Electronic Journal of Statistics
Vol. 11 (2017) 725–751
ISSN: 1935-7524
DOI: 10.1214/17-EJS1239

Some properties of the autoregressive-aided block bootstrap

Tobias Niebuhr

Fachbereich Mathematik
Universität Hamburg
e-mail: tobias.niebuhr@uni-hamburg.de

Jens-Peter Kreiss

Institut für Mathematische Stochastik
Technische Universität Braunschweig
e-mail: j.kreiss@tu-braunschweig.de

and

Efstathios Paparoditis

Mathematics and Statistics Department
University of Cyprus
e-mail: stathisp@ucy.ac.cy

Abstract: We investigate properties of a hybrid bootstrap procedure for general, strictly stationary sequences, called the autoregressive-aided block bootstrap which combines a parametric autoregressive bootstrap with a nonparametric moving block bootstrap. The autoregressive-aided block bootstrap consists of two main steps, namely an autoregressive model fit and an ensuing (moving) block resampling of residuals. The linear parametric model-fit prewhitenes the time series so that the dependence structure of the remaining residuals gets closer to that of a white noise sequence, while the moving block bootstrap applied to these residuals captures nonlinear features that are not taken into account by the linear autoregressive fit. We establish validity of the autoregressive-aided block bootstrap for the important class of statistics known as generalized means which includes many commonly used statistics in time series analysis as special cases. Numerical investigations show that the hybrid bootstrap procedure considered in this paper performs quite well, it behaves as good as or it outperforms in many cases the ordinary moving block bootstrap and it is robust against mis-specifications of the autoregressive order, a substantial advantage over the autoregressive bootstrap.

MSC 2010 subject classifications: Primary 62F40, 62M10; secondary 62P20, 91B84.

Keywords and phrases: Block bootstrap, weak ARMA, CARMA, low-frequency observations.

Received April 2016.

1. Introduction

As in many areas of statistics, also for bootstrapping time series, different nonparametric as well as parametric procedures have been proposed in the literature. On the nonparametric side one important approach is the moving block bootstrap (MBB) which is a general nonparametric resampling scheme for strictly stationary time series. It has been actively investigated since its introduction by [26] and [32]. See among others, [38, 27, 42, 33, 8, 10, 28, 35]. The MBB is a simple and important extension of Efron's i.i.d. bootstrap (see [15]) to the case of dependent observations without assuming a model structure for the underlying stochastic process. Therefore, the advantage of the MBB is its generality, i.e. the fact that it can be applied to complicated situations where parametric modeling does not provide an appropriate description of the stochastic mechanism generating the time series. Based on the idea of resampling blocks of observations rather than single observations, several variants of the MBB have been introduced and investigated in the literature; among them the non-overlapping block bootstrap (see [12]), the circular block bootstrap (see [38]), the matched block bootstrap (see [13]), and the tapered block bootstrap (see [36]). For an extensive account we refer to the monograph by [29].

On the other hand, parametric bootstrap procedures for time series, like the autoregressive bootstrap, have been also proposed and investigated by many researchers. The idea is to generate bootstrap pseudo-time series by reducing the time series at hand via a parametric model fit to a white noise sequence and then to use the fitted parametric model structure together with Efron's i.i.d. bootstrap applied to the estimated white noise innovations. See among others [18, 16, 4], and [22]. Clearly, the advantage of such a parametric bootstrap approach is its efficiency which relies of course on the assumption that the underlying process follows such a parametric model structure and that this structure has been (at least asymptotically) correctly specified for the purposes of the bootstrap.

The idea to combine the generality of nonparametric methods with the efficiency of parametric procedures in a hybrid bootstrap approach for time series is quite attractive and not new; see for instance, [14, 25] and [24]. In particular, [14] considered a hybrid or two-step bootstrap approach which is called in the following the autoregressive-aided block bootstrap (ARAB). The basic idea of the ARAB is to *prewhiten* the time series by using an autoregressive model fit, and then to apply a block bootstrap procedure to the estimated series of (centered) residuals instead of an i.i.d. resampling as this is the case of the pure autoregressive bootstrap. This procedure is motivated by the fact that while the autoregressive fit succeeds in mimicking the linear dependence structure of the process, the nonparametric resampling of the residuals captures model mis-specifications or nonlinear dependence features of the process that are not appropriately mimicked by the autoregressive fit. Although this seems to be a promising approach for bootstrapping a time series, little is known about the asymptotic properties of such a procedure, i.e. the class of statistics and processes for which it is valid and how it performs compared with the pure MBB or the pure AR bootstrap. The aim of this paper is to fill this gap.

We first investigate the asymptotic properties of the ARAB procedure and show, under rather general assumptions on the underlying stochastic process, its validity for the wide and important class of so-called generalized mean statistics. This class includes many statistics commonly used in time series analysis, like autocovariances, autocorrelations or partial autocorrelations, as special cases. Notice that the proof of the corresponding main result of this paper (see Theorem 1) is quite involved since it requires a proper truncation of infinite moving average representations that appropriately takes into account the block structure of the pseudo-innovations. We also investigate by means of simulations the finite sample behavior of the ARAB procedure and compare its performance with that of the pure block bootstrap and of the pure autoregressive bootstrap. Our numerical results show that the ARAB is able to outperform both bootstrap methods in several situations. More precisely, for all different simulated scenarios considered, only the ARAB performs equally well, while the pure block and the pure AR-sieve bootstrap, both, show a somehow changing behavior. In particular, the ARAB is able to improve the performance of the moving block bootstrap in several situations where the gains of such an improvement depend on the particular simulation set-up considered. Moreover, and interestingly enough, even for model scenarios that correspond to highly nonlinear models, the ARAB was not outperformed by the moving block bootstrap. This indicates that the ARAB is very well suited for resampling stationary time series without a strong indication of some parametric model structure. Furthermore, using the ARAB seems to be more favorable than using the pure moving block or the pure autoregressive bootstrap, especially when doubts are raised about the assumption of a pure linear autoregressive model structure.

The paper is organized as follows. Section 2 introduces some technical preliminaries and describes the ARAB procedure in detail. Its connection to already established bootstrap proposals is also briefly discussed. Section 3 states the main theorem of this paper which establishes validity of the ARAB for the class of generalized means under mild assumptions on the underlying stochastic process. Results of a simulation study comparing the performance of the different bootstrap methods considered are reported in Section 4 while our conclusions are given in Section 5. Technical details and proofs are deferred to Section 6.

2. AR-aided block bootstrap

2.1. Preliminaries and assumptions

For the time series $X_1, X_2, \dots, X_n$ we assume that it is a stretch of a process $\mathbf{X} = \{X_t, t \in \mathbb{Z}\}$ which fulfills the following assumption.

Assumption 1. *Let $\mathbf{X} = \{X_t : t \in \mathbb{Z}\}$ be a zero mean strictly stationary process with $E|X_t|^{4+\delta} < \infty$ for some $\delta > 0$. The autocovariance function $\gamma_X : \mathbb{Z} \to \mathbb{R}$, where $\gamma_X(h) = EX_tX_{t+h}$, is assumed to fulfill $\gamma_X(0) > 0$ and $\sum_{h\in\mathbb{Z}} |\gamma_X(h)| < \infty$.*

Notice that absolute summability of the autocovariance function assures that a continuous and bounded spectral density f_X of the underlying process $\mathbf{X}$ exists and is given by $f_X(\omega) = (2\pi)^{-1} \sum_{h\in\mathbb{Z}} \gamma_X(h)e^{-ih\omega}$, $\omega \in (-\pi, \pi]$.

In the following we consider the class of statistics known as generalized means, see [26], which includes many statistics used in time series analysis as special cases. Among others, the class of generalized means include versions of the sample mean, sample autocorrelations, sample autocovariances, sample partial autocorrelations and Yule-Walker estimators. Given observations $X_1, \dots, X_n$ stemming from $\mathbf{X}$, a generalized mean statistic T_n is given by

$$T_n = f\left(\frac{1}{n-m+1}\sum_{t=1}^{n-m+1} g(X_t, \dots, X_{t+m-1})\right), \tag{1}$$

where $g = (g_1, \dots, g_q)'$ and $g : \mathbb{R}^m \to \mathbb{R}^q$ for some $m \in \{1, \dots, n\}$ and $f : \mathbb{R}^q \to \mathbb{R}^{\tilde{q}}$, where $q, \tilde{q} \geq 1$. We may think of T_n as an estimator of a function of the parameter $\theta = E(g(X_1, X_2, \dots, X_m))$ associated with the process $\mathbf{X}$. Our goal is to investigate the capabilities of the ARAB to estimate the distribution of the statistic T_n. Toward this goal we assume that T_n satisfies the following assumption.

