SPECIAL ISSUE

# DEPENDENT WILD BOOTSTRAP FOR THE EMPIRICAL PROCESS

PAUL DOUKHAN[a] GABRIEL LANG[b] ANNE LEUCHT[c] AND MICHAEL H. NEUMANN[d*]

[a] *UMR 8088 Analyse, Géométrie et Modélisation 2, Université Cergy-Pontoise and French Universitary Institute (IUF), avenue Adolphe Chauvin, 95302 Cergy-Pontoise Cedex, France*
[b] *UMR 518 Mathématique et Informatiques Appliquées, AgroParisTech and INRA, 16 rue C. Bernard, 75231 Paris Cedex 05, France*
[c] *Institut für Mathematische Stochastik, Technische Universität Braunschweig, Pockelsstr. 14, 38106 Braunschweig, Germany*
[d] *Institut für Mathematik, Friedrich-Schiller-Universität Jena, Ernst-Abbe-Platz 2, 07743 Jena, Germany*

In this paper, we propose a model-free bootstrap method for the empirical process under absolute regularity. More precisely, consistency of an adapted version of the so-called dependent wild bootstrap, which was introduced by Shao (2010) and is very easy to implement, is proved under minimal conditions on the tuning parameter of the procedure. We show how our results can be applied to construct confidence intervals for unknown parameters and to approximate critical values for statistical tests. In a simulation study, we investigate the size properties of a bootstrap-aided Kolmogorov-Smirnov test and show that our method is competitive to standard block bootstrap methods in finite samples.

## 1. INTRODUCTION

Given real-valued observations $X_1, \ldots, X_n$ with a common cumulative distribution function (cdf) $F$, many important statistics $T_n$ can be rewritten as or approximated by functionals of the empirical process $G_n = (G_n(x))_{x \in \mathbb{R}}$, where $G_n(x) = \sqrt{n}(F_n(x) - F(x))$ and $F_n(x) = n^{-1} \sum_{t=1}^{n} \mathbb{1}(X_t \le x)$. A typical example is given by the Kolmogorov-Smirnov test statistic. When knowledge of the distribution of $T_n$ is required, for example, for the construction of confidence sets or the determination of critical values for tests, knowledge of the distributional properties of $G_n$ would help. In the case of independent and identically distributed (i.i.d.) random variables and a continuous cdf $F$, it is well known that the distribution of $(G_n(F^{-1}(u)))_{u \in (0,1)}$ does not depend on the particular $F$. As a consequence, the distribution of the Kolmogorov-Smirnov test statistic $T_n = \sup_{x \in \mathbb{R}} |G_n(x)|$ is invariant under $F$, which makes the choice of critical values quite easy. In the case of dependent random variables, however, this situation changes dramatically. It is well known (see also Theorem 2.1 in the succeeding text) that the distribution of $G_n$ and also its weak limit as $n$ tends to infinity depend on the particular dependence properties of the underlying process. Because these properties are usually not known in advance, it is important to have a method of estimating the distribution of $G_n$ at hand. It is known that, under certain conditions, blockwise bootstrap methods provide a consistent approximation; see, for example, Bühlmann (1994,1995) and Naik-Nimbalkar and Rajarshi (1994). In this paper, we derive results of this type for an alternative bootstrap method, the so-called dependent wild bootstrap. This approach was first proposed by Shao (2010) for functionals of the sample mean and is very easy to implement. This property is preserved in the case of missing data, where, in contrast, the algorithms for ordinary block-bootstrap methods have to be adjusted properly.

---

* Correspondence to: Michael H. Neumann, Friedrich-Schiller-Universität Jena, Institut für Mathematik, Ernst-Abbe-Platz 2, 07743 Jena, Germany. E-mail: michael.neumann@uni-jena.de

Dependent wild bootstrap methods have already been successfully applied in the field of hypothesis testing; see Shao (2011), Leucht and Neumann (2013) and Smeekes and Urbain (2014). Here, we show that an obvious adaptation of this approach to the empirical process is consistent under rather weak conditions on the original process $(X_t)_{t \in \mathbb{N}}$ and on a wide range for the tuning parameter of the bootstrap process. The tuning parameter of the dependent wild bootstrap plays a similar role as the block length for classical block-based methods. In the present case, the blocky structure refers to the covariances of the bootstrap variables rather than the data itself, which ensures that the dependence structure between two consecutive observations is captured by this resampling method.

In Section 4, we present applications of our general consistency results to statistics of different types, including the Kolmogorov-Smirnov statistic as well as degenerate and non-degenerate von Mises statistics. A small simulation study reported in Section 5 sheds some light on the finite sample behaviour of the bootstrap approximation, and it seems that the performance of the dependent wild bootstrap is comparable to that of the classical moving block bootstrap (MBB) introduced by Künsch (1989) and Liu and Singh (1992) and the tapered block bootstrap (TBB) of Paparoditis and Politis (2001).

## 2. ASSUMPTIONS, THE EMPIRICAL PROCESS

Suppose that we observe $X_1, \ldots, X_n$ from a (strictly) stationary and real-valued process $(X_t)_{t \in \mathbb{Z}}$. We denote by $F$ the common cdf of the $X_t$s and by $F_n$ the empirical distribution function, that is, $F_n(x) = n^{-1} \sum_{t=1}^{n} \mathbb{1}(X_t \leq x)$. For simplicity, we assume that $F$ is continuous, although we think that our results can be generalized to discontinuous cdf's. The empirical process $G_n = (G_n(x))_{x \in \mathbb{R}}$ is given by

$$G_n(x) = \sqrt{n}(F_n(x) - F(x)).$$

We assume

(A1) $(X_t)_{t \in \mathbb{Z}}$ is strictly stationary and absolutely regular ($\beta$-mixing) with mixing coefficients satisfying $\sum_{r=1}^{\infty} \beta_X(r) < \infty$. The cdf $F$ of $X_0$ is continuous.

The following result is a special case of Theorem 1 in Rio (1998).

**Theorem 2.1.** Suppose that (A1) is fulfilled. Then,

$$G_n \xrightarrow{d} G,$$

where $G = (G(x))_{x \in \mathbb{R}}$ is a Gaussian process with continuous sample paths, $EG(x) = 0$, and $\text{cov}(G(x), G(y)) = \sum_{r=-\infty}^{\infty} \text{cov}(\mathbb{1}(X_0 \leq x), \mathbb{1}(X_r \leq y))$. Here, convergence holds with respect to the supremum metric, that is, $\sup_{f \in \mathcal{F}_L} |Ef(G_n) - Ef(G)| \longrightarrow 0$ holds with $\mathcal{F}_L = \{f : \mathcal{F} := \{h : \mathbb{R} \to \mathbb{R} \text{ is càdlàg}\} \to \mathbb{R} \mid f \text{ bounded}, |f(h_1) - f(h_0)| \leq \|h_1 - h_0\|_\infty\}$.

**Remark 1.**
 (i) The aforementioned characterization of weak convergence can be found in van der Vaart and Wellner (2000, Section 1.12).
 (ii) Doukhan, Massart and Rio (1995, Section 1) discussed several notions of mixing and concluded that absolute regularity ($\beta$-mixing) is an appropriate condition in the context of the study of empirical processes due to Berbee's maximal coupling. Later, Rio (2000, Theorem 7.2) derived a uniform CLT for stationary and strong mixing ($\alpha$-mixing) processes under the condition $\alpha(r) = O(r^{-\kappa})$, for some $\kappa > 1$. We think that our results in the succeeding text may also be proved under alternative dependence conditions, such as strong mixing or weak dependence conditions from Doukhan and Louhichi (1999). For sake of definiteness, we restrict ourselves to the notion of absolute regularity here.

## 3. DEPENDENT WILD BOOTSTRAP FOR THE EMPIRICAL PROCESS

### 3.1. Methodology and basic assumptions

The so-called dependent wild bootstrap was introduced by Shao (2010) for smooth functions of the sample mean. In the case of weakly dependent and real-valued random variables $X_1, \ldots, X_n$, the idea of the dependent wild bootstrap is to construct the pseudo-observations as follows:

$$X_t^* = \bar{X}_n + (X_t - \bar{X}_n)\varepsilon_{t,n}^*, \quad t = 1, \ldots, n.$$

Here, $\bar{X}_n = n^{-1} \sum_{t=1}^n X_t$ and $\left(\varepsilon_{t,n}^*\right)_{t=1}^n$ is a triangular scheme of weakly dependent random variables that is independent of $X_1, \ldots, X_n$. Shao (2010) verified that under certain regularity conditions,

$$\sup_{x \in \mathbb{R}} \left| P\left(\sqrt{n}\left[H(\bar{X}_n) - H(EX_1)\right] \leq x\right) - P^*\left(\sqrt{n}\left[H\left(\bar{X}_n^*\right) - H(\bar{X}_n)\right] \leq x\right) \right| \xrightarrow{P} 0,$$

where $H$ is a smooth function and $\bar{X}_n^* = n^{-1} \sum_{t=1}^n X_t^*$.

In our case of the empirical process, the role of the $X_t$s earlier is taken by the processes $(\mathbb{1}(X_t \leq x))_{x \in \mathbb{R}}$. Following the idea of Shao (2010), we define bootstrap counterparts of $Y_t = \mathbb{1}(X_t \leq x)$ and of $F_n$ as

$$
\begin{aligned}
Y_t^* &= \bar{Y}_n + (Y_t - \bar{Y}_n)\varepsilon_{t,n}^* \\
&= F_n(x) + (\mathbb{1}(X_t \leq x) - F_n(x))\varepsilon_{t,n}^*
\end{aligned}
$$

and

$$F_n^*(x) = F_n(x) + \frac{1}{n}\sum_{t=1}^n (\mathbb{1}(X_t \leq x) - F_n(x))\varepsilon_{t,n}^* \tag{3.1}$$

respectively. This leads to the following bootstrap version of the empirical process:

$$
\begin{aligned}
G_n^*(x) &= \frac{1}{\sqrt{n}}\sum_{t=1}^n \left(Y_t^* - \bar{Y}_n\right) \\
&= \frac{1}{\sqrt{n}}\sum_{t=1}^n (\mathbb{1}(X_t \leq x) - F_n(x))\varepsilon_{t,n}^*.
\end{aligned}
$$

Note that the result of Shao (2010) remains valid if the $X_t$s are $\mathbb{R}^d$-valued random vectors. Therefore, it is clear that, under appropriate conditions, the distribution of $(G_n^*(x_1), \ldots, G_n^*(x_d))'$ consistently approximates that of $(G_n(x_1), \ldots, G_n(x_d))'$ for any $x_1, \ldots, x_d \in \mathbb{R}$, $d \in \mathbb{N}$. In fact, we show under a condition slightly stronger than (A1) and under simple conditions for the $\varepsilon_{t,n}^*$s that this is indeed the case. Moreover, we prove stochastic equicontinuity on the bootstrap side, which yields convergence of $(G_n^*)_{n \in \mathbb{N}}$ to the desired limit. To this end, we impose the following condition:

(A2) For all $n$, $\left(\varepsilon_{t,n}^*\right)_{t=1,\ldots,n}$ is a centered stationary Gaussian process with

$$\sum_{r=1}^n \left| cov\left(\varepsilon_{1,n}^*, \varepsilon_{r,n}^*\right) \right| = O(l_n) \text{ and } A_n(s,t) := \text{cov}\left(\varepsilon_{s,n}^*, \varepsilon_{t,n}^*\right) \to_{n \to \infty} 1.$$

The sequence $(l_n)_{n \in \mathbb{N}}$ has to be chosen such that

$$l_n \underset{n \to \infty}{\longrightarrow} \infty \qquad \text{and} \qquad l_n/n \underset{n \to \infty}{\longrightarrow} 0. \tag{3.2}$$

Similarly to Shao's work, the assumption of Gaussianity of the wild bootstrap variables $(\varepsilon_{t,n}^*)$ can be dropped. However, if one wants to go beyond $l_n$-dependence of these variables, this typically goes along with technical assumptions on moments and dependence structure of these bootstrap variables; see, for example, assumption (B2) in Leucht and Neumann (2013). Since the process $(G_n^*(x))_{x \in \mathbb{R}}$ is intended to mimic the stochastic behaviour of $(G_n(x))_{x \in \mathbb{R}}$, which is asymptotically Gaussian, and to simplify the mathematical part in the succeeding text, we stick to the assumption of normality here.

