

On the detection of changes in autoregressive time series, II. Resampling procedures¹

Marie Hušková¹), Claudia Kirch²), Zuzana Prášková³), Josef Steinebach⁴)

¹) Charles University of Prague, Department of Statistics, Sokolovská 83,
CZ-186 75 Praha 8, Czech Republic; huskova@karlin.mff.cuni.cz

²) Universität Kaiserslautern, Fachbereich Mathematik, Erwin-Schrödinger-Straße,
D-67 653 Kaiserslautern, Germany; ckirch@mathematik.uni-kl.de

³) Charles University of Prague, Department of Statistics, Sokolovská 83,
CZ-186 75 Praha 8, Czech Republic; praskova@karlin.mff.cuni.cz

⁴) Universität zu Köln, Mathematisches Institut, Weyertal 86-90,
D-50 931 Köln, Germany; jost@math.uni-koeln.de

Abstract: We study an autoregressive time series model with a possible change in the regression parameters. Approximations to the critical values for change-point tests are obtained through various bootstrapping methods. Theoretical results show that the bootstrapping procedures have the same limiting behavior as their asymptotic counterparts discussed in Hušková et al. [17]. In fact, a small simulation study illustrates that the bootstrap tests behave better than the original asymptotic tests if performance is measured by the α - and β -errors, respectively.

Mathematics Subject Classification 2000: Primary: 62G09
Secondary: 60F17, 62G20, 62M10

Keywords: Change point analysis, bootstrap, resampling, autoregressive time series

Acknowledgement. This work has been supported by grants GAČR 201/06/0186, PST.EAP.CLG980599 and MSM 0021620839.

¹This is a preprint of an article published in *J. Statist. Plann. Inference*, 138:1697-1721, 2008, Copyright 2008 by Elsevier.

1 Introduction

We consider the time series model with a change after an unknown time point m , i.e.

$$Y_i = Y_{i,n} = \mathbf{x}_i^T \boldsymbol{\beta} + \mathbf{x}_i^T \boldsymbol{\delta} I\{i > m\} + e_i, \quad i = p+1, \dots, n, \quad (1.1)$$

where

$$\mathbf{x}_i = \mathbf{x}_{i,n} = (Y_{i-1,n}, \dots, Y_{i-p,n})^T, \quad i = p+1, \dots, n,$$

$(p <) m = m_n (\leq n)$, $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^T$, $\boldsymbol{\delta} = \boldsymbol{\delta}_n = (\delta_{1n}, \dots, \delta_{pn})^T \neq \mathbf{0}$ are unknown parameters, and e_1, \dots, e_n are independent identically distributed (i.i.d.) random errors having a positive variance $\text{Var}(e_1) = \sigma^2$ and satisfying further conditions specified below.

The function $I\{A\}$ in (1.1) denotes the indicator of the set A . For the sake of convenience, we suppress the index n in the observations $Y_{i,n}$ as well as in the parameters m_n and $\boldsymbol{\delta}_n$ (and in variables depending on the latter) whenever possible, but keep in mind that in the limiting results below, as $n \rightarrow \infty$, both m_n and $\boldsymbol{\delta}_n$ may be changing when n is increasing.

The model (1.1) is a model for a change in autoregression. The parameter m is called the *change point*. The problem is to test

$$H_0 : m = n \quad \text{vs.} \quad H_1 : m < n.$$

In the first part [17] of this work the following statistics for our testing problem were introduced and their asymptotic properties, both under H_0 and H_1 , were studied:

$$T_n = \max_{p < k < n} \{ \mathbf{S}_k^T \mathbf{C}_k^{-1} \mathbf{C}_n (\mathbf{C}_k^0)^{-1} \mathbf{S}_k \} / \widehat{\sigma}_n^2, \quad (1.2)$$

$$T_n(\varepsilon) = \max_{n\varepsilon \leq k \leq n(1-\varepsilon)} \{ \mathbf{S}_k^T \mathbf{C}_k^{-1} \mathbf{C}_n (\mathbf{C}_k^0)^{-1} \mathbf{S}_k \} / \widehat{\sigma}_n^2, \quad (1.3)$$

$$T_n(q) = \sup_{0 < t < 1} \left\{ \frac{\mathbf{S}_{\lfloor (n+1)t \rfloor}^T \mathbf{C}_n^{-1} \mathbf{S}_{\lfloor (n+1)t \rfloor}}{q^2(t) \widehat{\sigma}_n^2} \right\}, \quad (1.4)$$

where $\varepsilon \in (0, \frac{1}{2})$ is fixed, $\lfloor a \rfloor$ denotes the integer part of a , and \mathbf{S}_k , $k = p+1, \dots, n$, are partial sums of weighted residuals, i.e.,

$$\mathbf{S}_k = \sum_{i=p+1}^k \mathbf{x}_i (Y_i - \mathbf{x}_i^T \widehat{\boldsymbol{\beta}}_n), \quad k = p+1, \dots, n, \quad (1.5)$$

with

$$\widehat{\boldsymbol{\beta}}_n = \mathbf{C}_n^{-1} \sum_{i=p+1}^n \mathbf{x}_i Y_i, \quad (1.6)$$

$$\mathbf{C}_k = \sum_{i=p+1}^k \mathbf{x}_i \mathbf{x}_i^T, \quad \mathbf{C}_k^0 = \mathbf{C}_n - \mathbf{C}_k, \quad k = p+1, \dots, n. \quad (1.7)$$

Note that $\widehat{\boldsymbol{\beta}}_n$ defined in (1.6) is the ordinary least squares estimator of the vector parameter $\boldsymbol{\beta}$ in the model (1.1) with $m = n$ (“no change”).

Furthermore $\widehat{\sigma}_n^2$ is a consistent estimator of σ^2 (under H_0), i.e.,

$$\widehat{\sigma}_n^2 \xrightarrow{P} \sigma^2 \quad (n \rightarrow \infty), \quad (1.8)$$

possibly satisfying some additional assumptions, and q is a positive weight function of the form

$$q(t) = q_\beta(t) = (t(1-t))^\beta, \quad t \in (0, 1), \quad \beta \in [0, 1/2]. \quad (1.9)$$

Remark 1.1. Similar to the discussion in Hušková et al. [17], we could allow for more general weight functions q , being positive on $(0, 1)$, nondecreasing in a neighborhood of 0, nonincreasing in a neighborhood of 1, with $\inf\{q(t); t \in (\eta, 1 - \eta)\} > 0$ for all $\eta \in (0, 1/2)$ and

$$\int_0^1 \frac{1}{s(1-s)} \exp\left\{-\frac{cq^2(s)}{s(1-s)}\right\} ds < \infty$$

for all $c > 0$, but in order to avoid technicalities in the proofs we confine ourselves to functions q as in (1.9).

Note that large values of either of the test statistics in (1.2)–(1.4) indicate that the null hypothesis is violated. Therefore the critical regions have the form:

$$T_n \geq t_n(\alpha), \quad (1.10)$$

where $t_n(\alpha)$ is determined in such a way that the test has level α . Approximations to critical values can be obtained either through the limit distribution of the considered test statistics under H_0 or through resampling. In Part I of this work [17], the first approach was studied, while we will now consider the possible application of different bootstrapping methods.

The paper is organized as follows. In the next section, for the sake of comparison below, we summarize the asymptotic distributions of the above test statistics under the null hypothesis. In Section 3, we introduce two types of bootstrap, namely the regression bootstrap and the pair bootstrap, to derive critical values. It will turn out that the bootstrap statistics, given the observations Y_1, \dots, Y_n , have the same limit

behavior under the null hypothesis as the original statistics. Under alternatives they converge in distribution to a similar but different limit distribution. Yet this still guarantees consistency of the corresponding test procedures. In Section 4, a small simulation study will be presented which indicates that the bootstrap tests are better than the corresponding asymptotic ones, if the performance is measured by α - and β -errors, respectively. The proofs of our results are given in Section 5.

2 Limiting distributions under H_0

Here we briefly summarize the assumptions and results of the first part [17] of our work. This is mainly for the sake of comparison, because it will turn out that the bootstrap statistics – at least under the null hypothesis – have the same asymptotic behavior.

We assume that the sequence $\{Y_i, i = p + 1, \dots, n\}$ satisfies either of the following assumptions:

(A.1) The observations Y_{p+1}, \dots, Y_n follow the model (1.1) with $m = n$; the initial values Y_1, \dots, Y_p are independent of e_{p+1}, \dots, e_n ; $\beta_p \neq 0$, and the roots of the polynomial $t^p - \beta_1 t^{p-1} - \dots - \beta_p$ are less than one in absolute value.

(A.2) The vector $\mathbf{x}_{p+1} = (Y_p, \dots, Y_1)^T$ of initial observations satisfies

$$\mathbf{x}_{p+1} = \sum_{j=0}^{\infty} \mathbf{B}^j \mathbf{e}_{p-j}, \quad (2.1)$$

where

$$\mathbf{B} = \begin{pmatrix} \beta_1, \dots, \beta_p \\ \mathbf{I}_{p-1} & \mathbf{0} \end{pmatrix} \quad \text{and} \quad \mathbf{e}_k = (e_k, 0, \dots, 0)^T, \quad (2.2)$$

with \mathbf{I}_{p-1} denoting the $(p-1)$ -dimensional unit matrix.

(A.3) The observations Y_{p+1}, \dots, Y_n follow the model (1.1) with $m = \lfloor n\theta \rfloor$, $\theta \in (0, 1)$; Y_1, \dots, Y_p are independent of e_{p+1}, \dots, e_n ; $\beta_p \neq 0$, the roots of the polynomial $t^p - \beta_1 t^{p-1} - \dots - \beta_p$ are less than one in absolute value, $\beta_p + \delta_p \neq 0$, and the roots of $t^p - (\beta_1 + \delta_1)t^{p-1} - \dots - (\beta_p + \delta_p)$ are also less than one in absolute value, $\boldsymbol{\delta} = (\delta_1, \dots, \delta_p)^T \neq \mathbf{0}$.

Note that (A.1) corresponds to H_0 , under which the observations form an autoregressive sequence, while (A.3) corresponds to an alternative model, in which the observations before some change follow an autoregressive model with parameter $\boldsymbol{\beta}$, and after the change they follow a different autoregressive model with parameter $\boldsymbol{\beta} + \boldsymbol{\delta}$ ($\boldsymbol{\delta} \neq \mathbf{0}$).

The distributions of the error terms $\{e_i\}$ satisfy either of the following assumptions:

(B.1) $\{e_i, i = 0, \pm 1, \dots\}$ are i.i.d. random variables having zero mean, positive variance σ^2 , finite moment $\mathbb{E}|e_i|^{4+\eta} < \infty$ for some $\eta > 0$, and a density f such that

$$\sup_{-\infty < a < \infty} \frac{1}{|a|} \int_{-\infty}^{\infty} |f(x+a) - f(x)| dx < \infty.$$

[This assumption has been used in [17] to show that the sequence $\{Y_i\}$ fulfills a strong mixing condition.]

(B.2) $\{e_i, i = 0, \pm 1, \dots\}$ are i.i.d. random variables having zero mean, positive variance σ^2 and finite fourth moment $\mathbb{E}|e_i|^4 < \infty$.

Next, we recall the main results from [17] on the limiting behavior of our test statistics under the null hypothesis.

Theorem 2.1. a) *Let assumptions (A.1), (A.2) and (B.1) be satisfied. Let $\hat{\sigma}_n^2$ be an estimator of σ^2 satisfying*

$$\hat{\sigma}_n^2 - \sigma^2 = o_{\mathbb{P}}((\log \log n)^{-1}) \quad (n \rightarrow \infty). \quad (2.3)$$

Then,

$$\lim_{n \rightarrow \infty} \mathbb{P}(a_n T_n^{1/2} - b_n \leq t) = \exp\{-2e^{-t}\}, \quad t \in \mathbb{R}, \quad (2.4)$$

where $a_n = a(\log n)$, $b_n = b_p(\log n)$, and

$$a(y) = (2 \log y)^{1/2}, \quad b_p(y) = 2 \log y + \frac{p}{2} \log \log y - \log \Gamma\left(\frac{p}{2}\right), \quad y > 1, \quad (2.5)$$

with $\Gamma(p) = \int_0^{\infty} t^{p-1} e^{-t} dt$, $p > 0$.

b) *Let assumptions (A.1), (A.2) and (B.2) be satisfied. Let $\hat{\sigma}_n^2$ be a consistent estimator of σ^2 , i.e., $\hat{\sigma}_n^2 \xrightarrow{\mathbb{P}} \sigma^2$ as $n \rightarrow \infty$. Then, for any $\varepsilon \in (0, 1/2)$,*

$$(T_n(\varepsilon))^{1/2} \xrightarrow{\mathbb{D}} \sup_{\varepsilon \leq t \leq 1-\varepsilon} \frac{\|\mathbf{B}(t)\|}{\sqrt{t(1-t)}} \quad (n \rightarrow \infty), \quad (2.6)$$

where $\|\cdot\|$ denotes the p -dimensional Euclidean norm and $\{\mathbf{B}(t), t \in [0, 1]\}$ is a p -dimensional (standard) Brownian bridge process, i.e., $\mathbf{B}(t) = (B_1(t), \dots, B_p(t))^T$ with independent (1-dimensional) Brownian bridges $\{B_j(t), t \in [0, 1]\}$.

c) *Under (1.9) and the assumptions of b), also*

$$(T_n(q))^{1/2} \xrightarrow{\mathbb{D}} \sup_{0 < t < 1} \frac{\|\mathbf{B}(t)\|}{q(t)} \quad (n \rightarrow \infty), \quad (2.7)$$

with $\{\mathbf{B}(t), t \in [0, 1]\}$ as in b).

