SFB 823

Bootstrap for the sample mean and for *U*-Statistics of stationary processes

Olimjon Sh. Shapirov, Martin Wendler

Nr. 35/2009

Discussion Paper

BOOTSTRAP FOR THE SAMPLE MEAN AND FOR U-STATISTICS OF STATIONARY PROCESSES

OLIMJON SH. SHARIPOV AND MARTIN WENDLER

Abstract. The validity of various bootstrapping methods has been proved for the sample mean of strongly mixing data. But in many applications, there appear nonlinear statistics of processes that are not strongly mixing. We investigate the nonoverlapping block bootstrap for functionals of absolutely regular processes, which occur from chaotic dynamical systems. We establish the strong consistency of the bootstrap distribution estimator not only for the sample mean, but also for U-statistics, which include examples as Gini's mean difference or the χ^2-test statistic.

keywords: strongly mixing sequences, functionals of absolutely regular sequences, U-statistics, block bootstrap
AMS 2000 subject classification: 62G09, 60G10

1. Introduction

1.1. **U-Statistics.** U-statistics play an important role in nonparametric statistics because many estimators and test statistics can be written at least asymptotically as U-statistics. Well-known examples include the sample variance, Gini's mean difference, and the χ^2 goodness of fit test statistic. A more recent example is the Grassberger-Procaccia dimension estimator. U-statistics can be described as generalized means, i.e. means of the values of a kernel function $h(X_{i_1}, \ldots, X_{i_k})$. For simplicity of notation, we concentrate on the case of bivariate U-statistics:

Definition 1.1. *A U-statistic with a symmetric and measurable kernel $h : \mathbb{R}^2 \to \mathbb{R}$ is defined as*

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

The key tool in the analysis of U-statistics is the Hoeffding-decomposition [13] of $U_n(h)$ into a so-called linear part and a degenerate part

$$U_n(h) = \theta + \frac{2}{n} \sum_{i=1}^{n} h_1(X_i) + U_n(h_2)$$

with

$$\begin{aligned} \theta &:= Eh(X, Y), \\ h_1(x) &:= Eh(x, Y) - \theta, \\ h_2(x, y) &:= h(x, y) - h_1(x) - h_1(y) - \theta. \end{aligned}$$

Date: November 27, 2009.
Olimjon Sharipov was supported by DFG (German Research Foundation). Martin Wendler was supported by the Studienstiftung des deutschen Volkes.

for X, Y independent and with the same distribution as X_1.

Under independence, the summands of the degenerate part $U_n(h_2)$ are uncorrelated, so that $\mathrm{Var} U_n(h_2) = O\left(n^{-2}\right)$ and the asymptotic behavior of U_n is dominated by the linear part $\frac{2}{n}\sum h_1(X_i)$. For independent data, second moments of the kernel are required. For mixing data, one needs higher moments, and in the case of strong mixing and functionals of absolutely regular processes a continuity condition:

Definition 1.2. *Let $(X_n)_{n\in\mathbb{N}}$ be a stationary process*

(1) *A kernel has uniform r-moments, if there is a $M>0$ such that for all $k\in\mathbb{N}_0$*

$$E\left|h(X_1,X_k)\right|^r \leq M \quad \text{and} \quad E\left|h(X,Y)\right|^r \leq M$$

for X, Y independent and with the same distribution as X_1.

(2) *A kernel h is called P-Lipschitz-continuous with constant $L>0$, if*

$$E\left[\left|h(X,Y)-h(X',Y)\right|\mathbb{1}_{\{|X-X'|\leq\epsilon\}}\right] \leq L\epsilon$$

for every $\epsilon>0$, every pair (X,Y) with the same common distribution as (X_1,X_k) for some $k\in\mathbb{N}$ or independent with the same distribution as X_1 and (X',Y) also with one of these common distributions. With $\mathbb{1}_A$, we denote the indicator function of a set A.

It is clear that every Lipschitz-continuous kernel function is P-Lipschitz-continuous, but the above definition covers more examples:

Example 1.3 (Variance estimation)**.** *Consider stationary random variables with a finite first moment and the kernel $h(x,y)=\frac{1}{2}(x-y)^2$. The related U-statistic is the well known variance estimator*

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

Example 1.4 (Gini's mean difference)**.** *Let $h(x_1,x_2)=|x_1-x_2|$. Then the corresponding U-statistic is*

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

also known as Gini's mean difference.

The P-Lipschitz-continuity of these two kernels follows by an easy lemma:

Lemma 1.5. (1) *Let h be a polynomial kernel of degree d, that is*

$$h(x,y) = \sum_{i=0}^{d}\sum_{j=0}^{d-i} c_{ij}\left(x^i y^j + x^j y^i\right).$$

If $E\left|X_1\right|^{d-1}<\infty$, then h is P-Lipschitz-continuous.

(2) *Let h be a P-Lipschitz-continuous kernel and $f:\mathbb{R}\to\mathbb{R}$ a Lipschitz-continuous function. Then $g\circ h$ is P-Lipschitz-continuous.*

Proof. (1) We can concentrate on an expression of the form $g(x,y) = x^i y^j$, $i+j \le d$:

$$E\left[\left|g(X,Y) - g(X',Y)\right| \mathbb{1}_{\{|X-X'|\le\epsilon\}}\right]$$
$$=E\left[\left|(X-X')\left(X^{i-1} + X^{i-2}X' + \ldots + XX'^{i-2} + X'^{i-1}\right)Y^j\right| \mathbb{1}_{\{|X-X'|\le\epsilon\}}\right]$$
$$\le \epsilon E\left[\left|\left(X^{i-1} + X^{i-2}X' + \ldots + XX'^{i-2} + X'^{i-1}\right)Y^j\right|\right] \le \epsilon i E\left|X_1\right|^{i+j}$$

(2) This is obvious.

□

Example 1.6 (χ^2 goodness of fit test)**.** *Let $(X_n)_{n\in\mathbb{N}}$ be a stationary process such that X_1 can take only the values $t_1,\ldots,t_k$. Furthermore, let be $p_1,\ldots,p_k > 0$ with $\sum_{i=1}^n p_i = 1$ and*

$$h(x,y) = \sum_{i=1}^k \frac{1}{p_i}\left(\mathbb{1}_{\{x=t_i\}} - p_i\right)\left(\mathbb{1}_{\{y=t_i\}} - p_i\right).$$

Then h is P-Lipschitz-continuous, as

$$E\left[\left|h(X,Y) - h(X',Y)\right| \mathbb{1}_{\{|X-X'|\le\epsilon\}}\right] = 0,$$

if $\epsilon < \min_{i\neq j}|t_i - t_j|$. The related U-statistic is

$$U_n(h) = \frac{1}{n(n-1)}\sum_{l=1}^k\left(\frac{1}{p_l}\left(\sum_{i=1}^n\left(\mathbb{1}_{\{X_i=t_l\}} - p_l\right)\right)^2 - \frac{1}{p_l}\sum_{i=1}^n\left(\mathbb{1}_{\{X_i=t_l\}} - p_l\right)^2\right)$$
$$= \frac{1}{n-1}\chi^2 - \frac{1}{n(n-1)}\sum_{l=1}^k\frac{1}{p_l}\sum_{i=1}^n\left(\mathbb{1}_{\{X_i=t_l\}} - p_l\right)^2.$$

χ^2 is used for testing the hypothesis that $P[X_1 = t_l] = p_i$ for $l = 1,\ldots,k$.

Example 1.7 (Dimension estimation)**.** *Let $t > 0$. The kernel $h(x,y) = \mathbb{1}_{\{|x-y|<t\}}$ is related to the Grassberger-Procaccia dimension estimator (see Borovkova et al.* [3] *for details). It is P-Lipschitz-continuous, if there is an $L > 0$, such that for all $\epsilon > 0$ and every common distribution of X and Y from Definition 1.2:*

$$P\left[t - \epsilon \le |X - Y| \le t + \epsilon\right] \le L\epsilon$$

For proof: See Dehling, Wendler [7]*.*

1.2. **Mixing Random Variables.** In many statistical applications the data does not come from an independent stochastic process. A standard assumption of weak dependence is given by the strong mixing condition:

Definition 1.8. *Let $(X_n)_{n\in\mathbb{N}}$ be a stationary process. Then the strong mixing coefficient is given by*

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

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

For further information on strong mixing and a detailed description of other mixing conditions see Doukhan [10] and Bradley [4]. However, this class of weak dependent processes excludes examples like linear processes with innovations that do not have a density or data from dynamical systems.

Example 1.9. *Let $(Z_n)_{n\in\mathbb{N}}$ be independent r.v.'s with $P[Z_n=1]=P[Z_n=0]=\frac{1}{2}$ and*

$$X_n=\sum_{k=n}^{\infty}\frac{1}{2^{k-n+1}}Z_k.$$

Then $(X_n)_{n\in\mathbb{N}}$ is not strong mixing, as

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

We will consider functionals of absolutely regular processes, as this class of dependent sequences covers the example above and data from other 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 (see Hofbauer, Keller [14]).

Definition 1.10. *Let $(X_n)_{n\in\mathbb{N}}$ be a stationary process. Then the absolute regularity coefficient is given by*

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

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

We call a sequence $(X_n)_{n\in\mathbb{Z}}$ a two-sided functional of $(Z_n)_{n\in\mathbb{Z}}$ if there is a measurable function defined on $\mathbb{R}^{\mathbb{Z}}$ such that

$$X_n=f\left((Z_{n+k})_{k\in\mathbb{Z}}\right).$$

Similarly we call a sequence $(X_n)_{n\in\mathbb{N}}$ a one-sided functional of $(Z_n)_{n\in\mathbb{N}}$ if there is a measurable f such that

$$X_n=f((Z_{n+k})_{k\geq 0}).$$

In addition we will assume that $(X_n)_{n\in\mathbb{Z}}$ satisfies the 1-approximation condition:

Definition 1.11. *We say that $(X_n)_{n\in\mathbb{Z}}$ satisfies the 1-approximating condition, if*

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

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

Example 1.12. *The process $(X_n)_{n\in\mathbb{N}}$ in Example 1.9 satisfies the 1-approximating condition, as*

$$\left\|X_1-E(X_1/\mathcal{F}_0^{l})\right\|_1=\left\|\sum_{k=l+1}^{\infty}\frac{1}{2^{k+1}}Z_k\right\|_1\leq\sum_{k=l+1}^{\infty}\frac{1}{2^{k+1}}=\frac{1}{2^l}=:a_l.$$

Whereas the limit theory for partial sums of weakly dependent processes is very well developed, much less attention has been paid to nonlinear statistics like U-statistics. The summands of $U_n(h_2)$ can be correlated, if the random variables $(X_n)_{n\in\mathbb{N}}$ are dependent, so one has to establish generalized covariance inequalities to derive moment bounds for $U_n(h_2)$. Yoshihara [27] considered absolutely regular processes, Denker and Keller [9] functionals of absolutely regular processes, and Dehling and Wendler [7] strongly mixing sequences.