Assumption 2. *It holds that*

$$\frac{1}{\sqrt{n-m+1}}\sum_{t=1}^{n-m+1}\left(g(X_t, \dots, X_{t+m-1}) - \theta\right) \xrightarrow{\mathcal{D}}_{n\to\infty} \mathcal{N}\left(0_q, \Sigma_{q\times q}\right), \tag{2}$$

where

$$\begin{aligned}\Sigma_{q\times q} &= \lim_{n\to\infty}(n-m+1)^{-1}\cdot Var\left(\sum_{t=1}^{n-m+1} g(X_t, \dots, X_{t-m+1})\right)\\ &= \left(\sum_{k=-\infty}^{\infty} Cov(g_u(X_0, \dots, X_{m-1}), g_v(X_k, \dots, X_{k+m-1}))\right)_{u,v=1,\dots,q}.\end{aligned}$$

Clearly, if the functions g and f fulfill certain smoothness conditions (cf. Assumption 3), then the weak convergence (2) and the delta-method yield

$$\sqrt{n}\left(T_n - f(\theta)\right) \xrightarrow{\mathcal{D}}_{n\to\infty} f'(\theta)\,\mathcal{N}\left(0_q, \Sigma_{q\times q}\right), \tag{3}$$

where $f'(\cdot) := Df(\cdot)$ is the Jacobi matrix of f. The following assumption specifies the conditions imposed on the functions f and g.

Assumption 3. *The function $f = (f_1, \dots, f_{\tilde{q}})$ has continuous partial derivatives $y \mapsto \sum_{i=1}^q \frac{\partial f_u}{\partial x_i}\big|_{x=y} y_i$, for all $u = 1, \dots, \tilde{q}$, $i = 1, \dots, m$, for y in a neighborhood of θ and the differentials at θ, $y \mapsto \sum_{i=1}^q \frac{\partial f_u}{\partial x_i}\big|_{x=\theta} y_i$, for all $u = 1, \dots, \tilde{q}$, do not vanish. Further, the function g has continuous partial derivatives of order h, where h is an integer greater than or equal to one. Furthermore,*

$$\left|\frac{\partial^h g_i(x)}{\partial x_{i_1}\dots\partial x_{i_h}}\Big|_{x=y} - \frac{\partial^h g_i(x)}{\partial x_{i_1}\dots\partial x_{i_h}}\Big|_{x=z}\right| \leq C_u\|y - z\|, \tag{4}$$

where $u = 1, \dots, q$; $1 \le i_1, \dots, i_h \le m$, C_u is some suitable constant and $\| \cdot \|$ denotes the Euclidean norm and $x, y, z \in \mathbb{R}^m$.

Assumptions 2 and 3 guarantee the asymptotic normality of $\sqrt{n-m+1}(T_n - f(\theta))$ which is a minimal requirement and it is satisfied if the underlying process $\mathbf{X}$ fulfills a variety of weak dependence conditions. Assumption 3 is in line with the assumptions imposed in [19] and [9] for dealing with this type of statistics in the context of the block bootstrap.

The next assumption requires the $\sqrt{n}$-consistency of a fixed number of sample autocovariances $\widehat{\gamma}(h) = n^{-1} \sum_{t=1}^{n-|h|} (X_t - \overline{X}_n)(X_{t+|h|} - \overline{X}_n)$, $0 \le h \le p$, where $\overline{X}_n = n^{-1} \sum_{t=1}^{n} X_t$ denotes the sample mean and it is also satisfied for a wide range of weakly dependent stationary processes.

Assumption 4. *The autocovariance estimator $\widehat{\gamma}(h)$ satisfies for every fixed $p \in \mathbb{N}$,*

$$\sqrt{n}(\widehat{\gamma}_X(h) - \gamma_X(h)) = \mathcal{O}_P(1) \text{ for } h = 0, \dots, p. \tag{5}$$

2.2. The AR-aided block bootstrap algorithm

We describe in this section the ARAB procedure. This bootstrap procedure generates replicates T_n^* of the statistic T_n which are obtained by replacing in (1) the observed time series $X_1, X_2, \dots, X_n$ by the pseudo-time series $X_1^*, \dots, X_n^*$ generated by the following algorithm. Notice that in this algorithm the positive integer l denotes the block length while b the total number of blocks. Furthermore, we assume that $bl \ge n + p$.

Step 1: *Fit an autoregressive process of order p to the time series $X_1, \dots, X_n$ and denote by $\hat{a}_1(p), \dots, \hat{a}_p(p)$ the estimated autoregressive parameters obtained by solving the Yule-Walker equations, i.e., $(\hat{a}_1(p), \dots, \hat{a}_p(p))' = \widehat{\mathbf{\Gamma}}(p)^{-1} \widehat{\gamma}_p$, where $\widehat{\mathbf{\Gamma}} = (\widehat{\gamma}(i-j))_{i,j=1,2,\dots,p}$ is a $p \times p$ matrix and $\widehat{\gamma}_p = (\widehat{\gamma}(1), \widehat{\gamma}(2), \dots, \widehat{\gamma}(p))'$. Let $\widehat{U}_t$, $t = p+1, p+2, \dots, n$, be the residuals of this fit, i.e.,*

$$\hat{U}_t = X_t - \sum_{j=1}^{p} \hat{a}_j(p) X_{t-j}. \tag{6}$$

Step 2: *Center the estimated residuals $\hat{U}_{p+1}, \dots, \hat{U}_n$ by*

$$\bar{\hat{U}} := \frac{1}{n-p-l+1} \left[\sum_{i=1+p}^{n} \hat{U}_i - \sum_{i=1}^{l-1} \frac{i}{l} \hat{U}_{p+l-i} - \sum_{i=1}^{l-1} \frac{i}{l} \hat{U}_{n-l+1+i} \right]. \tag{7}$$

and let $\hat{U}_t^c = \hat{U}_t - \bar{\hat{U}}$, $t = p+1, \dots, n$.

Step 3: *Generate bootstrap residuals $\left(U_{1-p}^*, U_{2-p}^*, \dots, U_n^*\right)$ by application of the moving block bootstrap to the sequence $\hat{U}_{p+1}^c, \dots, \hat{U}_n^c$. That is, draw with replacement b independent identically distributed random variables $i_1, \dots, i_b$ which* *have the same discrete uniform distribution on the set* $\{p, \ldots, n-l\}$ *and let*

$$\left(U_{1-p}^*, \ldots, U_n^*\right) = \left(\hat{U}_{i_1+1}^c, \ldots, \hat{U}_{i_1+l}^c, \hat{U}_{i_2+1}^c, \ldots, \hat{U}_{i_2+l}^c, \ldots, \ldots, \hat{U}_{i_b+l}^c\right). \quad (8)$$

For convenience we set in the following $U_t^* \equiv 0$ *for all* $t < 1-p$.

Step 4: *Compute a bootstrap time series* $X_1^*, X_2^*, \ldots, X_n^*$ *via*

$$X_t^* = \sum_{j=1}^{p} \hat{a}_j(p) X_{t-j}^* + U_t^*, \qquad t = 1, \ldots, n. \quad (9)$$

Obtain the bootstrap analogue of T_n *as*

$$T_n^* = f\left(\frac{1}{n-m+1} \sum_{t=1}^{n-m+1} g(X_t^*, \ldots, X_{t+m-1}^*)\right) \quad (10)$$

and approximate the distribution of $\sqrt{n}(T_n - f(\theta))$ *by that of* $\sqrt{n}(T_n^* - f(\theta^*))$, *where* $\theta^* = E^*(g(X_t^*, X_{t+1}^*, \ldots, X_{t+m-1}^*))$.

Some remarks are in order. Notice first that we do not attempt to reduce the time series to residuals which are then used to imitate i.i.d. innovations. Because of this, the present approach differs from purely autoregressive approaches based on i.i.d. resampling of estimated (and appropriately centered) residuals; cf. [18, 16, 4, 22]. In fact, the AR fit applied is in the spirit of the so-called *prewhitening* idea (see [40]) and is used to capture the autocovariance structure of the time series $X_1, X_2, \ldots, X_n$ at hand. Fitting an autoregression should be seen as a (convenient) approach for such a whitening although other parametric models may be considered for the same purpose, too. However, autoregressive fits together with the use of Yule-Walker estimators are rather convenient and have nice features including the property that all roots of the estimated polynomial $\hat{A}(z) = 1 - \sum_{j=1}^{p} \hat{a}_j(p) z^j$ lie outside the unit disc in the complex plane.

An important issue when using autoregressive fits is the choice of the order p. The motivation is to balance the complexity of the model and the volatility of the resulting residuals. The most convenient way to address this conflict is according to Akaike's information criterion (AIC) (see [1, 2]). While [7] stated that AIC tends to select orders that are too high, even asymptotically, [21] found a modification of the AIC which assures for consistency. However, since it is not necessary for the ARAB to assure for correctly specified model orders, we propose using the classical AIC.