The role of the parameter $l_n$ is similar to that of the block length in blockwise bootstrap methods. For a long time, these blockwise methods have been known to be consistent if the block length tends to infinity within a certain 'corridor', that is, $l_n \to \infty$ but $l_n = o(n^\delta)$, for $\delta \in (0, 1/2)$; see, for example, Bühlmann (1994,1995) and Naik-Nimbalkar and Rajarshi (1994). However, a recent result of Wieczorek (2014) shows that the weaker conditions of $l_n \to \infty$ and $l_n = o(n)$ are still sufficient for consistency. In our context, it is clear that the aforementioned assumptions on $(l_n)_{n \in \mathbb{N}}$ are some sort of minimal condition for the dependent wild bootstrap to work: the condition $l_n \to \infty$ takes care that the dependence structure of the original process $X_1, \ldots, X_n$ is asymptotically captured. On the other hand, $l_n/n \to 0$ implies that the conditional distribution of $G_n^*$ is non-degenerate. The question of the optimal choice of the tuning parameter $l_n$ is addressed in Shao (2010) for the mean. It turns out that the optimal rates $l_n = O(n^{1/3})$ and $l_n = O(n^{1/5})$ known for the block lengths of the MBB and the TBB respectively are also optimal for the tuning parameter of the dependent wild bootstrap for suitable choices of the covariance function $A_n$ of the bootstrap variables; compare the comments in the succeeding text Remark 2.1 and Corollary 4.1 in Shao (2010). These findings can be carried over to our results on the finite dimensional distributions of the empirical process. However, the question of uniform optimal rates for the whole process seems to be much more delicate and is left outside the scope of the paper.

**Remark 2.**

(i) A simple special case of a process satisfying the aforementioned conditions is given by defining $\varepsilon_{t,n}^* = U_{t/l_n}$ where $(U_t)_{t \geq 0}$ is an Ornstein-Uhlenbeck process, that is, a Gaussian process with continuous sample paths, $EU_t = 0$ and $\text{cov}(U_s, U_t) = \exp(-|s-t|) \; \forall s, t \geq 0$. In this case, the practical implementation is rather easy since a discrete sample of an Ornstein-Uhlenbeck process forms an AR(1) process, that is,

$$\varepsilon_{t,n}^* = e^{-1/l_n} \varepsilon_{t-1,n}^* + \sqrt{1 - e^{-2/l_n}} \; \varepsilon_t^*,$$

where $\varepsilon_{0,n}^*, \varepsilon_1^*, \ldots, \varepsilon_n^*$ are independent standard normal variables. Other choices of the covariance structure of $\varepsilon_{1,n}^*, \ldots, \varepsilon_{n,n}^*$ are considered in Section 5, too.

(ii) There are also other variants of the dependent wild bootstrap in the literature. Shao (2011) proposed a block-wise wild bootstrap procedure, where variables from blocks of length $l_n$ are multiplied with one and the same auxiliary random variable. To deal with heteroskedasticity in the context of unit root testing, Smeekes and Urbain (2014) proposed, besides the dependent wild bootstrap and the blockwise wild bootstrap as in Shao (2010, 2011), an autoregressive wild bootstrap. Of course, in view of (i), the latter is a special case of our variant of the dependent wild bootstrap.

Finally, we have to replace assumption (A1) by the following slightly stronger assumption:

(A3) $(X_t)_{t \in \mathbb{Z}}$ is strictly stationary and absolutely regular ($\beta$-mixing) with coefficients satisfying $\sum_{r=1}^\infty r^2 \beta_X(r) < \infty$.

### 3.2. Asymptotics for the bootstrapped empirical process

It turns out that the following representation of the bootstrapped empirical process simplifies the investigation of its asymptotics. We write

$$G_n^*(x) = G_n^{*,0}(x) - R_n^*(x),$$

where

$$G_n^{*,0}(x) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} (\mathbb{1}(X_t \leq x) - F(x))\varepsilon_{t,n}^*$$

and $R_n^*(x) = (F_n(x) - F(x)) \, n^{-1/2} \sum_{t=1}^{n} \varepsilon_{t,n}^*$.

**Remark 3.** Since $\sup_{x \in \mathbb{R}} |F_n(x) - F(x)| = O_P(n^{-1/2})$, we obtain that

$$\sup_{x \in \mathbb{R}} \left| R_n^*(x) \right| = \sup_{x \in \mathbb{R}} |F_n(x) - F(x)| \left| n^{-1/2} \sum_{t=1}^{n} \varepsilon_{t,n}^* \right| = O_{P^*}(\sqrt{l_n/n})$$

under mild assumptions stated in the succeeding text. (We write $Y_n^* = O_{P^*}(r_n)$ if $\forall \epsilon > 0 \, \exists K(\epsilon) < \infty$ such that $P\left(P^*\left(|Y_n^*/r_n| > K(\epsilon)\right) > \epsilon\right) \xrightarrow[n \to \infty]{} 0$.) Hence, we can analyze $G_n^{*,0}$ instead of $G_n^*$ in the sequel.

As usual, we have to prove convergence of the finite-dimensional distributions to the correct limits and stochastic equicontinuity of the processes $\left(G_n^{*,0}\right)_{n \in \mathbb{N}}$. The first task is rather easy since the finite-dimensional distributions are by construction centered Gaussian.

**Lemma 3.1.** Suppose that (A2) and (A3) *are* fulfilled. Then, for arbitrary $x, y \in \mathbb{R}$,

$$\operatorname{cov}^*\left(G_n^{*,0}(x), G_n^{*,0}(y)\right) \xrightarrow{P} \operatorname{cov}(G(x), G(y)).$$

This and Remark 3 yields convergence of the finite-dimensional distributions:

**Corollary 3.1.** Suppose that (A2) and (A3) are fulfilled. Then, for arbitrary $x_1, \ldots, x_k \in \mathbb{R}, k \in \mathbb{N}$,

$$\left(G_n^*(x_1), \ldots, G_n^*(x_k)\right)' \xrightarrow{d} \left(G(x_1), \ldots, G(x_k)\right)' \qquad \text{in probability.}$$

It turns out that the proof of stochastic equicontinuity of $\left(G_n^{*,0}\right)_{n \in \mathbb{N}}$ is more delicate than that of Lemma 3.1. We have to prove that for each $\epsilon > 0$ and $\eta > 0$, there exists a grid $-\infty = x_0 < x_1 < \ldots < x_{M-1} < x_M = \infty$ such that

$$P\left(P^*\left(\max_{1 \leq i \leq M} \sup_{x \in (x_{i-1}, x_i]} \left|G_n^{*,0}(x) - G_n^{*,0}(x_i)\right| > \epsilon\right) \leq \eta\right) \xrightarrow[n \to \infty]{} 1. \qquad (3.3)$$

(As usual, we set $G_n^{*,0}(-\infty) = G_n^{*,0}(\infty) = 0$.) To this end, we prove that

$$
\begin{aligned}
E\left[ P^*\left( \max_{1 \le i \le M} \sup_{x \in (x_{i-1}, x_i]} \left| G_n^{*,0}(x) - G_n^{*,0}(x_i) \right| > \epsilon \right) \right] \\
= E\ E^*\left[ \mathbb{1}\left( \max_{1 \le i \le M} \sup_{x \in (x_{i-1}, x_i]} \left| G_n^{*,0}(x) - G_n^{*,0}(x_i) \right| > \epsilon \right) \right] \le \eta^2
\end{aligned}
\tag{3.4}
$$

for all $n \ge n_0(\epsilon, \eta)$, which implies by Markov's inequality that

$$
P\left( P^*\left( \max_{1 \le i \le M} \sup_{x \in (x_{i-1}, x_i]} \left| G_n^{*,0}(x) - G_n^{*,0}(x_i) \right| > \epsilon \right) > \eta \right) \le \eta \qquad \forall n \ge n_0(\epsilon, \eta)
$$

and, therefore, (3.3).

**Remark 4.**
(i) Although we have to show (3.3), which is a result on the conditional distribution of $G_n^{*,0}$ given $X_1, \ldots, X_n$, we prove first the *unconditional* result (3.4) for the increments of $G_n^{*,0}$. Taking the expectation with respect to (w.r.t.) the original sample allows us to take advantage of the fixed dependence structure of $X_1, \ldots, X_n$, with $\sum_{r=1}^{\infty} r^2 \beta_X(r) < \infty$. In contrast, if the process $(U_t)_{t \ge 0}$ is absolutely regular with coefficients satisfying $\int_0^{\infty} \beta_U(r)\, dr < \infty$ and if $\varepsilon_{t,n}^* = U_{t/l_n}$, then the $\varepsilon_{t,n}^*$ are also absolutely regular, however, with mixing coefficients satisfying only $\sum_{r=1}^{\infty} \beta_{\varepsilon^*}(r) = O(l_n)$. Thus, working with conditional expectations alone would be more difficult and probably go along with additional assumptions on the tuning parameter $l_n$.
(ii) Some preliminary calculations suggest that we could employ Rio's (1998) approach to prove stochastic equicontinuity of the bootstrap process. Suppose that the variables $\varepsilon_{1,n}^*, \ldots, \varepsilon_{n,n}^*$ are obtained from a Gaussian process $(U_t)_{t \ge 0}$ via $\varepsilon_{t,n}^* = U_{t/l_n}$. If the process $(U_t)_{t \ge 0}$ is absolutely regular with coefficients $\beta_U(r), r > 0$, then it follows from independence of the $X_t$s and the $\varepsilon_{t,n}^*$s that the bivariate process $\left( (X_t, \varepsilon_{t,n}^*) \right)_{t=1,\ldots,n}$ is absolutely regular with coefficients $\beta_{X,\varepsilon^*}(r) \le \beta_X(r) + \beta_U(r/l_n)$. Unfortunately, although the $\beta_X(r)$s are summable we only obtain that $\sum_{r=0}^{n-1} \beta_{X,\varepsilon^*}(r) = O(l_n)$, which would require for the proof of stochastic equicontinuity of $(G_n^{*,0})_{n \in \mathbb{N}}$ an additional restriction on the sequence $(l_n)_{n \in \mathbb{N}}$ beyond the obviously necessary conditions $l_n \to \infty$ and $l_n/n \to 0$. In view of this, we have decided to use a different approach tailor-made for our problem at hand.

As a first step, the following lemma provides upper estimates for the fourth moment of increments of $G_n^{*,0}$ over certain intervals $I_{j,k} = (x_{j,k-1}, x_{j,k}]$. To find appropriate grid points $x_{j,k}$, we adapt an idea from Viennet (1997) for strictly stationary and absolutely regular processes $(\xi_t)_{t \in \mathbb{Z}}$ on $(\Omega, \mathcal{A}, P)$ with summable coefficients of absolute regularity. Using the representation $\beta(\sigma(\xi_0), \sigma(\xi_k)) = \frac{1}{2} E \| P^{\xi_k | \xi_0} - P^{\xi_k} \|_{Var}$, where $\| Q \|_{Var}$ denotes the total variation norm of a signed measure $Q$, she shows that there exists a nonnegative function $b \in L_1(P)$ such that $\mathrm{var}\left( n^{-1/2} \sum_{t=1}^{\infty} \psi(\xi_t) \right) \le 4 \int b(x) \psi^2(x)\, dP(x)$ holds for all $\psi \in L_2(P)$. This implies, for any choice of $-\infty < x_0 < x_1 \cdots < x_M < \infty, M \in \mathbb{N}$, that

$$
\sum_{k=1}^{M} \mathrm{var}\left( n^{-1/2} \sum_{t=1}^{n} \mathbb{1}_{(x_{k-1}, x_k]}(\xi_t) \right) \le 4 \int_{-\infty}^{\infty} b(x)\, dP(x) < \infty,
$$

Copyright © 2015 Wiley Publishing Ltd

that is, we obtain an upper bound not depending on the fineness of the decomposition of $\mathbb{R}$. In view of this, it becomes apparent that Viennet's idea is tailor-made for proving a result such as Lemma 3.2 in the succeeding text. Since we estimate fourth moments of the increments, we have to carry over this approach to higher moments; see the proof of the following lemma for details.