1.3. **Block Bootstrap.** In many statistical applications, for example in the determination of confidence bands, one faces the task to compute the distribution of a statistic $T_n = T_n(X_1, \dots, X_n)$. This is usually rather difficult, as the distribution F of X_i is unknown, so one often has to use approximation by the normal distribution. Efron [11] proposed the bootstrap as an alternative. For i.i.d. data, the validity of the bootstrap was established by Bickel and Freedman [2], and Singh [25]. Using Edgeworth expansion, one can often show that the bootstrap works better than normal approximation, see Hall [12] for details.

Computation of the distribution of T_n becomes even more difficult when the observations are dependent, e.g., in the case of the sample mean $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$, one gets for weakly dependent data under some technical assumptions

$$\sqrt{n}\left(\bar{X}_n - EX_1\right) \to N\left(0, \sigma^2\right)$$

in distribution, where $\sigma^2 = \operatorname{Var}[X_1] + 2\sum_{i=1}^{\infty} \operatorname{Cov}[X_1, X_{i+1}]$. So one has not only the variance to estimate, but also the autocovariances of the process. The naive bootstrap can fail under dependence, as Singh [25] mentioned. Therefore, block bootstrappings method are commonly used for nonparametric inference under dependence. There are different ways to resample blocks, for example the circular block bootstrap or the moving block bootstrap (for a detailed description of the different bootstrapping methods see Lahiri [17]). For the circular block bootstrap, Shao and Yu [24] have shown that under strong mixing the distribution of the block bootstrap version $\bar{X}_n^\star$ of the sample mean converges almost surely to the same distribution as the sample mean $\bar{X}_n$. Peligrad [19] has proved asymptotic normality of $\bar{X}_n^\star$ under another set of conditions, which does not necessarily imply the central limit theorem for $\bar{X}_n$. Radulovic [20] has established weak consistency under very weak conditions.

We consider the nonoverlapping bootstrap, proposed by Carlstein [5], for the sample mean and for U-statistics. Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of r.v.'s. Let $p \in \mathbb{N}$ be the block length such that $p = p(n) = o(n)$, $p \to \infty$ as $n \to \infty$. We introduce the following blocks of indices and r. v.'s:

$$\begin{aligned} \mathrm{I}_i &= \left(X_{(i-1)p+1}, \dots, X_{ip}\right), \\ \mathrm{B}_i &= \{(i-1)p+1, \dots, ip\}, \quad i = 1, \dots, k \end{aligned}$$

where $k = k(n) = \left[\frac{n}{p}\right]$ is the number of blocks. We consider a new sample $X_1^*, \dots, X_{kp}^*$, which is constructed by choosing randomly and independently blocks k times with

$$\mathrm{P}\left((X_1^*, \dots, X_p^*) = \mathrm{I}_i\right) = \frac{1}{k} \quad i = 1, 2, \dots, k.$$

As a bootstrap version of the sample mean we consider:

$$\bar{X}_{n,kp}^* = \frac{1}{kp}\sum_{i=1}^{kp} X_i^*.$$

With $P^\star$, $E^\star$, $\mathrm{Var}^\star$ we denote the probability, expectation and variance conditionally on $(X_n)_{n\in\mathbb{N}}$. Note that

$$\mathrm{E}^* \bar{X}^*_{n,kp} = \frac{1}{kp}\sum_{i=1}^{kp} X_i =: \bar{X}_{n,kp}.$$

The first aim of this paper is to prove the strong consistency of the nonoverlapping block bootstrap for functionals of absolutely regular sequences (Theorems 2.6 and 2.7), as this class of weak dependent processes covers examples that do not satisfy the strong mixing conditions. The proof of this is based on Theorems 2.1 to 2.3 for general stationary processes, which are similar to the results of Peligrad [19] and of Shao and Yu [24] for the circular block bootstrap. Additionally, we will show the strong consistency of the nonoverlapping block bootstrap for strongly mixing sequences (Theorems 2.4 and 2.5).

Our second aim is to prove the validity of the nonoverlapping block bootstrap for U-statistics. Although the estimation of the distribution for U-statistics is even more complicated than for the sample mean, there is only very little literature on the bootstrap for U-statistics. Bickel and Freedman [2] proved the validity of the bootstrap for nondegenerate U-statistcs of i.i.d. data, Arcones, Giné [1], Dehling, Mikosch [6], and Leucht, Neumann [18] for degenerate U-statistics of i.i.d. data. Dehling and Wendler [7] have shown that the bootstrap distributions of a nondegenerate U-statistics converges to the real distribution in probability for strongly mixing or absolutely regular sequences. We will not only extend this result to functionals of absolutely regular processes, but show the almost sure convergence (Theorem 2.8).

Instead of the nonoverlapping block bootstrap, one can use the following bootstrap based on Bernstein blocking: Let n, p be as above and q such that $q = q(n) = o(p), q \to \infty$ as $n \to \infty$. We introduce the following blocks of indices and r. v.'s:

$$\mathrm{I'}_i = \left(X_{(i-1)(p+q)+1}, \dots, X_{(i-1)(p+q)+p}\right)$$
$$\mathrm{B'}_i = \{(i-1)(p+q)+1, \dots, (i-1)(p+q)+p\}\,, \quad i = 1, \dots, k'$$

where k' is $k'(n) = \left[\frac{n}{p+q}\right]$. The bootstrap sample $X_1^*, \dots, X_{k'p}^*$ is constructed as before by choosing randomly and independently blocks k' times with

$$\mathrm{P}\left((X_1^*, \dots, X_p^*) = \mathrm{I'}_i\right) = \frac{1}{k'} \quad i = 1, 2, \dots k'$$

This bootstrapping method is better adapted to the typical way of proving the central limit theorem under mixing conditions. Especially when the sample size is big, we do fewer calculations than with other blocking methods (nonoverlapping, moving or circular). It is easy to see that Theorems 2.1 to 2.8 are also valid for the Bernstein block bootstrap without any changes.

2. Main Results

2.1. **Bootstrap for stationary sequences.** In this section, and in what follows, we denote by $\bar{X}_n$ the sample mean of the observations $X_1, \dots, X_n$, by $N(0, \sigma^2)$ a Gaussian r.v. with mean zero and variance σ^2 and by C a constant which may depend on several parameters and might have different values even in one chain of inequalities. First we will give theorems for general stationary sequences which are analogues to the results of Peligrad [19], and Shao and Yu [24].

Theorem 2.1. *Let $\{X_i, i \geq 1\}$ be a stationary sequence of r.v'.s such that $EX_1 = \mu$ and $\mathrm{Var}X_1 < \infty$. Assume that the following conditions hold*

$$\mathrm{Var} \quad n^{1/2}(\bar{X}_n - \mu) \to \sigma^2 > 0, \tag{1}$$

$$n^{1/2}(\bar{X}_n - \mu) \to N(0, \sigma^2) \quad \textit{in distribution}, \tag{2}$$

$$p^{1/2}(\bar{X}_{n,kp} - \mu) \to 0 \qquad a.s. \ , \tag{3}$$

$$\frac{1}{kp}\sum_{i=1}^{k}\left[\left(\sum_{j\in B_i}^{p}(X_j-\mu)\right)^2 - E\left(\sum_{j\in B_i}(X_j-\mu)\right)^2\right] \to 0 \qquad a.s. \ , \tag{4}$$

$$\frac{1}{kp}\sum_{i=1}^{k}\left(\sum_{j\in B_i}(X_j-\mu)\right)^2 \mathbb{1}_{\left\{\left|\sum_{j\in B_i}(X_j-\mu)\right|^2 > \epsilon kp\right\}} \to 0 \qquad a.s. \tag{5}$$

for any $\epsilon > 0$. Then the following takes place as $n \to \infty$

$$\mathrm{Var}^*(\sqrt{kp}\bar{X}^*_{n,kp}) \to \sigma^2 \qquad a.s.$$

$$\sup_{x\in\mathbb{R}}\left|P^*\left(\sqrt{kp}(\bar{X}^*_{n,kp} - \bar{X}_{n,kp}) \leq x\right) - P\left(\sqrt{n}(\bar{X}_n - \mu) \leq x\right)\right| \to 0 \qquad a.s.$$

Theorem 2.2. *Let $(X_n)_{n\in\mathbb{N}}$ be a stationary sequence of r.v.'s. with $EX_1 = \mu, \mathrm{Var}X_1 < \infty$. Assume that conditions (1), (2), (4) and for each fixed $x \in \mathbb{R}$*

$$\frac{1}{kp}\sum_{k=1}^{k}\left(\mathbb{1}_{\left\{\frac{1}{\sqrt{p}}\sum_{j\in B_i}(X_j-\mu)\leq x\right\}} - P\left(\frac{1}{\sqrt{p}}\sum_{i=1}^{p}(X_i-\mu) \leq x\right)\right) \to 0 \qquad a.s. \tag{6}$$

hold. Then the statement of Theorem 2.1 remains true.

Theorem 2.3. *Let $(X_n)_{n\in\mathbb{N}}$ be a stationary sequence of bounded almost surely r.v.'s with $EX_1 = \mu$. Assume that (3) and following conditions hold*

$$\frac{p^2}{n} \to 0 \qquad \textit{as } n \to \infty, \tag{7}$$

$$\frac{1}{n}\mathrm{Var}S_n \to \sigma^2 \qquad \textit{as } n \to \infty, \tag{8}$$

$$\frac{1}{kp}\sum_{i=1}^{k}\left(\sum_{j\in B_i}(X_j-\mu)\right)^2 \to \sigma^2 \qquad a.s. \ \textit{as } n \to \infty. \tag{9}$$

Then almost surely as $n \to \infty$

$$\mathrm{Var}^*(\sqrt{kp}\bar{X}^*_{n,kp}) \to \sigma^2, \tag{10}$$

$$\sqrt{kp}(\bar{X}^*_{n,kp} - \bar{X}_{n,kp}) \to N(0, \sigma^2) \ \textit{in distribution}. \tag{11}$$

2.2. **Bootstrap for strongly mixing sequences.** We formulate theorems under assumptions on the strong mixing coefficients which are analogues to the results of Peligrad [19] and Shao, Yu [24].

Theorem 2.4. *Let $(X_n)_{n\in\mathbb{N}}$ be a stationary sequence of strong mixing r.v.'s with $EX_1 = \mu$ and $(E|X_1|^{2+\delta})^{\frac{1}{2+\delta}} < \infty$ for some $0 < \delta \leq \infty$. Assume*

$$\alpha(n) \leq C \cdot n^{-r} \ \textit{for some } C > 0, r > \frac{2+\delta}{\delta},$$

$$p(n) \leq Cn^{\epsilon} \text{ for some } 0 < \epsilon < 1 \text{ and} \tag{12}$$

$$p(n) = p(2^l) \text{ for } 2^l < n \leq 2^{l+1}, \qquad l = 1, 2, \dots \tag{13}$$

Then $\sigma^2 = EX_1^2 + 2\sum_{i=2}^{\infty} \operatorname{Cov}(X_1, X_i) < \infty$ *and in the case* $\sigma^2 > 0$ *the statement of Theorem 2.1 holds.*

Theorem 2.5. *Let* $(X_n)_{n\in\mathbb{N}}$ *be a stationary sequence of almost surely bounded strongly mixing r.v.'s. Assume that (8), (13) and the following conditions hold*

$$\sum_{n=1}^{\infty} \frac{p^2(n)\alpha(p(n))}{n} < \infty, \tag{14}$$

$$\sum_{n=1}^{\infty} \frac{p^3(n)}{n^2} < \infty. \tag{15}$$

Then (10), (11) hold.

