u}|| \leq T} \left| \mathrm{E}\exp(i < \mathbf{u}, M_k^{-1/2}\eta_k >) - \exp(-\frac{1}{2} < \mathbf{u}, \Gamma_{d_k}\mathbf{u} >) \right| d\mathbf{u} \\
&\leq \; \frac{16d_k}{T}\log T + P(N(\mathbf{0}, \Gamma_{d_k}) > T/2) + CT^{d_k}\exp(-k^{1-\varepsilon})\int_{||\mathbf{u}|| \leq T} ||\mathbf{u}||^2 du \\
&\leq \; C\exp(-Ck^{\varepsilon})
\end{aligned}$$

since $\varepsilon < 1/4$.

Since $\hat{\eta}_k$, $k = 1, \ldots, M$ are independent, we can define independent random vectors $\zeta_1, \ldots, \zeta_M$ such that $M_k^{-1/2}\zeta_k \sim N(\mathbf{0}, \Gamma_{d_k})$ and

$$P(||M_k^{-1/2}\hat{\eta}_k - M_k^{-1/2}\zeta_k|| > C\exp(-k^{\varepsilon})) \leq C\exp(-k^{\varepsilon}) \; .$$

This yields

$$P(||M_k^{-1/2}\eta_k - M_k^{-1/2}\zeta_k|| > C\exp(-k^{\varepsilon})) \leq C\exp(-k^{\varepsilon})$$

and then by the Borel-Cantelli lemma

$$M_k^{-1/2}(\eta_k - \zeta_k) \leq \exp(-k^{\varepsilon}) \;\; \text{almost surely.} \tag{35}$$

For $k \geq 1$, define random vectors Z_k in $\mathbb{R}^{d_k}$ by

$$Z_k = (R(x_i, t_{k+1}) - R(x_i, t_k), i = 1, \ldots, d_k).$$

Since $Z_k - \eta_k$ is the sum of p_k random vectors in $\mathbb{R}^{d_k}$ with coordinates bounded by 1 and since $M_k \leq t_k$, $d_k^{1/2} p_k^{1/2} \leq k t_k^{\rho/2} \leq \exp(ck^{1-\varepsilon})$, $\varepsilon < 1/2$ we have by (35)

$$\begin{aligned} ||Z_k - \zeta_k|| &\leq ||Z_k - \eta_k|| + ||\eta_k - \zeta_k|| \\ &\leq C(d_k^{1/2} p_k^{1/2} + t_k^{1/2} \exp(-k^{\varepsilon})) \leq C t_k^{1/2} \exp(-Ck^{\varepsilon}) . \end{aligned} \tag{36}$$

Thus, the skeleton process $\{Z_k, k \geq 1\}$ can be approximated by the sequence $\{\zeta_k, k \geq 1\}$. The latter can be extended to a centered Gaussian process $\{K(x,t), x \in \mathbb{R}, t \geq 0\}$ with covariance $t \wedge t' \Gamma_{d_k}(s, x')$ such that

$$\zeta_{ki} = K(x_i, t_{k+1}) - K(x_i, t_k + p_k), \quad i = 1, \ldots, d_k.$$

Define $Y_k = (K(x_i, t_{k+1}) - K(x_i, t_k), i = 1, \ldots, d_k)$. Then

$$||\zeta_k - Y_k|| \leq C t_k^{1/2} \exp(-Ck^{\varepsilon}) \text{ almost surely}$$

by the last inequality on p. 807 in [3]. Thus, by (36)

$$||Z_k - Y_k|| \leq C t_k^{1/2} \exp(-Ck^{\varepsilon}) \text{ almost surely.}$$

5.2 Oscillations

Lemma 5.6 *Under the conditions of Theorem 3.1, for $n \geq 1$, $\lambda \geq n^{\frac{1}{2}}$ and any $-\infty < a < b < \infty$ we have*

$$P\left(\sup_{a \leq x \leq x' \leq b, 0 \leq k \leq n} \sum_{i=1}^{k} Y_i(x, x') \geq \lambda \right) \leq C \exp\left(-C\lambda^2/(n(b-a)^2)\right) + \frac{C}{n^{\eta}},$$

where $\eta > 0$.