**Lemma 3.2.** Suppose that (A2) and (A3) are fulfilled. Then, there exists a dyadic sequence of grid points $-\infty = x_{j,0} < x_{j,1} < \ldots < x_{j,2^j} = \infty$, $j \in \mathbb{N}$, with $x_{j,k} = x_{j+1,2k}$ such that, for all $j \in \mathbb{N}, k \in \{1, \ldots, 2^j\}$,

$$EE^* \left[ \left( G_n^{*,0}(x_{j,k-1}) - G_n^{*,0}(x_{j,k}) \right)^4 \right] \leq K_0 \left( 2^{-2j} + n^{-1} 2^{-j} \right),$$

for some $K_0 < \infty$.

With the grid points chosen in the proof of Lemma 3.2, we can prove similarly to Theorem 15.6 in Billingsley (1968) that we have the desired stochastic equicontinuity (in probability) for the bootstrap processes:

**Corollary 3.2.** Suppose that (A2) and (A3) are fulfilled. Then, for each $\epsilon > 0$ and $\eta > 0$, there exists a grid $-\infty = x_0 < x_1 < \ldots < x_{M-1} < x_M = \infty$ such that

$$P \left( P^* \left( \max_{1 \leq i \leq M} \sup_{x \in (x_{i-1}, x_i]} \left| G_n^{*,0}(x) - G_n^{*,0}(x_i) \right| > \epsilon \right) \leq \eta \right) \underset{n \to \infty}{\longrightarrow} 1.$$

As a consequence of the Corollaries 3.1 and 3.2, we obtain the convergence of the bootstrap processes $G_n^*$ to the same limit as for the original processes $G_n$.

**Theorem 3.1.** Suppose that (A2) and (A3) are fulfilled. Then,

$$G_n^* \overset{d}{\longrightarrow} G \qquad \text{in probability.}$$

Here, the convergence holds with respect to the supremum metric with the additional qualification 'in probability', that is, $\sup_{f \in \mathcal{F}_L} \left| E^* f \left( G_n^* \right) - E f(G) \right| \overset{P}{\longrightarrow} 0$ holds.

## 4. APPLICATIONS

We discuss some specific applications of our results earlier in this section. Theorems 2.1 and 3.1 act as master theorems that imply bootstrap consistency in some particular cases of interest.

### 4.1. Quantile estimation

Quantile estimation plays an important role in financial risk management since several risk measures like the value-at-visk or the expected shortfall can be represented as functions of quantiles.

For $q \in (0, 1)$, the $q$-quantile of $F$ is defined as $t_q = F^{-1}(q) = \inf\{x : F(x) \geq q\}$. This can be conveniently estimated by its empirical counterpart,

$$t_{n,q} = F_n^{-1}(q).$$

We impose the following additional condition:

(A4)  $F$ is continuously differentiable at $t_q$ and $F'(t_q) > 0$.

For $\sqrt{n}(t_{n,q} - t_q)$, Sun and Lahiri (2006) and Sharipov and Wendler (2013) proved consistency of the block bootstrap in the case of strong mixing processes. The next theorem follows immediately as a special case of the Theorems 1 and 2 in Sharipov and Wendler (2013).

**Theorem 4.1** (Sharipov and Wendler (2013))**.**   Suppose that (A1) and (A4) are fulfilled. Then,

(i)   $t_{n,q} - t_q = \frac{q - F_n(t_q)}{F'(t_q)} + o_P(n^{-1/2})$,

(ii)   $\sqrt{n}(t_{n,q} - t_q) \xrightarrow{d} Z_q \sim \mathcal{N}(0, \operatorname{var}(G(t_q))/(F'(t_q))^2)$.

On the bootstrap side, we define

$$t_{n,q}^* = F_n^{*-1}(q),$$

where $F_n^*$ is defined as in (3.1). Note that a non-standard feature in this context is that $F_n^*$ is not monotonously non-decreasing. Therefore, $\sqrt{n}\left(t_{n,q}^* - t_{n,q}\right) \leq x$ is not equivalent to $F_n^*(t_{n,q} + x/\sqrt{n}) \geq q$ but to $\inf\{F_n^*(s) : s \leq t_{n,q} + x/\sqrt{n}\} \geq q$. In view of this, we cannot obtain the asymptotic distribution of $\sqrt{n}\left(t_{n,q}^* - t_{n,q}\right)$ directly from the asymptotics of $P^*\left(F_n^*(t_{n,q} + x/\sqrt{n}) \geq q\right)$. The following theorem states first the validity of a Bahadur representation for $t_{n,q}^*$ which eventually leads to the limit distribution for $\sqrt{n}\left(t_{n,q}^* - t_{n,q}\right)$.

**Theorem 4.2.**   Suppose that (A2), (A3) and (A4) are fulfilled. Then,

(i)   $t_{n,q}^* - t_{n,q} = \frac{F_n(t_q) - F_n^*(t_q)}{F'(t_q)} + o_{P^*}(n^{-1/2})$,
   where we write $R_n^* = o_{P^*}(a_n)$ if $P^*\left(|R_n^*|/|a_n| > \varepsilon\right) \xrightarrow{P} 0, \ \forall \varepsilon > 0$.

(ii)   $\sqrt{n}\left(t_{n,q}^* - t_{n,q}\right) \xrightarrow{d} Z_q \sim \mathcal{N}(0, \operatorname{var}(G(t_q))/(F'(t_q))^2)$   in probability.

**Corollary 4.1.**   Suppose that (A2), (A3) and (A4) are fulfilled. If additionally $\operatorname{var}(G(t_q)) > 0$, then

(i)  $\sup_{x \in \mathbb{R}} \left| P^*\left(t_{n,q}^* - t_{n,q} \leq x\right) - P\left(t_{n,q} - t_q \leq x\right)\right| \xrightarrow{P} 0$

(ii)  With $c_\gamma^* := \inf\left\{c : P^*\left(|t_{n,q}^* - t_{n,q}| \leq c\right) \geq \gamma\right\}, 0 < \gamma < 1$,

$$P\left(t_q \in \left[t_{n,q} - c_\gamma^*, t_{n,q} + c_\gamma^*\right]\right) \xrightarrow[n \to \infty]{} 1 - \gamma.$$

## 4.2.  Kolmogorov-Smirnov test

A classical test problem in mathematical statistics is given by

$$H_0 : \quad F = F_0 \qquad \text{vs.} \qquad H_1 : \quad F \neq F_0.$$

Based on observations $X_1, \ldots, X_n \sim F$, we give a decision rule with nominal size $\gamma \in (0, 1)$ based on the Kolmogorov-Smirnov test statistic,

$$T_n = \sup_{x \in \mathbb{R}} \sqrt{n}|F_n(x) - F_0(x)|.$$

The null hypothesis is rejected if the value of the test statistic is larger than the $(1 - \gamma)$-quantile of the distribution of $T_n$, which in turn depends on the dependence structure of the data. Even in case the latter is not completely specified our bootstrap procedure can be successfully applied to approximate these quantiles. To this end, we define a bootstrap version of the test statistic $T_n^* = \sup_{x \in \mathbb{R}} \sqrt{n} \, |F_n^*(x) - F_n(x)|$ and the corresponding bootstrap quantile

$$t_\gamma^* = \inf \left\{ x : P^* \left( T_n^* > x \right) \geq \gamma \right\}.$$

**Theorem 4.3.** Assume that (A2) and (A3) are fulfilled and that there exists some $x \in \mathbb{R}$ with $\mathrm{var}(G(x)) > 0$. Then, if $F = F_0$,

$$P_0 \left( T_n > t_\gamma^* \right) \xrightarrow[n \to \infty]{} \gamma.$$

**Remark 5.** Our master theorems, Theorems 2.1 and 3.1, can also be invoked to set up a two-sample test. Based on observations $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ of two independent absolutely regular and strictly stationary processes $(X_n)_n$ and $(Y_n)_n$, one aims to decide whether the marginal distribution of these processes $P^X$ and $P^Y$ are identical, that is

$$H_0 : \quad P^X = P^Y \qquad \text{versus} \qquad H_1 : \quad P^X \neq P^Y.$$

The Kolmogorov-Smirnov type test statistic is then given by

$$\widehat{T}_n = \sup_{x \in \mathbb{R}} \sqrt{n} \left| F_n^{(X)}(x) - F_n^{(Y)}(x) \right|,$$

where $F_n^{(X)}$ and $F_n^{(Y)}$ denote the empirical distribution functions based on $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ respectively. Denoting the corresponding empirical processes by $G_n^{(X)}$ and $G_n^{(Y)}$ and their independent limits by $G^{(X)}$ and $G^{(Y)}$, we get

$$\widehat{T}_n = \sup_{x \in \mathbb{R}} \left| G_n^{(X)} - G_n^{(Y)} \right| \xrightarrow{d} \sup_{x \in \mathbb{R}} |G^{(X)} - G^{(Y)}|,$$

and the dependent wild bootstrap method can again be applied to derive critical values of the test statistic.

### 4.3. von Mises statistics

A von Mises ($V$-) statistic based on $X_1, \ldots, X_n$ is defined as

$$V_n = \frac{1}{n^2} \sum_{s,t=1}^n h(X_s, X_t). \tag{4.1}$$

It is well known that many important statistics are of the form (4.1). Simple examples are the usual variance estimator (with $h(x, y) = (x - y)^2/2$), Gini's mean difference ($h(x, y) = |x - y|$) and, more importantly, test statistics of $L_2$-type such as the Cramér-von Mises test statistic. We assume that the kernel $h : \mathbb{R}^2 \to \mathbb{R}$ satisfies the following condition.

(A5) $h$ is continuous, bounded and symmetric w.r.t. permutation of its arguments, that is, $h(x, y) = h(y, x)$. Moreover, let $h$, $h_F(\cdot) := \int h(\cdot, y) dF(y)$, and $h(x, \cdot)$ have bounded variation (uniformly in $x$).

Beutner and Zähle (2014) proposed a partial integration approach to derive limit distributions of $V$-statistics based on results on convergence of empirical processes in (weighted) sup-norms. Under (A5) the statistic $V_n$ can be represented as a Stieltjes integral

$$V_n = \iint h(x, y) \, dF_n(x) \, dF_n(y).$$

Note that, with $V = \iint h(x, y) \, dF(x) \, dF(y)$,

$$\begin{aligned} V_n - V &= \iint h(x, y) \, d(F_n - F)(x) \, d(F_n - F)(y) \\ &\quad + 2 \int h_F(x) \, d(F_n - F)(x). \end{aligned}$$

It follows from Lemmas 3.4 and 3.6 in Beutner and Zähle (2014) that we can apply integration by parts and obtain

$$\begin{aligned} V_n - V &= \iint (F_n - F)(x-)(F_n - F)(y-) \, dh(x, y) \\ &\quad - 2 \int (F_n - F)(x-) \, dh_F(x), \end{aligned} \tag{4.2}$$

where $g(z-)$ denotes the limit from the left of a function $g$ at point $z$. This representation allows to infer from a convergence result for the empirical process the asymptotic behaviour of the $V$-statistic, both in the degenerate (with $h_F \equiv 0$) and the non-degenerate case. The following result is an immediate consequence of Theorem 3.15 in Beutner and Zähle (2014) and our Theorem 2.1.

**Theorem 4.4.** Suppose that (A1) and (A5) hold. Then,

(i) $\sqrt{n} \, (V_n - V) \xrightarrow{d} -2 \int G(x) \, dh_F(x)$.

(ii) If $V_n$ is degenerate, that is, if $h_F \equiv 0$, then

$$n \, V_n \xrightarrow{d} \iint G(x) \, G(y) \, dh(x, y).$$

Both limit distributions depend on the covariance structure of the process $G$, which might be unknown in applications. Thus, quantiles of the (asymptotic) distributions (e.g., to derive critical values of the Cramér-von Mises statistic for data with unspecified dependence structure) cannot be determined analytically. This difficulty can be circumvented by the application of the bootstrap method of Section 3.