We mention that the centering in Step 2 of the bootstrap algorithm is based on the expected value of U_t^* with respect to the ensuing moving block bootstrap. Thus, expression (7) differs from the common sample mean (to which it is reduced for $l = 1$) and ensures that the block bootstrapped residuals have zero mean (with respect to the moving block bootstrap distribution) to avoid an unnecessary bias; see Step 3.

3. Bootstrap validity

We now derive the main result of this paper which shows that the ARAB is valid for the class of generalized means described in (1). Recall that the ARAB procedure generates pseudo-time series $X_1^*, \dots, X_n^*$ by means of (9). The estimated autoregressive polynomial $1-\sum_{j=1}^{p} \hat{a}_j(p)z^j$, has a theoretical counterpart given by

$$A(z) = 1 - \sum_{j=1}^{p} a_j(p)z^j, \tag{11}$$

where the coefficients $\hat{a}_j(p)$ are determined by

$$(a_1(p), \dots, a_p(p))' = \underset{(c_1,\dots,c_p)}{\operatorname{argmin}} E\left[\left(X_t - \sum_{j=1}^{p} c_j X_{t-j}\right)^2\right], \tag{12}$$

i.e., $(a_1(p), \dots, a_p(p))' = \mathbf{\Gamma}(p)^{-1}\gamma_p$, where $\mathbf{\Gamma} = (\gamma(i-j))_{i,j=1,2,\dots,p}$ is a $p \times p$ matrix and $\gamma_p = (\gamma(1), \gamma(2), \dots, \gamma(p))'$. For the proofs, it is more convenient to use an infinite moving-average representation of X_t which is based on the inverse polynomial

$$A^{-1}(z) = \left(1 - \sum_{j=1}^{p} a_j(p)z^j\right)^{-1} = \sum_{j=0}^{\infty} \psi_j(p)z^j =: \Psi(z), \tag{13}$$

where $\psi_0(p) = 1$. Notice that an analogue expression appears using the inverse $\widehat{A}^{-1}(z)$ of the estimated autoregressive polynomial $\widehat{A}(z) = 1-\sum_{j=1}^{p} \widehat{a}_j(p)z^j$. Let $\hat{\psi}_j(p)$ be the coefficients appearing in this inverse polynomial. In the following we use for brevity the notation ψ_j for $\psi_j(p)$ and $\hat{\psi}_j$ for $\hat{\psi}_j(p)$ respectively.

Notice that the inverse autoregressive polynomial $\Psi(z)$ given in (13) is bounded away from zero for all $|z| \le 1$, $z \in \mathbb{C}$, and that the coefficients ψ_j fulfill the summability condition $\sum_{j=0}^{\infty} j^r|\psi_j| < \infty$ for some $r \ge 0$. This follows directly using Cauchy's inequality for holomorphic functions (e.g. [17], Theorem 6.1).

Application of the filter $A(L) = 1 - \sum_{j=1}^{p} a_j(p)L^j$ to the series $\mathbf{X}$, where L denotes the backshift operator leads to the filtered time series $\{U_t : t \in \mathbb{Z}\}$ with $U_t = X_t - \sum_{j=1}^{p} a_j(p)X_{t-j}$, which is the process analogue of the estimated residual series $\hat{U}_t$ given in (6). The following assumption specifies our requirements on the moments of U_t and their sample estimators.

Assumption 5. *The filtered process* $\{U_t : t \in \mathbb{Z}\}$ *fulfills*

$$E\left[U_t^{4h}\right] < \infty, \tag{14}$$

where h is the number of continuous partial derivatives of the function g; see Assumption 3 for details. Furthermore,

$$\frac{1}{n-p}\sum_{t=1}^{n-p}\hat{U}_t^{\alpha} \xrightarrow{P} E\left[U_t^{\alpha}\right], \quad \text{for every} \quad \alpha \leq 4h \tag{15}$$

and for F_U denoting the distribution function of $\{U_t : t \in \mathbb{Z}\}$ and $\hat{F}_U$ denoting the distribution function of its estimated counterpart, it holds $\hat{F}_U \xrightarrow{\mathcal{D}} F_U$, as $n \to \infty$.

The following theorem establishes validity of the autoregressive-aided block bootstrap in estimating the distribution of generalized means.

Theorem 1. *Suppose that Assumptions 1–5 hold with $\delta \geq 4h-4$. T_n^* is defined as in (10), and $\Sigma_{q\times q}$ as in (2). Further assume that $l \to \infty$ and $l^{2+2/\delta}/n \to 0$, as $n \to \infty$. Then, as $n \to \infty$,*

$$(n-m+1)^{-1}\, Var^*\left(\sum_{t=1}^{n-m+1}\left(g(X_t^*,\ldots,X_{t+m-1}^*)-\theta^*\right)\right) \to \Sigma_{q\times q} \tag{16}$$

and

$$\sqrt{n}(T_n^* - f(\theta^*)) \xrightarrow{\mathcal{D}} f'(\theta)\,\mathcal{N}\left(0_q, \Sigma_{q\times q}\right), \tag{17}$$

in probability.

Notice that (3) and (17) together with the continuity of the distribution function of the Gaussian distribution, leads under the assumptions of the above theorem to the result that

$$\sup_{x\in\mathbb{R}^{\tilde{q}}}\left|P^*(\sqrt{n}(T_n^*-f(\theta^*)) \leq x) - P(\sqrt{n}(T_n - f(\theta)) \leq x)\right| = o_P(1), \tag{18}$$

as $n \to \infty$.

According to the above theorem, validity of the ARAB holds as long as $l^{2+2/\delta}/n \to 0$ as $n \to \infty$ is fulfilled. The parameter δ depends on characteristics of the underlying process $\mathbf{X}$ and is the same as the one appearing in the moment condition $E|X_t|^{4+\delta} < \infty$ in Assumption 1. Thus, the condition on the block length holds true if $l = n^{1/2-\varepsilon}$ for any ε satisfying $1/(2+2\delta) < \varepsilon < 1/2$.

The ARAB is tailor-made but not restricted to AR processes with non-independent identically distributed innovations, so called weak AR or more generally weak (autoregressive moving-average) ARMA processes. Such processes arise in several situations like of instance, strong ARMA processes with non-normal distributed noise observed at lower frequencies, see [34], processes obtained from observing only the first component of a vector-valued AR process or equidistant sampling of continuous-time ARMA processes (e.g. [6]).

Furthermore, the usual model-based bootstrap approaches typically assume that the model structure has been correctly specified. This is different for the

ARAB. There is neither an assumption of an autoregressive representation of the underlying process $\mathbf{X}$ nor it is assumed in the case that $\mathbf{X}$ has such an AR representation, that the autoregressive order has been correctly specified. The AR model fit solely *prewhitens* the time series at hand and, therefore, weakens its dependence structure leading to residuals to which a block bootstrap procedure is applied. Thus, even in the case that $\mathbf{X}$ has a finite order AR representation it is not essential for ARAB validity, that the autoregressive order has been correctly specified. In this sense, the ARAB makes the ordinary residual based AR-bootstrap robust against model mis-specifications.

One important problem in applying the block bootstrap is the choice of the block length l. This problem is common in all blocking methods, and the conditions stated for deriving the asymptotic results presented in Theorem 1 do not provide some guidance on how to choose l in practice. [29] considered the problem of determining an optimal block length for the class of smooth functions of means while further approaches for choosing this bootstrap parameter are given in [20, 11, 39, 30] or [37].

4. Numerical examples

In this section we investigate by means of numerical simulations the finite-sample behavior of the autoregressive-aided block bootstrap (ARAB) and compare its performance with that of the autoregressive residual based bootstrap (AR) and of the moving block bootstrap (MBB). The MBB and the ARAB are applicable to a wide class of stationary and weak dependent time series while the AR bootstrap is tailor-made for but restricted to autoregressive processes driven by i.i.d. innovations. Taking this into account, we consider several scenarios with different models, namely finite order autoregressive, non-linear autoregressive and autoregressive conditional heteroscedastic (ARCH) models, and consider the estimation of quantiles of the distribution of the second order autocorrelation estimator $\hat{\rho}(2) = \hat{\gamma}(2)/\hat{\gamma}(0)$. Regarding the bootstrap estimates, the three different methods are compared by presenting boxplots of the results together with the true finite sample quantiles. The true quantiles, which are the target parameters of interest, have been estimated using $1,000,000$ replications of the models considered (or $100,000$ replications in the ARCH case respectively) and are shown in the figures presented by a dashed vertical line. We report results for different choices of the block size l while the autoregressive order used has been selected by means of AIC (see [1]).