The limit distribution in Theorem 2.1a) belongs to the extreme value types and it is known that the convergence is fairly slow. Although approximations to the critical values can be easily calculated, they are only reasonable for large n . The explicit form of the limit distribution in Theorem 2.1c) is only known for $q = q_0 \equiv 1$. Otherwise limit distributions have to be simulated. Therefore we are interested in alternative methods to derive critical values via bootstrapping procedures.

3 Bootstrapping

It appears that bootstrap procedures, both with and without replacement, provide reasonable approximations for the critical values of test statistics constructed to detect changes in location and linear regression models with *independent* observations (see, e.g., Antoch et al. [5], Antoch and Hušková [2],[3], Hušková [14], Hušková and Picek [15],[16]).

In the present section we discuss some bootstrap methods to get approximations for critical values related to the test statistics $T_n, T_n(\varepsilon)$ and $T_n(q)$ defined in (1.2), (1.3) and (1.4), respectively, i.e., for *dependent* data. There exist various methods to implement bootstrap procedures for time series data; for a comprehensive discussion see, e.g., Lahiri [20], Franke et al. [10] or Hidalgo [11]. In contrast to Kirch [18], who used blockwise resampling and resampling in the frequency domain to detect a change in a location model with errors generated by a linear process, we focus here on resampling methods based on residuals, which have been modified for our setup. More precisely, we propose a *regression bootstrap* and a *pair bootstrap* to determine critical values for the above test statistics.

We should point out that we wish to obtain an approximation for the null distribution function of the test statistics introduced in Section 1, however, we do not know whether the available data follow the null hypothesis or an alternative. Certain parameters are present under alternatives only. Therefore we have to adjust existing bootstrapping methods to our setup.

We shall prove that the limit distributions of our bootstrap statistics, given the observations Y_1, \dots, Y_n , are the same as for the original statistics, at least under the null hypothesis. This shows that the bootstrap test holds the chosen level asymptotically. Under alternatives the bootstrap statistics, given the observations, still converge in distribution, but in some cases to a somewhat different limit distribution. Since Theorem 3.1 of the first part [17] of this work shows that the original statistics converge (in a P-stochastic sense) to infinity under alternatives, the bootstrap tests are also consistent.

One can develop bootstrapping procedures based either on “residuals under H_0 ” or on “residuals taking possible changes into account”. Here are the residuals we use:

$$\widehat{e}_i = Y_i - \mathbf{x}_i^T \widehat{\boldsymbol{\beta}}_n, \quad i = p + 1, \dots, n, \quad (3.1)$$

and

$$\widetilde{e}_i = \begin{cases} \widehat{e}_i, & \min(\widehat{m}, n - \widehat{m}) \leq d_{1n}, \quad i = p + 1, \dots, n, \\ Y_i - \mathbf{x}_i^T \widehat{\boldsymbol{\beta}}_{\widehat{m}}, & \min(\widehat{m}, n - \widehat{m}) > d_{1n}, \quad i = p + 1, \dots, \widehat{m}, \\ Y_i - \mathbf{x}_i^T \widehat{\boldsymbol{\beta}}_{\widehat{m}}^0, & \min(\widehat{m}, n - \widehat{m}) > d_{1n}, \quad i = \widehat{m} + 1, \dots, n, \end{cases} \quad (3.2)$$

with a sequence $\{d_{1n}\}$ to be specified below, and with $\widehat{\boldsymbol{\beta}}_{\widehat{m}}$ and $\widehat{\boldsymbol{\beta}}_{\widehat{m}}^0$ denoting the ordinary least squares estimators of $\boldsymbol{\beta}$ based on $Y_1, \dots, Y_{\widehat{m}}$ and $Y_{\widehat{m}+1}, \dots, Y_n$, respectively. Furthermore, \widehat{m} is supposed to be an *integer-valued* estimator of m satisfying

$$\widehat{m} - m = o_{\mathbb{P}}(d_{2n}) \quad (n \rightarrow \infty), \quad (3.3)$$

under H_1 , for a sequence $\{d_{2n}\}$ such that

$$d_{2n} \rightarrow \infty, \quad d_{2n}/n \rightarrow 0 \quad (n \rightarrow \infty). \quad (3.4)$$

Proposition 3.1 in the first part [17] of this work gives some basic limit properties for the estimators

$$\hat{m}_1 = \min \{k \in (p, n) : \mathbf{S}_k^T \mathbf{C}_k^{-1} \mathbf{C}_n (\mathbf{C}_k^0)^{-1} \mathbf{S}_k = \max_{p < \ell < n} \mathbf{S}_\ell^T \mathbf{C}_\ell^{-1} \mathbf{C}_n (\mathbf{C}_\ell^0)^{-1} \mathbf{S}_\ell\} \quad (3.5)$$

and

$$\hat{m}_2 = \min \{k \in (p, n) : \mathbf{S}_k^T \mathbf{C}_n^{-1} \mathbf{S}_k = \max_{p < \ell < n} \mathbf{S}_\ell^T \mathbf{C}_n^{-1} \mathbf{S}_\ell\}. \quad (3.6)$$

Under (A.2), (A.3) and (B.1), both estimators have the property (3.3) for any sequence $\{d_{2n}\}$ satisfying (3.4).

The quantities \hat{e}_i , $i = p + 1, \dots, n$, are the residuals under H_0 , while the residuals \tilde{e}_i , $i = p + 1, \dots, n$, take a possible change into account. The latter residuals are more convenient for our situation and will be used in the following two subsections. The former ones can also be taken, but one possibly loses some power.

To develop proper bootstrap versions of our test statistics it is useful to note that, under H_0 ,

$$\mathbf{S}_k = \sum_{i=p+1}^k \mathbf{x}_i e_i - \mathbf{C}_k \mathbf{C}_n^{-1} \sum_{j=p+1}^n \mathbf{x}_j e_j, \quad k = p + 1, \dots, n, \quad (3.7)$$

$$\mathbf{S}_n = \mathbf{0}, \quad (3.8)$$

and also that, by Lemma 4.5 in [17], under H_0 , the sequence of processes

$$\{\mathbf{C}_n^{-1/2} \mathbf{V}_n(t), t \in [0, 1]\}, \quad n = p + 1, \dots,$$

with

$$\mathbf{V}_n(t) = \frac{1}{\tilde{\sigma}_n \sqrt{n}} \sum_{i=p+1}^{\lfloor nt \rfloor} \mathbf{x}_i e_i, \quad k = p + 1, \dots, n, \quad (3.9)$$

$$\tilde{\sigma}_n^2 = \frac{1}{n-p} \sum_{i=p+1}^n (\tilde{e}_i - \bar{\tilde{e}}_n)^2, \quad (3.10)$$

$$\bar{\tilde{e}}_n = \frac{1}{n-p} \sum_{i=p+1}^n \tilde{e}_i, \quad (3.11)$$

converges in $D[0, 1]$ to a p -dimensional Wiener process $\{\mathbf{W}(t), t \in [0, 1]\}$ with independent components, i.e., $\mathbf{W}(t) = (W_1(t), \dots, W_p(t))^T$ with independent (1-dimensional) standard Wiener processes $\{W_j(t), t \in [0, 1]\}$. The test statistics T_n , $T_n(\varepsilon)$ and $T_n(q)$ are functionals of the process $\{\mathbf{V}_n(t), t \in [0, 1]\}$.

We shall study a bootstrap with replacement. However, one can also consider a bootstrap without replacement. It has the same limit performance as the respective bootstrap with replacement, but the proofs of the results become somewhat more technical.

In the sequel \mathbf{P}^* , \mathbf{E}^* , \mathbf{Var}^* , and \mathbf{Cov}^* denote the conditional probability, expectation, variance, and covariance, respectively, given the observations Y_1, \dots, Y_n . The indices R or P are used to distinguish between *regression* bootstrap and *pair* bootstrap.

The proofs of the results of the following two subsections are postponed to Section 5.

3.1 Regression bootstrap

The idea of the regression bootstrap relies on a similarity of the considered model (1.1) to the linear regression model with given design points \mathbf{x}_i , $i = p + 1, \dots, n$ (see, e.g., Csörgő and Horváth [7] or Antoch and Hušková [3] among others). Then one works with the bootstrap sample with replacement from $\tilde{e}_{p+1}, \dots, \tilde{e}_n$. Denote such a sample by $\tilde{e}_{p+1}^*, \dots, \tilde{e}_n^*$.

This leads to the regression bootstrap versions $\mathbf{S}_{k,R}^*$ of \mathbf{S}_k , $k = p + 1, \dots, n$, defined as

$$\mathbf{S}_{k,R}^* = \sum_{i=p+1}^k \mathbf{x}_i(\tilde{e}_i^* - \bar{\tilde{e}}_n) - \mathbf{C}_k \mathbf{C}_n^{-1} \sum_{j=p+1}^n \mathbf{x}_j(\tilde{e}_{jn}^* - \bar{\tilde{e}}_n), \quad k = p + 1, \dots, n, \quad (3.12)$$

where $\bar{\tilde{e}}_n$ is given in (3.11).

The corresponding bootstrap versions of the considered test statistics are defined as the respective test statistics with \mathbf{S}_k being replaced by $\mathbf{S}_{k,R}^*$, $k = p + 1, \dots, n$, e.g., the bootstrap version of T_n is defined as

$$T_{n,R}^* = \max_{p < k < n} \{ \mathbf{S}_{k,R}^{*T} \mathbf{C}_k^{-1} \mathbf{C}_n (\mathbf{C}_k^0)^{-1} \mathbf{S}_{k,R}^* \} / \tilde{\sigma}_n^2, \quad (3.13)$$

and similarly for the other test statistics.

Next we describe the conditional limit behavior of the bootstrap version of the test statistic T_n .

Theorem 3.1. *Let assumption (A.1) or (A.3) in combination with (A.2), (B.2) and (2.3) be satisfied, and let \hat{m} be either of the estimators defined in (3.5) and (3.6). Then, as $n \rightarrow \infty$,*

$$\mathbf{P}^* (a_n (T_{n,R}^*)^{1/2} - b_n \leq t) \xrightarrow{\mathbf{P}} \exp\{-2e^{-t}\}, \quad t \in \mathbb{R}, \quad (3.14)$$

with a_n and b_n as defined in Theorem 2.1.

Before stating assertions on the limit behavior of $T_n(\varepsilon)$ and $T_n(q)$, we investigate the process

$$\mathbf{V}_{n,R}^*(t) = \frac{1}{\tilde{\sigma}_n \sqrt{n}} \sum_{i=p+1}^{\lfloor nt \rfloor} \mathbf{x}_i(\tilde{e}_i^* - \tilde{e}_n), \quad t \in [0, 1], \quad (3.15)$$

with $\tilde{\sigma}_n^2$ from (3.10). The process $\{\mathbf{V}_{n,R}^*(t), t \in [0, 1]\}$, given Y_1, \dots, Y_n , is a process with

$$\mathbf{E}^* \mathbf{V}_{n,R}^*(t) = \mathbf{0}, \quad t \in [0, 1], \quad (3.16)$$

$$\text{Cov}^*(\mathbf{V}_{n,R}^*(t_1), \mathbf{V}_{n,R}^*(t_2)) = \frac{1}{n} \mathbf{C}_{\lfloor nt_1 \rfloor}, \quad 0 \leq t_1 \leq t_2 \leq 1. \quad (3.17)$$

Recall that $m = \lfloor n\theta \rfloor$, $\theta \in (0, 1]$, where $\theta \in (0, 1)$ corresponds to alternatives while $\theta = 1$ does to the null hypothesis. From a functional central limit theorem for processes with independent increments together with the properties of the matrices \mathbf{C}_k we have:

Theorem 3.2. *Let assumption (A.1) or (A.3) in combination with (A.2) and (1.8) be satisfied, and let \hat{m} be either of the estimators defined in (3.5) and (3.6). Then, as $n \rightarrow \infty$,*

$$\mathbf{P}^*(f(\mathbf{V}_{n,R}^*(\cdot)) \leq x) \xrightarrow{\mathbf{P}} \mathbf{P}(f(\mathbf{V}_\theta(\cdot)) \leq x), \quad (3.18)$$

for any $x \in \mathbb{R}$ and any continuous, real-valued f on $D[0, 1]$, where $\{\mathbf{V}_\theta(t), t \in [0, 1]\}$ is a Gaussian process with zero mean, covariance

$$\text{Cov}(\mathbf{V}_\theta(t_1), \mathbf{V}_\theta(t_2)) = \mathbf{Q}_{t_1}, \quad 0 \leq t_1 \leq t_2 \leq 1, \quad (3.19)$$

and

$$\mathbf{C} \stackrel{\mathbf{P}}{=} \lim_{m \rightarrow \infty} \frac{1}{m} \mathbf{C}_m, \quad \mathbf{C}^0 \stackrel{\mathbf{P}}{=} \lim_{n-m \rightarrow \infty} \frac{1}{n-m} \mathbf{C}_m^0, \quad (3.20)$$

$$\mathbf{Q}_t = \min(t, \theta) \mathbf{C} + (t - \theta) I\{t > \theta\} \mathbf{C}^0, \quad t \in [0, 1]. \quad (3.21)$$

This theorem in combination with the properties of the matrices \mathbf{C}_k implies:

Theorem 3.3. *Let assumption (A.1) or (A.3) in combination with (A.2), (B.2) and (1.8) be satisfied, and let \hat{m} be either of the estimators defined in (3.5) and (3.6). Set*

$$\mathbf{S}_\theta(t) = \mathbf{V}_\theta(t) - \mathbf{Q}_t \mathbf{Q}_1^{-1} \mathbf{V}_\theta(1), \quad t \in [0, 1].$$

Then, as $n \rightarrow \infty$, for any $x \in \mathbb{R}$,

a)

$$\mathbf{P}^*(T_{n,R}^*(\varepsilon) \leq x) \xrightarrow{\mathbf{P}} \mathbf{P}\left(\max_{\varepsilon \leq t \leq 1-\varepsilon} \mathbf{S}_\theta^T(t) \mathbf{Q}_t^{-1} \mathbf{Q}_1 (\mathbf{Q}_t^0)^{-1} \mathbf{S}_\theta(t) \leq x\right), \quad (3.22)$$

with $\mathbf{Q}_t^0 = \mathbf{Q}_1 - \mathbf{Q}_t$.

b)

$$\mathbb{P}^*(T_{n,R}^*(q) \leq x) \xrightarrow{\mathbb{P}} \mathbb{P}\left(\max_{0 < t < 1} \frac{\mathbf{S}_\theta^T(t) \mathbf{Q}_t^{-1} \mathbf{S}_\theta(t)}{q^2(t)} \leq x\right). \quad (3.23)$$

Clearly, in view of Theorem 3.1, the regression bootstrap “works” for the statistic T_n , i.e., the approximation to critical values based on the regression bootstrap are asymptotically correct, no matter whether the observations follow the null hypothesis or some alternative. However, by Theorems 2.1 and 3.3, the regression bootstrap for $T_n(\varepsilon)$ and $T_n(q)$ only has the correct asymptotics, when the observations follow the null hypothesis or some *local* alternative. It converges to a different limit distribution, when the data follow a *fixed* alternative. This is due to the fact that, while under H_0 or local alternatives,

$$\mathbf{Q}_t = t\mathbf{C}, \quad \mathbf{Q}_t^{-1} \mathbf{Q}_1 (\mathbf{Q}_t^0)^{-1} = (t(1-t))^{-1} \mathbf{C}^{-1}, \quad t \in [0, 1],$$

the latter is not true under fixed alternatives. For details we refer to Lemma 4.2 in [17]. Nevertheless, as mentioned earlier, the corresponding bootstrap test is still consistent.

Remark 3.1. One can propose modifications of the above test statistics such that the bootstrap also works under fixed alternatives, however, the procedure becomes more involved. For example, let us define

$$\mathbf{V}_{n,R}^{**}(t) = \begin{cases} \frac{1}{\tilde{\sigma}_n} \mathbf{C}_{\hat{m}}^{-1/2} \sqrt{\frac{\hat{m}}{n}} \sum_{i=p+1}^{\lfloor nt \rfloor} \mathbf{x}_i (\tilde{e}_i^* - \tilde{e}_n), & 0 \leq t \leq \frac{\hat{m}}{n}, \\ \frac{1}{\tilde{\sigma}_n} \mathbf{C}_{\hat{m}}^{-1/2} \sqrt{\frac{\hat{m}}{n}} \sum_{i=p+1}^{\hat{m}} \mathbf{x}_i (\tilde{e}_i^* - \tilde{e}_n) \\ \quad + \frac{1}{\tilde{\sigma}_n} (\mathbf{C}_{\hat{m}}^0)^{-1/2} \sqrt{\frac{n-\hat{m}}{n}} \sum_{i=\hat{m}+1}^{\lfloor nt \rfloor} \mathbf{x}_i (\tilde{e}_i^* - \tilde{e}_n), & \frac{\hat{m}}{n} < t \leq 1, \end{cases} \quad (3.24)$$

where \hat{m} is either of the estimators in (3.5) or (3.6). Clearly, for given Y_1, \dots, Y_n , $\{\mathbf{V}_{n,R}^{**}(t), t \in [0, 1]\}$ is a process with

$$\begin{aligned} \mathbf{E}^* \mathbf{V}_{n,R}^{**}(t) &= \mathbf{0}, \quad t \in [0, 1], \\ \text{Var}^*(\mathbf{V}_{n,R}^{**}(t)) &= \begin{cases} \frac{\hat{m}}{n} \mathbf{C}_{\hat{m}}^{-1/2} \mathbf{C}_{\lfloor nt \rfloor} \mathbf{C}_{\hat{m}}^{-1/2}, & 0 \leq t \leq \frac{\hat{m}}{n}, \\ \frac{\hat{m}}{n} \mathbf{I} + \frac{n-\hat{m}}{n} (\mathbf{C}_{\hat{m}}^0)^{-1/2} (\mathbf{C}_{\lfloor nt \rfloor} - \mathbf{C}_{\hat{m}}) (\mathbf{C}_{\hat{m}}^0)^{-1/2}, & \frac{\hat{m}}{n} < t \leq 1. \end{cases} \end{aligned} \quad (3.25)$$

The corresponding regression bootstrap test statistics have the form:

$$\begin{aligned} T_{n,R}^{**}(\varepsilon) &= \max_{\varepsilon \leq t \leq 1-\varepsilon} \left\{ \frac{\|\mathbf{V}_{n,R}^{**}(t) - t\mathbf{V}_{n,R}^{**}(1)\|^2}{t(1-t)} \right\}, \\ T_{n,R}^{**}(q) &= \max_{0 < t < 1} \left\{ \frac{\|\mathbf{V}_{n,R}^{**}(t) - t\mathbf{V}_{n,R}^{**}(1)\|^2}{q^2(t)} \right\}. \end{aligned}$$

These bootstrap versions provide asymptotically correct approximations for the limit distributions of $T_n(\varepsilon)$ and $T_n(q)$, regardless whether the data follow the null or alternative hypotheses.

3.2 Pair bootstrap

This bootstrap was introduced by Künsch [19] (see also Shao and Tu [22] for more details). We take a bootstrap sample with replacement from the pairs $(\mathbf{x}_i, \tilde{e}_i)$, $i = p + 1, \dots, n$, and denote it by $(\mathbf{x}_i^*, \tilde{e}_i^*)$, $i = p + 1, \dots, n$. Then the corresponding bootstrap partial sums are defined as

$$\mathbf{S}_{k,P}^* = \mathbf{V}_{n,P}^*(k/n) - \mathbf{C}_k^* \mathbf{C}_n^{*-1} \mathbf{V}_{n,P}^*(1), \quad k = p + 1, \dots, n, \quad (3.26)$$

where

$$\mathbf{C}_k^* = \sum_{i=p+1}^k \mathbf{x}_i^* \mathbf{x}_i^{*T} + a_n \mathbf{I}(I\{p < k \leq a_n\} + I\{n - a_n \leq k \leq n\}), \quad k = p + 1, \dots, n, \quad (3.27)$$

with $a_n = (\log \log n)^{\frac{1}{2} - \kappa}$ for some $0 < \kappa < \frac{1}{2}$, and

$$\mathbf{V}_{n,P}^*(t) = \frac{1}{\tilde{\sigma}_n} \sum_{i=p+1}^{\lfloor nt \rfloor} \mathbf{x}_i^* (\tilde{e}_i^* - \tilde{e}_n), \quad 0 \leq t \leq 1. \quad (3.28)$$

Note that $\mathbf{V}_{n,P}^*(k/n)$, $k = p + 1, \dots, n$, are sums of i.i.d. random vectors. By direct calculations we have

$$\mathbb{E}^* \mathbf{V}_{n,P}^*(k/n) = \frac{1}{\tilde{\sigma}_n} \frac{k}{n-p} \sum_{i=p+1}^n \mathbf{x}_i^* \tilde{e}_n = \frac{k}{n} \mathcal{O}_{\mathbb{P}}(1), \quad p < k \leq n,$$

$$\text{Var}^*(\mathbf{V}_{n,P}^*(k/n)) = k \hat{\Sigma}_n, \quad k = p + 1, \dots, n,$$

with

$$\hat{\Sigma}_n = \frac{1}{\tilde{\sigma}_n^2} \frac{1}{n-p} \sum_{i=p+1}^n \mathbf{x}_i \mathbf{x}_i^T (\tilde{e}_i - \tilde{e}_n)^2. \quad (3.29)$$

Also

$$\mathbb{E}^* \mathbf{C}_k^* = \frac{k-p}{n-p} \mathbf{C}_n, \quad k = p + 1, \dots, n. \quad (3.30)$$

The pair bootstrap version of T_n is defined as

$$T_{n,P}^* = \max_{p < k < n} \{ \mathbf{S}_{k,P}^{*T} (\mathbf{C}_k^*)^{-1} \mathbf{C}_n^* (\mathbf{C}_k^{0*})^{-1} \mathbf{S}_{k,P}^* \},$$

and $T_{n,P}^*(\varepsilon)$ and $T_{n,P}^*(q)$ are defined accordingly. All three bootstrap versions of the test statistics have the same limit distributions, given the observations Y_1, \dots, Y_n , as the respective test statistics under H_0 . Hence they provide suitable approximations to the critical values.

Theorem 3.4. *Let assumption (A.1) or (A.3) in combination with (A.2), (B.2) and (2.3) be satisfied, and let \hat{m} be either of the estimators defined in (3.5) and (3.6). Then, as $n \rightarrow \infty$,*

$$\mathbf{P}^* \left(a_n (T_{n,P}^*)^{1/2} - b_n \leq t \right) \xrightarrow{\mathbf{P}} \exp(-2e^{-t}), \quad t \in \mathbb{R}, \quad (3.31)$$

with a_n and b_n as given in Theorem 2.1.

Theorem 3.5. *Let assumption (A.1) or (A.3) in combination with (A.2), (B.2) and (1.8) be satisfied, and let \hat{m} be either of the estimators defined in (3.5) and (3.6). Then, as $n \rightarrow \infty$, for any $x \in \mathbb{R}$,*

a)

$$\mathbf{P}^* \left((T_{n,P}^*(\varepsilon))^{1/2} \leq x \right) \xrightarrow{\mathbf{P}} \mathbf{P} \left(\sup_{\varepsilon \leq t \leq 1-\varepsilon} \frac{\|\mathbf{B}(t)\|}{\sqrt{t(1-t)}} \leq x \right); \quad (3.32)$$

b)

$$\mathbf{P}^* \left((T_{n,P}^*(q))^{1/2} \leq x \right) \xrightarrow{\mathbf{P}} \mathbf{P} \left(\sup_{0 < t < 1} \frac{\|\mathbf{B}(t)\|}{q(t)} \leq x \right). \quad (3.33)$$

Remark 3.2. One can introduce simplified pair bootstrap test statistics based on

$$\mathbf{S}_{k,P}^{**} = \mathbf{V}_{n,P}^*(k/n) - \frac{k}{n} \mathbf{V}_{n,P}^*(1), \quad p < k \leq n,$$

and define

$$\begin{aligned} T_{n,P}^{**} &= \max_{p < k < n} \left\{ \frac{1}{\hat{\sigma}_n^2} \frac{n^2}{k(n-k)} \mathbf{S}_{k,P}^{**T} (\mathbf{C}_n^*)^{-1} \mathbf{S}_{k,P}^{**} \right\}, \\ T_{n,P}^{**}(\varepsilon) &= \max_{n\varepsilon < k < n(1-\varepsilon)} \left\{ \frac{1}{\hat{\sigma}_n^2} \frac{n^2}{k(n-k)} \mathbf{S}_{k,P}^{**T} (\mathbf{C}_n^*)^{-1} \mathbf{S}_{k,P}^{**} \right\}, \\ T_{n,P}^{**}(q) &= \max_{p < k < n} \left\{ \frac{1}{q^2(k/n) \hat{\sigma}_n^2} \mathbf{S}_{k,P}^{**T} (\mathbf{C}_n^*)^{-1} \mathbf{S}_{k,P}^{**} \right\}. \end{aligned}$$

These bootstrap statistics also have the same limit distributions, given the observations, as the respective test statistics under H_0 .

4 Simulations

In this section we present a simulation study to illustrate the performance of the bootstrapping methods for small sample sizes n and to compare them with the asymptotic tests. Let $q = q_0 \equiv 1$ and $\varepsilon = \varepsilon_0 = 0.1$. We take \hat{m}_2 from (3.6) as an estimator for the unknown change-point m .

statistic	90%	95%	97.5%
$T_n^{1/2}$	2.689	3.013	3.199
$(T_n(\varepsilon_0))^{1/2}$	2.655	2.862	3.119
$(T_n(q_0))^{1/2}$	1.177	1.334	1.466

Table 1: Simulated quantiles under the null hypothesis (based on 1000 rep.)

statistic	90%	95%	97.5%
$T_n^{1/2}$	3.265	3.659	4.045
$(T_n(\varepsilon_0))^{1/2}$	2.81	3.073	3.312
$(T_n(q_0))^{1/2}$	1.224	1.358	1.48

Table 2: Asymptotic quantiles under the null hypothesis, $n = 200$

The model is an AR(1)-model of length 200 with a parameter change at 100, $\beta = 0.3$, $b := \beta + \delta = 0.3, 0.6, 0.8$, and with standard normally distributed errors. Note that $b = 0.3$ corresponds to the null hypothesis H_0 . Using the corresponding negative parameter values turned out to give similar results.