Remark. The condition (14) implies $\sum_{n=1}^{\infty} \frac{\alpha(n)}{n} < \infty$. We can reformulate Theorem 2.5 under above condition on mixing coefficients instead of conditions (14) and (15) claiming that there is a sequence $(p\,(n))$ that the statement of Theorem 2.5 holds, as it was done in Peligrad [19].

2.3. **Bootstrap for functionals of absolutely regular sequences.** To prove the validity of the nonoverlapping block bootstrap for functionals of absolutely regular sequences, we need assumptions not only on the decay of mixing coefficients, but also on the decay of the approximation constants. Our mixing conditions are the same as for the central limit theorem in Borovkova et al. [3].

Theorem 2.6. *Let* $(X_n)_{n\in\mathbb{Z}}$ *be a 1-approximating (with constants* $(a_l)_{l\in\mathbb{N}}$*) functional of a stationary absolutely regular process* $(Z_n)_{n\in\mathbb{Z}}$. *Assume that (12), (13) and the following conditions hold for some* $\delta > 0$

$$E\,|X_1|^{4+\delta} < \infty, \ \sum_{k=0}^{\infty} k^2 (a_k^{\frac{\delta}{3+\delta}} + (\beta(k))^{\frac{\delta}{4+\delta}}) < \infty.$$

Then $\sigma^2 = EX_1^2 + 2\sum_{i=2}^{\infty} \operatorname{Cov}(X_1, X_i) < \infty$ *and in the case* $\sigma^2 > 0$ *the statement of Theorem 2.1 holds.*

Theorem 2.7. *Let* $(X_n)_{n\in\mathbb{Z}}$ *be a 1-approximating (with constants* $\{a_l\}$*) functional of stationary and absolutely regular process* $(Z_n)_{n\in\mathbb{Z}}$. *Assume that (12), (13) and the following conditions hold*

$$|X_1| \leq C \text{ a.s. for some } C > 0,$$
$$\sum_{k=0}^{\infty} k^2 (a_k + \beta(k)) < \infty.$$

Then (10) and (11) hold.

2.4. **Bootstrap for U-Statistics.** To bootstrap a U-statistic under dependence, one can apply the nonoverlapping block bootstrap and plug the observations $X_1^\star, \dots, X_n^\star$ in:

$$U_n^\star(h) = \frac{2}{pk(pk-1)} \sum_{1 \le i < j \le pk} h\left(X_i^\star, X_j^\star\right)$$
$$= \theta + \frac{2}{pk} \sum_{i=1}^{pk} h_1\left(X_i^\star\right) + \frac{2}{pk(pk-1)} \sum_{1 \le i < j \le pk} h_2\left(X_i^\star, X_j^\star\right).$$

Theorem 2.8. *Let $(X_n)_{n\in\mathbb{N}}$ be a stationary process and h a kernel with uniform $2+\delta$-moments for a $\delta > 0$. Assume that (12), (13), and one of the following three conditions hold:*

1. *$(X_n)_{n\in\mathbb{N}}$ is absolutely regular and $\beta(n) = O(n^{-\rho})$ for a $\rho > \frac{2+\delta}{\delta}$,*
2. *$(X_n)_{n\in\mathbb{N}}$ is strongly mixing, $E|X_1|^\gamma < \infty$ for a $\gamma > 0$, h is P-Lipschitz-continuous with constant $L > 0$ and $\alpha(n) = O(n^{-\rho})$ for a $\rho > \frac{3\gamma\delta+\delta+5\gamma+2}{2\gamma\delta}$,*
3. *$(X_n)_{n\in\mathbb{N}}$ is a 1-approximating functional of an absolutely regular process, h is P-Lipschitz-continuous with constant $L > 0$ and $E|h_1(X_1)|^{4+\delta}$. For $\alpha_L = \sqrt{2\sum_{i=L}^\infty a_i}$: $\sum_{k=0}^n k^2 \left(\beta^{\frac{\delta}{3+\delta}}(k) + \alpha_k^{\frac{\delta}{4+\delta}}\right) < \infty$,*

then a.s. as $n \to \infty$

$$\operatorname{Var}^\star\left[\sqrt{pk}U_n^\star(h)\right] - \operatorname{Var}\left[\sqrt{n}U_n(h)\right] \to 0, \tag{16}$$

$$\sup_{x\in\mathbb{R}}\left|P^\star\left[\sqrt{pk}\left(U_n^\star(h) - E^\star\left[U_n^\star(h)\right]\right) \le x\right] - P\left[\sqrt{n}\left(U_n(h) - \theta\right) \le x\right]\right| \to 0. \tag{17}$$

3. Preliminary results

3.1. **Central Limit Theorem, Moment and Maximum Inequalities for Partial Sums.** In this subsection we will give some known results which will be used in the next section in the proofs of the theorems. We set

$$S_n = \sum_{i=1}^n X_i.$$

Lemma 3.1 (Ibragimov [16]). *Let $(X_n)_{n\in\mathbb{N}}$ be a stationary sequence of strongly mixing r.v.'s with $EX_1 = \mu$ and $(E|X_1|^{2+\delta})^{\frac{1}{2+\delta}} < \infty$ some $0 < \delta \le \infty$. Assume that*

$$\sum_{n=1}^\infty \alpha^{\frac{\delta}{2+\delta}}(n) < \infty.$$

Then $\sigma^2 = \operatorname{Var}X_1 + 2\sum_{k=2}^\infty \operatorname{Cov}(X, X) < \infty$ and $\frac{\operatorname{Var}S_n}{n} \to \sigma^2$. If in addition $\sigma^2 > 0$, then

$$n^{1/2}(\bar{X}_n - \mu) \to N(0, \sigma^2) \quad \textit{in distribution.}$$

Lemma 3.2 (Shao [23]). *Let $(\xi_n)_{n\in\mathbb{N}}$ be a strongly mixing sequence of r. v.'s with $E\xi_i = 0$ and $(E|\xi_i|^s)^{1/s} \le \mathcal{D}_n$ for $1 \le i \le n$ and for some $1 < s \le \infty$. Assume that*

$$\alpha(i) \le C_0 i^{-\theta} \textit{ for some } C_0 > 1 \textit{ and } \theta > 0.$$

Then there exists a constant $K = K(C_0, \theta, s)$, such that for any $x \geq K\mathcal{D}_n n^{1/2}\log n$

$$P(\max_{i\leq n}\left|\sum_{j=1}^{i}\xi_j\right| \geq x) \leq Kn(\frac{\mathcal{D}_n}{x})^{\frac{s(\theta+1)}{s+\theta}}(\log\frac{x}{\mathcal{D}_n})^{\theta}.$$

Lemma 3.3 (Yokoyama [26]). *Let $(X_n)_{n\in\mathbb{N}}$ be a stationary strongly mixing sequence of r.v.'s with $EX_1 = \mu$ and $(E|X_1|^{2+\delta})^{\frac{1}{2+\delta}} < \infty$ for some $0 < \delta \leq \infty$ suppose that $2 \leq s < 2+\delta$ and*

$$\sum_{n=1}^{\infty} n^{\frac{s}{2}-1}(\alpha(n))^{(2+\delta-s)/(2+\delta)} < \infty.$$

Then there exists a constant C depending only on s, δ and the mixing coefficients $(\alpha(n))_{n\in\mathbb{N}}$ such that

$$E\left|\sum_{i=1}^{n}(X_i-\mu)\right|^s \leq Cn^{s/2}(E|X_1|^{2+\delta})^{\frac{s}{2+\delta}}.$$

Lemma 3.4 (Rio [21], Peligrad [19]). *Let $(X_n)_{n\in\mathbb{N}}$ be a strongly mixing sequence of r. v.'s with $EX_i = 0$ and $|X_i| \leq C$ a.s. Then there is a universal constant K such that for every $x > 0$ and $n \geq 1$*

$$P\left(\max_{1\leq i\leq n}|S_i| > x\right) \leq Kx^{-2}\left(\sum_{i=1}^{n}EX_i^2 + C^2\cdot n\sum_{i=1}^{n}\alpha(i)\right).$$

Lemma 3.5 (Borovkova et al. [3]). *Let $(X_n)_{n\in\mathbb{N}}$ be a 1-approximating functional with constants $(a_k)_{k\in\mathbb{N}}$ of an absolutely regular stationary process $(Z_n)_{n\in\mathbb{Z}}$ with mixing coefficients $(\beta(k))_{k\in\mathbb{N}}$. Suppose that one of the following conditions holds*

(1) $EX_0 = 0, E|X_0|^{4+\delta} < \infty$,
$\sum_{k=1}^{\infty}k^2(a_k^{\frac{\delta}{3+\delta}} + (\beta(k))^{\frac{\delta}{4+\delta}}) < \infty$ *for some* $\delta > 0$.
(2) X_0 *is bounded a.s.*, $EX_0 = 0$,
$\sum_{k=1}^{\infty}k^2(a_k+\beta(k)) < \infty$.

Then $\sigma^2 = EX_0^2 + 2\sum_{k=1}^{\infty}EX_0X_k < \infty$ and in the case $\sigma^2 > 0$ we have

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}X_i \to N(0,\sigma^2) \quad \text{in distribution as } n\to\infty.$$

Lemma 3.6 (Borovkova et al. [3]). *Let $(X_n)_{n\in\mathbb{Z}}$ be a 1-approximating functional with constants $(a_k)_{k\in\mathbb{N}}$ of an absolutely regular stationary process $(Z_n)_{n\in\mathbb{Z}}$ with mixing coefficients $(\beta(k))_{k\in\mathbb{N}}$. Assume that one of the conditions of Lemma 3.5 holds. Then there exists a constant C such that*

$$ES_n^4 \leq Cn^2.$$

Lemma 3.7 (Borovkova et al. [3]). *Let $(X_n)_{n\in\mathbb{Z}}$ be a 1-approximating functional with constants $(a_k)_{k\in\mathbb{N}}$ of an absolutely regular process $(Z_n)_{n\in\mathbb{Z}}$ with mixing coefficients $(\beta(k))_{k\in\mathbb{N}}$. Assume that $EX_0 = 0$ and one of the following two conditions holds:*

(1) X_0 *is bounded a. s. and* $\sum_{k=0}^{\infty}(a_k+\beta(k)) < \infty$,
(2) $E|X_0|^{2+\delta} < \infty$ *and* $\sum_{k=0}^{\infty}(a_k^{\frac{\delta}{1+\delta}} + (\beta(k))^{\frac{\delta}{2+\delta}}) < \infty$.