Proof. Define

$$M_{u,v} = \max_{0 \leq i \leq 2^u, 0 \leq j \leq 2^v} \left| \sum_{k=nj2^{-v}}^{n(j+1)2^{-v}} Y_k(bi2^{-u}, b(i+1)2^{-u}) \right| .$$

Then for any integer $L \geq 1$,

$$\left|\sum_{i=1}^{k} Y_i(0,x)\right| \leq \sum_{1\leq u,v\leq L} M_{u,v} + \frac{2n}{2^L} .$$

Choose

$$0 < \varepsilon < (\alpha - 2)/(2C_6) \tag{37}$$

for some constant C_6 to be specified later. Since $n^{1/2} \leq \lambda$, $u, v \leq L$ we have by Lemma 3.4

$$\begin{aligned}
&P(M_{u,v} \geq \lambda 2^{-\varepsilon(u+v)}) \\
&\leq \; 2^{u+v} C \left\{ (\lambda 2^{-\varepsilon(u+v)})^{-\alpha} + \exp\left(-C\lambda^2 2^{-2\varepsilon(u+v)}/(n2^{-v}(b2^{-u})^2)\right) + \right. \\
&\quad \left. \exp\left(-C\lambda 2^{-\varepsilon(u+v)}/(n2^{-v})^{\rho}\right) \right\} \\
&\leq \; 2^{u+v} C \left\{ n^{-\frac{1}{2}\alpha} 2^{2\varepsilon L\alpha} + \exp\left(-C\lambda^2 2^{u(1-2\varepsilon)} 2^{v(1-2\varepsilon)}/(nb^2)\right) + \right. \\
&\quad \left. \left(n^{(\frac{1}{2}-\rho)} 2^{-\varepsilon(u+v)} 2^{v\rho}\right)^{-r} \right\}
\end{aligned}$$

for any $r > 0$. Choose $r = \frac{\alpha}{2(1/2-\rho))}$. Since $u + v \leq 2L$,

$$\begin{aligned}
&P(M_{u,v} \geq \lambda 2^{-\varepsilon(u+v)}) \\
&\leq \; 2^{u+v} C \left\{ n^{-\alpha/2} 2^{2C\varepsilon L\alpha} + \exp\left(-C\lambda^2 2^{u(1-2\varepsilon)} 2^{v(1-2\varepsilon)}/(nb^2)\right) \right\}.
\end{aligned}$$

For

$$A = \{M_{u,v} \geq \lambda 2^{-\epsilon(u+v)} \text{ for some } 1 \leq u, v \leq L\}$$

we have

$$\begin{aligned}
P(A) &\leq \sum_{1\leq u,v\leq L} 2^{u+v} \exp\left(-C\lambda^2 2^{u(1-2\varepsilon)} 2^{v(1-2\varepsilon)}/(nb^2)\right) + 2^{2L+2} 2^{CL\varepsilon} n^{-\alpha/2} \\
&\leq \left(\sum_{1\leq u\leq L} \exp\left(u - C\lambda^2 2^{u(1-2\varepsilon)}/(nb^2)\right)\right)^2 + 2^{L(2+C\varepsilon)+2} n^{-\alpha/2} \\
&\leq C\exp\left(-C\lambda^2/(nb^2)\right) + C2^{L(2+C_6\varepsilon)} n^{-\alpha/2}.
\end{aligned}$$

Choose L so that $n^{1/2} < 2^L < 2n^{1/2}$, then by (37) the second part in the latter expression is bounded from above by $Cn^{-\eta}$ with some $\eta > 0$. Further, on A^c, we have $\sum_{1\leq u,v\leq L} M_{u,v} \leq C_7\lambda$ and

$$\sum_{1\leq u,v\leq L} M_{u,v} + \frac{2n}{2^L} \leq C_7\lambda + 4\lambda = C_8\lambda.$$

Therefore, the result follows.