4.1. Linear AR time series

We generate time series from the AR(2) process

$$X_t = -0.9X_{t-1} - 0.7X_{t-2} + e_t, \tag{19}$$

where e_t are i.i.d., standard Gaussian distributed. We are interested in estimating the 5% and 95% quantile of the distribution of

$$\sqrt{n}(\hat{\rho}_X(2) - E[\hat{\rho}_X(2)]). \tag{20}$$

For this, $K = 500$ time series of length $n = 100$ of the AR(2) model have been generated and for each one of them the three different bootstrap approaches have been applied with $B = 1,000$ bootstrap repetitions. For the MBB the block lengths $l = 8, 10, 12$ and 14 have been chosen and for the ARAB the block sizes $l = 3, 5, 8$ and 10 have been used.

The simulation results are presented in Figure 1. As it can be seen, all approaches capture the target values quite satisfactory when the block lengths are appropriately chosen. Since the autoregressive model fit captures (at least to some extend) the dependence structure of the process, the pre-whitened series becomes closer to an i.i.d. sequence. This is the reason why the ARAB requires shorter blocks than the MBB to capture the time series dependence. Recall that therefore this model set-up is tailor-made for the AR and the ARAB bootstrap and, their performances are not surprising. It is, however, remarkable that the ARAB is less biased than the AR approach for some choices of l.

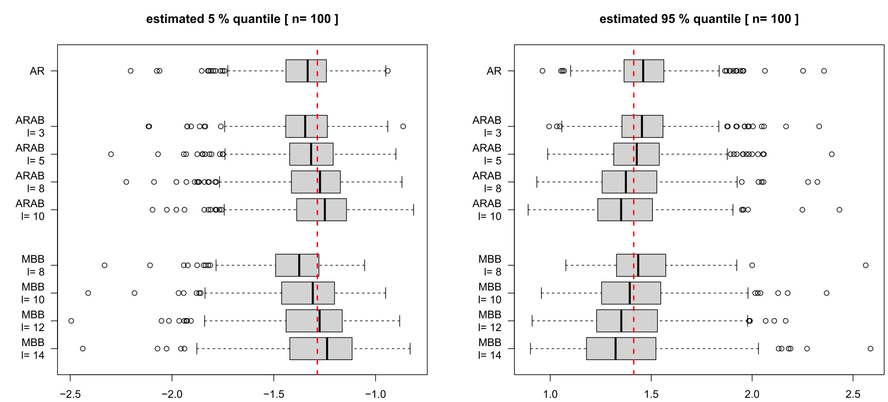

FIG 1. *AR time series: Boxplots of bootstrap quantile estimates using the AR, the ARAB and the MBB bootstrap. The true finite sample quantile is presented by the dashed red line.*

4.2. Nonlinear time series

We next consider the following non-linear autoregressive model also used in [36] and [43]:

$$X_t = 0.6 \sin(X_{t-1}) + e_t, \quad t \in \mathbb{Z}, \tag{21}$$

where the e_t's are i.i.d. standard Gaussian distributed random variables. As a quantity of interest we again consider estimating the distribution (20) and in particular the 5% and 95% quantiles of this distribution with $n = 200$. Notice that for this nonlinear autoregressive model, the linear pre-whitening cannot capture the entire dependence structure of the time series. For this reason the block lengths have been selected to be $l = 3, 5, 8, 10$ for the ARAB and the MBB. The further simulation parameters read the same as in the previous example. The simulation results are presented in Figure 2. As it is seen, the linear AR bootstrap behaves quite satisfactory even in this case. The MBB seems to be prone to underestimate the values of the quantiles while the preliminary autoregressive fit moves the ARAB estimates in the right direction.

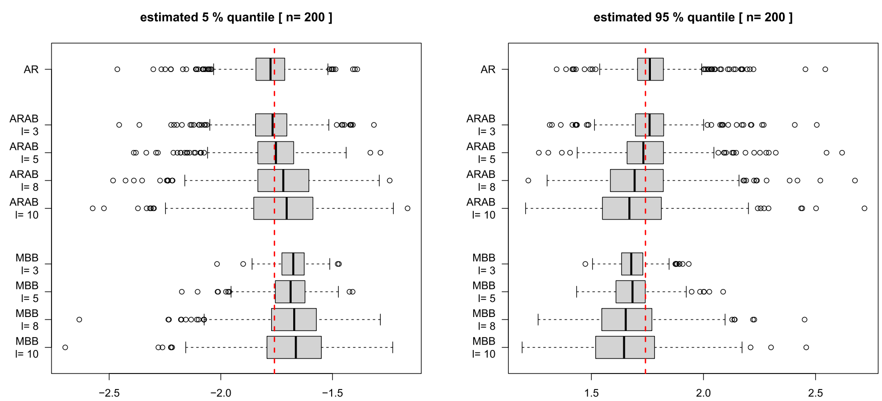

FIG 2. *Nonlinear AR Time Series: Boxplots of bootstrap quantile estimates using the AR, the ARAB and the MBB bootstrap. The true finite sample quantile is presented by the dashed red line.*

4.3. ARCH time series

We finally consider the following ARCH model

$$\begin{aligned} X_t &= \sigma_t \varepsilon_t \\ \sigma_t^2 &= 1 + 0.25 \cdot X_{t-1}^2, \qquad t \in \mathbb{Z}, \end{aligned} \tag{22}$$

for which an autoregressive fit is not appropriate to capture its dependence structure due to the white noise property of its realizations. Here the e_t's are i.i.d. standard Gaussian distributed innovations. We consider estimating the 5% and 95% quantiles of the distribution (20). $K = 400$ replications of length $n = 2,000$ of this model have been generated and to each of them the three different bootstrap methods have been applied using $B = 800$ bootstrap replications. For the same arguments as for the nonlinear autoregressive model, the block sizes $l = 8, 12, 16$ and $l = 20$ have been considered for both block bootstrap approaches. The simulation results are shown in Figure 3. Since the ARCH model contains no linear part, the AR bootstrap cannot capture the target values. However, the ARAB bootstrap performs competitive to the MBB also in this case and leads to estimates which are not outperformed by the MBB bootstrap.

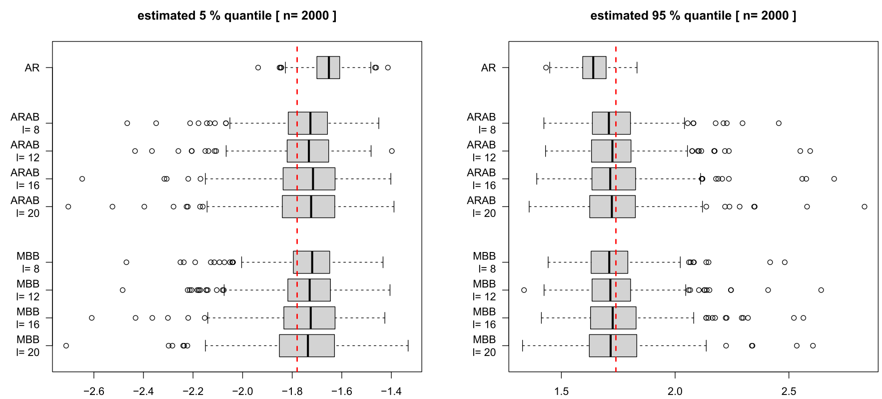

Fig 3. *ARCH Time Series: Boxplots of bootstrap quantile estimates using the AR, the ARAB and the MBB bootstrap. The true finite sample quantile is presented by a dashed red line.*

5. Conclusions

This paper investigated properties of a hybrid resampling method for time series, called the autoregressive-aided block bootstrap (ARAB), which consists of a linear autoregressive bootstrap combined with a block resampling of residuals. It has been shown that this bootstrap procedure is valid for general weak dependent stationary processes and for a wide class of statistics. The method shares the generality of nonparametric resampling procedures. Moreover, the nonparametric resampling part of the ARAB makes this resampling procedure robust against mis-specifications of the model structure, a problem from which purely parametric bootstrap methods suffer. On the other hand, the parametric fit makes the ARAB procedure less sensitive to the choice of nonparametric resampling parameters like the block size. These attractive features of the ARAB have been demonstrated by means of simulations with the ARAB procedure showing a very good finite sample performance illustrating its aforementioned attractive features.

Acknowledgements

The authors are indebted to the editor, the associate editor and two anonymous referees for detailed comments which lead to an improved version of the paper.

6. Proofs and auxiliary results

The main strategy for proving the validity of the ARAB bootstrap is the following. In a first step, we state Lemma 1 which allows for a special truncation of the process' (weak) moving average representation. An ordinary truncation of the moving average representation at some fixed $M \in \mathbb{N}$ is not sufficient for the proof (i.e. Proposition 6.3.9 of [5] cannot be applied). Due to the underlying block structure of the innovations the present proof requires a truncation point that depends on n and - more important - on the time point t in relation to its position in the block of pseudo-innovations. This turns out to be highly relevant for the computation of the limiting covariance matrix of the estimator.