Then there exists a constant C such that

$$ES_n^2 \leq Cn.$$

Lemma 3.8 (Borovkova et al. [3]). *Let $(X_n)_{n\in\mathbb{Z}}$ be a 1-approximating functional with constants $(a_k)_{k\in\mathbb{N}}$ of an absolutely regular process $(Z_n)_{n\in\mathbb{Z}}$ with mixing coefficients $(\beta(k))_{k\in\mathbb{N}}$.*
Then

(1) *if $|X_0| \leq M$ a. s. for all non-negative integers $i \leq j < k \leq l$, we have*

$$|E(X_iX_jX_kX_l) - E(X_iX_j)E(X_kX_l)| \leq$$
$$\left(4(\beta([\frac{k-j}{3}]))^{\frac{\delta}{2+\delta}}(E|X_0|^{2+\delta})^{\frac{2}{2+\delta}} + 8(a_{[\frac{k-j}{3}]})^{\frac{\delta}{1+\delta}}(E|X_0|^{2+\delta})^{\frac{1}{1+\delta}}\right) \cdot M^2.$$

(2) *if $E|X_0|^{4+\delta} < \infty$ for all non-negative $i \leq j < k \leq l$, we have*

$$|E(X_iX_jX_kX_l) - E(X_iX_j)E(X_kX_l)|$$
$$\leq 4(\beta([\frac{k-j}{3}]))^{\frac{\delta}{4+\delta}}(E|X_0|^{4+\delta})^{\frac{4}{4+\delta}} + 8(a_{[\frac{k-j}{3}]})^{\frac{\delta}{3+\delta}}(E|X_0|^{4+\delta})^{\frac{3}{3+\delta}}.$$

3.2. **Moment Inequalities for U-Statistics.** To control the moments of degenerate U-statistics, we need bounds for the covariance of h_2. Recall that h_2 is defined as

$$h_2(x,y) := h(x,y) - h_1(x) - h_1(y) - \theta.$$

In the following three lemmas, let be $m = \max\left\{i_{(2)} - i_{(1)}, i_{(4)} - i_{(3)}\right\}$, where $i_{(1)} \leq i_{(2)} \leq i_{(3)} \leq i_{(4)}$ and $\{i_1, i_2, i_3, i_4\} = \left\{i_{(1)}, i_{(2)}, i_{(3)}, i_{(4)}\right\}$.

Lemma 3.9 (Yoshihara [27]). *Let h be a kernel with uniform $2+\delta$-moments for a $\delta > 0$. If $(X_n)_{n\in\mathbb{N}}$ is a stationary process, then there is a constant C, such that*

$$|E[h_2(X_{i_1}, X_{i_2})h_2(X_{i_3}, X_{i_4})]| \leq C\beta^{\frac{\delta}{2+\delta}}(m).$$

Lemma 3.10 (Dehling, Wendler [7]). *Let h be a P-Lipschitz-continuous kernel with constant L and with uniform $2+\delta$-moments for some $\delta > 0$, $(X_n)_{n\in\mathbb{N}}$ a stationary sequence of random variables. If there is a $\gamma > 0$ with $E|X_k|^\gamma < \infty$, then there exists a constant C, such that the following inequality holds:*

$$|E[h_2(X_{i_1}, X_{i_2})h_2(X_{i_3}, X_{i_4})]| \leq C\alpha^{\frac{2\gamma\delta}{3\gamma\delta+\delta+5\gamma+2}}(m).$$

Lemma 3.11 (Dehling, Wendler [8]). *Let h be a P-Lipschitz-continuous kernel with constant L and with uniform $2+\delta$-moments for some $\delta > 0$, and $(X_n)_{n\in\mathbb{N}}$ a 1-approximating functional of an absolutely regular process with constants a_l. Define α_L as $\alpha_L = \sqrt{2\sum_{i=L}^{\infty} a_i}$ and $(\beta(n))_{n\in\mathbb{N}}$ as the mixing coefficents of $(Z_n)_{n\in\mathbb{N}}$. Then:*

$$|E[h_2(X_{i_1}, X_{i_2})h_2(X_{i_3}, X_{i_4})]| \leq C\beta^{\frac{\delta}{2+\delta}}\left(\lfloor\frac{m}{3}\rfloor\right) + C\alpha_{\lfloor\frac{m}{3}\rfloor}^{\frac{\delta}{2+\delta}}.$$

Yoshihara [27] deduced the following moment bound under condition 1. of the lemma below with the help of Lemma 3.9. The result follows from conditions 2. and 3. in the same way using the Lemmas 3.10 and 3.11 instead.

Lemma 3.12. *Let $(X_n)_{n\in\mathbb{N}}$ be a stationary process and h a kernel with uniform $2+\delta$-moments for a $\delta > 0$. Let be $\tau \geq 0$ such that one of the following three conditions holds:*

(1) $(X_n)_{n\in\mathbb{N}}$ *is absolutely regular and* $\sum_{k=0}^{n} k\beta^{\frac{\delta}{2+\delta}}(k) = O(n^{\tau})$.
(2) $(X_n)_{n\in\mathbb{N}}$ *is strongly mixing,* $E|X_1|^{\gamma} < \infty$ *for a* $\gamma > 0$, h *is* P*-Lipschitz-continuous with constant* $L > 0$ *and* $\sum_{k=0}^{n} k\alpha^{\frac{2\gamma\delta}{\gamma\delta+\delta+5\gamma+2}}(k) = O(n^{\tau})$.
(3) $(X_n)_{n\in\mathbb{N}}$ *is a 1-approximating functional of an absolutely regular process,* h *is* P*-Lipschitz-continuous with constant* $L > 0$ *and for* $\alpha_L = \sqrt{2\sum_{i=L}^{\infty} a_i}$: $\sum_{k=0}^{n} k\left(\beta^{\frac{\delta}{2+\delta}}(k) + \alpha_k^{\frac{\delta}{2+\delta}}\right) = O(n^{\tau})$.

Then:

$$\sum_{i_1,i_2,i_3,i_4=1}^{n} |E[h_2(X_{i_1}, X_{i_2}) h_2(X_{i_3}, X_{i_4})]| = O\left(n^{2+\tau}\right).$$

4. Proofs of the theorems

4.1. **Bootstrap for stationary sequences.** The proofs of Theorems 2.1 - 2.5 are mainly based on the methods developed in Peligrad [19] and Shao, Yu [24]. We will give full proofs for completeness.

Proof of Theorem 2.1. We note that

$$\sqrt{kp}(\bar{X}^{\star}_{n,kp} - \bar{X}_{n,kp}) = \frac{1}{\sqrt{k}} \sum_{i=1}^{k} Z^{*}_{i,n}$$

where $Z^{*}_{i,n} = \frac{1}{\sqrt{p}} \sum_{j\in B_i} (X^{*}_j - \bar{X}_{n,kp})$, $i = 1, \ldots, k$ are i.i.d. r.v.'s (conditionally on $(X_n)_{n\in\mathbb{N}}$). By simple calculations we have

$$E^{*} Z^{*2}_{1,n} = \frac{1}{kp} \sum_{i=1}^{k} (\sum_{j\in B_i} (X_j - \bar{X}_{n,kp}))^2 = \frac{1}{kp} \sum_{i=1}^{k} (\sum_{j\in B_i} (X_j - \mu))^2 - p(\bar{X}_{n,kp} - \mu)^2$$

$$= \frac{1}{kp} \sum_{i=1}^{k} \left((\sum_{j\in B_i} (X_j - \mu))^2 - E(\sum_{j\in B_1} (X_j - \mu))^2 \right) + \frac{1}{p} \operatorname{Var} S_p - p(\bar{X}_{n,kp} - \mu)^2.$$

Conditions (1), (3) and (4) imply that a.s. as $n \to \infty$

$$E^{*} Z^{*2}_{1,n} \to \sigma^2, \quad \text{and consequently } \operatorname{Var}^{\star}\left[\sqrt{kp} X^{\star}_{n,kp}\right] \to \sigma^2. \tag{18}$$

For any $\epsilon > 0$, we have

$$\begin{aligned}
E^{*} Z^{*2}_{i,n} \mathbb{1}_{\left\{Z^2_{i,n} > \epsilon k\right\}} &= \frac{1}{kp} \sum_{i=1}^{k} (\sum_{j\in B_i} (X_j - \bar{X}_{n,kp}))^2 \mathbb{1}_{\left\{\left|\sum_{j\in B_i}(X_j - \bar{X}_{n,kp})\right|^2 > kp\epsilon\right\}} \\
&\leq \frac{4}{kp} \sum_{i=1}^{k} (\sum_{j\in B_i} (X_j - \mu))^2 \mathbb{1}_{\left\{(\sum_{j\in B_i}(X_j-\mu))^2 > \frac{\epsilon kp}{4}\right\}} \\
&\quad + \frac{4}{kp} \sum_{i=1}^{k} p^2 (\bar{X}_{n,kp} - \mu)^2 \mathbb{1}_{\left\{p(\bar{X}_{n,kp}-\mu)^2 > \frac{\epsilon kp}{4}\right\}} \\
&\leq \frac{4}{kp} \sum_{i=1}^{k} (\sum_{j\in B_i} (X_j - \mu))^2 \mathbb{1}_{\left\{(\sum_{j\in B_i}(X_j-\mu))^2 > \frac{\epsilon kp}{4}\right\}} + 4p(\bar{X}_{n,kp} - \mu)^2.
\end{aligned}$$

Now (3) and (5) imply that

$$E^* Z_{1,n}^{*2} \mathbb{1}_{\{(Z_{1,n}^{*2} > \epsilon k\}} \to 0 \text{ a. s. as } n \to \infty \tag{19}$$

what means that $Z_{i,n}^*, i = 1, 2 \ldots$ satisfies the Lindeberg condition. Thus (18) and (19) imply the statement of the theorem. □

Proof of Theorem 2.2. We define

$$\widetilde{F}_n(x) = \frac{1}{k} \sum_{i=1}^{k} \mathbb{1}_{\left\{\frac{1}{p}\sum_{j=(i-1)p+1}^{ip}(X_j^* - \mu) \le x\right\}}.$$

It is easy to see that the r.v.'s

$$\left\{ \frac{1}{\sqrt{p}} \sum_{j=(i-1)p+1}^{ip} (X_j^* - \mu) \right\}, i = 1, \ldots, k$$

are i.i.d. with distribution function $\widetilde{F}_n(x)$. We denote by $\widetilde{F}_n^{(m)}$ the distribution function of

$$(kp)^{1/2}(\bar{X}_{n,kp}^* - \bar{X}_{n,kp}) = \frac{1}{\sqrt{k}} \sum_{i=1}^{k} \left(\frac{1}{\sqrt{p}} \sum_{j=(i-1)p+1}^{ip} ((X_j^* - \mu) - E^*(X_j^* - \mu)) \right).$$

Note that

$$\begin{aligned}\int x^2 d\widetilde{F}_n(x) =& \frac{1}{kp} \sum_{i=1}^{k} \left((\sum_{j \in B_i} (X_j - \mu))^2 - E(\sum_{j=1}^{p} (X_j - \mu))^2 + E(\sum_{j=1}^{p} (X_j - \mu))^2 \right) \\ =& \frac{1}{kp} \sum_{i=1}^{k} \left((\sum_{j \in B_i} (X_j - \mu))^2 - E(\sum_{j=1}^{p} (X_j - \mu))^2 \right) + \frac{1}{p} \mathrm{Var} S_p.\end{aligned}$$

From (1) and (4) we have

$$\int x^2 d\widetilde{F}_n(x) \to \sigma^2 \qquad \text{a.s. as } n \to \infty.$$

Now conditions (2) and (6) imply

$$\widetilde{F}_n(x) \to N(0, \sigma^2) \text{ a.s.}$$

The rest of the proof is the same as in the proof of Theorem 2.2 of Shao and Yu [24]. □

Proof of Theorem 2.3. First we will prove the following proposition which is analogue of Proposition 3.1 of Peligrad [19] (In this proposition we assume that $(x_n)_{n\in\mathbb{N}}$ is a fixed realization of $(X_n)_{n\in\mathbb{N}}$).