In the non-degenerate case, we mimic $V_n - V$ by $V_n^* - V_n$, where, because of $F_n^*(x) = n^{-1}\sum_{t=1}^n \mathbb{1}(X_t \le x)(1 + \varepsilon_{t,n}^* - \bar{\varepsilon}_n^*)$,

$$V_n^* = \iint h(x,y)\,dF_n^*(x)\,dF_n^*(y)$$

$$= \frac{1}{n^2}\sum_{s,t=1}^n h(X_s, X_t)\left(1 + \varepsilon_{s,n}^* - \bar{\varepsilon}_n^*\right)\left(1 + \varepsilon_{t,n}^* - \bar{\varepsilon}_n^*\right).$$

We obtain that

$$V_n^* - V_n = \frac{1}{n^2}\sum_{s,t=1}^n h(X_s, X_t)\left(\varepsilon_{s,n}^* - \bar{\varepsilon}_n^*\right)\left(\varepsilon_{t,n}^* - \bar{\varepsilon}_n^*\right) + \frac{2}{n}\sum_{s=1}^n h_{F_n}(X_s)\left(\varepsilon_{s,n}^* - \bar{\varepsilon}_n^*\right)$$

$$= \iint \left(F_n^* - F_n\right)(x-)\left(F_n^* - F_n\right)(y-)\,dh(x,y) - 2\int\left(F_n^* - F_n\right)(x-)dh_F(x) + r_n^*,$$

where $r_n^* = -2n^{-1}\sum_{s=1}^n (h_{F_n}(X_s) - h_F(X_s))\left(\varepsilon_{s,n}^* - \bar{\varepsilon}_n^*\right)$ and $h_{F_n}(\cdot) = \int h(x,\cdot)dF_n(x)$. It turns out that $r_n^*$ is asymptotically negligible and Theorem 3.1 eventually yields consistency of the bootstrap approximation in the non-degenerate case.

In the degenerate case, where the right normalizing factor is $n$ rather than $\sqrt{n}$, we have to proceed in a different way. It can be conjectured from recent results from Leucht and Neumann (2013) (under some variant of Doukhan and Louhichi's (1999) weak dependence instead of $\beta$-mixing) that the term $n^{-1}\sum_{s,t=1}^n h(X_s, X_t)\left(\varepsilon_{s,n}^* - \bar{\varepsilon}_n^*\right)\left(\varepsilon_{t,n}^* - \bar{\varepsilon}_n^*\right)$ converges to the correct limit. On the other hand, the additional term $2\sum_{s=1}^n h_{F_n}(X_s)\left(\varepsilon_{s,n}^* - \bar{\varepsilon}_n^*\right)$ is of the same order and disturbs the intended convergence. In fact, we have to take into account that $h_F \equiv 0$, which also implies $V = 0$. Therefore, (4.2) simplifies to

$$V_n = \iint (F_n - F)(x-)(F_n - F)(y-)\,dh(x,y),$$

which suggests the bootstrap approximation

$$V_n^{**} = \iint \left(F_n^* - F_n\right)(x-)\left(F_n^* - F_n\right)(y-)\,dh(x,y)$$

$$= \frac{1}{n^2}\sum_{s,t=1}^n h(X_s, X_t)\left(\varepsilon_{s,n}^* - \bar{\varepsilon}_n^*\right)\left(\varepsilon_{t,n}^* - \bar{\varepsilon}_n^*\right).$$

Asymptotic validity of this approximation has been shown in Leucht and Neumann (2013) under conditions different from those imposed here while consistency of a block bootstrap method for non-degenerate $U$-statistics was proved in Dehling and Wendler (2010). In our context, consistency follows again from Theorem 3.1. All consistency results are summarized in the following theorem.

**Theorem 4.5.** Suppose that (A2), (A3) and (A5) hold. Then,

(i) $\sqrt{n}\left(V_n^* - V_n\right) \xrightarrow{d} -2\int G(x)\,dh_F(x)$     in probability.

(ii) If $V_n$ is degenerate, that is, if $h_F \equiv 0$, then

$$nV_n^{**} \xrightarrow{d} \iint G(x)\,G(y)\,dh(x,y) \qquad \text{in probability.}$$

## 5. SIMULATIONS

To provide some idea of the finite sample properties of the different bootstrap methods, we report the results of a small simulation study.

We investigated the size of the Kolmogorov-Smirnov test, with a nominal size chosen as $\gamma = 0.05, 0.1$. Data were generated from a stationary AR(1)-process,

$$X_t = \theta\, X_{t-1} + \eta_t, \qquad t \in \mathbb{N},$$

where $\theta = 0, 0.5, 0.7$ and $\eta_t \sim \mathcal{N}(0, 1 - \theta^2)$ are independent. With this choice, the $X_t$s have a standard normal distribution.

Our primary intention was to compare the performance of the dependent wild bootstrap with that of well-established block bootstrap methods. We have chosen two variants of the block bootstrap methodology, the MBB of Künsch (1989) and Liu and Singh (1992), which consists of independently drawing blocks of observations and then patching them together to a bootstrap time series, and the TBB by Paparoditis and Politis (2001). The latter method has superior bias properties than the classical block bootstrap; see Section 2 in Paparoditis and Politis (2001) for details. These methods are compared with three versions of the dependent wild bootstrap (DWB1–3). Although all of them are clearly in the spirit of the original proposal by Shao (2010), the first one is the special case of an autoregressive wild bootstrap also employed in Leucht and Neumann (2013) and Smeekes and Urbain (2014). The autocovariance function of the wild bootstrap variables $\varepsilon_{t,n}^*$ obeys Assumption 2.2 in Shao (2010) with $q = 1$ in the first two cases and with $q = 2$ in the third case. According to Remark 2.1 in that paper, the third variant shares the superior asymptotic bias properties with the tapered block bootstrap, whereas the first two variants have inferior bias properties comparable to those of the MBB. Finally, to show the necessity of not neglecting the dependence of the data, we also included Wu's (1986) (independent) wild bootstrap. Here is a summary of the technical details:

- WB: Wu's (1986) (independent) wild bootstrap

$$\varepsilon_{t,n}^* \sim \mathcal{N}(0, 1) \quad \text{i.i.d.}$$

- DWB1: Discretely sampled Ornstein-Uhlenbeck process (autoregressive wild bootstrap)

$$\varepsilon_{t,n}^* = e^{-1/l_n} \varepsilon_{t-1,n}^* + \zeta_t, \qquad t \in \mathbb{N},$$

  where $\zeta_t \sim \mathcal{N}(0, 1 - e^{-2/l_n})$ are i.i.d.
- DWB2: MA-process, rectangular weight function

$$\varepsilon_{t,n}^* = \zeta_t + \cdots + \zeta_{t-l_n+1},$$

  where $\zeta_t \sim \mathcal{N}(0, 1/l_n)$ are i.i.d.
- DWB3: MA-process, triangular weight function

$$\varepsilon_{t,n}^* = c_{n,1} \zeta_t + \cdots + c_{n,l_n} \zeta_{t-l_n+1},$$

  where $c_{n,k} = 0.5 - |(k - 0.5)/l_n - 0.5|$, $\zeta_t \sim \mathcal{N}(0, 1/c_n)$ are i.i.d. with $c_n = c_{n,1}^2 + \cdots + c_{n,l_n}^2$.
- MBB

  The original time series $Y_1, \ldots, Y_n$ ($Y_t = \mathbb{1}(X_t \leq x)$, here) is split in nonoverlapping blocks of length $l_n$. From these blocks, bootstrap blocks are generated by drawing with replacement; then, these blocks are patched together to a bootstrap series $Y_1^*, \ldots, Y_n^*$.

- TBB

  To reduce bias problems, Paparoditis and Politis (2001) proposed to split $Y_1, \ldots, Y_n$ in blocks of length $l_n$, apply a taper to these blocks, that is,

  $$Z_{(i-1)l_n+k} = c_{n,k} Y_{(i-1)l_n+k} / \sqrt{c_n},$$

  where $c_{n,k}$ and $c_n$ are chosen as mentioned earlier. From these new blocks, a bootstrap version is generated by drawing with replacement.

In the latter five cases, the tuning parameter $l_n$ plays a similar role. For simplicity, we have used the same values $l_n = 8, 10, 12, 15, 20, 30$ for all methods. To avoid having an incomplete block with the blockwise methods, we have chosen sample sizes $n = 240, 480, 960$ that are multiples of the $l_n$. We repeated the simulations $N = 1000$ times, each with $B = 1000$ bootstrap resamplings. The implementation was carried out with the aid of the statistical software package $R$; see R Core Team (2012). The results are reported in Tables I to IX below.

In the case of independent observations ($\theta = 0$), the classical (independent) wild bootstrap has quite a similar performance as the five time series bootstraps; however, it fails drastically in the two cases of dependence ($\theta = 0.5, 0.7$). The three versions of the dependent wild bootstrap showed a similar performance as the block bootstrap methods, whereas, as expected in view if the asymptotic results for the bias, DBW3 and TBB are slightly better than the other competitors. It is quite apparent that the empirical size is in almost all cases higher than the nominal one. This is due to the fact that, for all five bootstrap schemes, covariances are systematically underestimated. In our case of an AR(1)-process with all covariances positive, this effect explains the oversizing of the test.

Although we do not give any theoretical results on the behaviour of the test under the alternative, we illustrate the finite sample power properties under two alternative scenarios. To this end, we consider samples of size $n = 480$ of AR(1) processes with $\theta = 0.5$ and Gaussian observations such that the marginals are $\mathcal{N}(0.25, 1)$ and $\mathcal{N}(0, 1.5)$ respectively; see Tables X and XI. Again, we observe a comparable behaviour of the dependent wild and the block bootstrap. As usual, we see that more oversized tests have better power properties.

## 6. PROOFS

*Proof of Lemma 3.1*
We define

$$T_{n,1} = \text{cov}^* \left( G_n^{*,0}(x), G_n^{*,0}(y) \right) = \frac{1}{n} \sum_{s,t=1}^{n} (\mathbb{1}(X_s \leq x) - F(x))(\mathbb{1}(X_t \leq y) - F(y)) A_n(s,t)$$

Table I. Empirical size ($n = 240, \theta = 0$)

| | WB | | DWB1 | | DWB2 | | DWB3 | | MBB | | TBB | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ |
| $l_n = 8$ | 0.064 | 0.113 | 0.079 | 0.133 | 0.068 | 0.130 | 0.069 | 0.128 | 0.069 | 0.126 | 0.064 | 0.119 |
| $l_n = 10$ | | | 0.082 | 0.154 | 0.068 | 0.140 | 0.069 | 0.128 | 0.068 | 0.137 | 0.065 | 0.127 |
| $l_n = 12$ | | | 0.086 | 0.158 | 0.071 | 0.134 | 0.070 | 0.135 | 0.079 | 0.137 | 0.063 | 0.126 |
| $l_n = 15$ | | | 0.085 | 0.169 | 0.077 | 0.129 | 0.071 | 0.125 | 0.077 | 0.131 | 0.068 | 0.110 |
| $l_n = 20$ | | | 0.092 | 0.173 | 0.059 | 0.129 | 0.060 | 0.123 | 0.067 | 0.133 | 0.054 | 0.112 |
| $l_n = 30$ | | | 0.131 | 0.225 | 0.087 | 0.148 | 0.076 | 0.138 | 0.095 | 0.149 | 0.073 | 0.122 |

WB, wild bootstrap; DWB, dependent WB; MBB, moving block bootstrap; TBB, tapered block bootstrap.

Table II. Empirical size ($n = 240, \theta = 0.5$)

| | WB | | DWB1 | | DWB2 | | DWB3 | | MBB | | TBB | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ |
| $l_n = 8$ | 0.257 | 0.334 | 0.086 | 0.174 | 0.085 | 0.163 | 0.087 | 0.161 | 0.080 | 0.163 | 0.081 | 0.160 |
| $l_n = 10$ | | | 0.087 | 0.171 | 0.084 | 0.157 | 0.086 | 0.157 | 0.086 | 0.158 | 0.079 | 0.143 |
| $l_n = 12$ | | | 0.092 | 0.176 | 0.077 | 0.157 | 0.076 | 0.151 | 0.084 | 0.154 | 0.071 | 0.140 |
| $l_n = 15$ | | | 0.098 | 0.187 | 0.073 | 0.155 | 0.076 | 0.139 | 0.079 | 0.152 | 0.067 | 0.132 |
| $l_n = 20$ | | | 0.101 | 0.177 | 0.078 | 0.140 | 0.073 | 0.125 | 0.076 | 0.138 | 0.059 | 0.112 |
| $l_n = 30$ | | | 0.141 | 0.240 | 0.083 | 0.166 | 0.073 | 0.143 | 0.094 | 0.169 | 0.070 | 0.130 |

WB, wild bootstrap; DWB, dependent WB; MBB, moving block bootstrap; TBB, tapered block bootstrap.