Table 1 shows the simulated quantiles for the original statistics under the null hypothesis based on 1000 repetitions. The critical values of the limit distribution can be found in Table 2. Here and in the sequel we use the approximation of the distribution function of $\sup_{\varepsilon_0 \leq t \leq 1-\varepsilon_0} \{B^2(t)/(t(1-t))\}$ given in (3.6) of Antoch et al. [4]. The tables show that the asymptotic critical values are too conservative in all cases.

Furthermore we simulated quantiles of the bootstrap statistics based on one (fixed) underlying sequence of errors, which result in three samples of observations ($b = 0.3, 0.6, 0.8$). To calculate the bootstrap statistic we use 1000 random bootstrap samples of the above observations. The results can be found in Tables 3 and 4, respectively. The values turn out to provide better approximations of the actual null quantiles than the asymptotic critical values. Furthermore the values are quite stable for different alternatives in case of the *regression* bootstrap. In case of the *pair* bootstrap they become somewhat larger the greater the parameter change is. Yet this does not seem to influence the power of the test greatly as the size-power curves indicate below.

To get a better idea of how well the bootstrap quantiles match the null quantiles we created QQ-plots of the one against the other. Again we used 1000 simulations to approximate the null quantiles and 1000 bootstrap samples (of one fixed underlying sequence of observations) to obtain an approximation of the bootstrap statistic given these observations. Observations under different alternatives are based on the same underlying error sequence for the sake of comparison. We indicate the 90%-, 95%-, and 97.5%-quantiles to point out the relevant area of the QQ-plots. The plots can be found in Figures 1 and 2, respectively. Again the plots show a very good match of the actual null quantiles with the bootstrap quantiles, especially for the *regression* bootstrap. For the *pair* bootstrap the quality of the QQ-plot depends on whether the observations follow the null hypothesis or some alternative. Note that deviations in the upper tail just occur for a very small percentage of the total sample of 1000 simulated values of the test statistics.

Finally we demonstrate the actual level and the power of the tests based on boot-

statistic	$\beta + \delta$	90%	95%	97.5%
$(T_{nR}^*)^{1/2}$	0.3	2.761	2.967	3.199
$(T_{nR}^*)^{1/2}$	0.6	2.721	2.984	3.18
$(T_{nR}^*)^{1/2}$	0.8	2.685	2.957	3.215
$(T_{nR}^*(\varepsilon_0))^{1/2}$	0.3	2.592	2.897	3.053
$(T_{nR}^*(\varepsilon_0))^{1/2}$	0.6	2.575	2.879	3.043
$(T_{nR}^*(\varepsilon_0))^{1/2}$	0.8	2.525	2.81	3.047
$(T_{nR}^*(q_0))^{1/2}$	0.3	1.164	1.295	1.414
$(T_{nR}^*(q_0))^{1/2}$	0.6	1.155	1.304	1.414
$(T_{nR}^*(q_0))^{1/2}$	0.8	1.131	1.27	1.417

Table 3: Simulated quantiles of the regression bootstrap (based on 1000 bootstrap samples each for the same sequence of observations)

statistic	$\beta + \delta$	90%	95%	97.5%
$(T_{nP}^*)^{1/2}$	0.3	2.705	2.981	3.215
$(T_{nP}^*)^{1/2}$	0.6	2.864	3.12	3.371
$(T_{nP}^*)^{1/2}$	0.8	3.03	3.306	3.591
$(T_{nP}^*(\varepsilon_0))^{1/2}$	0.3	2.582	2.817	3.053
$(T_{nP}^*(\varepsilon_0))^{1/2}$	0.6	2.723	2.977	3.244
$(T_{nP}^*(\varepsilon_0))^{1/2}$	0.8	2.922	3.175	3.385
$(T_{nP}^*(q_0))^{1/2}$	0.3	1.146	1.265	1.388
$(T_{nP}^*(q_0))^{1/2}$	0.6	1.181	1.302	1.437
$(T_{nP}^*(q_0))^{1/2}$	0.8	1.282	1.396	1.542

Table 4: Simulated quantiles for the pair bootstrap (based on 1000 bootstrap samples for the same sequence of observations)

strapped quantiles respectively asymptotic critical values. For doing so, we created plots of the empirical distribution function of the p -values of the statistic of 1000 simulated time series according to the model with respect to the bootstrap respectively asymptotic distribution. The bootstrap distribution is calculated for each of these time series based on 1000 bootstrap samples.

What we get is a plot that shows the actual α -errors respectively $1 - (\beta$ -errors), i.e. empirical size and power, on the y -axis, for the chosen quantiles on the x -axis. So, the graph for the null hypothesis should be close to the diagonal (which is given by the dotted line), and for the alternatives it should be as steep as possible. The results can be found in Figures 3 and 4, respectively, and are denoted size-power curves (SPC), which differ from size-power trade-off curves as e.g. in Horváth et al. [12]. Note that, from bottom to top, the curves correspond to the null hypothesis ($b = 0.3$), to the alternative $b = 0.6$, and to $b = 0.8$, both for the asymptotic (dashed line) as well as for the bootstrap test (solid line). The plots show the empirical size and power of the test for any level between $(0, 0.1)$, for $\alpha = 0.05$ one obtains, e.g., the values in Table 5.

For all statistics the bootstrap tests work better than the asymptotic tests, the latter possessing too low levels. For example, given a nominal level of 10%, the asymptotic test has an actual level of approximately 5% for statistic T_n .

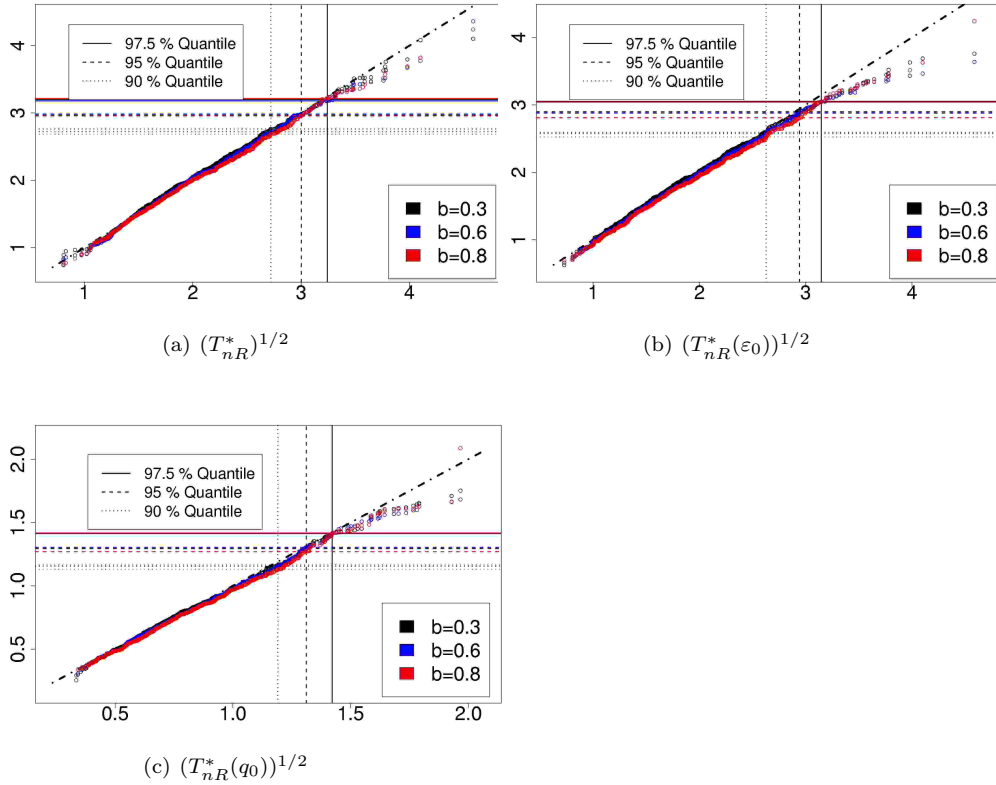


Figure 1: QQ-Plots for the regression bootstrap based on 1000 samples

statistic	asymptotic			bootstrap		
	$b = 0.3$	$b = 0.6$	$b = 0.8$	$b = 0.3$	$b = 0.6$	$b = 0.8$
Regression bootstrap						
$(T_{nR}^*)^{1/2}$	0.004	0.175	0.809	0.049	0.425	0.947
$(T_{nR}^*(\epsilon_0))^{1/2}$	0.028	0.391	0.938	0.051	0.459	0.954
$(T_{nR}^*(q_0))^{1/2}$	0.042	0.479	0.952	0.058	0.528	0.965
Pair bootstrap						
$(T_{nR}^*)^{1/2}$	0.004	0.156	0.825	0.05	0.419	0.944
$(T_{nR}^*(\epsilon_0))^{1/2}$	0.038	0.415	0.942	0.058	0.516	0.965
$(T_{nR}^*(q_0))^{1/2}$	0.04	0.486	0.956	0.074	0.559	0.965

Table 5: Empirical level and power for nominal level $\alpha = 0.05$

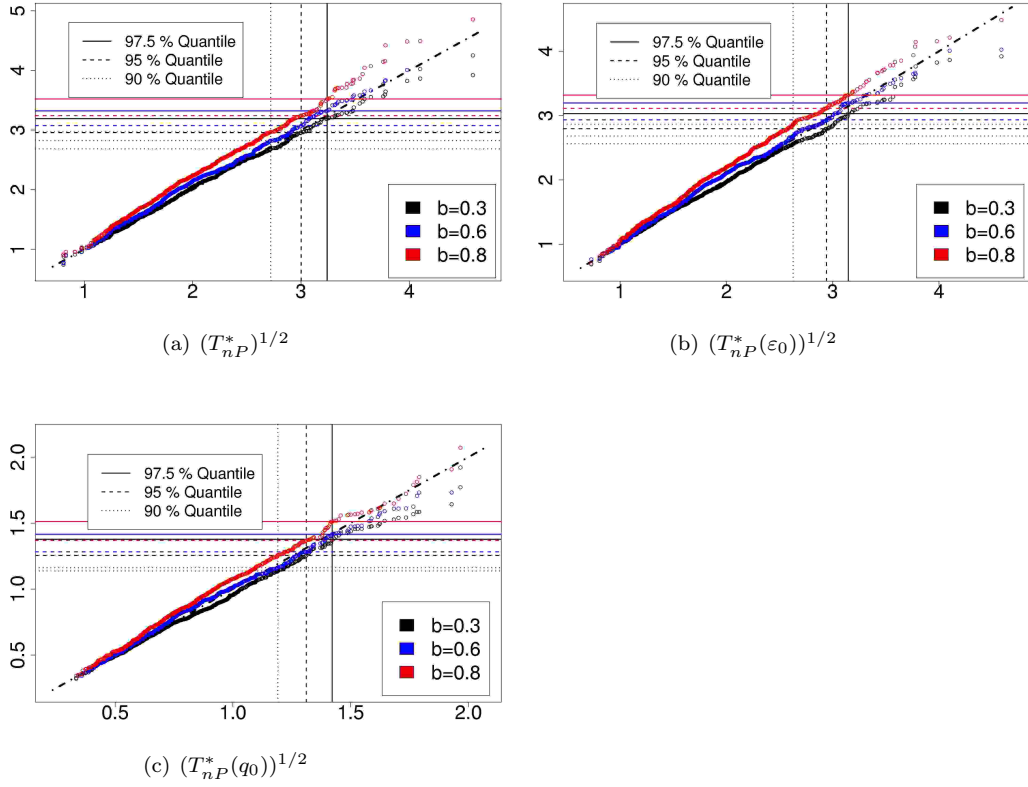


Figure 2: QQ-Plots for the pair bootstrap based on 1000 samples

5 Proofs

We start this section with a lemma, which will be useful for both types of bootstrap procedures.