Proposition 4.1. *Let $(x_n)_{n\in\mathbb{N}}$ be a bounded sequence of real numbers. For each n, let $T_{n1}, T_{n2}, \ldots T_{nk}$ be independent r.v.'s uniformly distributed on $\{1, 2, \ldots, k\}$. Assume that (7) and*

$$V_n = \frac{p}{k} \sum_{i=1}^{k} (\bar{x}_{pi} - \bar{x}_{n,kp})^2 \to \sigma^2 > 0 \tag{20}$$

hold, where $\bar{x}_{pi} = \frac{1}{p}\sum_{j\in B_i} x_j$, $\bar{x}_{n,pk} = \frac{1}{kp}\sum_{i=1}^{kp} x_i$. *Then*

$$\sqrt{kp}(\bar{X}^*_{n,kp} - \bar{x}_{n,kp}) = \sqrt{kp}(\frac{1}{k}\sum_{j=1}^{k}\sum_{i=1}^{k} \mathbb{1}_{\{T_{nj}=i\}}\bar{x}_{pi} - \bar{x}_{n,kp}) \to N(0,\sigma^2)$$

in distribution.

Proof. We have

$$\mathrm{Var}\bar{X}^*_{n,kp} = \frac{1}{k^2}\sum_{i=1}^{k}(\bar{x}_{pi} - \bar{x}_{n,kp})^2$$

and by (2) obtain

$$\mathrm{Var}(\sqrt{kp}\bar{X}^*_{n,kp}) = \frac{p}{k}\sum_{i=1}^{k}(\bar{x}_{pi} - \bar{x}_{n,kp})^2 \to \sigma^2.$$

Note that

$$\sqrt{kp}(\bar{X}^*_{n,kp} - \bar{x}_{n,kp}) = \sqrt{kp}(\frac{1}{k}\sum_{j=1}^{k}\sum_{i=1}^{k} \mathbb{1}_{\{T_{nj}=i\}}\bar{x}_{pi} - \bar{x}_{n,kp})$$
$$= \sqrt{kp}(\frac{1}{k}\sum_{j=1}^{k}\sum_{i=1}^{k} \mathbb{1}_{\{T_{nj}=i\}}(\bar{x}_{pi} - \bar{x}_{n,kp})) = \sum_{j=i}^{k} U_{nj},$$

where

$$U_{nj} = \frac{\sqrt{p}}{\sqrt{k}}\sum_{i=1}^{k} \mathbb{1}_{\{T_{nj}=i\}}(\bar{x}_{pi} - \bar{x}_{n,kp}).$$

In our case Lindeberg condition holds if

$$E\max_{1\le j\le k} U^2_{nj} \to 0 \qquad \text{as } n\to\infty.$$

Taking into account that r. v.'s are bounded we have

$$|U_{nj}| \le \sqrt{\frac{p}{k}}\max_{i\le i\le k}|\bar{x}_{pi} - \bar{x}_{n,kp}| = O(\sqrt{\frac{p}{k}}).$$

and

$$|U_{nj}|^2 = O(\frac{p}{k}) = O(\frac{p^2}{n})$$

Now the condition (7) implies the statement of the proposition. □

In order to prove Theorem 2.3 we will show that

$$V_n = \frac{p}{k}\sum_{i=1}^{k}(\bar{X}_{pi} - \bar{X}_{n,kp})^2 \to \sigma^2 \text{ a.s. as } n\to\infty.$$

W.l.g. assume that $\mu = EX_1 = 0$. Set $S_{pi} = \sum_{j\in B_i} X_j$. Then

$$V_n = \frac{p}{k}\sum_{i=1}^{k}\bar{X}^2_{pi} - p\bar{X}^2_{n,kp} = \frac{p}{k}\sum_{i=1}^{k}\frac{S^2_{pi}}{p^2} - p\bar{X}^2_{n,kp}$$
$$= \frac{1}{k}\sum_{i=1}^{k}\frac{S^2_{pi} - ES^2_{pi}}{p} + \frac{\mathrm{Var}S_p}{p} - p\bar{X}^2_{n,kp}.$$

Conditions (3), (8) and (9) imply

$$V_n \to \sigma^2 \text{ a.s. as } n \to \infty.$$

This completes the proof of Theorem 2.3. □

4.2. **Bootstrap for strongly mixing sequences.** Before giving the proofs of Theorems 2.4 and 2.5, we want to remind the reader of a well-known fact that we will use very often: In order to prove $S_n = X_1 + \ldots + X_n \to 0$ a.s. it suffices to show that

$$\max_{2^k < n \le 2^{k+1}} |S_n| \to 0 \text{ a.s. as } k \to \infty.$$

Proof of Theorem 2.4. The proof is based on Theorem 2.2. Lemma 3.1 implies that conditions (1) and (2) of Theorem 2.2 are satisfied. It remains to prove (4) and (6). W.l.g. we can assume that $\mu = 0$ and we will prove

$$\frac{1}{kp}\sum_{i=1}^{k}\left((\sum_{j=1}^{p} X_{(i-1)p+j})^2 - E(\sum_{i=1}^{p} X_i)^2\right) \to 0 \text{ a.s. as } n \to \infty.$$

By the Borel-Cantelli lemma, it suffices to show that

$$\sum_{l=1}^{\infty} P\left(\max_{2^l < n \le 2^{l+1}} \left|\frac{1}{kp}\sum_{i=1}^{k}\left\{(\sum_{j=1}^{p} X_{(i-1)p+j})^2 - ES_p^2\right\}\right| > \epsilon\right) < \infty.$$

Taking into account that $kp \backsim n$ we have

$$\begin{aligned} I_l &:= P\left(\max_{2^l < n \le 2^{l+1}} \left|\sum_{i=1}^{k(n)}\left((\sum_{j=1}^{p(2^l)} X_{(i-1)p+j})^2 - ES_p^2\right)\right| > \epsilon k(2^l)p(2^l)\right) \\ &\le P\left(\max_{n \le 2^{l+1}} \left|\sum_{i=1}^{k(n)}\left((\sum_{j=1}^{p(2^l)} X_{(i-1)p+j})^2 - ES_{p(2^l)}^2\right)\right| > \epsilon C 2^l\right) \\ &\le P\left(\max_{m \le k(2^{l+1})} \left|\sum_{i=1}^{m}\left((\sum_{j=1}^{p(2^l)} X_{(i-1)p+j})^2 - ES_{p(2^l)}^2\right)\right| > \epsilon C 2^l\right). \end{aligned}$$

From Lemma 3.3 we have for $s > 1$ that

$$(E\left|(\sum_{j=1}^{p(2^l)} X_{(i-1)p+j})^2 - ES_{p(2^l)}^2\right|^s)^{1/s} \le C(E\left|S_{p(2^l)}\right|^{2s})^{1/s} \le Cp(2^l). \tag{21}$$

Now using Lemma 3.2 and taking into account (21), we obtain

$$I_l \le Ck(2^{l+1})(\frac{p(2^l)}{2^l})^{\frac{s(r+1)}{s+r}} \log^r(\frac{2^l}{p(2^l)}) \le C(\frac{p(2^l)}{2^l})^{\frac{(s-1)r}{s+r}} \log^r(\frac{2^l}{p(2^l)}).$$

From the condition (12), it follows that

$$\sum_{l=1}^{\infty} P\left(\max_{2^l < n \le 2^{l+1}} \left|\frac{1}{kp}\sum_{i=1}^{k}\left((\sum_{j=1}^{p} X_{(i-1)p+j})^2 - ES_p^2\right)\right| > \epsilon\right) \le \sum_{l=1}^{\infty} I_l < \infty.$$

It remains prove (6), i.e.

$$\frac{1}{kp}\sum_{i=1}^{k}\left(\mathbb{1}_{\left\{\frac{1}{\sqrt{p}}\sum_{j=1}^{p}X_{(i-1)p+j}\leq x\right\}}-P\left(\frac{1}{\sqrt{p}}\sum_{i=1}^{p}X_i\leq x\right)\right)\to 0 \text{ a.s.}$$

Because of the Borel-Cantelli lemma, it suffices to show that

$$\sum_{l=1}^{\infty}P\left(\max_{2^l<n\leq 2^{l+1}}\left|\sum_{i=1}^{k}\left(\mathbb{1}_{\left\{\frac{1}{\sqrt{p}}\sum_{j=1}^{p}X_{(i-1)p+j}\leq x\right\}}\right.\right.\right.$$
$$\left.\left.\left.-P\left(\frac{1}{\sqrt{p}}\sum_{i=1}^{p}X_i\leq x\right)\right)\right|>\epsilon k(2^l)p(2^l)\right)<\infty.$$

Using Lemma 3.4 we conclude

$$\begin{aligned}II_l :=&P\left(\max_{2^l<n\leq 2^{l+1}}\left|\sum_{i=1}^{k(n)}\left(\mathbb{1}_{\left\{\frac{1}{\sqrt{p(2^l)}}\sum_{j=1}^{p(2^l)}X_{(i-1)p+j}\leq x\right\}}\right.\right.\right.\\
&\left.\left.\left.-P(\frac{1}{\sqrt{p(2^l)}}\sum_{i=1}^{p(2^l)}X_i\leq x)\right)\right|>\epsilon k(2^l)p(2^l)\right)\\
\leq&P\left(\max_{m\leq k(2^{l+1})}\left|\sum_{i=1}^{m}\left(\mathbb{1}_{\left\{\frac{1}{\sqrt{p(2^l)}}\sum_{j=1}^{p(2^l)}X_{(i-1)p+j}\leq x\right\}}\right.\right.\right.\\
&\left.\left.\left.-\ P(\frac{1}{\sqrt{p((2^l)}}\sum_{i=1}^{p(2^l)}X_i\leq x)\right)\right|>\epsilon k(2^l)p(2^l)\right)\\
\leq&\frac{C(k(2^{l+1})+k(2^{l+1})\sum_{i=1}^{k(2^{l+1})}\bar{\alpha}(i))}{\epsilon^2k^2(2^l)p^2(2^l)}.\end{aligned}$$

where $\bar{\alpha}(i)=\alpha\left((i-1)p(2^l)+1\right)$. As $\sum_{i=1}^{\infty}\alpha\left(i\right)<\infty$:

$$II_l\leq\frac{Ck(2^{l+1})}{\epsilon^2k^2(2^l)p^2(2^l)}.$$

From (12) we get that

$$\sum_{l=1}^{\infty}II_l<\infty.$$

This completes the proof of Theorem 2.4. □

Proof of Theorem 2.5. W.l.g. we assume that $\mu=EX_1=0$. Because of Theorem 2.3 we need to prove (3) and

$$\frac{1}{k}\sum_{i=1}^{k}\frac{S_{p,i}^2-ES_{p,i}^2}{p}\to 0 \text{ a.s. as } n\to\infty \tag{22}$$

where $S_{p,i}=\sum_{j\in B_i}X_j$ is the sum of the i-th block. We note that since $|X_i|<C$ a.s. $i=1,2\dots$ we have $|S_{p,i}|\leq C\cdot p$ and the sequences $S_{p,2i}$ and $S_{p,(2i-1)}$, $i=1,2,\dots$ are strongly mixing with coefficients

$$\tilde{\alpha}(i)=\alpha((2i-1)p+1),\qquad i=1,2,\dots$$

Lemma 3.4 implies that

$$
\begin{aligned}
(23)\quad & P(\max_{1\le i\le k}\left|\sum_{j=1}^{i}(S_{p,j}^2-ES_{p,j}^2)\right|>2x) \\
&\le P(\max_{1\le i\le k}\left|\sum_{\substack{j=1\\ j\text{ odd}}}^{i}(S_{p,j}^2-ES_{p,j}^2)\right|>x)+P(\max_{1\le i\le k}\left|\sum_{\substack{j=1\\ j\text{ even}}}^{i}(S_{p,j}^2-ES_{p,j}^2)\right|>x) \\
&\qquad\qquad\le Cx^{-2}\left(kES_{p,1}^4+p^4k\sum_{1\le i\le\frac{k}{2}}\tilde{\alpha}(i)\right)
\end{aligned}
$$

for some constant $C>0$. In order to establish (22) it suffices to show

$$\sum_{l=1}^{\infty}P(\max_{2^l<n\le 2^{l+1}}\frac{1}{k\,(n)}\left|\sum_{i=1}^{k(n)}\frac{S_{p,i}^2-ES_{p,i}^2}{p}\right|>2\epsilon)<\infty.$$

Using (23) and (13) we obtain for any $\epsilon>0$

$$
\begin{aligned}
III_l &:= P(\max_{2^l<n\le 2^{l+1}}\frac{1}{k}\left|\sum_{i=1}^{k}\frac{S_{p,i}^2-ES_{p,i}^2}{p}\right|>2\epsilon) \\
&\le P(\max_{m\le k(2^{l+1})}\left|\sum_{i=1}^{m}(S_{p(2^l),i}^2-ES_{p(2^l),i}^2)\right|>\epsilon p(2^l)k(2^l)) \\
&\le \frac{C(k(2^{l+1})p^4(2^l)+p^4(2^l)\cdot k(2^{l+1})\sum_{i=1}^{k(2^{k+1})}\tilde{\alpha}(i)}{\epsilon^2p^2(2^l)k^2(2^l)} \\
&\le \frac{C(k(2^{l+1})p^4(2^l)+p^4(2^l)k^2(2^{l+1})\alpha(p(2^l)+1)}{\epsilon^2p^2(2^l)k^2(2^{l+1})} \\
&\le C\left(\frac{p^2(2^l)}{\epsilon^2k(2^{l+1})}+\frac{p^2(2^l)\alpha(p(2^l)+1)}{\epsilon^2}\right).
\end{aligned}
$$

By (14) and (15) we have that

$$\sum_{l=1}^{\infty}III_l<\infty.$$

It remains to prove (3), i. e.

$$\sqrt{p}\bar{X}_{n,kp}\to 0 \text{ a.s. as } n\to\infty.$$

It suffices to prove that for any $\epsilon>0$

$$\sum_{l=1}^{\infty}P(\max_{2^l<n\le 2^{l+1}}\left|\sum_{i=1}^{k}S_{p(2^l),i}\right|\ge\epsilon k(2^l)\sqrt{p(2^l)})<\infty.$$

Using Lemma 3.4 we obtain

$$\begin{aligned}
IV_l :=& P\left(\max_{2^l<n\le 2^{l+1}}\left|\sum_{i=1}^{k} S_{p(2^l),i}\right| \ge \epsilon k(2^l)p\sqrt{p(2^l)}\right)\\
\le& P\left(\max_{2^l<n\le 2^{l+1}}\left|\sum_{\substack{i=1\\ i \text{ odd}}}^{k} S_{p(2^l),i}\right| \ge \frac{\epsilon}{2}k(2^l)\sqrt{p(2^l)}\right)\\
&+ P\left(\max_{2^l<n\le 2^{l+1}}\left|\sum_{\substack{i=1\\ i \text{ even}}}^{k} S_{p(2^l),i}\right| \ge \frac{\epsilon}{2}k(2^l)\sqrt{p(2^l)}\right)\\
\le& \frac{C(k(2^{l+1})p^2(2^l) + p^2(2^l)k(2^{l+1})\sum_{i=1}^{k(2^{l+1})}\tilde{\alpha}(i)}{\epsilon k^2(2^l)p(2^l)}\\
\le& C\left(\frac{p(2^l)}{\epsilon k(2^l)} + \frac{p(2^l)\alpha(p(2^l)+1)}{\epsilon}\right).
\end{aligned}$$

By conditions (14), (15) we have $\sum_{l=1}^{\infty} IV_l < \infty$. (3) follows by the Borel-Cantelli lemma and Theorem 2.5 is proved. □

4.3. **Bootstrap for functionals of absolutely regular sequences.**

Proof of Theorem 2.6. The proof is based on Theorem 2.1. Lemma 3.5 implies the conditions (1) and (2). W.l.g. we will assume that $EX_1 = \mu = 0$ and first we will prove (5). In order to do that it suffices to show that for any $\epsilon > 0$ and $\epsilon_1 > 0$

$$\sum_{l=1}^{\infty} P\left(\max_{2^l<n\le 2^{l+1}}\left|\frac{1}{kp}\sum_{i=1}^{k} S_{p,i}^2 \mathbb{1}_{\{|S_{p,i}|^2>\epsilon kp\}}\right| > \epsilon_1\right) < \infty \tag{24}$$

where $S_{p,i} = \sum_{j\in B_i} X_j$. Using Markov, Hölder, Chebyshev inequalities, (13), and Lemma 3.6 we obtain

$$\begin{aligned}
&P(\max_{2^l<n\le 2^{l+1}}\left|\frac{1}{kp}\sum_{i=1}^{k} S_{p,i}^2 \mathbb{1}_{\{|S_{p,i}|^2>\epsilon kp\}}\right| > \epsilon_1)\\
\le& P(\max_{2^l<n\le 2^{l+1}}\left|\sum_{i=1}^{k} S_{p,i}^2 \mathbb{1}_{\{|S_{p,i}|^2>\epsilon k(2^l)p(2^l)\}}\right| > \epsilon_1 k(2^l)p(2^l))\\
\le& \frac{E(\max_{2^l<n\le 2^{l+1}}\left|\sum_{i=1}^{k} S_{p,i}^2 \mathbb{1}_{\{|S_{p,i}|^2>\epsilon k(2^l)p(2^l)\}}\right|}{\epsilon_1 k(2^l)p(2^l)}\\
\le& \frac{\sum_{i=1}^{k(2^{l+1})} ES_{p,i}^2 \mathbb{1}_{\{|S_{p,i}|^2>\epsilon k(2^l)p(2^l)\}}}{\epsilon_1 k(2^l)p(2^l)} \le \frac{k(2^{l+1})ES_p^2 \mathbb{1}_{\{(|S_p|^2>\epsilon k(2^l)p(2^l)\}}}{\epsilon_1 k(2^l)p(2^l)}\\
\le& \frac{C(ES_p^4)^{1/2}(P(|S_p|^2 > \epsilon k(2^l)p(2^l))^{1/2}}{\epsilon_1 p(2^l)} \le \frac{C \quad ES_p^4}{\epsilon_1 \cdot \epsilon k(2^l)p^2(2^l)}\\
\le& \frac{C \quad p^2(2^l)}{\epsilon_1 \epsilon k(2^l)p^2(2^l)} \le \frac{C}{\epsilon_1 \cdot \epsilon k(2^l)}.
\end{aligned}$$

The latter implies (24) and thus (5) is proved. Now we will prove (3). Note that by stationarity and Lemma 3.7 we have for any $a \geq 1$ and some $C > 0$

$$E\left|\sum_{i=1}^{a} S_{p,i}\right|^2 \leq Cap.$$

Theorem A of Serfling [22] implies that

$$E\left[\max_{1\leq a\leq m}\left|\sum_{i=1}^{a} S_{p,i}\right|^2\right] \leq Cm(\log_2 2m)^2 p. \tag{25}$$

In order to prove (3) it suffices to show that for any $\epsilon > 0$

$$\sum_{l=1}^{\infty} P(\max_{2^l<n\leq 2^{l+1}}\left|p^{1/2}\bar{X}_{n,kp}\right| > \epsilon) < \infty. \tag{26}$$

By Chebyshev inequality and (25), it follows that

$$\begin{aligned}
&P(\max_{2^l<n\leq 2^{l+1}}\left|p^{1/2}\bar{X}_{n,kp}\right| > \epsilon) \leq P(\max_{2^l<n\leq 2^{l+1}}\left|\sum_{i=1}^{k} S_{p,i}\right| > k(2^l)p^{1/2}(2^l)\epsilon)\\
\leq&\frac{C\epsilon(\max_{1\leq j\leq k(2^{l+1})}\left|\sum_{i=1}^{k} S_{p,i}\right|^2)}{k^2(2^l)p(2^l)\epsilon} \leq \frac{Ck(2^{l+1})(\log(2\cdot k(2^{l+1})))^2 p(2^l)}{k^2(2^l)p(2^l)\epsilon}\\
\leq&\frac{C(\log(2\cdot k(2^{l+1})))^2}{k(2^l)\epsilon}.
\end{aligned}$$

The latter implies (26) and hence (3) is proved. It remains to prove (4). First we will prove the existence of the constant $C > 0$ such that

$$E\left|\sum_{i=1}^{m}(S_{p,i}^2 - ES_p^2)\right|^2 \leq Cmp^2 \tag{27}$$

for $m \geq 1$. Stationarity and Lemmas 3.6, 3.7 and 3.8 imply

$$\begin{aligned}
&E\left|\sum_{i=1}^{m}(S_{p,i}^2 - ES_p^2)\right|^2 \leq 4E\left|\sum_{\substack{i=1\\ i \text{ odd}}}^{m}(S_{p,i}^2 - ES_p^2)\right|^2\\
=&4mE\left(S_p^2 - ES_p^2\right)^2 + 8\sum_{i=3}^{m-1}(m-i+1)E(S_p^2 - ES_p^2)\cdot(S_{p,i}^2 - ES_p^2)\\
\leq&Cmp^2 + Cm\sum_{i=3}^{m-1}\sum_{\substack{i_1,i_2\in B_1\\ j_1,j_2\in B_2}}|EX_{i_1}X_{i_2}X_{j_1}X_{j_2} - EX_{i_1}X_{i_2}EX_{j_1}X_{j_2}|\\
\leq&Cmp^2 + Cm\sum_{i=2}^{m-1} p^4((a_{[\frac{(i-2)p}{3}]})^{\frac{\delta}{3+\delta}} + (\beta([\frac{(i-2)p}{3}]))^{\frac{\delta}{4+\delta}})\\
\leq&Cmp^2 + Cmp^2\sum_{k=1}^{\infty} k^2((a_{[\frac{k}{3}]})^{\frac{\delta}{3+\delta}} + (\beta([\frac{k}{3}]))^{\frac{\delta}{4+\delta}}) \leq Cmp^2.
\end{aligned}$$