Table III. Empirical size ($n = 240, \theta = 0.7$)

| | WB | | DWB1 | | DWB2 | | DWB3 | | MBB | | TBB | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ |
| $l_n = 8$ | 0.397 | 0.487 | 0.122 | 0.204 | 0.127 | 0.211 | 0.147 | 0.224 | 0.133 | 0.207 | 0.141 | 0.215 |
| $l_n = 10$ | | | 0.117 | 0.195 | 0.116 | 0.191 | 0.124 | 0.208 | 0.112 | 0.190 | 0.115 | 0.200 |
| $l_n = 12$ | | | 0.115 | 0.192 | 0.108 | 0.186 | 0.112 | 0.186 | 0.108 | 0.183 | 0.105 | 0.177 |
| $l_n = 15$ | | | 0.119 | 0.202 | 0.101 | 0.179 | 0.094 | 0.179 | 0.101 | 0.177 | 0.086 | 0.166 |
| $l_n = 20$ | | | 0.119 | 0.190 | 0.090 | 0.159 | 0.081 | 0.145 | 0.091 | 0.156 | 0.077 | 0.131 |
| $l_n = 30$ | | | 0.151 | 0.254 | 0.098 | 0.186 | 0.083 | 0.164 | 0.108 | 0.196 | 0.073 | 0.142 |

WB, wild bootstrap; DWB, dependent WB; MBB, moving block bootstrap; TBB, tapered block bootstrap.

Table IV. Empirical size ($n = 480, \theta = 0$)

| | WB | | DWB1 | | DWB2 | | DWB3 | | MBB | | TBB | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ |
| $l_n = 8$ | 0.054 | 0.116 | 0.057 | 0.122 | 0.054 | 0.122 | 0.053 | 0.119 | 0.058 | 0.124 | 0.056 | 0.114 |
| $l_n = 10$ | | | 0.061 | 0.112 | 0.051 | 0.102 | 0.052 | 0.103 | 0.056 | 0.103 | 0.051 | 0.099 |
| $l_n = 12$ | | | 0.084 | 0.152 | 0.077 | 0.139 | 0.073 | 0.136 | 0.084 | 0.139 | 0.076 | 0.132 |
| $l_n = 15$ | | | 0.065 | 0.149 | 0.057 | 0.131 | 0.057 | 0.128 | 0.062 | 0.134 | 0.052 | 0.127 |
| $l_n = 20$ | | | 0.101 | 0.161 | 0.083 | 0.152 | 0.081 | 0.147 | 0.083 | 0.149 | 0.075 | 0.136 |
| $l_n = 30$ | | | 0.091 | 0.157 | 0.061 | 0.127 | 0.063 | 0.123 | 0.068 | 0.133 | 0.058 | 0.117 |

WB, wild bootstrap; DWB, dependent WB; MBB, moving block bootstrap; TBB, tapered block bootstrap.

Table V. Empirical size ($n = 480, \theta = 0.5$)

| | WB | | DWB1 | | DWB2 | | DWB3 | | MBB | | TBB | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ |
| $l_n = 8$ | 0.238 | 0.325 | 0.095 | 0.143 | 0.094 | 0.147 | 0.099 | 0.150 | 0.094 | 0.149 | 0.094 | 0.146 |
| $l_n = 10$ | | | 0.094 | 0.151 | 0.094 | 0.150 | 0.092 | 0.147 | 0.096 | 0.148 | 0.086 | 0.140 |
| $l_n = 12$ | | | 0.084 | 0.159 | 0.077 | 0.146 | 0.068 | 0.141 | 0.077 | 0.145 | 0.069 | 0.144 |
| $l_n = 15$ | | | 0.084 | 0.159 | 0.072 | 0.138 | 0.068 | 0.131 | 0.077 | 0.140 | 0.064 | 0.127 |
| $l_n = 20$ | | | 0.088 | 0.167 | 0.074 | 0.142 | 0.063 | 0.133 | 0.076 | 0.144 | 0.063 | 0.129 |
| $l_n = 30$ | | | 0.101 | 0.159 | 0.091 | 0.129 | 0.087 | 0.129 | 0.092 | 0.132 | 0.076 | 0.122 |

WB, wild bootstrap; DWB, dependent WB; MBB, moving block bootstrap; TBB, tapered block bootstrap.

Table VI. Empirical size ($n = 480, \theta = 0.7$)

| | WB | | DWB1 | | DWB2 | | DWB3 | | MBB | | TBB | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ |
| $l_n = 8$ | 0.420 | 0.504 | 0.120 | 0.177 | 0.133 | 0.191 | 0.146 | 0.199 | 0.132 | 0.188 | 0.142 | 0.197 |
| $l_n = 10$ | | | 0.101 | 0.173 | 0.099 | 0.177 | 0.109 | 0.189 | 0.101 | 0.178 | 0.106 | 0.187 |
| $l_n = 12$ | | | 0.102 | 0.165 | 0.099 | 0.164 | 0.102 | 0.171 | 0.098 | 0.159 | 0.100 | 0.163 |
| $l_n = 15$ | | | 0.105 | 0.169 | 0.097 | 0.154 | 0.095 | 0.150 | 0.093 | 0.150 | 0.091 | 0.141 |
| $l_n = 20$ | | | 0.106 | 0.160 | 0.090 | 0.146 | 0.085 | 0.146 | 0.088 | 0.151 | 0.083 | 0.139 |
| $l_n = 30$ | | | 0.113 | 0.179 | 0.098 | 0.150 | 0.095 | 0.144 | 0.099 | 0.154 | 0.081 | 0.131 |

WB, wild bootstrap; DWB, dependent WB; MBB, moving block bootstrap; TBB, tapered block bootstrap.

Table VII. Empirical size ($n = 960, \theta = 0$)

| | WB | | DWB1 | | DWB2 | | DWB3 | | MBB | | TBB | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ |
| $l_n = 8$ | 0.064 | 0.129 | 0.063 | 0.131 | 0.062 | 0.126 | 0.064 | 0.126 | 0.062 | 0.129 | 0.063 | 0.124 |
| $l_n = 10$ | | | 0.071 | 0.129 | 0.064 | 0.126 | 0.067 | 0.124 | 0.071 | 0.128 | 0.065 | 0.125 |
| $l_n = 12$ | | | 0.057 | 0.109 | 0.053 | 0.109 | 0.053 | 0.101 | 0.054 | 0.107 | 0.053 | 0.105 |
| $l_n = 15$ | | | 0.069 | 0.123 | 0.061 | 0.118 | 0.063 | 0.111 | 0.062 | 0.114 | 0.064 | 0.111 |
| $l_n = 20$ | | | 0.061 | 0.120 | 0.053 | 0.110 | 0.050 | 0.108 | 0.051 | 0.117 | 0.047 | 0.104 |
| $l_n = 30$ | | | 0.086 | 0.140 | 0.074 | 0.126 | 0.068 | 0.124 | 0.074 | 0.125 | 0.064 | 0.115 |

WB, wild bootstrap; DWB, dependent WB; MBB, moving block bootstrap; TBB, tapered block bootstrap.

Table VIII. Empirical size ($n = 960, \theta = 0.5$)

| | WB | | DWB1 | | DWB2 | | DWB3 | | MBB | | TBB | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ |
| $l_n = 8$ | 0.254 | 0.361 | 0.081 | 0.141 | 0.079 | 0.140 | 0.083 | 0.140 | 0.080 | 0.143 | 0.079 | 0.142 |
| $l_n = 10$ | | | 0.082 | 0.150 | 0.086 | 0.145 | 0.089 | 0.148 | 0.084 | 0.146 | 0.083 | 0.141 |
| $l_n = 12$ | | | 0.080 | 0.133 | 0.071 | 0.132 | 0.075 | 0.128 | 0.077 | 0.134 | 0.075 | 0.128 |
| $l_n = 15$ | | | 0.059 | 0.136 | 0.051 | 0.121 | 0.046 | 0.120 | 0.055 | 0.126 | 0.054 | 0.116 |
| $l_n = 20$ | | | 0.071 | 0.119 | 0.065 | 0.116 | 0.059 | 0.114 | 0.065 | 0.116 | 0.058 | 0.111 |
| $l_n = 30$ | | | 0.085 | 0.147 | 0.064 | 0.138 | 0.056 | 0.126 | 0.067 | 0.140 | 0.061 | 0.122 |

WB, wild bootstrap; DWB, dependent WB; MBB, moving block bootstrap; TBB, tapered block bootstrap.

Table IX. Empirical size ($n = 960, \theta = 0.7$)

| | WB | | DWB1 | | DWB2 | | DWB3 | | MBB | | TBB | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ |
| $l_n = 8$ | 0.445 | 0.524 | 0.107 | 0.177 | 0.122 | 0.197 | 0.136 | 0.220 | 0.116 | 0.199 | 0.134 | 0.217 |
| $l_n = 10$ | | | 0.097 | 0.172 | 0.097 | 0.174 | 0.112 | 0.189 | 0.102 | 0.173 | 0.110 | 0.190 |
| $l_n = 12$ | | | 0.089 | 0.162 | 0.091 | 0.166 | 0.092 | 0.168 | 0.086 | 0.164 | 0.090 | 0.168 |
| $l_n = 15$ | | | 0.075 | 0.154 | 0.074 | 0.147 | 0.078 | 0.147 | 0.074 | 0.151 | 0.076 | 0.153 |
| $l_n = 20$ | | | 0.076 | 0.127 | 0.073 | 0.122 | 0.067 | 0.119 | 0.068 | 0.126 | 0.067 | 0.120 |
| $l_n = 30$ | | | 0.084 | 0.160 | 0.066 | 0.129 | 0.061 | 0.121 | 0.067 | 0.136 | 0.062 | 0.118 |

WB, wild bootstrap; DWB, dependent WB; MBB, moving block bootstrap; TBB, tapered block bootstrap.

Table X. Empirical power ($\mathcal{N}(0.25, 1)$)

| | WB | | DWB1 | | DWB2 | | DWB3 | | MBB | | TBB | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ |
| $l_n = 8$ | 0.969 | 0.976 | 0.896 | 0.939 | 0.901 | 0.942 | 0.904 | 0.943 | 0.902 | 0.940 | 0.908 | 0.940 |
| $l_n = 10$ | | | 0.889 | 0.925 | 0.880 | 0.938 | 0.883 | 0.936 | 0.883 | 0.939 | 0.879 | 0.935 |
| $l_n = 12$ | | | 0.896 | 0.942 | 0.892 | 0.938 | 0.888 | 0.937 | 0.890 | 0.935 | 0.884 | 0.924 |
| $l_n = 15$ | | | 0.876 | 0.940 | 0.871 | 0.933 | 0.874 | 0.935 | 0.878 | 0.935 | 0.866 | 0.929 |
| $l_n = 20$ | | | 0.894 | 0.941 | 0.879 | 0.937 | 0.874 | 0.933 | 0.882 | 0.936 | 0.867 | 0.928 |
| $l_n = 30$ | | | 0.890 | 0.937 | 0.871 | 0.929 | 0.865 | 0.929 | 0.873 | 0.929 | 0.858 | 0.920 |

WB, wild bootstrap; DWB, dependent WB; MBB, moving block bootstrap; TBB, tapered block bootstrap.