Lemma 5.1. *Let $\tilde{\varepsilon}_i$, $i = p + 1, \dots, n$, be the residuals defined in (3.2) with $\{d_{1n}\}$ satisfying $1 \leq d_{1n} \leq n$ and $d_{1n}/n \rightarrow \infty$, as $n \rightarrow \infty$, and let $\tilde{\varepsilon}_n$ and $\tilde{\sigma}_n^2$ be as in (3.10) and (3.11), respectively. Choose \hat{m} as either of the estimators in (3.5),(3.6). Then*

(i) *under (A.1), (A.2) and (B.2),*

$$\tilde{\varepsilon}_n = \mathcal{O}_{\mathbb{P}}(n^{-\frac{1}{2}}), \quad \tilde{\sigma}_n^2 = \sigma^2 + \mathcal{O}_{\mathbb{P}}(n^{-\frac{1}{2}}) \quad (5.1)$$

(independently of d_{1n});

(ii) *under (A.2), (A.3) and (B.2), with $\delta \neq \mathbf{0}$ fixed, for any $0 < \lambda < \frac{1}{4}$,*

$$\tilde{\varepsilon}_n = \mathcal{O}_{\mathbb{P}}(n^{-\frac{1}{2}} + n^{-\frac{3}{2}}d_{2n}^{1-\lambda} + n^{-\lambda-1}) \quad (5.2)$$

$$\tilde{\sigma}_n^2 = \sigma^2 + \mathcal{O}_{\mathbb{P}}(n^{-\frac{1}{2}} + d_{2n}n^{-1}) \quad (5.3)$$

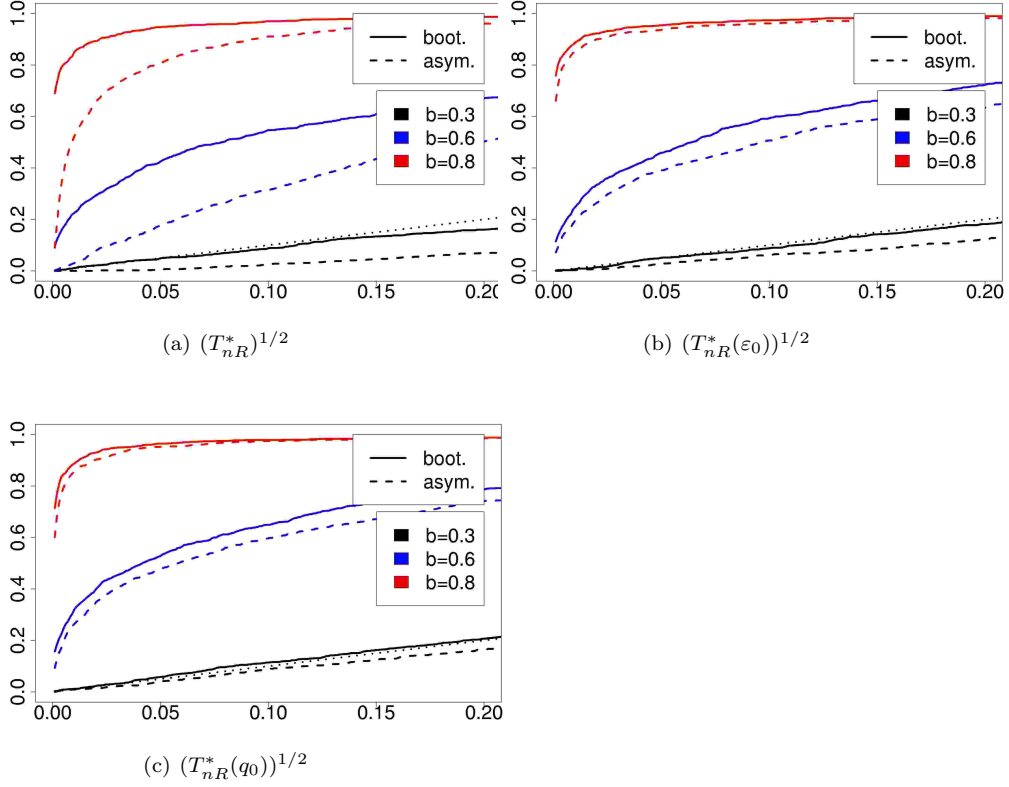


Figure 3: SPC-Plots for the regression bootstrap based on 1000 time series
 (curves correspond to $b = 0.3$ (H_0), $b = 0.6$, $b = 0.8$ from bottom to top for asymptotic
 (dashed) as well as bootstrap test (solid line))

for any sequence $\{d_{2n}\}$ satisfying (3.4).

Proof. We will not present here all technical details of the proof. Let us only notice that, for $\min(\widehat{m}, n - \widehat{m}) \leq d_{1n}$,

$$\tilde{e}_i = e_i - \mathbf{x}_i^T \mathbf{C}_n^{-1} \sum_{j=p+1}^n \mathbf{x}_j e_j + \mathbf{x}_i^T \delta I\{i > n\} - \mathbf{x}_i^T \mathbf{C}_n^{-1} \mathbf{C}_m^0 \delta, \quad i = p+1, \dots, n, \quad (5.4)$$

and, for $\min(\widehat{m}, n - \widehat{m}) > d_{1n}$, if we assume $\widehat{m} \leq m$ without loss of generality,

$$\tilde{e}_i = \begin{cases} e_i - \mathbf{x}_i^T \mathbf{C}_{\widehat{m}}^{-1} \sum_{j=p+1}^{\widehat{m}} \mathbf{x}_j e_j, & i \leq \widehat{m}, \\ e_i - \mathbf{x}_i^T (\mathbf{C}_{\widehat{m}}^0)^{-1} \sum_{j=\widehat{m}+1}^n \mathbf{x}_j e_j \\ \quad - \mathbf{x}_i^T (\mathbf{I} I\{\widehat{m} < i \leq m\} + (\mathbf{C}_{\widehat{m}}^0)^{-1} (\mathbf{C}_{\widehat{m}} - \mathbf{C}_m)) \delta, & i > \widehat{m}. \end{cases} \quad (5.5)$$

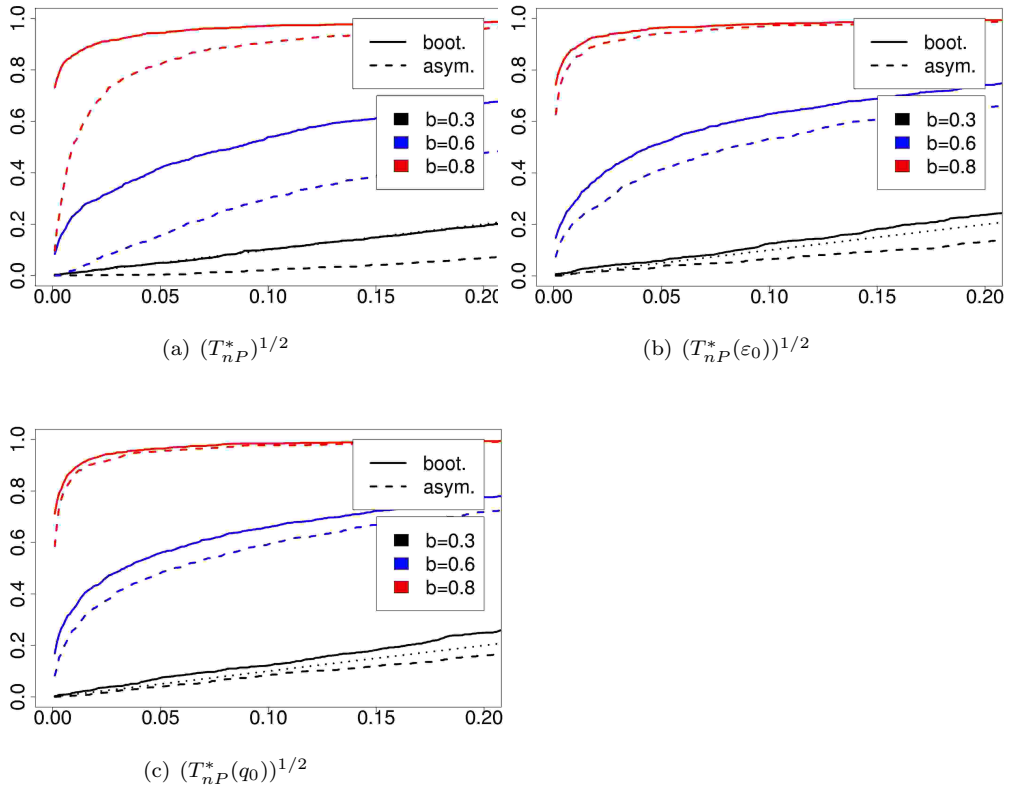


Figure 4: SPC-Plots for the pair bootstrap based on 1000 time series
 (curves correspond to $b = 0.3$ (H_0), $b = 0.6$, $b = 0.8$ from bottom to top for asymptotic (dashed) as well as bootstrap test (solid line))

□

To make the paper more readable, instead of lengthy calculations, we formulate the following auxiliary lemma, which provides all the necessary arguments also appearing in the other proofs.

Lemma 5.2. (i) *Let assumptions (A.1) or (A.3) in combination with (A.2) and (B.2) be satisfied. Let \hat{m} be one of the estimators of m defined in (3.5),(3.6). Then, for any*

$0 < \lambda < \frac{1}{4}$, as $n \rightarrow \infty$,

$$\widehat{m}^{-1+\lambda} \left\| \sum_{i=p+1}^{\widehat{m}} \mathbf{x}_i \right\| = \mathcal{O}_{\mathbb{P}}(1), \quad (n - \widehat{m})^{-1+\lambda} \left\| \sum_{i=\widehat{m}+1}^n \mathbf{x}_i \right\| = \mathcal{O}_{\mathbb{P}}(1); \quad (5.6)$$

$$\widehat{m}^{-1+\lambda} \left\| \sum_{i=p+1}^{\widehat{m}} (\mathbf{x}_i \mathbf{x}_i^T - \mathbb{E} \mathbf{x}_i \mathbf{x}_i^T) \right\| = \mathcal{O}_{\mathbb{P}}(1); \quad (5.7)$$

$$(n - \widehat{m})^{-1+\lambda} \left\| \sum_{i=\widehat{m}+1}^n (\mathbf{x}_i \mathbf{x}_i^T - \mathbb{E} \mathbf{x}_i \mathbf{x}_i^T) \right\| = \mathcal{O}_{\mathbb{P}}(1); \quad (5.8)$$

$$\widehat{m}^{-\frac{1}{2}} \left\| \sum_{i=p+1}^{\widehat{m}} \mathbf{x}_i e_i \right\| = \mathcal{O}_{\mathbb{P}}(\sqrt{\log n}); \quad (5.9)$$

$$(n - \widehat{m})^{-\frac{1}{2}} \left\| \sum_{i=\widehat{m}+1}^n \mathbf{x}_i e_i \right\| = \mathcal{O}_{\mathbb{P}}(\sqrt{\log n}). \quad (5.10)$$

(ii) Under assumptions (A.2), (A.3) and (B.2) only, with $\boldsymbol{\delta} \neq \mathbf{0}$,

$$\left\| \mathbf{C}_m - \mathbf{C}_{\widehat{m}} \right\| = \mathcal{O}_{\mathbb{P}}(|m - \widehat{m}|). \quad (5.11)$$

Proof. The assertions follow from Lemmas 4.2 and 4.3 in [17] and some immediate modifications and extensions.

For instance, (5.7) and (5.8) follow as a consequence of Lemma 4.2 in [17]. When considering, e.g., $\eta_j = Y_{j-\nu}$, $\nu = 1, \dots, p$, and proceeding in the same way as in the proof of Lemma 4.2 in [17], we get, together with (4.6) in [17],

$$\max_{p < k \leq n} k^{-1+\lambda} \left\| \sum_{i=p+1}^k \mathbf{x}_i \right\| = \mathcal{O}_{\mathbb{P}}(1); \quad (5.12)$$

$$\max_{p \leq k < n} (n - k)^{-1+\lambda} \left\| \sum_{i=k+1}^n \mathbf{x}_i \right\| = \mathcal{O}_{\mathbb{P}}(1). \quad (5.13)$$

which imply (5.6). Another application of Lemma 4.2 in [17] immediately yields

$$\max_{1 \leq k \leq n - j_n} k^{-1+\lambda} \left\| \sum_{i=j_n+1}^{j_n+k} (\mathbf{x}_i \mathbf{x}_i^T - \mathbb{E} \mathbf{x}_i \mathbf{x}_i^T) \right\| = \mathcal{O}_{\mathbb{P}}(1), \quad (5.14)$$

for all sequences $\{j_n\}$ with $j_n \leq n$, which implies (5.11). (5.9) and (5.10) follow from Lemma 4.3 in [17]. \square

5.1 Proofs for the regression bootstrap

Given \mathbf{x}_i , $i = 1, \dots, n$, the partial sums $\mathbf{V}_{n,R}^*(k/n)$ defined in (3.15) are sums of independent random vectors and hence our present situation corresponds to one of a linear regression with regressors \mathbf{x}_i , $i = 1, \dots, n$. The difference from the usual situation there is that we have a triangular array here and, moreover, while for the data under H_0 we have a so-called nontrending regression $[n^{-1}\mathbf{C}_{\lfloor nt \rfloor} \approx t\mathbf{C}$, $t \in (0, 1)]$, this is not the case under the alternative (see also the discussion after Theorem 3.3). The proofs follow the lines of the proofs for a linear regression setup, but need modifications in some places. For the sake of brevity, we focus on these modifications.