Now again using Theorem A of Serfling [22], we obtain

$$E\left[\max_{1\le a\le m}\left|\sum_{i=1}^{a}(S_{p,i}^2-ES_{p,i}^2)\right|^2\right]\le Cm(\log_2 2m)^2p^2. \tag{28}$$

If we can prove

$$\sum_{l=1}^{\infty}P(\max_{2^l<n\le 2^{l+1}}\left|\sum_{i=1}^{k}(S_{p,i}^2-ES_p^2)\right|>\epsilon kp)<\infty, \tag{29}$$

(4) follows by the Borel-Cantelli lemma. Using Chebyshev inequality and (28) we get

$$\begin{aligned}
&P(\max_{2^l<n\le 2^{l+1}}\left|\sum_{i=1}^{k}(S_{p,i}^2-ES_p^2)\right|>\epsilon kp)\\
\le &P(\max_{1\le m\le k(2^{l+1})}\left|\sum_{i=1}^{m}(S_{p,i}^2-ES_p^2)\right|>\epsilon k(2^l)p(2^l))\\
\le &\frac{Ck(2^{l+1})(\log_2(2k(2^{l+1})))^2\cdot p^2(2^l)}{\epsilon k^2(2^l)p^2(2^l)}\le\frac{C(\log_2(2k(2^{l+1})))^2}{\epsilon k(2^l)}.
\end{aligned}$$

The latter implies (29) and hence (4), so Theorem 2.6 is proved. □

Proof of Theorem 2.7. This theorem can be proved in the same way as Theorem 2.6. Therefore the proof is omitted. □

4.4. **Bootstrap for U-Statistics.**

Lemma 4.2. *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of r.v.'s and $A\subset\{1,\dots,n\}^4$. Then there is a constant C, such that:*

$$\begin{aligned}
&\left|EE^\star\left[\sum_{(i_1,i_2,i_3,i_4)\in A}h_2\left(X_{i_1}^\star,X_{i_2}^\star\right)h_2\left(X_{i_3}^\star,X_{i_4}^\star\right)\right]\right|\\
&\qquad\le C\sum_{i_1,i_2,i_3,i_4=1}^{n}\left|E\left[h_2\left(X_{i_1},X_{i_2}\right)h_2\left(X_{i_3},X_{i_4}\right)\right]\right|.
\end{aligned}$$

Proof of Lemma 4.2. By triangle inequality:

$$\begin{aligned}
&\left|EE^\star\left[\sum_{(i_1,i_2,i_3,i_4)\in A}h_2\left(X_{i_1}^\star,X_{i_2}^\star\right)h_2\left(X_{i_3}^\star,X_{i_4}^\star\right)\right]\right|\\
&\qquad\le\frac{1}{(pk)^2(pk-1)^2}\sum_{(i_1,i_2,i_3,i_4)\in A}\left|EE^\star\left[h\left(X_{i_1},X_{i_2}\right)h\left(X_{i_3},X_{i_4}\right)\right]\right|.
\end{aligned}$$

The bootstrapped expectation of $h_2\left(X_{i_1}^\star,X_{i_2}^\star\right)h_2\left(X_{i_3}^\star,X_{i_4}^\star\right)$ (conditionally on $(X_n)_{n\in\mathbb{N}}$) depends on the way the indices i_1,i_2,i_3,i_4 are allocated to the different blocks. First consider indices i_1,i_2,i_3,i_4 lying in four different blocks $B_{j_1},B_{j_2},B_{j_3},B_{j_4}$ (therefore, $X_{i_1}^\star,\dots,X_{i_4}^\star$ are independent for fixed $(X_n)_{n\in\mathbb{N}}$). From the construction of the bootstrap sample for any four different blocks $B_{j_1}, B_{j_2}, B_{j_3}, B_{j_4}$:

$$\begin{aligned}&\left|E\left[E^\star\left[h_2\left(X_{i_1}^\star, X_{i_2}^\star\right) h_2\left(X_{i_3}^\star, X_{i_4}^\star\right)\right]\right]\right| \\ &= \Big|E\Big[\frac{1}{k^4} \sum_{\substack{1\le i_1+k_1p\le n\\ 1\le i_2+k_2p\le n\\ 1\le i_3+k_3p\le n\\ 1\le i_4+k_4p\le n}} h_2\left(X_{i_1+k_1p}, X_{i_2+k_2p}\right) h_2\left(X_{i_3+k_3p}, X_{i_4+k_4p}\right)\Big]\Big| \\ &\le \frac{1}{k^4} \sum_{\substack{1\le i_1+k_1p\le n\\ 1\le i_2+k_2p\le n\\ 1\le i_3+k_3p\le n\\ 1\le i_4+k_4p\le n}} \left|E\left[h_2\left(X_{i_1+k_1p}, X_{i_2+k_2p}\right) h_2\left(X_{i_3+k_3p}, X_{i_4+k_4p}\right)\right]\right| \\ \Rightarrow & \sum_{\substack{(i_1,i_2,i_3,i_4)\\ \in(B_{j_1}\times B_{j_2}\times B_{j_3}\times B_{j_4})\cap A}} \left|EE^\star\left[h\left(X_{i_1}, X_{i_2}\right) h\left(X_{i_3}, X_{i_4}\right)\right]\right| \\ &\le \frac{1}{k^4} \sum_{i_1,i_2,i_3,i_4=1}^{n} \left|E\left[h\left(X_{i_1}, X_{i_2}\right) h\left(X_{i_3}, X_{i_4}\right)\right]\right|\end{aligned}$$

As there are less than k^4 possibilities to choose these four blocks, one gets:

$$\sum_{\substack{(i_1,i_2,i_3,i_4)\in A\\ \text{4 diff. blocks}}} \left|EE^\star\left[h\left(X_{i_1}, X_{i_2}\right) h\left(X_{i_3}, X_{i_4}\right)\right]\right| \le \sum_{i_1,i_2,i_3,i_4=1}^{n} \left|E\left[h\left(X_{i_1}, X_{i_2}\right) h\left(X_{i_3}, X_{i_4}\right)\right]\right|.$$

As an example, let i_1 and i_2 now lie in the same block with $i_2 - i_i = d < 0$, while i_3, i_4 lie in two further blocks. $X_{i_1}^\star$ and $X_{i_2}^\star$ are dependent, the value of $X_{i_2}^\star$ is determinded by the value of $X_{i_1}^\star$ (conditionally on $(X_n)_{n\in\mathbb{N}}$). To repair this, add up the expected values for all i_2 such that i_2 in the same block as i_1 and take into account that there are at most k^3 possibilities for i_1, i_3, i_4:

$$\begin{aligned}&\left|E\left[E^\star\left[h_2\left(X_{i_1}^\star, X_{i_2}^\star\right) h_2\left(X_{i_3}^\star, X_{i_4}^\star\right)\right]\right]\right| \\ &\le \frac{1}{k^3} \sum_{\substack{1\le i_1+k_1p\le n-d\\ 1\le i_3+k_3p\le n\\ 1\le i_4+k_4p\le n}} \left|E\left[h_2\left(X_{i_1}, X_{i_1+d}\right) h_2\left(X_{i_3}, X_{i_4}\right)\right]\right| \\ \Rightarrow & \sum_{\substack{i_2\\ (i_1,i_2,i_3,i_4)\in A\\ i_2 \text{ in same block as } i_1}} \left|E\left[E^\star\left[h_2\left(X_{i_1}^\star, X_{i_2}^\star\right) h_2\left(X_{i_3}^\star, X_{i_4}^\star\right)\right]\right]\right| \\ &\le \frac{1}{k^3} \sum_{i_1,i_2,i_3,i_4=1}^{n} \left|E\left[h_2\left(X_{i_1}, X_{i_2}\right) h_2\left(X_{i_3}, X_{i_4}\right)\right]\right| \\ \Rightarrow & \sum_{\substack{(i_1,i_2,i_3,i_4)\in A\\ i_2 \text{ in same block as } i_1}} \left|E\left[E^\star\left[h_2\left(X_{i_1}^\star, X_{i_2}^\star\right) h_2\left(X_{i_3}^\star, X_{i_4}^\star\right)\right]\right]\right| \\ &\le \sum_{i_1,i_2,i_3,i_4=1}^{n} \left|E\left[h_2\left(X_{i_1}, X_{i_2}\right) h_2\left(X_{i_3}, X_{i_4}\right)\right]\right|\end{aligned}$$

When the indices are allocated to the blocks in another way, analogous arguments can be used, which completes the proof. □

Proof of Theorem 2.8. We first show that

$$P\left[\sqrt{pk}U_n^\star(h_2)\to 0\right]=E\left[P^\star\left[\sqrt{pk}U_n^\star(h_2)\to 0\right]\right]=1.$$

With Fubinis theorem, we will then conclude that

$$P^\star\left[\sqrt{pk}U_n^\star(h_2)\to 0\right]=1\text{ a.s.}\tag{30}$$

We set

$$Q_n^\star=\sum_{1\leq i_1<i_2\leq pk}h_2\left(X_{i_1}^\star,X_{i_2}^\star\right)\quad\text{and}\quad b_n=\frac{1}{\sqrt{pk}\,(pk-1)}.$$

With the method of subsequences, it suffices to show that

$$b_{2^l}Q_{2^l}^\star(h_2)\to 0\text{ a.s.,}\tag{31}$$

$$\max_{2^{l-1}\leq n<2^l}\left|b_nQ_n^\star-b_{2^{l-1}}Q_{2^{l-1}}^\star\right|\to 0\text{ a.s.,}\tag{32}$$

as $l\to\infty$. By the condition 1. or 2. or 3. of the theorem and Lemma 3.12, there exists a $\eta>0$ such that

$$\sum_{i_1,i_2,i_3,i_4=1}^{n}\left|E\left[h_2\left(X_{i_1},X_{i_2}\right)h_2\left(X_{i_3},X_{i_4}\right)\right]\right|=O\left(n^{3-\eta}\right).\tag{33}$$

We use Chebyshev inequalitity and Lemma 4.2 to prove (31). For every $\epsilon>0$:

$$\sum_{l=1}^{\infty}P\left[\left|b_{2^l}Q_{2^l}^\star(h_2)\right|>\epsilon\right]\leq\frac{1}{\epsilon^2}\sum_{l=1}^{\infty}b_{2^l}^2EE^\star\left[Q_{2^l}^{\star 2}(h_2)\right]$$

$$\leq C\frac{1}{\epsilon^2}\sum_{l=1}^{\infty}b_{2^l}^2\sum_{i_1,i_2,i_3,i_4=1}^{2^l}\left|E\left[h_2\left(X_{i_1},X_{i_2}\right)h_2\left(X_{i_3},X_{i_4}\right)\right]\right|\leq C\frac{1}{\epsilon^2}\sum_{l=1}^{\infty}2^{-\eta l}<\infty.$$