Lemma 1. *Assume that $X_1^*, \dots, X_n^*$ is a bootstrap sequence generated as described in Section 2.2. Furthermore, let Assumption 1 with some $\delta \geq 4h-4$ and Assumptions 2–5 hold true. Then, the bootstrap time series has the representation*

$$X_t^* = \sum_{j=0}^{\infty} \hat{\psi}_j(p) U_{t-j}^* \quad \forall t = 1, \dots, n. \tag{23}$$

Define

$$\check{X}_t^* = \sum_{j=0}^{(t-1) \bmod l + M(n)} \hat{\psi}_j(p) U_{t-j}^* \tag{24}$$

as a truncated version of X_t^*, where $M(n)$ fulfills $M(n) \leq l$ and $M(n) \to \infty$, as $n \to \infty$. Then, it holds

$$E^* \left[\left(\frac{1}{\sqrt{n-m+1}} \sum_{t=1}^{n-m+1} \left(g\left(\underline{X}_t^*\right) - g\left(\underline{\check{X}}_t^*\right) \right) \right)^2 \right] \to_{n \to \infty} 0 \quad i.p. \tag{25}$$

Consider $\check{X}_t^*$ as given in (24) which is a modified version of X_t^* and truncated at some time point depending on the time index t. In particular, the truncation is done at a point being $M(n)$ steps in the past of the beginning of the block in which the present time point t is located. Thus, the truncated bootstrap observation at time point $t = (r-1)l + s$, where $s \in \{1, \dots, l\}$, is truncated $M(n)$ steps in front of the beginning of the $(r-1)$-th block. Thus, the infinite moving average representation reduces to $s + M(n)$ summands. Since $M(n) \leq l$ the truncated version only correlates with two blocks of residuals. This is of advantage especially for computing covariances as required in the proof of Theorem 1.

Proof of Lemma 1. Denote $\underline{X}_t^* = (X_t^*, \dots, X_{t+m-1}^*)$ and $\underline{\check{X}}_t^* = (\check{X}_t^*, \dots, \check{X}_{t+m-1}^*)$. Now expand $g\left(\underline{X}_t^*\right)$ at $g\left(\underline{\check{X}}_t^*\right)$ by using a Taylor series to obtain

$$\begin{aligned} & g\left(X_t^*, \dots, X_{t+m-1}^*\right) - g\left(\check{X}_t^*, \dots, \check{X}_{t+m-1}^*\right) \\ &= \sum_{1 \leq |\alpha| \leq h-1} \frac{1}{\alpha!} D^\alpha g\left(\underline{\check{X}}_t^*\right) \left(\underline{X}_t^* - \underline{\check{X}}_t^*\right)^\alpha + \sum_{|\alpha| = h} \frac{1}{\alpha!} D^\alpha g\left(\underline{\tau}\right) \left(\underline{X}_t^* - \underline{\check{X}}_t^*\right)^\alpha \end{aligned} \tag{26}$$

for some $\underline{\tau} = \underline{\check{X}}_t^* + \lambda\left(\underline{X}_t^* - \underline{\check{X}}_t^*\right)$, $\lambda \in [0,1]$, where $\alpha \in \mathbb{N}_0^m$, $|\alpha| = \alpha_1 + \ldots + \alpha_m$, $(\underline{x})^\alpha = x_1^{\alpha_1} \cdot \ldots \cdot x_m^{\alpha_m}$ and $D^\alpha g(\underline{x}) := \frac{\partial^{|\alpha|} g(\underline{y})}{\partial y_1^{\alpha_1} \cdot \ldots \cdot y_m^{\alpha_m}}|_{\underline{y}=\underline{x}}$. Next, denote by $\|\cdot\|_r$ the usual $\mathcal{L}_r$-norm with respect to the bootstrap distribution. From (26) we get

$$\begin{aligned}
&\left\|g\left(\underline{X}_t^*\right) - g\left(\underline{\check{X}}_t^*\right)\right\|_2 \\
&\leq \sum_{1\leq|\alpha|\leq h-1} \frac{1}{\alpha!}\left\|D^\alpha g\left(\underline{\check{X}}_t^*\right)\left(\underline{X}_t^* - \underline{\check{X}}_t^*\right)^\alpha\right\|_2 \\
&\quad + \sum_{|\alpha|=h} \frac{1}{\alpha!}\left\|D^\alpha g(\underline{\tau})\left(\underline{X}_t^* - \underline{\check{X}}_t^*\right)^\alpha\right\|_2 .
\end{aligned} \tag{27}$$

It should be noted that $\underline{\check{X}}_t^*$ and $\underline{X}_t^* - \underline{\check{X}}_t^*$ are not independent due to the underlying blocks of innovations. Using Cauchy Schwarz's inequality the summands of the first sum yield

$$\left\|D^\alpha g(\underline{X}_t^*)\left(\underline{X}_t^* - \underline{\check{X}}_t^*\right)^\alpha\right\|_2 \leq \left(\|D^\alpha g(\underline{X}_t^*)\|_4 \left\|\left(\underline{X}_t^* - \underline{\check{X}}_t^*\right)^\alpha\right\|_4\right)^{1/2} \tag{28}$$

By Assumption 3 the h-th derivative of $g(\cdot)$ is Lipschitz and, therefore, one obtains for the summands of the second sum

$$\begin{aligned}
&\left\|D^\alpha g(\underline{\tau})\left(\underline{X}_t^* - \underline{\check{X}}_t^*\right)^\alpha\right\|_2 \\
&\leq \left\|D^\alpha g(\underline{\check{X}}_t^*)\left(\underline{X}_t^* - \underline{\check{X}}_t^*\right)^\alpha\right\|_2 + \left\|\left(D^\alpha g(\underline{\tau}) - D^\alpha g(\underline{\check{X}}_t^*)\right)\left(\underline{X}_t^* - \underline{\check{X}}_t^*\right)^\alpha\right\|_2 \\
&\leq \left\|D^\alpha g\left(\underline{\check{X}}_t^*\right)\right\|_4^{1/2} \cdot \left\|\left(\underline{X}_t^* - \underline{\check{X}}_t^*\right)^\alpha\right\|_4^{1/2} \\
&\quad + \left(C_g \lambda \sum_{j=0}^{m-1} \left\|X_{t+j}^* - \check{X}_{t+j}^*\right\|_4 \cdot \left\|\left(\underline{X}_t^* - \underline{\check{X}}_t^*\right)^\alpha\right\|_4\right)^{1/2},
\end{aligned} \tag{29}$$

for some suitable constant C_g, e.g. $C_g := \max_{u=1,\ldots,q} C_u$. Using (28) and (29) gives

$$\begin{aligned}
&\left\|\frac{1}{\sqrt{n-m+1}} \sum_{t=1}^{n-m+1}\left(g\left(\underline{X}_t^*\right) - g\left(\underline{\check{X}}_t^*\right)\right)\right\|_2 \\
&\leq \frac{1}{\sqrt{n-m+1}} \sum_{i=1}^{b} \sum_{k=1}^{l-m+1} \\
&\quad \left(\sum_{1\leq|\alpha|\leq h-1} \frac{1}{\alpha!}\left\|D^\alpha g\left(\underline{\check{X}}_{(i-1)l+k}^*\right)\right\|_4^{1/2} \cdot \left\|\left(\underline{X}_{(i-1)l+k}^* - \underline{\check{X}}_{(i-1)l+k}^*\right)^\alpha\right\|_4^{1/2}\right. \\
&\quad + \sum_{|\alpha|=h} \frac{1}{\alpha!}\left\|D^\alpha g\left(\underline{\check{X}}_{(i-1)l+k}^*\right)\right\|_4^{1/2} \cdot \left\|\left(\underline{X}_{(i-1)l+k}^* - \underline{\check{X}}_{(i-1)l+k}^*\right)^\alpha\right\|_4^{1/2}
\end{aligned} \tag{30}$$

$$+\sum_{|\alpha|=h}\frac{1}{\alpha!}\left(C_g\lambda\sum_{v=0}^{m-1}\left\|\left(X^*_{(i-1)l+k+v}-\check{X}^*_{(i-1)l+k+v}\right)\right\|_4\right.$$
$$\left.\times\left\|\left(\underline{X}^*_{(i-1)l+k}-\underline{\check{X}}^*_{(i-1)l+k}\right)^\alpha\right\|_4\right)^{1/2}\Bigg)+o_P(1).$$