Table XI. Empirical power ($\mathcal{N}(0, 1.5)$)

| | WB | | DWB1 | | DWB2 | | DWB3 | | MBB | | TBB | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ |
| $l_n = 8$ | 0.831 | 0.915 | 0.480 | 0.678 | 0.484 | 0.672 | 0.492 | 0.684 | 0.479 | 0.680 | 0.480 | 0.681 |
| $l_n = 10$ | | | 0.480 | 0.675 | 0.457 | 0.661 | 0.464 | 0.663 | 0.462 | 0.670 | 0.448 | 0.650 |
| $l_n = 12$ | | | 0.499 | 0.712 | 0.471 | 0.692 | 0.462 | 0.683 | 0.478 | 0.694 | 0.459 | 0.674 |
| $l_n = 15$ | | | 0.462 | 0.666 | 0.437 | 0.645 | 0.416 | 0.637 | 0.434 | 0.644 | 0.397 | 0.611 |
| $l_n = 20$ | | | 0.505 | 0.714 | 0.457 | 0.674 | 0.438 | 0.655 | 0.461 | 0.679 | 0.429 | 0.642 |
| $l_n = 30$ | | | 0.502 | 0.685 | 0.440 | 0.629 | 0.413 | 0.609 | 0.449 | 0.639 | 0.397 | 0.591 |

WB, wild bootstrap; DWB, dependent WB; MBB, moving block bootstrap; TBB, tapered block bootstrap.

and

$$T_{n,2} = \text{cov}(G(x), G(y)) = \sum_{r=-\infty}^{\infty} E[(\mathbb{1}(X_0 \leq x) - F(x))(\mathbb{1}(X_r \leq y) - F(y))].$$

Then,

$$\left| \text{cov}^* \left( G_n^{*,0}(x), G_n^{*,0}(y) \right) - \text{cov}(G(x), G(y)) \right| \leq |T_{n,1} - E[T_{n,1}]| + |T_{n,2} - E[T_{n,1}]|.$$

Proposition 1 in Section 1.1 and Lemma 3 in Section 1.2 of Doukhan (1994) yield $|\text{cov}(\mathbb{1}(X_0 \leq x), \mathbb{1}(X_r \leq y))| \leq 2\beta_X(r)$, which in turn implies $\sum_{r=1}^{\infty} |\text{cov}(\mathbb{1}(X_0 \leq x), \mathbb{1}(X_r \leq y))| \leq 2\sum_{r=1}^{\infty} \beta_X(r) < \infty$. Now, we obtain by majorized convergence that

$$|T_{n,2} - ET_{n,1}| \leq \sum_{r=-\infty}^{\infty} |\text{cov}(\mathbb{1}(X_0 \leq x), \mathbb{1}(X_r \leq y))| (1 - (1 - |r|/n)_+ A_n(0, r)) \xrightarrow[n \to \infty]{} 0.$$

Denote $Z_{s,x} = \mathbb{1}(X_s \leq x) - F(x)$ and $Z_{t,y} = \mathbb{1}(X_t \leq y) - F(y)$. We have that

$$E[(T_{n,1} - ET_{n,1})^2] = \frac{1}{n^2} \sum_{s,t,u,v=1}^{n} A_n(s,t) A_n(u,v) \{E[Z_{s,x} Z_{t,y} Z_{u,x} Z_{v,y}] - E[Z_{s,x} Z_{t,y}] E[Z_{u,x} Z_{v,y}]\}$$

$$= \frac{1}{n^2} \sum_{s,t,u,v=1}^{n} A_n(s,t) A_n(u,v) \, \text{cum}(Z_{s,x}, Z_{t,y}, Z_{u,x}, Z_{v,y})$$

$$+ \frac{1}{n^2} \sum_{s,t,u,v=1}^{n} A_n(s,t) A_n(u,v) \{E[Z_{s,x} Z_{u,x}] E[Z_{t,y} Z_{v,y}]$$

$$+ E[Z_{s,x} Z_{v,y}] E[Z_{t,y} Z_{u,x}]\} =: T_{n,11} + T_{n,12},$$

$$(6.1)$$

where $\text{cum}(Z_1, Z_2, Z_3, Z_4) = E[Z_1 Z_2 Z_3 Z_4] - E[Z_1 Z_2]E[Z_3 Z_4] - E[Z_1 Z_3]E[Z_2 Z_4] - E[Z_1 Z_4]E[Z_2 Z_3]$ denotes the joint cumulant of real-valued and centered random variables $Z_1, \ldots, Z_4$. Let $1 \le s \le t \le u \le v \le n$, $r = \max\{t - s, u - t, v - u\}$. As a prerequisite to estimate $T_{n,11}$, we prove that

$$|\text{cum}(Z_s, Z_t, Z_u, Z_v)| \le 8\,\beta_X(r), \tag{6.2}$$

where $Z_s$ denotes either $Z_{s,x}$ or $Z_{s,y}$. To see this, we distinguish between three cases, $r = t - s$, $r = u - t$ and $r = v - u$.

(i) $r = t - s$

Note that $|Z_s|, |Z_t|, |Z_u|, |Z_v| \le 1$. We obtain again from Proposition 1 in Section 1.1 and Lemma 3 in Section 1.2 of Doukhan (1994) that $|E[Z_s Z_t Z_u Z_v]| = |\text{cov}(Z_s, Z_t Z_u Z_v)| \le 2\beta_X(r)$ and $\max\{|E[Z_s Z_t]|, |E[Z_s Z_u]|, |E[Z_s Z_v]|\} \le 2\beta_X(r)$, which implies

$$|\text{cum}(Z_s, Z_t, Z_u, Z_v)| \le 8\,\beta_X(r).$$

(ii) $r = u - t$

Here, we have $|\text{cov}(Z_s Z_t, Z_u Z_v)| \le 2\beta_X(r)$ and $\max\{|E[Z_s Z_u]|, |E[Z_s Z_v]|, |E[Z_t Z_u]|, |E[Z_t Z_v]|\} \le 2\beta_X(r)$, which yields

$$\begin{aligned}|\text{cum}(Z_s, Z_t, Z_u, Z_v)| &\le |\text{cov}(Z_s Z_t, Z_u Z_v)| + |E[Z_s Z_u]E[Z_t Z_v]| + |E[Z_s Z_v]E[Z_t Z_u]| \\ &\le 2\beta_X(r) + 4\,\beta_X^2(r).\end{aligned}$$

(iii) $r = v - u$

This case is analogous to (i).

Since $|A_n(s, t)| \le 1$, we obtain from (6.2) that

$$|T_{n,11}| \le \frac{4!}{n^2} \sum_{1 \le s \le t \le u \le v \le n} |\text{cum}(Z_s, Z_t, Z_u, Z_v)| = O(n^{-1}). \tag{6.3}$$

Moreover, we obtain again by a covariance inequality and since $\sum_{r=0}^{n-1} A_n(0, r) = O(l_n)$ that

$$|T_{n,12}| = O(l_n\,n^{-1}), \tag{6.4}$$

which completes the proof. □

*Proof of Lemma 3.2*
To find an appropriate grid, we have to take into account the impact of the dependence structure on sums of mixed fourth moments of the increments of the processes $G_n^{*,0}$. Since the dependence between the bootstrap random variables $\varepsilon_{1,n}^*, \ldots, \varepsilon_{n,n}^*$ gets stronger as $n \to \infty$, we do not lose much by estimating the fourth moments of the

increment of $G_n^{*,0}$ over the interval $(x, y]$ as

$$E\, E^* \left[ \left( G_n^{*,0}(x) - G_n^{*,0}(y) \right)^4 \right] = \frac{1}{n^2} \sum_{s,t,u,v=1}^{n} E\left[ \tilde{Z}_s \tilde{Z}_t \tilde{Z}_u \tilde{Z}_v \right]\, E^*\left[ \varepsilon_{s,n}^* \varepsilon_{t,n}^* \varepsilon_{u,n}^* \varepsilon_{v,n}^* \right]$$

$$\leq \frac{3}{n^2} \sum_{s,t,u,v=1}^{n} \left| E\left[ \tilde{Z}_s \tilde{Z}_t \tilde{Z}_u \tilde{Z}_v \right] \right|,$$

where $\tilde{Z}_w = \mathbb{1}(X_w \in (x, y]) - P(X_w \in (x, y])$.

For arbitrary $s_1 \leq \ldots \leq s_u \leq t_1 \leq \ldots \leq t_v, u, v \in \mathbb{N}$, let $P^{X_{t_1}, \ldots, X_{t_v} | X_{s_1} = x_1, \ldots, X_{s_u} = x_u}(B)$, defined for $x_1, \ldots, x_u \in \mathbb{R}$ and $B \in \mathcal{B}^v$, denote a regular conditional distribution of $(X_{t_1}, \ldots, X_{t_v})'$ given $X_{s_1}, \ldots, X_{s_u}$. For $1 \leq s_1 \leq s_2 \leq s_3 \leq s_4 \leq n$, we use the estimates

$$\left| \text{cov}\left( \tilde{Z}_{s_1}, \tilde{Z}_{s_2} \tilde{Z}_{s_3} \tilde{Z}_{s_4} \right) \right| = \left| E\left[ \mathbb{1}(X_{s_1} \in (x, y])\, \left( \tilde{Z}_{s_2} \tilde{Z}_{s_3} \tilde{Z}_{s_4} - E\left[ \tilde{Z}_{s_2} \tilde{Z}_{s_3} \tilde{Z}_{s_4} \right] \right) \right] \right|$$

$$\leq 4 \int_{(x, y]} \sup_{B \in \mathcal{B}^3} \left| P^{X_{s_2}, X_{s_3}, X_{s_4} | X_{s_1} = z}(B) - P^{X_{s_2}, X_{s_3}, X_{s_4}}(B) \right|\, P^{X_{s_1}}(dz) \tag{6.5}$$

and, for $u = 2, 3$,

$$\left| \text{cov}\left( \tilde{Z}_{s_1} \cdots \tilde{Z}_{s_u}, \tilde{Z}_{s_{u+1}} \cdots \tilde{Z}_{s_4} \right) \right|$$

$$\leq \left| E\left[ \prod_{w=1}^{u-1} \tilde{Z}_{s_w}\, \mathbb{1}(X_{s_u} \in (x, y])\, \left( \tilde{Z}_{s_{u+1}} \cdots \tilde{Z}_{s_4} - E[\tilde{Z}_{s_{u+1}} \cdots \tilde{Z}_{s_4}] \right) \right] \right|$$

$$+ P\left( X_{s_u} \in (x, y] \right) \left| E\left[ \prod_{w=1}^{u-1} \tilde{Z}_{s_w}\, \left( \tilde{Z}_{s_{u+1}} \cdots \tilde{Z}_{s_4} - E[\tilde{Z}_{s_{u+1}} \cdots \tilde{Z}_{s_4}] \right) \right] \right| \tag{6.6}$$

$$\leq 4 \int_{\mathbb{R}^{u-1} \times (x, y]} \sup_{B \in \mathcal{B}^{4-u}} \left| P^{X_{s_{u+1}}, \ldots, X_{s_4} | (X_{s_1}, \ldots, X_{s_u})' = (z_1, \ldots, z_u)'}(B) - P^{X_{s_{u+1}}, \ldots, X_{s_4}}(B) \right|$$

$$P^{X_{s_1}, \ldots, X_{s_u}}(dz_1, \ldots, dz_u)$$

$$+ P^{X_0}((x, y])\, 2\, \beta_X(s_{u+1} - s_{u-1}).$$

On the other hand, according to Equation (2) on page 3 in Doukhan (1994), we have

$$\int_{\mathbb{R}^u} \sup_{B \in \mathcal{B}^v} \left| P^{X_{t_1}, \ldots, X_{t_v} | (X_{s_1}, \ldots, X_{s_u})' = (z_1, \ldots, z_u)'}(B) - P^{X_{t_1}, \ldots, X_{t_v}}(B) \right|\, P^{X_{s_1}, \ldots, X_{s_u}}(dz_1, \ldots, dz_u)$$

$$= \beta(\sigma(X_{s_1}, \ldots, X_{s_u}), \sigma(X_{t_1}, \ldots, X_{t_v})) \leq \beta_X(t_1 - s_u). \tag{6.7}$$

Inspired by (6.5), (6.6) and (6.7), we define

$$\Delta_r(x) = \beta_X(r)\, F(x)$$

$$+ \sup_{m,n \in \mathbb{N}} \int_{\mathbb{R}^{u-1} \times (-\infty, x]} \sup_{B \in \mathcal{B}^v} \left| P^{X_r, \ldots, X_{r+n-1} | (X_{-m+1}, \ldots, X_0)' = (z_m, \ldots, z_1)'}(B) - P^{X_r, \ldots, X_{r+n-1}}(B) \right|$$

$$P^{X_{-m+1}, \ldots, X_0}(dz_m, \ldots, dz_1)$$

and

$$\Delta(x) = \sum_{r=0}^{\infty} \Delta_r(x).$$

It is clear that $\Delta$ is monotonously non-decreasing and that $D := \lim_{x\to\infty} \Delta(x) \le 2 \sum_{r=0}^{\infty} \beta_X(r) < \infty$. Moreover, it follows from $|\Delta(x) - \Delta(y)| \le 2K|F(x) - F(y)| + 2\sum_{r=K}^{\infty} \beta_X(r)$ for all $x, y \in \mathbb{R}$ and $K \in \mathbb{N}$ that $\Delta$ is a continuous function. According to the previous considerations, to prove the assertion of the lemma, we construct a dyadic system of intervals related to $\Delta$ as follows.