(31) follows with the Borel-Cantelli Lemma. To prove (32), we first have to find a bound for the second moments, using a well known chaining technique:

$$\max_{2^{l-1}\leq n<2^l}\left|b_nQ_n^\star-b_{2^{l-1}}Q_{2^{l-1}}^\star\right|$$

$$\leq\sum_{d=1}^{l}\max_{i=1,\ldots,2^{l-d}}\left|b_{2^{l-1}+i2^{d-1}}Q_{2^{l-1}+i2^{d-1}}^\star-b_{2^{l-1}+(i-1)2^{d-1}}Q_{2^{l-1}+(i-1)2^{d-1}}^\star\right|.$$

As for any random variables $Y_1,\ldots,Y_n$: $E\left(\max|Y_i|\right)^2\leq\sum EY_i^2$, it follows that

$$EE^\star\left[\left(\max_{2^{l-1}\leq n<2^l}\left|b_nQ_n^\star-b_{2^{l-1}}Q_{2^{l-1}}^\star\right|\right)^2\right]$$

$$\leq l\sum_{d=1}^{l}\sum_{i=1}^{2^{l-d}}EE^\star\left[\left(b_{2^{l-1}+i2^{d-1}}Q_{2^{l-1}+i2^{d-1}}^\star-b_{2^{l-1}+(i-1)2^{d-1}}Q_{2^{l-1}+(i-1)2^{d-1}}^\star\right)^2\right]$$

$$\leq l \sum_{d=1}^{l} \sum_{i=1}^{2^{l-d}} 2b^2_{2^{l-1}+i2^{d-1}} EE^\star \left[\left(Q^\star_{2^{l-1}+i2^{d-1}} - Q^\star_{2^{l-1}+(i-1)2^{d-1}} \right)^2 \right]$$

$$+ l \sum_{d=1}^{l} \sum_{i=1}^{2^{l-d}} 2 \left(b_{2^{l-1}+i2^{d-1}} - b_{2^{l-1}+(i-1)2^{d-1}} \right)^2 EE^\star \left[Q^{\star 2}_{2^{l-1}+(i-1)2^{d-1}} \right]$$

$$= \sum_{d=1}^{l} 2b^2_{2^{l-1}+i2^{d-1}} EE^\star \left[\sum_{i=1}^{2^{l-d}} \left(Q^\star_{2^{l-1}+i2^{d-1}} - Q^\star_{2^{l-1}+(i-1)2^{d-1}} \right)^2 \right]$$

$$+ l \sum_{d=1}^{l} \sum_{i=1}^{2^{l-d}} 2 \left(b_{2^{l-1}+i2^{d-1}} + b_{2^{l-1}+(i-1)2^{d-1}} \right) \left(b_{2^{l-1}+i2^{d-1}} - b_{2^{l-1}+(i-1)2^{d-1}} \right)$$

$$\cdot EE^\star \left[Q^{\star 2}_{2^{l-1}+(i-1)2^{d-1}} \right]$$

$$\leq l^2 6 b^2_{2^{l-1}} \sum_{i_1,i_2,i_3,i_4=1}^{2^l} \left| E \left[h_2 \left(X_{i_1}, X_{i_2} \right) h_2 \left(X_{i_3}, X_{i_4} \right) \right] \right| \leq C l^2 2^{-\eta l}.$$

In the last line we used the fact that the sequence $(b_n)_{n\in\mathbb{N}}$ is decreasing, Lemma 4.2 and (33). It now follows for all $\epsilon > 0$

$$\sum_{l=1}^{\infty} P \left[\max_{2^{l-1} \leq n < 2^l} \left| a_n Q_n - a_{2^{l-1}} Q_{2^{l-1}} \right| > \epsilon \right] \leq \frac{C}{\epsilon^2} \sum_{l=1}^{\infty} l^2 2^{-\eta l} < \infty,$$

the Borel-Cantelli Lemma completes the proof of (32). Furthermore, we have that

$$\left(E \left[Q^\star_n \right] \right)^2 \leq E \left[Q^{\star 2}_n \right]$$

and conclude that $\frac{1}{\sqrt{pk}(pk-1)} E \left[Q^\star_n \right] \to 0$ a.s. We use now the Hoeffding-decomposition

$$\sqrt{pk} \left(U^\star_n(h) - E^\star \left[U^\star_n(h) \right] \right) = \frac{2}{\sqrt{pk}} \sum_{i=1}^{pk} \left(h_1 \left(X^\star_i \right) - E^\star \left[h_1 \left(X^\star_i \right) \right] \right)$$

$$+ \frac{2}{\sqrt{pk}\,(pk-1)} \left(\sum_{1 \leq i < j \leq pk} h_2 \left(X^\star_i, X^\star_j \right) - E^\star \left[\sum_{1 \leq i < j \leq pk} h_2 \left(X^\star_i, X^\star_j \right) \right] \right).$$

By Theorem 2.4 (for absolutely regular or strongly mixing sequences) or 2.6 (for functionals of absolutely regular sequences), we have that

$$\sup_{x\in\mathbb{R}} \left| P^\star \left[\frac{2}{\sqrt{pk}} \sum_{i=1}^{pk} \left(h_1 \left(X^\star_i \right) - E^\star h_1 \left(X^\star_i \right) \right) \leq x \right] - P \left[\frac{2}{\sqrt{n}} \sum_{i=1}^{n} h_1 \left(X_i \right) \leq x \right] \right|$$

$$\to 0 \text{ a.s.}$$

and by Theorem 2.1 of Dehling, Wendler [8]

$$\sqrt{n} U_n \left(h_2 \right) \to 0 \text{ a.s.}$$

Since $\sqrt{pk} U^\star_n \left(h_2 \right) \to 0$, $\sqrt{pk} E^\star U^\star_n \left(h_2 \right) \to 0$ a.s. have been already proved, (17) follows with the Lemma of Slutzky. To prove (16), first recall that by Theorem 2.4 or 2.6

$$\mathrm{Var}^\star \left[\sqrt{pk} \sum_{i=1}^{kp} h_1 \left(X^\star_i \right) \right] - \mathrm{Var} \left[\sqrt{n} \sum_{i=1}^{n} h_1 \left(X_i \right) \right] \to 0 \text{ a.s.},$$

and by Lemma 3.12 $\mathrm{Var}\sqrt{n}U_n\left(h_2\right) \to 0$. Similar to the proof of (30), one can show that $\mathrm{Var}^\star\sqrt{pk}U_n^\star\left(h_2\right) \to 0$ a.s., so (16) follows, which completes the proof. □

Acknowledgment

We are very grateful to Herold Dehling, who encouraged us to study this topic and discussed it with us many times.

References

[1] Arcones, M.A. and Giné, E. (1992). On the bootstrap for U and V statistics. *Ann. Statist.* **20** 655-674.
[2] Bickel, P.J. and Freedman, D.A. (1981). Some asymptotic theory for the bootstrap. *Ann. Statist.* **9** 1196-1217.
[3] Borovkova, S., Burton, R. and Dehling, H. (2001). Limit theorems for functionals of mixing processes with applications to U-statistics and dimension estimation, *Trans. Amer. Math. Soc.* **353** 4261–4318.
[4] Bradley, R.C. (2007). *Introduction to strong mixing conditions.* Vol. 1-3, Kendrick Press, Heber City, Utah.
[5] Carlstein, E. (1986). The use of subseries values for estimating the variance of a general statistic from stationary sequence. *Ann. Statist.* **14** 1171-1179.
[6] Dehling, H. and Mikosch, T. (1994). Random quadratic forms and the bootstrap for U-statistics. *J. Multivariate Anal.* **51** 392-413.
[7] Dehling, H. and Wendler, M. (2010). Central limit theorem and the bootstrap for U-statistics of strongly mixing data. *J. Multivariate Ana.* **101** 126-137.
[8] Dehling, H. and Wendler, M. (2009). Law of the iterated logarithm for U-statistics of weakly dependent observations, *prepint* arXiv:0911.1200.
[9] Denker, M. and Keller, G. (1986). Rigorous statistical procedures for data from dynamical systems, *J. Stat. Phys.* **44** 67-93.
[10] Doukhan, P. (1994). *Mixing.* Springer, New York.
[11] Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Ann. Statist.* **7** 1-26.
[12] Hall, P. (1992). *The bootstrap and Edgeworth expansions.* Springer, New York.
[13] Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Stat.* **19** (1948) 293-325.
[14] Hofbauer, F. and Keller, G. (1982). Ergodic properties of invariant measures for piecewise monotonic transformations. *Math. Z.* **180** 119-142.
[15] Ibragimov, I.A. (1962). Some limit theorems for stationary processes. *Theory Prob. Appl.* **7** 349-382.
[16] Ibragimov, I.A. and Linnik, Y.V. (1971). *Independent and stationary sequences of random variables.* Wolters-Noordhoff, Groningen.
[17] Lahiri, S.N. (2003). *Resampling methods for dependent data.* Springer, New York.
[18] Leucht, A. and Neumann, M.H. (2009). Consistency of general bootstrap methods for degenerate U-type and V-type statistics. *J. Multivariate Anal.* **100** 1622-1633.
[19] Peligrad, M. (1998). On the blockwise bootstrap for empirical processes for stationary sequences. *Ann. Probab.* **2** 877-901.
[20] Radulovic, D. (1996). The bootstrap of the mean for strong mixing sequences under minimal conditions. *Statist. Probab. Lett.* **28** 65-72.
[21] Rio, E. (1995). A maximal inequality and dependent Marcinkiewicz-Zygmund strong laws. *Ann. Probab.* **23** 918-937.
[22] Serfling, R. J. (1970). Moment inequalities for the maximum cumulative sum, *Ann. Math. Statist.* **41** 1227-1234.
[23] Shao, Q.M. (1993). Complete convergence for α-mixing sequences. *Statist. Probab. letters* **16** 279-287.
[24] Shao, Q.M. and Yu, H. (1993). Bootstrapping the sample means for stationary mixing sequences. *Stochastic Process. Appl.* **48** 175-190.
[25] Singh, K. (1981). On the asymptotic accuracy of Efron's bootstrap, *Ann. Statist.* **9** 1187-1195.
[26] Yokoyama, R. (1980). Moment bounds for stationary mixing sequences. *Z. Wahrsch. verw. Gebiete* **52** 45-57.

[27] Yoshihara, K. (1976). Limiting behavior of U-statistics for stationary, absolutely regular processes. *Z. Wahrsch. verw. Gebiete* **35** 237-252.

Institute of Mathematics and Information Technologies, Uzbek Academy of Sciences, 29 Dormon Yoli Str., Tashkent, 100125, Uzbekistan

Fakultät für Mathematik, Ruhr-Universität Bochum, 44780 Bochum, Germany, Email address: Martin.Wendler@rub.de