Furthermore, by Assumptions 1 and 3 and by the definitions of $D^\alpha g(\cdot)$ and $\underline{\check{X}}^*_t$, it holds

$$\left\|D^\alpha g\left(\underline{\check{X}}^*_t\right)\right\|_4=\mathcal{O}_P(1) \tag{31}$$

for any $|\alpha|\leq h$. The same holds true for $\underline{\check{X}}^*_t$ in (31) replaced by $\underline{X}^*_t$. Next we consider the further terms in (30). We have

$$\begin{aligned}&\left\|\left(\underline{X}^*_{(i-1)l+k}-\underline{\check{X}}^*_{(i-1)l+k}\right)^\alpha\right\|_4\\&=\left\|\left(\sum_{j=((i-1)l+k)\bmod l+M(n)+1}^{\infty}\hat{\psi}_j\underline{U}^*_{(i-1)l+k-j}\right)^\alpha\right\|_4\end{aligned}\tag{32}$$

and the modulo condition can be simplified since $((i-1)l+k) \bmod l+M(n)+1=k+M(n)+1$. Furthermore, it is well-known (e.g. [23]) that the coefficients ψ_j (and $\hat{\psi}_j$) of the inverse autoregressive polynomial (and its estimated counterpart) uniformly yield

$$|\psi_j|\leq C\rho^j,\ \ \forall j\in\mathbb{Z},\ C>0 \tag{33}$$

for some $0<\rho<1$ which depends on the autoregressive parameters. By Assumption 5 and noting that $|\alpha|\leq h$, (32) further computes to

$$\begin{aligned}&\left\|\left(\sum_{j=k+M(n)+1}^{\infty}\hat{\psi}_j\underline{U}^*_{(i-1)l+k-j}\right)^\alpha\right\|_4\\&\leq\prod_{w=1}^{m}\left\|\left(\sum_{j=k+M(n)+1}^{\infty}\hat{\psi}_jU^*_{(i-1)l+k-j}\right)^{\alpha_w}\right\|_4\\&\leq\prod_{w=1}^{m}\sum_{j_1,\dots,j_{\alpha_w}=k+M(n)+1}^{\infty}\rho^{j_1}\cdot\ldots\cdot\rho^{j_{\alpha_w}}\\&\qquad\times\left\|U^*_{(i-1)l+k-j_1}\cdot\ldots\cdot U^*_{(i-1)l+k-j_{\alpha_w}}\right\|_4\cdot\mathcal{O}(1)\\&\leq\prod_{w=1}^{m}\sum_{j_1,\dots,j_{\alpha_w}=k+M(n)+1}^{\infty}\rho^{j_1}\cdot\ldots\cdot\rho^{j_{\alpha_w}}E^*\left[(U^*_1)^{4|\alpha|}\right]\cdot\mathcal{O}(1)\\&=\rho^{(k+M(n)+1)|\alpha|}\cdot C'\cdot\mathcal{O}_P(1),\end{aligned}\tag{34}$$

where the constant C' can be easily obtained. The further terms in (30) can be handled along the same lines, where the corresponding constants may change. Thus we get

$$
\begin{aligned}
&\left\| \frac{1}{\sqrt{n-m+1}} \sum_{t=1}^{n-m+1} \left(g\left(\underline{X}_t^*\right) - g\left(\underline{\check{X}}_t^*\right)\right) \right\|_2 \qquad (35) \\
&\leq \frac{1}{\sqrt{n-m+1}} \sum_{i=1}^{b} \sum_{k=1}^{l-m+1} \Bigg(\sum_{1\leq|\alpha|\leq h-1} \frac{1}{\alpha!} \cdot \mathcal{O}_P(1) \cdot \rho^{(k+M(n)+1)|\alpha|\cdot\frac{1}{2}} \\
&\quad + \sum_{|\alpha|=h} \frac{1}{\alpha!} \cdot \mathcal{O}_P(1) \cdot \rho^{(k+M(n)+1)|\alpha|\cdot\frac{1}{2}} \\
&\quad + \sum_{|\alpha|=h} \frac{1}{\alpha!} \left(\mathcal{O}(1) \cdot \sum_{v=0}^{m-1} \rho^{k+v+M(n)+1} \cdot \rho^{(k+v+M(n)+1)|\alpha|} \right)^{1/2} \Bigg) + o_P(1) \\
&= \mathcal{O}_P\left(\frac{b}{\sqrt{n}}\, \rho^{M(n)} \right)
\end{aligned}
$$

which converges to zero in probability as $M(n) \to \infty$ and concludes the proof. □

Proof of Theorem 1. By Lemma 1 it suffices to show that the assertions of the theorem hold true for T_n^*, where X_t^* is replaced by $\check{X}_t^*$. For the truncated series we choose $M(n)$ such that $M(n)^2 l^{-1} \to 0$ as $n \to \infty$.

First, consider (16), the limiting covariance matrix. As discussed at the beginning of the proof of Lemma 1, each $\check{X}_t^*$ only is correlated with at least two blocks of residuals. Thus, two bootstrap realizations $\check{X}_t^*$ and $\check{X}_s^*$ either stem from the same block, or $\check{X}_t^*$ stems from the block before or after $\check{X}_s^*$, or they are uncorrelated. Hence,

$$
\begin{aligned}
&Cov^*\left(\frac{1}{\sqrt{n-m+1}} \sum_{t=1}^{n-m+1} g_u\left(\underline{\check{X}}_t^*\right), \frac{1}{\sqrt{n-m+1}} \sum_{s=1}^{n-m+1} g_v\left(\underline{\check{X}}_s^*\right) \right) \qquad (36) \\
&\quad = \frac{1}{n-m+1}\, Cov^*\Bigg(\sum_{i_1=1}^{b} \sum_{k_1=1}^{l-m+1} g_u\left(\underline{\check{X}}_{(i_1-1)l+k_1}^*\right), \\
&\qquad\qquad \sum_{i_2=1}^{b} \sum_{k_2=1}^{l-m+1} g_v\left(\underline{\check{X}}_{(i_2-1)l+k_2}^*\right) \Bigg) + o_P(1) \\
&\quad =: R_1 + R_2 + R_3 + o_P(1),
\end{aligned}
$$

where

$$
R_1 \quad := \quad \frac{b}{n-m+1} \sum_{k_1=1}^{l-m+1} \sum_{k_2=1}^{l-m+1} Cov\left(g_u\left(\underline{X}_{k_1}\right), g_v\left(\underline{X}_{k_2}\right)\right) \qquad (37)
$$

$$+\frac{b}{n-m+1}\sum_{k_1=1}^{l-m+1}\sum_{k_2=1}^{l-m+1} Cov\left(g_u\left(\underline{\check{X}}_{k_1}\right)-g_u\left(\underline{X}_{k_1}\right), g_v\left(\underline{X}_{k_2}\right)\right)$$
$$+\frac{b}{n-m+1}\sum_{k_1=1}^{l-m+1}\sum_{k_2=1}^{l-m+1} Cov\left(g_u\left(\underline{\check{X}}_{k_1}\right), g_v\left(\underline{\check{X}}_{k_2}\right)-g_v\left(\underline{X}_{k_2}\right)\right)$$
$$+\frac{b}{n-m+1}\sum_{k_1=1}^{l-m+1}\sum_{k_2=1}^{l-m+1}\Big(Cov^*\left(g_u\left(\underline{\check{X}}^*_{k_1}\right), g_v\left(\underline{\check{X}}^*_{k_2}\right)\right) - Cov\left(g_u\left(\underline{\check{X}}_{k_1}\right), g_v\left(\underline{\check{X}}_{k_2}\right)\right)\Big)$$

and R_2 (and R_3) is of same form as R_1 but with k_1 replaced by k_1+l (k_2 replaced by k_2+l) and where b is replaced by $b-1$. Denote by $R_{i,j}$, $i=1,\ldots,3$, $j=1,\ldots,4$, the j-th term of R_i (compare (37)). Consider first $R_{1,1}+R_{2,1}+R_{3,1}$. By using formulae (A.11) of [23] one directly computes

$$\begin{aligned}
&R_{1,1}+R_{2,1}+R_{3,1} \qquad (38)\\
&=\frac{b}{n}\Bigg[\sum_{h=-l}^{l} l\, Cov(g_u(\underline{X}_0), g_v(\underline{X}_h))\\
&\qquad+\sum_{h=l+1}^{2l-m}(2l-m+2-h)\, Cov(g_u(\underline{X}_0), g_v(\underline{X}_h))\\
&\qquad+\sum_{h=-(l+1)}^{-(2l-m)}(2l-m+1-|h|)\, Cov(g_u(\underline{X}_0), g_v(\underline{X}_h))\Bigg]+o(1)\\
&\to_{n\to\infty}\sum_{h=-\infty}^{\infty} Cov(g_u(\underline{X}_0), g_v(\underline{X}_h)) = \left(\Sigma_{q\times q}\right)_{u,v},
\end{aligned}$$