□

Let

$$\tilde{R}_{i,k} = \sup_{s\in[x_i,x_{i+1})} \sup_{t\in[t_k,t_{k+1}]} |R(s,t) - R(x_i,t_k)|.$$

Using Lemma 5.6 we obtain the following result.

Lemma 5.7 *Under the conditions of Theorem 3.1, for some $\varepsilon^* > 0$,*

$$\max_{1\le i\le d_k} |\tilde{R}_{i,k}| \le C t_k^{1/2} (\log k)^{-\varepsilon^*} \quad \text{almost surely.}$$

5.3 Conclusion of the proof

For $s \in [x_i, x_{i+1})$, $t \in [t_k, t_{k+1}]$ write

$$|R(s,t)-K(s,t)| \le |R(s,t)-R(x_i,t_k)|+|K(s,t)-K(x_i,t_k)|+|R(x_i,t_k)-K(x_i,t_k)|.$$

Both the first and the second part are bounded almost surely by $Ct_k^{1/2}(\log k)^{-\varepsilon^*}$, the first one by Lemma 5.7, the second one by (2.55) in [3]. The second part is bounded in the same way as (2.56) in [3].

Acknowledgement.

The comments of Neville Weber and both referees are greatly appreciated. This work was done partially during my stay at Carleton University. I am thankful to Professors Barbara Szyszkowicz and Miklos Csörgő for their support and helpful remarks.

References

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

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

[3] Berkes, I. and Horváth, L. (2001). Strong approximation of the empirical process of GARCH sequences. *Ann. Appl. Probab.* **11**, 789–809.

[4] Csörgő, M. (1983). *Quantile Processes with Statistical Applications.* CBMS-NSF Regional Conference Series in Applied Mathematics.

[5] Csörgő, M. and Révész, P. (1978). Strong approximation of the quantile process. *Ann. Statist.* **6**, 882–894.

[6] Doukhan, P. (1984). *Mixing: Properties and Examples.* Lecture notes in Statisitcs. Springer-Verlag.

[7] Fotopoulos, S.B. and Ahn, S.K. (1994). Strong Approximation of the Qunatile Processes and Its Applications under Strong Mixing Properties. *J. Multivariate Anal.* **51**, 17–45.

[8] Hesse, C.H. (1990). A Bahadur-Type Representation for Empirical Quantiles of a Large Class of Stationary, Possibly Infinite-Variance, Linear Processes. *Ann. Statist.* **18**, 1188–1202.

[9] Ho, H.-C. and Hsing, T. (1997). Limit theorems for functionals of moving averages. *Ann. Probab.* **25**, 1636–1669.

[10] Hsing, T. and Wu, W.B. (2004) On weighted U-statisitcs for stationary sequences. *Ann. Probab.* **32**, 1600–1631.

[11] Kiefer, J. (1967). On Bahadur's representation of sample qunatiles. *Ann. Math. Statist.* **38**, 1323–1342.

[12] Kiefer, J. (1970). Deviations between the sample qunatile process and the sample df. *nonparametric Techniques in Statistical Inference*, 299–319, M.L. Puri, ed., Cambridge University Press.

[13] Kulik, R. (2006). Bahadur-Kiefer theory for weakly dependent random variables. *Technical Report* **430**. Laboratory for Research in Statistics and Probability. Carleton University and University of Ottawa.

[14] Petrov, V.V. (1975). *Sums of Independent Random Variables.* Springer-Verlag.

[15] Rio, E. (1995). The functional law of the iterated logarithm for stationary strongly mixing sequences. *Ann. Probab.* **23**, 1188–1203.

[16] Withers, C. S. (1981). Conditions for linear processes to be strong-mixing. *Z. Wahrsch. Verw. Gebiete* **57**, 477–480.

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

[18] Wu, W. B. and Mielniczuk, J. (2002). Kernel density estimation for linear processes. *Ann. Statist.* **30**, 1441–1459.

[19] Yu, H. (1996). A note on strong approximation for quantile processes of strong mixing sequences. *Statist. Probab. Lett.* **30**, 1–7.