We need the following lemma:

Lemma 5.3. *Let assumptions (A.1) or (A.3) in combination with (A.2) and (B.1) be satisfied. Then, for any $0 < \lambda < \frac{1}{4}$, as $n \rightarrow \infty$,*

$$\max_{k_{1n} \leq k \leq k_{2n}} k^{-1+\lambda} \|\mathbf{C}_k - k\mathbf{C}\| = o_P(1), \quad (5.15)$$

$$\max_{k_{1n} \leq n-k \leq k_{2n}} (n-k)^{-1+\lambda} \|\mathbf{C}_k^0 - (n-k)\mathbf{C}^0\| = o_P(1), \quad (5.16)$$

$$\sup_{\varepsilon < t < 1-\varepsilon} \left\| \frac{1}{n} \mathbf{C}_{\lfloor nt \rfloor} - \mathbf{Q}_t \right\| = o_P(1), \quad (5.17)$$

where \mathbf{C} , \mathbf{C}^0 , \mathbf{Q}_t are defined in (3.20) and (3.21), respectively, and $\{k_{1n}\}, \{k_{2n}\}$ are sequences of integers satisfying

$$1 \leq k_{1n} \leq k_{2n} < n, \quad k_{1n} \rightarrow \infty, \quad k_{2n}/n \rightarrow 0.$$

Proof. Relations (5.15) and (5.17) are immediate consequences of Lemma 4.2 in [17] and the properties of autoregressive sequences. Concerning (5.16) one has to be slightly more careful. By Lemma 4.2 in [17] we have

$$\max_{k_{1n} \leq n-k \leq k_{2n}} (n-k)^{-1+\lambda} \|\mathbf{C}_k^0 - \mathbf{E}\mathbf{C}_k^0\| = o_P(1).$$

For the proof of

$$\max_{k_{1n} \leq n-k \leq k_{2n}} (n-k)^{-1+\lambda} \|\mathbf{E}\mathbf{C}_k^0 - (n-k)\mathbf{C}^0\| = o(1)$$

it suffices to notice (see Section 4 in [17]) that $\mathbf{E}(\mathbf{x}_{i+1}\mathbf{x}_{i+1}^T) = \mathbf{C}$, $i = p, \dots, m$, while for $i \geq m+1$ we have

$$\begin{aligned} \mathbf{E}(\mathbf{x}_{i+1}\mathbf{x}_{i+1}^T) &= \sum_{j=0}^{i-m-1} \Delta^j \mathbf{J}(\Delta^T)^j + \Delta^{i-m} \mathbf{C}(\Delta^T)^{i-m} \\ &= \mathbf{C}^0 - \sum_{j=i-m}^{\infty} \Delta^j \mathbf{J}(\Delta^T)^j + \Delta^{i-m} \mathbf{C}(\Delta^T)^{i-m}, \end{aligned}$$

where $\mathbf{J} = \mathbf{E}(\mathbf{e}_1 \mathbf{e}_1^T)$ with $\|\mathbf{J}\| = \sigma^2$ (see also (2.2)) and $\mathbf{\Delta}$ is the matrix defined in Section 4 of [17] with all the eigenvalues less than one in absolute value. The result then follows directly by utilizing the fact that $\|C\|$ remains constant and $\|\mathbf{\Delta}^j\| = \mathcal{O}(\rho^j)$, where $0 < |\lambda_{\max}| < \rho < 1$, $|\lambda_{\max}|$ being the largest in modulus eigenvalue of $\mathbf{\Delta}$. \square

Proof of Theorem 3.2. It follows from the above lemma together with Lemma 5.1, Lemma 5.2 and the functional central limit theorem (confer, e.g., Davidson (1994), Theorem 27.17). \square

Proof of Theorem 3.3. It is then an immediate consequence of Theorem 3.2. \square

Proof of Theorem 3.1. We follow the usual pattern of the proofs of these types of assertions (see, e.g., Csörgő and Horváth [7], Davis et al. [8]). However, it has to be taken into account that we have a triangular array here. So, we apply Theorem 6.2 below, which provides a suitable weak invariance principle with the required rate, together with Lemma 5.3. \square

5.2 Proofs for the pair bootstrap

In the present situation $\mathbf{V}_{n,P}^*(k/n)$, $k = p+1, \dots, n$, defined in (3.28) are sums of i.i.d. random vectors which somehow simplifies the situation, however, there is a problem with the random matrices \mathbf{C}_k^* . Their elements are sums of i.i.d. random variables.

We shall make use of the following three lemmas and a weak invariance principle with a suitable rate (cf. Theorem 6.1 below). According to the first lemma, \mathbf{C}_k^* is sufficiently close to its expectation and therefore it suffices to prove the results for the considered statistics with \mathbf{C}_k^* replaced by its expectation, i.e., by $(k/n)\mathbf{C}_n$. The second lemma enables us to replace the variance matrix $\hat{\Sigma}_n$ by its asymptotic counterpart, while the third lemma gives an upper bound for $\mathbf{E}^* \|\mathbf{x}_i^* \tilde{\mathbf{e}}_i^*\|^2$, which is needed for an application of the invariance principle.

Lemma 5.4. *Under the assumptions of Theorem 3.4, for any $A > 0$, $\gamma \in [0, \frac{1}{2})$, and $p < k_n \leq n$,*

$$\begin{aligned} & \mathbf{P}^* \left(\max_{k_n < k \leq n} \frac{1}{k^\gamma} \left\| \mathbf{C}_k^* - \frac{k}{n} \mathbf{C}_n \right\| \geq A \right) \\ & \leq DA^{-2} \frac{1}{n} \sum_{i=p+1}^n \|\mathbf{x}_i\|^2 (k_n^{-2\gamma+1} + n^{-2\gamma+1}) \end{aligned} \quad (5.18)$$

and

$$\begin{aligned} & \mathbf{P}^* \left(\max_{p < k \leq n-k_n} \frac{1}{(n-k)^\gamma} \left\| \mathbf{C}_k^{*0} - \frac{n-k}{n} \mathbf{C}_n \right\| \geq A \right) \\ & \leq DA^{-2} \frac{1}{n} \sum_{i=p+1}^n \|\mathbf{x}_i\|^2 (k_n^{-2\gamma+1} + n^{-2\gamma+1}), \end{aligned} \quad (5.19)$$

with some $D > 0$.

Proof. Denoting by x_{ij} and x_{ij}^* ($i = p+1, \dots, n$; $j = 1, \dots, p$) the components of \mathbf{x}_i and \mathbf{x}_i^* , respectively, we get that, for $1 \leq j, v \leq p$, and any $p < k_n \leq n$,

$$\mathbb{E}^* x_{ij}^* x_{iv}^* = \frac{1}{n-p} \sum_{i=p+1}^n x_{ij} x_{iv},$$

and

$$\text{Var}^*(x_{ij}^* x_{iv}^*) \leq \frac{1}{n-p} \sum_{i=p+1}^n (x_{ij} x_{iv})^2 \leq \frac{2}{n} \sum_{i=p+1}^n \|\mathbf{x}_i\|^4.$$

This in combination with the Hájek-Rényi inequality implies that, for each $A > 0$,

$$\begin{aligned} & \mathbb{P}^* \left(\max_{k_n < k \leq n} \frac{1}{k^\gamma} \left| \sum_{i=p+1}^k (x_{ij}^* x_{iv}^* - \mathbb{E}^* x_{ij}^* x_{iv}^*) \right| \geq A \right) \\ & \leq A^{-2} \left(\sum_{i=k_n}^n \frac{\text{Var}^*(x_{ij}^* x_{iv}^*)}{k^{2\gamma}} + \frac{1}{k_n^{2\gamma}} \sum_{i=p+1}^{k_n} \text{Var}^*(x_{ij}^* x_{iv}^*) \right) \\ & \leq DA^{-2} \frac{1}{n} \sum_{i=k_n}^n \|\mathbf{x}_i\|^4 \left(\sum_{k=k_n}^n k^{-2\gamma} + k_n^{1-2\gamma} \right), \end{aligned}$$

with some $D > 0$ and for any $\gamma \in [0, \frac{1}{2})$. This immediately implies (5.18). The assertion (5.19) can be proved in the same way and is hence omitted. \square

Lemma 5.5. *Under the assumptions of Theorem 3.4, as $n \rightarrow \infty$,*

$$\widehat{\Sigma}_n = \frac{1}{n} \mathbf{C}_n + \mathcal{O}_{\mathbb{P}}(n^{-1/2} + (d_{2n} n^{-1})^{1/2}),$$

with some $\lambda > 0$ and for any sequence $\{d_{2n}\}$ satisfying (3.4).

Proof. We focus on the case of fixed alternatives, i.e., $\boldsymbol{\delta} \neq \mathbf{0}$ does not depend on n . First we assume that $m - d_{2n} \leq \widehat{m} \leq m$ with $\{d_{2n}\}$ satisfying (3.4).

Due to Lemmas 5.1 and 5.2, it suffices to investigate

$$\frac{1}{n-p} \sum_{i=p+1}^n \mathbf{x}_i^T \mathbf{x}_i \widetilde{\mathbf{e}}_i^2. \quad (5.20)$$

We split the sum in (5.20) into two sums for the indices $i \leq \widehat{m}$ and $i > \widehat{m}$. For the sum over the indices $i > \widehat{m}$, we have

$$\frac{1}{n-p} \sum_{i=\widehat{m}+1}^n \mathbf{x}_i^T \mathbf{x}_i \widetilde{e}_i^2 = A_{n1} - 2A_{n2} + A_{n3}, \quad (5.21)$$

where

$$\begin{aligned} A_{n1} &= \frac{1}{n-p} \sum_{i=\widehat{m}+1}^n \mathbf{x}_i^T \mathbf{x}_i (e_i - \mathbf{x}_i^T (\mathbf{C}_{\widehat{m}}^0)^{-1} \sum_{j=\widehat{m}+1}^n \mathbf{x}_j e_j)^2, \\ A_{n2} &= \frac{1}{n-p} \sum_{i=\widehat{m}+1}^n \mathbf{x}_i^T \mathbf{x}_i (e_i - \mathbf{x}_i^T (\mathbf{C}_{\widehat{m}}^0)^{-1} \sum_{j=\widehat{m}+1}^n \mathbf{x}_j e_j) \\ &\quad \times (\mathbf{I} I\{\widehat{m} < i \leq m\} + (\mathbf{C}_{\widehat{m}}^0)^{-1} (\mathbf{C}_{\widehat{m}}^0 - \mathbf{C}_m)) \boldsymbol{\delta}, \\ A_{n3} &= \frac{1}{n-p} \sum_{i=\widehat{m}+1}^n \mathbf{x}_i^T \mathbf{x}_i (\{\mathbf{x}_i^T (\mathbf{I} I\{\widehat{m} < i \leq m\} + (\mathbf{C}_{\widehat{m}}^0)^{-1} (\mathbf{C}_{\widehat{m}}^0 - \mathbf{C}_m)) \boldsymbol{\delta}\}^2). \end{aligned}$$

By Lemmas 5.1 and 5.2, we get

$$\begin{aligned} A_{n1} &= \frac{1}{n-p} \sum_{i=\widehat{m}+1}^n \mathbf{x}_i^T \mathbf{x}_i e_i^2 + \mathcal{O}_{\mathbb{P}}(n^{-1/2}), \\ A_{n3} &= \mathcal{O}_{\mathbb{P}}\left(\|\boldsymbol{\delta}\|^2 \left(\frac{1}{n} \sum_{j=\widehat{m}+1}^m \|\mathbf{x}_j\|^4 + \frac{1}{n^3} \sum_{j=\widehat{m}+1}^n \|\mathbf{x}_j\|^2 \|\mathbf{C}_{\widehat{m}}^0 - \mathbf{C}_m\|^2\right)\right) \\ &= \mathcal{O}_{\mathbb{P}}\left(\|\boldsymbol{\delta}\|^2 d_{2n}/n\right), \\ |A_{n2}| &\leq \sqrt{A_{n1} A_{n3}} = \mathcal{O}_{\mathbb{P}}(\|\boldsymbol{\delta}\| (d_{2n}/n)^{1/2}). \end{aligned}$$

Hence

$$\frac{1}{n-p} \sum_{i=\widehat{m}+1}^n \mathbf{x}_i^T \mathbf{x}_i \widetilde{e}_i^2 = \frac{1}{n-p} \sum_{i=\widehat{m}+1}^n \mathbf{x}_i^T \mathbf{x}_i e_i^2 + \mathcal{O}_{\mathbb{P}}(n^{-1/2} + \|\boldsymbol{\delta}\| (d_{2n}/n)^{1/2}), \quad (5.22)$$

and similarly,

$$\frac{1}{n-p} \sum_{i=p+1}^{\widehat{m}} \mathbf{x}_i^T \mathbf{x}_i \widetilde{e}_i^2 = \frac{1}{n-p} \sum_{i=p+1}^{\widehat{m}} \mathbf{x}_i^T \mathbf{x}_i e_i^2 + \mathcal{O}_{\mathbb{P}}(n^{-1/2}).$$

The last two relations lead to

$$\frac{1}{n-p} \sum_{i=p+1}^n \mathbf{x}_i^T \mathbf{x}_i \widetilde{e}_i^2 = \frac{1}{(n-p)} \sum_{i=p+1}^n \mathbf{x}_i^T \mathbf{x}_i e_i^2 + \mathcal{O}_{\mathbb{P}}(n^{-1/2} + \|\boldsymbol{\delta}\| (d_{2n}/n)^{1/2}).$$

We get quite the same results for $m \leq \widehat{m} \leq m + d_{2n}$. Since the considered estimator \widehat{m} has the property (3.3), for any sequence $\{d_{2n}\}$ satisfying (3.4), the proof can be easily completed. \square

We need one more assertion:

Lemma 5.6. *Under the assumptions of Theorem 3.4, as $n \rightarrow \infty$,*

$$\mathbb{E}^* \|\mathbf{x}_i^* \widetilde{\mathbf{e}}_i^*\|^{2+\eta} = \mathcal{O}_{\mathbb{P}}(1).$$

Proof. We investigate only the situation under H_A and with $m - d_{2n} \leq \widehat{m} \leq m$. Clearly,

$$\mathbb{E}^* \|\mathbf{x}_i^* \widetilde{\mathbf{e}}_i^*\|^{2+\eta} = \frac{1}{n-p} \sum_{i=p+1}^n \|\mathbf{x}_i \widetilde{\mathbf{e}}_i\|^{2+\eta} \leq D(Q_{1n} + Q_{2n} + Q_{3n} + Q_{4n}),$$

with some $D > 0$, where

$$\begin{aligned} Q_{1n} &= \frac{1}{n-p} \sum_{i=p+1}^{\widehat{m}} \left\| \mathbf{x}_i (e_i - \mathbf{x}_i^T \mathbf{C}_{\widehat{m}}^{-1} \sum_{j=p+1}^{\widehat{m}} \mathbf{x}_j e_j) \right\|^{2+\eta}, \\ Q_{2n} &= \frac{1}{n-p} \sum_{i=\widehat{m}+1}^n \left\| \mathbf{x}_i (e_i - \mathbf{x}_i^T \mathbf{C}_{\widehat{m}}^{0-1} \sum_{j=\widehat{m}+1}^n \mathbf{x}_j e_j) \right\|^{2+\eta}, \\ Q_{3n} &= \frac{1}{n-p} \sum_{i=\widehat{m}+1}^m \|\mathbf{x}_i\|^{4+2\eta} \|\mathbf{C}_{\widehat{m}}^{0-1} \mathbf{C}_m^0\|^{2+\eta} \|\boldsymbol{\delta}\|^{2+\eta}, \\ Q_{4n} &= \frac{1}{n-p} \sum_{i=\widehat{m}+1}^n \|\mathbf{x}_i\|^{4+2\eta} \|\mathbf{C}_{\widehat{m}}^{0-1} (\mathbf{C}_m^0 - \mathbf{C}_{\widehat{m}}^0)\|^{2+\eta} \|\boldsymbol{\delta}\|^{2+\eta}. \end{aligned}$$