Hence, it only remains to show that all other terms converge to zero as $n\to\infty$. To prove this, the arguments for $R_{1,2}, R_{2,2}, R_{3,2}$ as well as for $R_{1,3}, R_{2,3}, R_{3,3}$ are very similar and therefore we only consider $R_{1,2}$ here. Applying a Taylor expansion to the first term in $R_{1,2}$ and further using the linearity of the covariance and Cauchy Schwarz's inequality leads to

$$\begin{aligned}
&R_{1,2} \qquad (39)\\
&\le\frac{b}{n-m+1}\sum_{k_1=1}^{l-m+1}\sum_{k_2=1}^{l-m+1}\\
&\Bigg(\sum_{1\le|\alpha|\le h-1}\frac{1}{\alpha!}E\left[\left(D^\alpha g_u(\underline{X}_{k_1})\left(\underline{\check{X}}_{k_1}-\underline{X}_{k_1}\right)^\alpha\right)^2\right]^{1/2}\cdot E\left[g_v\left(\underline{X}_{k_2}\right)^2\right]^{1/2}\\
&\quad+\sum_{|\alpha|=h}\frac{1}{\alpha!}E\left[\left(D^\alpha g_u(\underline{\tau})\left(\underline{\check{X}}_{k_1}-\underline{X}_{k_1}\right)^\alpha\right)^2\right]^{1/2}\cdot E\left[g_v\left(\underline{X}_{k_2}\right)^2\right]^{1/2}\Bigg)
\end{aligned}$$

$$
\begin{aligned}
&\leq \frac{b}{n-m+1} \sum_{k_1=1}^{l-m+1} \sum_{k_2=1}^{l-m+1} \\
&\left(\sum_{1 \leq |\alpha| \leq h-1} \frac{1}{\alpha!} E\left[\left(D^\alpha g_u(\underline{X}_{k_1}) \right)^4 \right]^{1/4} E\left[\left(\underline{\check{X}}_{k_1} - \underline{X}_{k_1} \right)^{4\alpha} \right]^{1/4} \right. \\
&\qquad\qquad \times E\left[g_v \left(\underline{X}_{k_2} \right)^2 \right]^{1/2} \\
&\quad + \sum_{|\alpha|=h} \frac{1}{\alpha!} E\left[\left(D^\alpha g_u(\underline{\tau}) \right)^4 \right]^{1/4} E\left[\left(\underline{\check{X}}_{k_1} - \underline{X}_{k_1} \right)^{4\alpha} \right]^{1/4} \\
&\qquad\qquad \left. \times E\left[g_v \left(\underline{X}_{k_2} \right)^2 \right]^{1/2} \right),
\end{aligned}
$$

where the above expectations are similar to the terms investigated in the proof of Lemma 1. By the same strategy as in (30)-(35), one is lead to $R_{1,2} = \mathcal{O}_P\left(\rho^{M(n)}\right)$, which converges to zero as $n \to \infty$. Analogously one proofs that $R_{2,2}, R_{3,2}, R_{1,3}, R_{2,3}$ and $R_{3,3}$ vanish asymptotically.

Next consider expression $R_{1,4}$ which validates that the bootstrap covariance of the truncated series converges to its counterpart based on the original time series. For that, decompose

$$
\begin{aligned}
R_{1,4} \quad = \quad & \frac{b}{n-m+1} \sum_{k_1=M(n)+1}^{l-m+1} \sum_{k_2=M(n)+1}^{l-m+1} \left(Cov^*\left(g_u\left(\underline{\check{X}}_{k_1}^* \right), g_v\left(\underline{\check{X}}_{k_2}^* \right) \right) \right. \qquad (40) \\
& \qquad \left. - Cov\left(g_u\left(\underline{\check{X}}_{k_1} \right), g_v\left(\underline{\check{X}}_{k_2} \right) \right) \right) \\
& + \frac{b}{n-m+1} \sum_{k_1=1}^{M(n)} \sum_{k_2=1}^{M(n)} \left(Cov^*\left(g_u\left(\underline{\check{X}}_{k_1}^* \right), g_v\left(\underline{\check{X}}_{k_2}^* \right) \right) \right. \\
& \qquad \left. - Cov\left(g_u\left(\underline{\check{X}}_{k_1} \right), g_v\left(\underline{\check{X}}_{k_2} \right) \right) \right) \\
& + \frac{b}{n-m+1} \sum_{k_1=1}^{M(n)} \sum_{k_2=M(n)+1}^{l-m+1} \left(Cov^*\left(g_u\left(\underline{\check{X}}_{k_1}^* \right), g_v\left(\underline{\check{X}}_{k_2}^* \right) \right) \right. \\
& \qquad \left. - Cov\left(g_u\left(\underline{\check{X}}_{k_1} \right), g_v\left(\underline{\check{X}}_{k_2} \right) \right) \right) \\
& + \frac{b}{n-m+1} \sum_{k_1=M(n)+1}^{l-m+1} \sum_{k_2=1}^{M(n)} \left(Cov^*\left(g_u\left(\underline{\check{X}}_{k_1}^* \right), g_v\left(\underline{\check{X}}_{k_2}^* \right) \right) \right. \\
& \qquad \left. - Cov\left(g_u\left(\underline{\check{X}}_{k_1} \right), g_v\left(\underline{\check{X}}_{k_2} \right) \right) \right).
\end{aligned}
$$

To show that (40) asymptotically vanishes in probability, we mainly follow the lines of the proof of Lemma 5.5 in [9] and write

$$
\check{X}_t^* \quad = \quad \sum_{j=0}^{M(n)} \psi_j U_{t-j}^* + \sum_{j=0}^{M(n)} (\hat{\psi}_j - \psi_j) U_{t-j}^* + \sum_{j=M(n)+1}^{M(n)+(t-1) \bmod l} \hat{\psi}_j U_{t-j}^*. \qquad (41)
$$

We now show that, for $M(n) \to \infty$ as $n \to \infty$, the second and third sum in (41) asymptotically vanish, while the first sum appropriately imitates X_t. The asymptotic negligibility of the second and of the third sum is mainly due to the consistent estimation of the moving average coefficients and to their exponential decay.

Let $x \in \mathbb{R}$ be a continuity point of the distribution function of X_t. Furthermore, let $\gamma > 0$ be arbitrary. Then, as in the proof of Slutsky's Theorem we have

$$\begin{aligned} & P^*\left[\check{X}_t^* \le x\right] \\ & \le P^*\left[\sum_{j=0}^{M(n)} \psi_j U_{t-j}^* \le x+\gamma\right] + P^*\left[\left|\sum_{j=0}^{M(n)} \left(\hat{\psi}_j - \psi_j\right) U_{t-j}^*\right| > \frac{\gamma}{2}\right] \\ & \quad + P^*\left[\left|\sum_{j=M(n)+1}^{M(n)+(t-1) \bmod l} \hat{\psi}_j U_{t-j}^*\right| > \frac{\gamma}{2}\right]. \end{aligned} \tag{42}$$

Markov's inequality then gives

$$P^*\left[\left|\sum_{j=0}^{M(n)} \left(\hat{\psi}_j - \psi_j\right) U_{t-j}^*\right| > \frac{\gamma}{2}\right] \le 2 \sum_{j=0}^{M(n)} \left|\hat{\psi}_j - \psi_j\right| E^*\left|U_{t-j}^*\right| \frac{1}{\gamma} \tag{43}$$

and

$$P^*\left[\left|\sum_{j=M(n)+1}^{M(n)+(t-1) \bmod l} \hat{\psi}_j U_{t-j}^*\right| > \frac{\gamma}{2}\right] \le 2 \sum_{j=M(n)+1}^{M(n)+(t-1) \bmod l} \left|\hat{\psi}_j\right| E^*\left|U_{t-j}^*\right| \frac{1}{\gamma}. \tag{44}$$

Furthermore, by Lemma 2.2 of [22] it holds $\sum_{j=0}^{M(n)} \left|\hat{\psi}_j - \psi_j\right| = o_P(1)$, and by (33) it is $\sum_{j=M(n)+1}^{M(n)+(t-1) \bmod l} \left|\hat{\psi}_j\right| \le C\rho^{M(n)}$, for a suitable constant $C > 0$. Then, using $E^*\left|U_{t-j}^*\right| \le \sqrt{E^*(U_{t-j}^*)^2}$ and Assumption 5, one obtains for n sufficiently large and for any $\kappa > 0$ that, in probability,

$$P^*\left[\left|\sum_{j=0}^{M(n)} \left(\hat{\psi}_j - \psi_j\right) U_{t-j}^*\right| > \frac{\gamma}{2}\right] \le \frac{\kappa}{2} \tag{45}$$

and

$$P^*\left[\left|\sum_{j=M(n)+1}^{M(n)+(t-1) \bmod l} \hat{\psi}_j U_{t-j}^*\right| > \frac{\gamma}{2}\right] \le \frac{\kappa}{2}. \tag{46}$$

Hence, using (45) and (46) in (42) yields, in probability,

$$P^*\left[\check{X}_t^* \le x\right] \le P^*\left[\sum_{j=0}^{M(n)} \psi_j U_{t-j}^* \le x+\gamma\right] + \kappa \tag{47}$$

and in direct analogy

$$P^*\left[\check{X}_t^* \leq x\right] \geq P^*\left[\sum_{j=0}^{M(n)} \psi_j U_{t-j}^* \leq x - \gamma\right] - \kappa, \tag{48}$$

which indicates that the first sum in (41) is the dominating term.