For $j \in \mathbb{N}_0$, $0 \le k \le 2^j$, we define

$$
x_{j,k} = \begin{cases} \Delta^{-1}(D\,k\,2^{-j}), & \text{for } 1 \le k < 2^j, \\ -\infty, & \text{for } k = 0, \\ \infty, & \text{for } k = 2^j \end{cases}
$$

Let $j \in \mathbb{N}_0$ and $1 \le k \le n$ be arbitrary. Let, for the time being, $\tilde{Z}_s = \mathbb{1}(X_s \in (x_{j,k-1}, x_{j,k}]) - P(X_s \in (x_{j,k-1}, x_{j,k}])$. We have

$$
E E^* \left[ \left( G_n^{*,0}(x_{j,k-1}) - G_n^{*,0}(x_{j,k}) \right)^4 \right]
$$

$$
\le \frac{3}{n^2} \sum_{s,t,u,v=1}^{n} \left| E\left[ \tilde{Z}_s \tilde{Z}_t \tilde{Z}_u \tilde{Z}_v \right] \right|
$$

$$
\le \frac{3\,4!}{n^2} \sum_{1 \le s \le t \le u \le v \le n} \left| \mathrm{cum}(\tilde{Z}_s, \tilde{Z}_t, \tilde{Z}_u, \tilde{Z}_v) \right|
$$

$$
+ \frac{3}{n^2} \sum_{s,t,u,v=1}^{n} \left| E\left[ \tilde{Z}_s \tilde{Z}_t \right] E\left[ \tilde{Z}_u \tilde{Z}_v \right] + E\left[ \tilde{Z}_s \tilde{Z}_u \right] E\left[ \tilde{Z}_t \tilde{Z}_v \right] + E\left[ \tilde{Z}_s \tilde{Z}_v \right] E\left[ \tilde{Z}_t \tilde{Z}_u \right] \right|.
$$

According to (6.5) and (6.6), the first term on the right-hand side is of order $O(n^{-1} 2^{-j})$. The second one is of order $O(2^{-2j})$, which yields the assertion. □

*Proof of Corollary 3.2*

We prove (3.4), which implies the assertion of the corollary. According to our dyadic grid points $x_{j,k}$, we define projections $\Pi_j$ as

$$
\Pi_j g(x) = g(x_{j,k}) \qquad \text{if } x \in I_{j,k} = (x_{j,k-1}, x_{j,k}].
$$

Let $J_n$ be such that $2^{J_n} \le n < 2^{J_n+1}$. We have, for $0 \le J_0 < J_n$,

$$
\max_{1 \le k \le 2^{J_0}} \sup_{x \in (x_{J_0,k-1}, x_{J_0,k}]} \left| G_n^{*,0}(x) - G_n^{*,0}(x_{J_0,k}) \right|
$$

$$
\le \sum_{j=J_0+1}^{J_n} \left\| \Pi_j G_n^{*,0} - \Pi_{j-1} G_n^{*,0} \right\|_\infty \quad + \quad \left\| G_n^{*,0} - \Pi_{J_n} G_n^{*,0} \right\|_\infty. \tag{6.8}
$$

We choose any $\alpha \in (0, 1/4)$ and define thresholds $\lambda_j = 2^{-j\alpha}$. We obtain by Lemma 3.2 and Markov's inequality that

$$
E\left[ P^*\left( \left\| \Pi_j G_n^{*,0} - \Pi_{j-1} G_n^{*,0} \right\|_\infty > \lambda_j \right) \right] \le \sum_{k=1}^{2^{j-1}} E\left[ P^*\left( \left| G_n^{*,0}(x_{j,2k-1}) - G_n^{*,0}(x_{j,2k}) \right| > \lambda_j \right) \right]
$$

$$
\le 2^{j-1} \frac{K_0 \left( 2^{-2j} + n^{-1} 2^{-j} \right)}{\lambda_j^4} = K_0\, 2^{j(4\alpha-1)},
$$

which implies that

$$
\begin{aligned}
&E\left[P^*\left(\left\|\Pi_j G_n^{*,0} - \Pi_{j-1} G_n^{*,0}\right\|_\infty > \lambda_j \quad \text{for some } j \in \{J_0+1,\dots,J_n\}\right)\right] \\
&\leq K_0 \sum_{j=J_0+1}^\infty 2^{j(4\alpha-1)} \leq \frac{\eta^2}{2},
\end{aligned}
\tag{6.9}
$$

if $J_0$ is sufficiently large. Moreover,

$$
\sum_{j=J_0+1}^\infty \lambda_j \leq \frac{\epsilon}{2},
\tag{6.10}
$$

again for sufficiently large $J_0$.

Furthermore, we use the rough estimate

$$
\begin{aligned}
\left\|G_n^{*,0} - \Pi_{J_n} G_n^{*,0}\right\|_\infty &\leq \max_{1\leq t\leq n}\left\{|\varepsilon_{t,n}^*|\right\} \max_{1\leq k\leq 2^{J_n}}\left\{\frac{1}{\sqrt{n}}\sum_{t=1}^n |\mathbb{1}(X_t \in I_{J_n,k}) - P(X_t \in I_{J_n,k})|\right\} \\
&\leq \max_{1\leq t\leq n}\left\{|\varepsilon_{t,n}^*|\right\} \max_{1\leq k\leq 2^{J_n}}\left\{|G_n(x_{J_n,k}) - G_n(x_{J_n,k-1})| + 2\sqrt{n}P(X_0 \in I_{J_n,k})\right\}
\end{aligned}
$$

and obtain that

$$
E\left[P^*\left(\left\|G_n^{*,0} - \Pi_{J_n} G_n^{*,0}\right\|_\infty > \frac{\epsilon}{2}\right)\right] \leq \frac{\eta^2}{2},
\tag{6.11}
$$

which completes, in conjunction with (6.8), (6.9) and (6.10), the proof. □

*Proof of Theorem 4.2*

(i) According to Theorem 2.1, $(G_n)_{n\in\mathbb{N}}$ converges (w.r.t. the supremum metric) to the process $G$, which posesses continuous sample paths. Therefore,

$$
F_n(t_{n,q}) = q + o_P(n^{-1/2})
\tag{6.12}
$$

and, since $F'(t_q) > 0$,

$$
t_{n,q} \xrightarrow{P} t_q.
\tag{6.13}
$$

Furthermore, by Theorem 3.1, $(G_n^*)_{n\in\mathbb{N}}$ converges in probability to the same limit $G$. Therefore, the largest jump of $F_n^*$ is of order $o_{P^*}(n^{-1/2})$ and we obtain

$$
F_n^*\left(t_{n,q}^*\right) = q + o_{P^*}(n^{-1/2}).
\tag{6.14}
$$

Since $F'$ is continuously differentiable and $F'(t_q) > 0$, we also obtain

$$
t_{n,q}^* \xrightarrow{P^*} t_q.
\tag{6.15}
$$

Armed with these prerequisites, we can now derive the Bahadur representation for $t_{n,q}^*$. Stochastic equicontinuity of $(G_n^{*,0})_{n\in\mathbb{N}}$ stated in Corollary 3.2 and $\sup_{x\in\mathbb{R}}\left|G_n^*(x) - G_n^{*,0}(x)\right| = O_{P^*}(\sqrt{l_n/n})$ imply in conjunction with (6.15) that

$$
F_n^*\left(t_{n,q}^*\right) - F_n\left(t_{n,q}^*\right) = F_n^*(t_q) - F_n(t_q) + o_{P^*}(n^{-1/2}).
\tag{6.16}
$$

On the other hand, it follows from (6.12) and (6.14) that

$$F_n^* \left(t_{n,q}^*\right) - F_n \left(t_{n,q}^*\right) = F_n(t_{n,q}) - F_n \left(t_{n,q}^*\right) + o_{P*}(n^{-1/2}).$$

Furthermore, we obtain from stochastic equicontinuity of $(G_n)_{n\in\mathbb{N}}$, (6.13) and (6.15) that

$$\left(F_n(t_{n,q}) - F_n \left(t_{n,q}^*\right)\right) - \left(F(t_{n,q}) - F \left(t_{n,q}^*\right)\right) = n^{-1/2} \left(G_n(t_{n,q}) - G_n \left(t_{n,q}^*\right)\right) = o_{P*}(n^{-1/2}).$$

These two approximations lead to

$$\begin{aligned} F_n^* \left(t_{n,q}^*\right) - F_n(t_{n,q}^*) &= F(t_{n,q}) - F \left(t_{n,q}^*\right) + o_{P*}(n^{-1/2}) \\ &= \left(t_{n,q} - t_{n,q}^*\right) \left(F'(t_q) + o_{P*}(1)\right) + o_{P*}(n^{-1/2}). \end{aligned} \tag{6.17}$$

Rearranging terms, we obtain from (6.16) and (6.17) that

$$t_{n,q}^* - t_{n,q} = \frac{F_n(t_q) - F_n^*(t_q)}{F'(t_q)} + o_{P*}(n^{-1/2}).$$

(ii) This is an immediate consequence of (i) and Theorem 3.1. □

*Proof of Theorem 4.3*

We obtain from the Theorems 2.1 and 3.1 and the continuous mapping theorem that

$$T_n \xrightarrow{d} T := \sup_{x\in\mathbb{R}} |G(x)| \tag{6.18}$$

and

$$T_n^* \xrightarrow{d} T \quad \text{in probability.} \tag{6.19}$$

Absolute continuity of the distribution of $T$ will be derived from a result from Lifshits (1984). First, we compactify the domain of the limit process. Define

$$\tilde{G}(y) = \begin{cases} 0, & \text{if } y = 0, \\ G\left(F_0^{-1}(y)\right), & \text{if } 0 < y < 1, \\ 0, & \text{if } y = 1 \end{cases}$$

It is obvious that $\sup_{x\in\mathbb{R}} |G(x)| = \sup_{y\in[0,1]} |\tilde{G}(y)|$. The process $((\tilde{G}(y))_{y\in[0,1]}$ is a centered Gaussian process defined on a compact set and with continuous sample paths. Hence, Proposition 3 of Lifshits (1984) can be applied and it follows that $\sup_{y\in[0,1]} \tilde{G}(y)$ is absolutely continuous w.r.t. Lebesgue measure on $(0,\infty)$. For the same reason, the distribution of $\sup_{y\in[0,1]}(-\tilde{G}(y))$ is also absolutely continuous on $(0,\infty)$. Hence, the distribution of $\sup_{y\in[0,1]} |\tilde{G}(y)|$, and therefore also that of $T$ has not an atom unequal to 0. However, since $P(T \neq 0) = 1$, we obtain that the distribution of $T$ is absolutely continuous. Therefore, we obtain from (6.18)

$$\sup_{x\in\mathbb{R}} |P(T_n \leq x) - P(T \leq x)| \xrightarrow[n\to\infty]{} 0, \tag{6.20}$$

and from (6.19)

$$\sup_{x\in\mathbb{R}} \left|P^* \left(T_n^* \leq x\right) - P(T \leq x)\right| \xrightarrow{P} 0, \tag{6.21}$$

by Polya's theorem. This implies

$$
\begin{aligned}
\left| P\left(T_n > t_\alpha^*\right) - \alpha \right| &\le \left| P\left(T_n > t_\alpha^*\right) - P^*\left(T_n^* > t_\alpha^*\right) \right| + \left| P^*\left(T_n^* > t_\alpha^*\right) - \alpha \right| \\
&\le \sup_x \left| P\left(T_n > x\right) - P^*\left(T_n^* > x\right) \right| + o_P(1).
\end{aligned}
$$

Therefore, we obtain that $P(T_n > t_\alpha^*) \longrightarrow_{n\to\infty} \alpha$, as required. $\qquad\square$

*Proof of Theorem 4.4*

The assertions follow from Theorem 3.15 in Beutner and Zähle (2014) and it remains to validate its prerequisites (a) to (c) with $a_n = n$.

(a) We have to check the assumptions of their Lemmas 3.4 and 3.6. Condition (a) of Lemma 3.4 is satisfied due to the comments after their Assumption 3.2. Moreover, $h_F$ is continuous and has bounded variation under (A5), which yields (b) and (c). The last assumption of Lemma 3.4 follows from boundedness of $h$. Condition (a) of Lemma 3.6 is contained in condition (a) of Lemma 3.4. Assumptions (b) to (d) follow immediately from (A5).

(b) This assertion follows from their Remark 3.16.