By Lemmas 5.1 and 5.2 and since

$$\left\| \sum_{i=p+1}^n \mathbf{x}_i e_i \right\|^{2+\eta} = \mathcal{O}_{\mathbb{P}}(n)$$

we get that

$$\begin{aligned} Q_{1n} + Q_{2n} &= \mathcal{O}_{\mathbb{P}}(1 + n^{-1/2}) \\ Q_{3n} &= \mathcal{O}_{\mathbb{P}}(\|\boldsymbol{\delta}\|^{2+\eta} (d_{2n}/n)^{2+\eta}) \\ Q_{4n} &= \mathcal{O}_{\mathbb{P}}(\|\boldsymbol{\delta}\|^{2+\eta} (d_{2n}/n)^{2+\eta} n^{\eta/2}) = \mathcal{O}_{\mathbb{P}}(\|\boldsymbol{\delta}\|^{2+\eta} d_{2n}^{2+\eta} n^{-2-\eta/2}), \end{aligned}$$

where we used also the inequality

$$\sum_{p+1}^n \|\mathbf{x}_i\|^{4+2\eta} \leq \left(\sum_{p+1}^n \|\mathbf{x}_i\|^4 \right)^{1+\eta/2} = \mathcal{O}_{\mathbb{P}}(n^{1+\eta/2}).$$

Hence, for $m - d_{2n} \leq \widehat{m} \leq m$,

$$\mathbb{E}^* \|\mathbf{x}_i^* \widetilde{\mathbf{e}}_i^*\|^{2+\eta} = \mathcal{O}_{\mathbb{P}}(1 + d_{2n}^{2+\eta} n^{-(4+\eta)/(4+2\eta)}),$$

and the same assertion can be proved for the case $m \leq \widehat{m} \leq m + d_{2n}$. An appropriate choice of the sequence $\{d_{2n}\}$ completes the proof. \square

Proof of Theorems 3.4 and 3.5. It follows from Lemmas 5.4–5.6 above in combination with Theorem 6.1 below.

6 Auxiliary results

Here we derive suitable versions of the invariance principles that play the crucial role in the proofs of our theorems in Sections 3 and 5. It is a consequence of the results proved by Sakhanenko [21] and later reproved by a student of Einmahl [9].

In our setup we need invariance principles for partial sums of independent random vectors that form a triangular array. The last fact makes it impossible to use the results by Csörgő and Horváth in their book [7] and a number of other papers.

We need two versions of invariance principles. One for the sums of i.i.d. random vectors that is needed for the proof of the results for the pair bootstrap (Theorems 3.4 and 3.5), and more general, for sums of independent but nonidentically distributed random vectors.

The problem here is the multidimensional situation and the fact that we have a triangular array.

Let us start with the i.i.d. situation.

Theorem 6.1. *Let $\mathbf{Z}_{1,n}, \dots, \mathbf{Z}_{n,n}$ be i.i.d. p -dimensional random vectors with zero mean, variance matrix Σ_n , and $\limsup_{n \rightarrow \infty} \mathbb{E} \|\mathbf{Z}_{1,n}\|^{2+\eta} < \infty$ with some $\eta > 0$. Then, on a suitable probability space, there exists a sequence of p -dimensional standard Wiener processes $\mathbf{W}_n = (W_{1,n}, \dots, W_{p,n})^T$ on $[0, \infty)$ such that, as $n \rightarrow \infty$,*

$$\max_{k_n \leq k \leq n} k^{-\beta} \left\| \sum_{i=1}^k \mathbf{Z}_{i,n} - \Sigma_n^{1/2} \mathbf{W}_n(k) \right\| = \mathcal{O}_{\mathbb{P}}(k_n^{-\lambda}), \quad (6.1)$$

with some $0 < \beta < 1/2$ and $\lambda > 0$ and for any $1 \leq k_n \leq n$, $k_n \rightarrow \infty$.

Proof. Without loss of generality we may assume that $\eta \in (0, 1)$. Let us denote

$$n_0 = 0, \quad n_j = \lfloor \exp\{j^\alpha\} \rfloor \quad (j = 1, \dots, L_n - 1), \quad n_{L_n} = n, \quad (6.2)$$

for $\alpha \in (0, 1)$, with L_n being the largest integer satisfying

$$\exp\{(L_n - 1)^\alpha\} < n. \quad (6.3)$$

Also denote L_n^0 the largest integer satisfying

$$\exp\{(L_n^0)^\alpha\} \leq k_n.$$

Then, by the results of Ampe [1] and Einmahl [9], we have that, for any $\beta > 0$, $x > 0$ and any $n \geq n_0$ with some $n_0 > 0$, there exist a universal positive constant D (depending neither on x nor on n) and a sequence of p -dimensional standard Wiener process $\mathbf{W}_n = (W_{1,n}, \dots, W_{p,n})^T$ on $[0, \infty)$ (the process depends x and n) such that for $j \geq 2$

$$\begin{aligned} & \mathbb{P}\left(\max_{n_{j-1} < k \leq n_j} \left\| \sum_{i=n_{j-1}+1}^k \mathbf{Z}_{i,n} - \Sigma_n^{1/2}(\mathbf{W}_n(k) - \mathbf{W}_n(n_{j-1})) \right\| \geq x n_{j-1}^\beta\right) \\ & \leq D x^{-(2+\eta)} n_{j-1}^{-\beta(2+\eta)} (n_j - n_{j-1}) \mathbb{E} \|\mathbf{Z}_{1,n}\|^{2+\eta}. \end{aligned} \quad (6.4)$$

Hence

$$\begin{aligned} & \sum_{j=L_n^0+1}^{L_n} \mathbb{P}\left(\max_{n_{j-1} < k \leq n_j} \left\| \sum_{i=n_{j-1}+1}^k \mathbf{Z}_{i,n} - \Sigma_n^{1/2}(\mathbf{W}_n(k) - \mathbf{W}_n(n_{j-1})) \right\| \geq x n_{j-1}^\beta\right) \\ & \leq D x^{-(2+\eta)} \sum_{j=L_n^0+1}^{L_n} n_{j-1}^{-\beta(2+\eta)} (n_j - n_{j-1}) \\ & \leq D_1 x^{-(2+\eta)} \sum_{j=L_n^0}^{L_n} \exp\{-(j-1)^\alpha(\beta(2+\eta) - 1)\} \\ & \leq D_2 \exp\{-(L_n^0 - 1)^\alpha(\beta(2+\eta) - 1)\} x^{-(2+\eta)} \\ & \leq D_3 k_n^{\beta(2+\eta)-1} x^{-(2+\eta)}, \end{aligned} \quad (6.5)$$

where we choose $\beta \in (1/(2+\eta), 1/2)$ and we used the inequalities

$$\exp\{j^\alpha - (j-1)^\alpha\} - 1 \leq \exp\{(j-1)^{\alpha-1}\alpha\} - 1 \leq (j-1)^{\alpha-1}\alpha. \quad (6.6)$$

Next note that, for $\beta \in [0, 1/2)$,

$$\begin{aligned} & \left\{ \max_{1 \leq k \leq n} k^{-\beta} \left\| \sum_{i=1}^k \mathbf{Z}_{i,n} - \Sigma_n^{1/2} \mathbf{W}_n(k) \right\| \geq x \right\} \\ & \subseteq \bigcup_{j=L_n^0}^{L_n} \left\{ \max_{n_{j-1} < k \leq n_j} \left\| \sum_{i=1}^k \mathbf{Z}_{i,n} - \Sigma_n^{1/2} \mathbf{W}_n(k) \right\| \geq x n_{j-1}^\beta \right\}, \end{aligned}$$

and also, for $n_{j-1} < k \leq n_j$,

$$\begin{aligned} \sum_{i=1}^k \mathbf{Z}_{i,n} - \Sigma_n^{1/2} \mathbf{W}_n(k) &= \sum_{v=1}^{j-1} \left(\sum_{i=n_{v-1}+1}^{n_v} \{ \mathbf{Z}_{i,n} - \Sigma_n^{1/2} (\mathbf{W}_n(n_v) - \mathbf{W}_n(n_{v-1})) \} \right) \\ &\quad + \sum_{i=n_{j-1}+1}^k \{ \mathbf{Z}_{i,n} - \Sigma_n^{1/2} (\mathbf{W}_n(k) - \mathbf{W}_n(n_{j-1})) \}. \end{aligned}$$

The last two relations further imply

$$\begin{aligned} &\left\{ \max_{k_n \leq k \leq n} k^{-\beta} \left\| \sum_{i=1}^k \mathbf{Z}_{i,n} - \Sigma_n^{1/2} \mathbf{W}_n(k) \right\| \geq x \right\} \\ &\subseteq \bigcup_{j=L_n^0}^{L_n} \left\{ \max_{n_{j-1} < k \leq n_j} \left\| \sum_{i=n_{j-1}+1}^k \mathbf{Z}_{i,n} - \Sigma_n^{1/2} \mathbf{W}_n(k) \right\| \geq D_1 x n_{j-1}^{\beta^*} \right\} \\ &\cup \left\{ \left\| \sum_{i=1}^{L_n^0} \mathbf{Z}_{i,n} - \Sigma_n^{1/2} \mathbf{W}_n(L_n^0) \right\| \geq D_1 x L_n^{0\beta^*} \right\}, \end{aligned}$$

for any $\beta^* < \beta$ and some $D_1 > 0$. Here we used

$$\sum_{v=1}^i n_{v-1}^{\beta^*} = \sum_{v=1}^i \exp\{(v-1)\alpha\beta^*\} \leq \exp\{(i-1)\alpha\beta^* + \log i\} \leq D \exp\{(i-1)\alpha\beta\},$$

for any $\beta^* < \beta$ and some $D > 0$. Hence

$$\mathbb{P} \left(\max_{k_n \leq k \leq n} k^{-\beta^*} \left\| \sum_{i=1}^k \mathbf{Z}_{i,n} - \Sigma_n^{1/2} \mathbf{W}_n(k) \right\| \geq x \right) \leq D x^{-2-\eta} k_n^{-\lambda},$$

for any $1/(2+\eta) < \beta^* < \beta$, some $D > 0$ and some $\lambda > 0$, which easily implies the assertion of our theorem. \square

Next, we are interested in a similar assertion for the partial sums

$$\mathbf{S}_{k,n} = \sum_{i=1}^k \mathbf{x}_{i,n} e_{i,n}, \quad k = p+1, \dots, n,$$

where

(C.1) $e_{1,n}, \dots, e_{n,n}$ are i.i.d. errors with mean 0, unit variance and $\mathbb{E}|e_{1,n}|^{2+\eta} \leq D$, with some $0 < \eta < 1$ and $D > 0$;

(C.2) there exists a sequence of symmetric matrices Σ_n such that

$$\max_{1 \leq k \leq n} k^\gamma \left\| \frac{1}{k} \mathbf{C}_{k,n} - \Sigma_n \right\| = \mathcal{O}(1),$$

with some $\gamma \in [0, 1/2)$ and

$$0 < \liminf_{n \rightarrow \infty} \Sigma_n,$$

where

$$\mathbf{C}_{k,n} = \sum_{i=1}^k \mathbf{x}_{in} \mathbf{x}_{in}^T, \quad k = 1, \dots, n;$$

(C.3) for some $\eta > 0$,

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \|\mathbf{x}_{in}\|^{2+\eta} < \infty.$$

Theorem 6.2. *Under the assumptions (C.1)–(C.3) there exists a sequence of p -dimensional standard Wiener process $\mathbf{W}_n = (W_{1,n}, \dots, W_{p,n})^T$ on $[0, \infty)$ such that, as $n \rightarrow \infty$,*

$$\max_{k_n \leq k \leq n} k^{-1/2} (\log k)^\beta \|\mathbf{S}_{k,n} - \Sigma_n^{1/2} \mathbf{W}_n(k)\| = \mathcal{O}_{\mathbb{P}}((\log k_n)^{-\lambda}),$$

for any $\beta > 0$, $\lambda > 0$, and for any $1 \leq k_n \leq n$, $k_n \rightarrow \infty$.

Proof. We follow the lines of the proof of Theorem 6.1 with some modifications. For chosen β , we take α such that $0 < \alpha(1 + 2\beta) < 1$. Particularly, since our summands are not i.i.d., we have instead of (6.4), for $j \geq 2$,

$$\begin{aligned} & \mathbb{P} \left(\max_{n_{j-1} < k \leq n_j} \left\| (\mathbf{S}_{k,n} - \mathbf{S}_{n_{j-1},n}) - \sum_{i=n_{j-1}+1}^k \mathbf{Q}_{ni} \right\| \geq x n_{j-1}^\beta \right) \\ & \leq D x^{-(2+\eta)} n_{j-1}^{-\beta(2+\eta)} \sum_{i=n_{j-1}+1}^{n_j} \|\mathbf{x}_{in}\|^{2+\eta}, \end{aligned} \quad (6.7)$$

where \mathbf{Q}_{ni} , $i = 1, \dots, n$, are independent normally distributed random vectors with zero means and $\text{Var } \mathbf{Q}_{ni} = \mathbf{x}_{in} \mathbf{x}_{in}^T$.