(c) Convergence of the empirical process to a Gaussian process with continuous paths follows from our Theorem 2.1. $\qquad\square$

*Proof of Theorem 4.5*

(i) We first show that $r_n^*$, defined after Theorem 4.4, is of order $o_{P^*}(n^{-1/2})$. We have

$$
-\frac{1}{2} r_n^* = \frac{1}{n^2} \sum_{s=1}^n W_{s,n} \left( \varepsilon_{s,n}^* - \bar\varepsilon_n^* \right),
$$

where $W_{s,n} = \sum_{t=1}^n [h(X_s, X_t) - \int h(X_s, x)\, dF(x)]$. Note that

$$
E[W_{s,n}^2] = \sum_{t_1, t_2 = 1}^n E W_{s, t_1, t_2},
$$

where $W_{s,t_1,t_2} = (h(X_s, X_{t_1}) - \int h(X_s, x)\, dF(x))(h(X_s, X_{t_2}) - \int h(X_s, x)\, dF(x))$. Let, w.l.o.g., $t_1 \le t_1 + r = t_2$. If $t_1 \le s \le t_2$, then $\max\{|t_1 - s|, |t_2 - s|\} \ge r/2$. In the case of $t_2 - s \ge r/2$, Berbee's lemma allows us to choose $\tilde X_{t_2}$ independent of $X_{t_1}, X_s$ such that $\tilde X_{t_2} \overset{d}{=} X_{t_2}$ and $P\left(\tilde X_{t_2} \ne X_{t_2}\right) \le \beta_X([r/2])$. This implies

$$
\begin{aligned}
\left| E W_{s,t_1,t_2} \right| &= \left| E(h(X_s, X_{t_1}) - \int h(X_s, x)\, dF(x))(h(X_s, X_{t_2}) - h(X_s, \tilde X_{t_2})) \right| \\
&\le 4 \|h\|_\infty^2\, \beta_X([r/2]).
\end{aligned}
$$

Analogously, we obtain in the case of $s - t_1 \ge r/2$ that

$$
\left| E W_{s,t_1,t_2} \right| \le 4 \|h\|_\infty^2\, \beta_X([r/2]).
$$

If $s \le t_1 \le t_2$ or $t_1 \le t_2 \le s$, we can proceed similarly. If, for example, $s \le t_1 \le t_1 + r = t_2$, then we can choose $\tilde X_{t_2}$ independent of $X_{t_1}, X_s$ such that $\tilde X_{t_2} \overset{d}{=} X_{t_2}$ and $P\left(\tilde X_{t_2} \ne X_{t_2}\right) \le \beta_X(r)$. This leads to

$$
\left| E W_{s,t_1,t_2} \right| = \left| E(h(X_s, X_{t_1}) - \int h(X_s, x)\, F(dx))(h(X_s, X_{t_2}) - h(X_s, \tilde X_{t_2})) \right| \le 4 \|h\|_\infty^2\, \beta_X(r).
$$

Since $\sum_{r=1}^{\infty} r^2 \beta_X(r) < \infty$, we obtain from the aforementioned estimates that

$$\max_{1 \leq s \leq n} E\left[W_{s,n}^2\right] = O(n). \tag{6.22}$$

Since

$$\sum_{s_1,s_2=1}^{n} \left|E^*\left[\left(\varepsilon_{s_1,n}^* - \bar{\varepsilon}_n^*\right)\left(\varepsilon_{s_2,n}^* - \bar{\varepsilon}_n^*\right)\right]\right| = O(n\, l_n), \tag{6.23}$$

we obtain that

$$EE^*\left[\left(\sum_{s=1}^{n} W_{s,n}\left(\varepsilon_{s_1,n}^* - \bar{\varepsilon}_n^*\right)\right)^2\right] = O(n^2 l_n),$$

which implies that

$$r_n^* = o_{P^*}(n^{-1/2}). \tag{6.24}$$

Recall that

$$V_n^* - V_n = \iint h(x,y)\, d\left(F_n^* - F_n\right)(x)\, d\left(F_n^* - F_n\right)(y) + 2\int h_F(x)\, d\left(F_n^* - F_n\right)(x) + r_n^*.$$

If we could validate the corresponding formulae of partial integration, we would end up with

$$V_n^* - V_n = \iint \left(F_n^* - F_n\right)(x-)\left(F_n^* - F_n\right)(y-)\, dh(x,y)$$
$$- 2\int \left(F_n^* - F_n\right)(x-)\, dh_F(x) + o_{P^*}(n^{-1/2}). \tag{6.25}$$

It then follows from the proof of Theorem 4.4 that the function $\Phi : (\{g \in D(\bar{\mathbb{R}}) : \|g\|_\infty < \infty\}, \|\cdot\|_\infty) \to \mathbb{R}$ given by $\Phi(f) = -2\int f(x-)\, dh_F(x) + \iint f(y-)\, dh(x,y)$ is continuous (this is equivalent to checking assumption (b) of Theorem 3.15 in Beutner and Zähle (2014)). Hence, the assertion follows from Theorem 3.1 and the continuous mapping theorem provided that (6.25) holds. Since $F_n^* - F_n$ and $h_F$ are bounded càdlàg functions of bounded variation by assumption (A5) and since $\lim_{x \to \pm\infty}\left(F_n^* - F_n\right)(x) = 0$,

$$\int h_F(x)\, d\left(F_n^* - F_n\right)(x) = -\int \left(F_n^* - F_n\right)(x-)\, dh_F(x)$$

can be deduced from Lemma B.1 in Beutner and Zähle (2013). Finally,

$$\iint h(x,y)\, d\left(F_n^* - F_n\right)(x)\, d\left(F_n^* - F_n\right)(y)$$
$$= -\iint \left(F_n^* - F_n\right)(x-)\left(F_n^* - F_n\right)(y-)\, dh(x,y) \tag{6.26}$$

can be verified in a similar manner as Lemma 3.6 in Beutner and Zähle (2014). Since $\left(F_n^* - F_n\right)\left(F_n^* - F_n\right)$ and $h$ are of bounded variation and continuous respectively we first get from Gill *et al.* (1995, Lemma 2.2)

that

$$\int_{a_1}^{a_2} \int_{b_1}^{b_2} h(x, y) \, d \left( F_n^* - F_n \right) (x) \, d \left( F_n^* - F_n \right) (y)$$

$$= \int_{a_1}^{a_2} \int_{b_1}^{b_2} \left( F_n^* - F_n \right) (x-) \left( F_n^* - F_n \right) (y-) \, dh(x, y)$$

$$- \int_{a_1}^{a_2} \left( F_n^* - F_n \right) (x-) \left( F_n^* - F_n \right) (b_2) \, dh(x, b_2) - \int_{b_1}^{b_2} \left( F_n^* - F_n \right) (y-) \left( F_n^* - F_n \right) (a_2) \, dh(y, a_2)$$

$$+ \int_{a_1}^{a_2} \left( F_n^* - F_n \right) (x-) \left( F_n^* - F_n \right) (b_1) \, dh(x, b_1) + \int_{b_1}^{b_2} \left( F_n^* - F_n \right) (y-) \left( F_n^* - F_n \right) (a_1) \, dh(y, a_1)$$

$$+ \left( F_n^* - F_n \right) (a_2) \left( F_n^* - F_n \right) (b_2) h(a_2, b_2) - \left( F_n^* - F_n \right) (a_2) \left( F_n^* - F_n \right) (b_1) h(a_2, b_1)$$

$$- \left( F_n^* - F_n \right) (a_1) \left( F_n^* - F_n \right) (b_2) h(a_1, b_2) + \left( F_n^* - F_n \right) (a_1) \left( F_n^* - F_n \right) (b_1) h(a_1, b_1)$$

for finite intervals $(a_1, a_2]$ and $(b_1, b_2]$. Obviously, the last four summands tend to zero as $-a_1, -a_2, b_1, b_2 \to \infty$. The same holds true for the summands two to five since $h(\cdot, x)$ is of bounded variation uniformly in $x$ under (A5). Noting that $h$ generates a finite signed measure on $\mathbb{R}^2$, we can deduce (6.26) from continuity from below of finite measures as in the proof of Lemma B.1 in Beutner and Zähle (2013).

(ii) This result follows from (6.26) and Theorem 3.1. □

## REFERENCES

Beutner E, Zähle H. 2013. Deriving the asymptotic distribution of $U$- and $V$-statistics of dependent data using weighted empirical processes. *Bernoulli* **18**: 803–822.

Beutner E, Zähle H. 2014. Continuous mapping approach to the asymptotics of $U$- and $V$-statistics. *Bernoulli* **20**: 846–877.

Billingsley P. 1968. *Convergence of Probability Measures*. New York: Wiley.

Bühlmann P. 1994. Blockwise bootstrapped empirical process for stationary sequences. *Annals of Statistics* **22**: 995–1012.

Bühlmann P. 1995. The blockwise bootstrap for general empirical processes of stationary sequences. *Stochastic Processes and their Applications* **58**: 247–265.

Dehling H, Wendler M. 2010. Central limit theorem and the bootstrap for $U$-statistics of strongly mixing data. *Journal of Multivariate Analysis* **100**: 126–137.

Doukhan P. 1994. *Mixing. Properties and Examples*, Lecture Notes in Statistics, vol. 85. New York: Springer.

Doukhan P, Louhichi S. 1999. A new weak dependence condition and application to moment inequalities. *Stochastic Processes and their Applications* **84**: 313–342.

Doukhan P, Massart P, Rio E. 1995. Invariance principles for absolutely regular processes. *Annales de l'Institute Henri Poincaré, Probabilités et Statistiques* **31**: 393–427.

Gill RD, van der Laan MJ, Wellner JA. 1995. Inefficient estimators of the bivariate survival function for three models. *Annales de l'Institute Henri Poincaré, Probabilités et Statistiques.* **31**: 545–597.

Künsch HR. 1989. The jackknife and the bootstrap for general stationary observations. *Annals of Statistics* **17**: 1217–1241.

Leucht A, Neumann MH. 2013. Dependent wild bootstrap for degenerate $U$- and $V$-statistics. *Journal of Multivariate Analysis* **117**: 257–280.

Lifshits MA. 1984. Absolute continuity of functionals of "supremum" type for Gaussian processes. *Journal of the Soviet Mathematics* **27**: 3103–3112.

Liu RY, Singh K. 1992. Moving blocks jackknife and bootstrap capture weak dependence. In *Exploring the Limits of Bootstrap*, LePage R, Billard L (eds.) Wiley: New York, pp. 225–248.

Naik-Nimbalkar UV, Rajarshi MB. 1994. Validity of blockwise bootstrap for empirical processes with stationary observations. *Annals of Statistics* **22**: 980–994.

Paparoditis E, Politis DN. 2001. Tapered block bootstrap. *Biometrika* **88**: 1105–1119.

R Core Team. 2012. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. ISBN 3-900051-07-0.

Rio E. 1998. Processus empiriques absolument réguliers et entropie universelle. *Probability Theory and Related Fields* **111**: 585–608.

Rio E. 2000. *Théorie Asymptotique des Processus Aléatoires Faiblement Dépendants*, Mathématiques et Applications, vol. 31. Paris: Springer.

Shao X. 2010. The dependent wild bootstrap. *Journal of the American Statistical Association* **105**: 218–235.

Shao X. 2011. A bootstrap-assisted spectral test of white noise under unknown dependence. *Journal of Econometrics* **162**: 213–224.

Sharipov OSh, Wendler M. 2013. Normal limits, nonnormal limits, and the bootstrap for quantiles of dependent data. *Statistics and Probability Letters* **83**: 1028–1035.

Smeekes S, Urbain J-P. 2014. A multivariate invariance principle for modified wild bootstrap methods with an application to unit root testing. *GSBE Research Memorandum RM/14/008*, Maastricht University.

Sun S, Lahiri SN. 2006. Bootstrapping the sample quantile of a weakly dependent sequence. *Sankhya* **68**: 130–166.

Van der Vaart AW, Wellner JA. 2000. *Weak Convergence and Empirical Processes*, With Applications to Statistics. New York: Springer.

Viennet G. 1997. Inequalities for absolutely regular sequences: application to density estimation. *Probability Theory and Related Fields* **107**: 467–492.

Wieczorek B. 2014. Blockwise bootstrap of the estimated empirical process based on $\psi$-weakly dependent observations. Preprint.

Wu CFJ. 1986. Jackknife, bootstrap and other resampling methods in regression analysis (with discussion). *Annals of Statistics* **14**: 1261–1350.