Hence

$$\begin{aligned} & \mathbb{P} \left(\max_{k_n \leq k \leq n} k^{-\beta} \left\| (\mathbf{S}_{k,n} - \sum_{i=1}^k \mathbf{Q}_{ni}) \right\| \geq x \right) \\ & \leq D x^{-2-\eta} \sum_{j=L_n^0}^{L_n} n_{j-1}^{-\beta(2+\eta)} \sum_{i=n_{j-1}+1}^{n_j} \|\mathbf{x}_{in}\|^{2+\eta}. \end{aligned} \quad (6.8)$$

By the Abel summation formula,

$$\begin{aligned} & \sum_{j=L_n^0}^{L_n} n_{j-1}^{-\beta(2+\eta)} \sum_{i=n_{j-1}+1}^{n_j} \|\mathbf{x}_{in}\|^{2+\eta} \\ & \leq \sum_{j=L_n^0}^{L_n-1} \sum_{i=n_{L_n^0}}^{n_j} \|\mathbf{x}_{in}\|^{2+\eta} \left(n_{j-1}^{-\beta(2+\eta)} - n_j^{-\beta(2+\eta)} \right) + \sum_{i=n_{L_n^0}}^n \|\mathbf{x}_{in}\|^{2+\eta} n^{-\beta(2+\eta)}. \end{aligned}$$

From assumption (C.3) we get that the r.h.s. of (6.8) is smaller than

$$D_1 \left(\sum_{j=L_n^0}^{L_n-1} n_j \left(n_{j-1}^{-\beta(2+\eta)} - n_j^{-\beta(2+\eta)} \right) + n^{-\beta(2+\eta)+1} \right),$$

with some $D_2 > 0$, if we choose again $\beta \in (1/(2+\eta), 1/2)$. Then, following the proof of the previous theorem, we obtain

$$\mathbb{P} \left(\max_{k_n \leq k \leq n} k^{-\beta} \left\| \mathbf{S}_{k,n} - \sum_{i=1}^k \mathbf{Q}_{ni} \right\| \geq x \right) \leq D x^{-2-\eta} k_n^{-\lambda},$$

with some $\lambda > 0$. By the above arguments, we have just switched to the sums of independent normally distributed random vectors. We still need one more step get to a Wiener process approximation.

Toward this we define

$$\mathbf{Z}_{nj} = \sqrt{n_j - n_{j-1}} \boldsymbol{\Sigma}_n^{1/2} (\mathbf{C}_{n_j} - \mathbf{C}_{n_{j-1}})^{-1/2} \mathbf{M}_{nj}$$

with

$$\mathbf{M}_{nj} = \sum_{i=n_{j-1}+1}^{n_j} \mathbf{Q}_{ni}$$

Clearly,

$$\text{Var } \mathbf{Z}_{nj} = (n_j - n_{j-1}) \boldsymbol{\Sigma}_n$$

and

$$\text{Var} (\mathbf{Z}_{nj} - \mathbf{M}_{nj}) = \left((\mathbf{C}_{n_j} - \mathbf{C}_{n_{j-1}})^{1/2} - \sqrt{n_j - n_{j-1}} \boldsymbol{\Sigma}_n^{1/2} \right)^2$$

Notice that, by assumption (C.2), the \liminf of the eigenvalues of $\boldsymbol{\Sigma}_n$ is bounded away from 0. Then, in view of (C.2),

$$\max_{1 \leq j \leq L_n} \left\| (\mathbf{C}_{n_j} - \mathbf{C}_{n_{j-1}})^{1/2} - \sqrt{n_j - n_{j-1}} \boldsymbol{\Sigma}_n^{1/2} \right\| n_{j-1}^{-\gamma} = \mathcal{O}(1),$$

so that, after some calculations, we get

$$\left\| \text{Var} (\mathbf{Z}_{nj} - \mathbf{M}_{nj}) \right\| \leq D n_{j-1}^{2\gamma-1}, \quad j = 2, \dots,$$

with some $D > 0$.

We make use of the Feller inequality (applied to each component of $\mathbf{Z}_{nj} - \mathbf{M}_{nj}$) and recall that, if X has a normal $(0, 1)$ distribution, then $\mathbb{P}(|X| > x) \leq \exp\{-x^2/2\}d/x$, with some $d > 0$. This results in the estimate

$$\mathbb{P}\left(\|\mathbf{Z}_{nj} - \mathbf{M}_{nj}\| \geq x n_{j-1}^{\beta-\gamma+1/2}\right) \leq D \exp\left\{-n_{j-1}^{2(\beta-\gamma+1/2)}x^2/2\right\}x^{-1}n_{j-1}^{-\beta+\gamma-1/2},$$

for any $x > 0$ and some $D > 0$. Hence

$$\begin{aligned} & \mathbb{P}\left(\max_{L_n^0 \leq j \leq L_n} \|\mathbf{Z}_{nj} - \mathbf{M}_{nj}\| n_{j-1}^{-\beta} \geq x\right) \\ & \leq D \sum_{j=L_n^0}^{L_N} \exp\left\{-n_{j-1}^{2(\beta-\gamma+1/2)}x^2/2\right\}x^{-1}n_{j-1}^{-\beta+\gamma-1/2} \\ & \leq D_1 \exp\left\{-k_n^{2(\beta-\gamma+1/2)}x^2/2\right\}x^{-1}k_n^{-\beta+\gamma-1/2} \leq D_3 k_n^{-\lambda_1}, \end{aligned}$$

for any $\lambda_1 > 0$ and some positive constants D, D_1, D_2 , where we choose β such that $\beta - \gamma + 1/2 > 0$. It follows that

$$\max_{L_n^0 \leq j \leq L_n} \|\mathbf{Z}_{nj} - \mathbf{M}_{nj}\| n_{j-1}^{-\beta} = o_{\mathbb{P}}(k_n^{-\lambda_2}),$$

for some $\lambda_2 > 0$.

Finally we show that, for any $\beta > 0, \lambda > 0$ and $0 < \alpha < (1 + 2\beta)^{-1}$,

$$\max_{L_n^0 \leq j \leq L_n} \max_{n_{j-1} \leq k \leq n_j} k^{-1/2}(\log k)^\beta \left\| \sum_{i=n_{j-1}+1}^k \mathbf{Q}_i \right\| = o_{\mathbb{P}}((\log k_n)^{-\lambda}).$$

We prove this estimate componentwise. Recall that we have independent random variables Q_i with a normal $(0, \sigma_i^2)$ distribution, $i = 1, \dots, N$, such that $\sum_{i=1}^N \sigma_i^2 \leq BN$ with some $B > 0$. Then, by Example 9, p. 256, in Chow and Teicher [6], we have

$$\mathbb{P}\left(\max_{1 \leq k \leq N} k^{-1/2}(\log k)^\beta \left| \sum_{i=1}^k Q_i \right| \geq A\right) \leq \exp\{-At\} \mathbb{E}\left(\exp\left\{t \sum_{i=1}^N Q_i\right\} + \exp\left\{-t \sum_{i=1}^N Q_i\right\}\right)$$

for any $A > 0$ and $t > 0$. Since the Q_i 's have a normal distribution,

$$\mathbb{E} \exp\left\{t \sum_{i=1}^N Q_i\right\} = \exp\left\{\frac{t^2}{2} \sum_{i=1}^N \sigma_i^2\right\} \leq \exp\left\{\frac{t^2}{2} NB\right\}.$$

Hence

$$\mathbb{P}\left(\max_{1 \leq k \leq N} k^{-1/2}(\log k)^\beta \left| \sum_{i=1}^k Q_i \right| \geq A\right) \leq 2 \exp\left\{-At + \frac{t^2}{2} NB\right\}.$$

On choosing $t = A/(NB)$ we get

$$\mathbb{P}\left(\max_{1 \leq k \leq N} \left| \sum_{i=1}^k Q_i \right| \geq A\right) \leq 2 \exp\{-A^2/(NB2)\}.$$

Applying this to our situation we obtain, for some $B > 0$,

$$\begin{aligned} & \sum_{j=L_n^0}^{L_n} \mathbb{P}\left(\max_{n_{j-1} \leq k \leq n_j} k^{-1/2}(\log k)^\beta \left\| \sum_{i=n_{j-1}+1}^k Q_i \right\| \geq A\right) \\ & \leq \sum_{j=L_n^0}^{L_n} \mathbb{P}\left(\max_{n_{j-1} \leq k \leq n_j} k^{-1/2}(\log k)^\beta \left\| \sum_{i=n_{j-1}+1}^k Q_i \right\| \geq A n_{j-1}^{1/2}(\log n_{j-1})^{-\beta}\right) \\ & \leq D \sum_{j=L_n^0}^{L_n} \exp\left\{-\frac{A^2 n_{j-1}(\log n_{j-1})^{-2\beta}}{2B(n_j - n_{j-1})}\right\}. \end{aligned} \quad (6.9)$$

Since

$$(n_j - n_{j-1})/n_{j-1} \leq (j-1)^{\alpha-1}\alpha,$$

the r.h.s. of (6.9) can further be bounded by

$$\begin{aligned} & \leq D \sum_{j=L_n^0}^{L_n} \exp\left\{-A^2 D_1 (j-1)^{-\alpha(1+2\beta)+1}\right\} \\ & \leq D \exp\left\{-A^2 D_1 (L_n^0 - 1)^{-\alpha(1+2\beta)+1}\right\} \leq D(\log k_n)^{-\lambda}, \end{aligned}$$

for any $\lambda > 0$, which completes the proof of our theorem. For similar assertions confer also the paper by Horváth and Shao [13]. \square

References

- [1] Ampe, M. Strong approximation under low moment assumptions and applications. Master's thesis, Free University of Brussels, 2005.
- [2] Antoch, J., and Hušková, M. Permutation tests for change point analysis. *Statist. Probab. Letters*, 53:37–46, 2001.
- [3] Antoch, J., and Hušková, M. Detection of structural changes in regression. *Tatra Mt. Math. Publ.*, 26:201–215, 2003.
- [4] Antoch, J., Hušková, M., and Jarušková, D. Change point detection. In: Lauro, C., et al. (eds.) *Lecture Notes of the 5th IASC Summer School*, pp. 1–75. ISI, Voorburg, 2000.

- [5] Antoch, J., Hušková, M., and Veraverbeke, N. Change-point and bootstrap. *Nonparam. Statist.*, 5:123–144, 1995.
- [6] Chow, Y.S., and Teicher, H. *Probability Theory*. Springer, New York, 2003. Third edition.
- [7] Csörgő, M., and Horváth, L. *Limit Theorems in Change-Point Analysis*. Wiley, Chichester, 1997.
- [8] Davis, R.A., Huang, D., and Yao, Y.-C. Testing for a change in the parameter values and order of an autoregressive model. *Ann. Statist.*, 23:282–304, 1995.
- [9] Einmahl, U. Personal communication, 2005.
- [10] Franke, J., Kreiss, H.P., and Mammen, E. Bootstrap of kernel smoothing in nonlinear time series. *Bernoulli*, 8:1–37, 2002.
- [11] Hidalgo, J. An alternative bootstrap to moving blocks for time series regression models. *J. Econometrics*, 117:369–399, 2003.
- [12] Horváth, L., Kokoszka, P., and Teyssière, G. Bootstrap misspecification tests for ARCH based on the empirical process of squared residuals. *J. Stat. Comput. Simul.*, 74:469–485, 2004.
- [13] Horváth, L., and Shao, Q.-M. Limit theorems for the union-intersection test. *J. Statist. Plann. Infer.*, 44:133–148, 1995.
- [14] Hušková, M. Permutation principle and bootstrap in change point analysis. In: Horváth, L., and Szyszkowicz, B. (eds.) *Asymptotic Methods in Stochastics. Fields Institute Communications*, 44:273–291, 2004.
- [15] Hušková, M., and Picek, J. Some remarks on permutation type tests in linear models. *Discuss. Math. Probab. Stat.*, 24:151–182, 2004.
- [16] Hušková, M., and Picek, J. Bootstrap in detection of changes in linear regression. *Sankhya*, 67:200–226, 2005.
- [17] Hušková, M., Prášková, Z., and Steinebach, J. On the detection of changes in autoregressive time series, I. Asymptotics. *J. Statist. Plann. Infer.*, 137:1243–1259, 2007.
- [18] Kirch, C. *Resampling Methods for the Change Analysis of Dependent Data*. PhD thesis, Universität zu Köln, 2006. <http://kups.ub.uni-koeln.de/volltexte/2006/1795/>.
- [19] Künsch, H.R. The jackknife and the bootstrap for general stationary observations. *Ann. Statist.*, 17:1217–1241, 1989.
- [20] Lahiri, S.N. *Resampling Methods for Dependent Data*. Springer, New York, 2003.
- [21] Sakhnenko, A.I. A new way to obtain estimates in the invariance principle. In: Giné, E., Mason, D., and Wellner, J. (eds.) *High dimensional probability II*. Progress in Probability, vol. 47, pp. 223–245. Birkhäuser, Boston, 2000.
- [22] Shao, J., and Tu, D.S. *The Jackknife and Bootstrap*. Springer, New York, 1995.