Now we take a deeper look at the first sum of (41) and especially at its corresponding probability from (42). Consider

$$\sum_{j=0}^{M(n)} \psi_j U_{t-j}^* = \sum_{j=0}^{K} \psi_j U_{t-j}^* + \sum_{j=K+1}^{M(n)} \psi_j U_{t-j}^* \tag{49}$$

for some $K < M(n)$. Using (33), find $\left|\sum_{j=K+1}^{M(n)} \psi_j U_{t-j}^*\right| = \mathcal{O}_P(\rho^K)$, which means that for any $\varepsilon, \delta > 0$, we have

$$P^*\left[\left|\sum_{j=K+1}^{M(n)} \psi_j U_{t-j}^*\right| > \varepsilon\right] \leq \delta \tag{50}$$

as K is sufficiently large. Hence, in probability,

$$P^*\left[\sum_{j=0}^{M(n)} \psi_j U_{t-j}^* \leq x + \gamma\right] \leq P^*\left[\sum_{j=0}^{K} \psi_j U_{t-j}^* \leq x + \gamma + \varepsilon\right] + \delta \tag{51}$$

and

$$P^*\left[\sum_{j=0}^{M(n)} \psi_j U_{t-j}^* \leq x + \gamma\right] \geq P^*\left[\sum_{j=0}^{K} \psi_j U_{t-j}^* \leq x + \gamma - \varepsilon\right] - \delta \tag{52}$$

which indicates the first sum on the right hand side in (49) as the determining part of the left hand side. In the same manner these findings hold true for the non-bootstrap analogue $\sum_{j=0}^{M(n)} \psi_j U_{t-j}$. Now, define the set

$$\tau_n := \{k \,|\, M(n) + 1 \leq k \leq l\}. \tag{53}$$

Then, for $t \in \tau_n$, the sum $\sum_{j=0}^{M(n)} \psi_j U_{t-j}^*$ refers to the values of U_t^* within the same block of innovations and we therefore have

$$\begin{aligned} &\left|P^*\left[\sum_{j=0}^{M(n)} \psi_j U_{t-j}^* \leq x + \gamma\right] - P\left[\sum_{j=0}^{M(n)} \psi_j U_{t-j} \leq x + \gamma\right]\right| \\ &\leq \left|P^*\left[\sum_{j=0}^{K} \psi_j U_{t-j}^* \leq x + \gamma + \varepsilon\right] - P\left[\sum_{j=0}^{K} \psi_j U_{t-j} \leq x + \gamma - \varepsilon\right]\right| + 2\delta \end{aligned} \tag{54}$$

$$\leq \left| \frac{1}{n-l+1} \sum_{k=1}^{n-l+1} \mathbb{1}_{(-\infty,x+\gamma+\varepsilon]} \left(\sum_{j=0}^{K} \psi_j U_{k+l-j} \right) \right.$$
$$\left. - P \left[\sum_{j=0}^{K} \psi_j U_{t-j} \leq x+\gamma+\varepsilon \right] \right|$$
$$+ P \left[x+\gamma-\varepsilon < \sum_{j=0}^{K} \psi_j U_{t-j} \leq x+\gamma+\varepsilon \right] + 2\delta$$

where the first term is the difference of a probability and its empirical version, and thus, for sufficiently large n, it can be made arbitrarily small. The remainder terms also can be made arbitrarily small since ε and δ are arbitrarily chosen.

This is a rather important point in the proof. As discussed at the beginning of the proof of Lemma 1, it is important to assure for $M(n) \to \infty$ in order to incorporate completely the process structure. However, for (54) to be valid, the bootstrap innovations have to belong to the same block and hence, the (truncated) series should not overlap from one block of innovations to another; otherwise the desired non-bootstrap analogue would not be mimicked appropriately since the dependence structure of the time series is destroyed at the joining points of the independent blocks. This is the reason why the set τ_n is introduced and furthermore, why the rate $M(n)$ has to increase slower than the block length l. The following computations will clarify in more detail why $M(n)^2 l^{-1} \to 0$ has to be fulfilled.

Reconsider (54) which states that the approximation error of the distribution of $\check{X}_t^*$ to the distribution of $\check{X}_t$ vanishes for $n \to \infty$, for all $\forall t \bmod l \in \tau_n$. Write

$$\mathcal{L}(\check{X}_t^*) = \mathcal{L}(\check{X}_t) + o_P(1) \ \ \forall t \bmod l \in \tau_n, \tag{55}$$

and observe that convergence is guaranteed for all time points t with $\check{X}_t^*$ relying only on innovations stemming from one single block. Hence, by the Cramér-Wold device one has, for arbitrary $d \in \mathbb{N}$, and for any $(t_1 \bmod l, \ldots, t_d \bmod l) \in \tau_n^d$, that

$$\mathcal{L}\left(\check{X}_{t_1}^*, \ldots, \check{X}_{t_d}^*\right) = \mathcal{L}\left(\check{X}_{t_1}, \ldots, \check{X}_{t_d}\right) + o_P(1). \tag{56}$$

It is crucial to notice that (56) only holds for bootstrap random variables stemming from the set τ_n. Bootstrap realizations stemming from the first $M(n)$ positions of a block have an overlapping moving average representation to previous blocks of innovations. Such bootstrap realizations are seperately treated later on (see (63) and the following).

For the moment, we stay with the set τ_n. As a next step the function g and its component functions is investigated. Truncate g_u via

$$\tilde{g}_u(x) = g_u(x)\mathbb{1}_{|g_u(x)|\leq K}(x) + K\, sign(g_u(x))\mathbb{1}_{|g_u(x)|>K}, \ \ K>0. \tag{57}$$

Then $\tilde{g}_u(\cdot)$ is continuous and bounded and one immediately has

$$Cov^*\left(\tilde{g}_u\left(\underline{\check{X}}_t^*\right), \tilde{g}_v\left(\underline{\check{X}}_s^*\right)\right) \tag{58}$$

$$= Cov\left(\tilde{g}_u\left(\underline{\check{X}}_t\right), \tilde{g}_v\left(\underline{\check{X}}_s\right)\right) + o_P(1) \;\; \forall t, s \bmod l \in \tau_n$$

by (56). Now we show that the effect of truncating the functions is asymptotically negligible. By Hölder's inequality

$$\begin{aligned} & E^* \left| g_u\left(\underline{\check{X}}_t\right) \mathbb{1}_{|g_u(x)|>K} \right|^2 \qquad (59) \\ & \leq \left(E^* \left| g_u\left(\underline{\check{X}}_t\right)\right|^{2(h+2)/(h+1)} \right)^{(h+1)/(h+2)} \left(P^* \left(\left| g_u\left(\underline{\check{X}}_t\right)\right| > K \right)\right)^{1/(h+2)} \\ & = \mathcal{O}_P(1) K^{-2/(h+1)}. \end{aligned}$$

To see that the expected value is $\mathcal{O}_P(1)$ approach as in the proof of Lemma 1. Then, for arbitrary $\kappa > 0$, we can choose $K = K(\kappa, n)$ such that for n sufficiently large and in probability, for all $t, s \bmod l \in \tau_n$,

$$\left| \; Cov^*\left(\tilde{g}_u\left(\underline{\check{X}}^*_t\right), \tilde{g}_v\left(\underline{\check{X}}^*_s\right)\right) - \; Cov^*\left(g_u\left(\underline{\check{X}}^*_t\right), g_v\left(\underline{\check{X}}^*_s\right)\right)\right| \leq \frac{\kappa}{N(n)} \qquad (60)$$

and in complete analogy, still for all $t, s \bmod l \in \tau_n$ only,

$$\left| \; Cov\left(\tilde{g}_u\left(\underline{\check{X}}_t\right), \tilde{g}_v\left(\underline{\check{X}}_s\right)\right) - \; Cov\left(g_u\left(\underline{\check{X}}_t\right), g_v\left(\underline{\check{X}}_s\right)\right)\right| \leq \frac{\kappa}{N(n)} \qquad (61)$$

where $N(n) \to \infty$, as $n \to \infty$, is chosen such that $l/N(n) \to 0$. Hence, it follows, from (58) by (60) and (61)

$$\begin{aligned} & \left| \; Cov^*\left(g_u\left(\underline{\check{X}}^*_t\right), g_v\left(\underline{\check{X}}^*_s\right)\right) - \; Cov\left(g_u\left(\underline{\check{X}}_t\right), g_v\left(\underline{\check{X}}_s\right)\right)\right| \qquad (62) \\ & = \Big| \; Cov^*\left(g_u\left(\underline{\check{X}}^*_t\right), g_v\left(\underline{\check{X}}^*_s\right)\right) - \; Cov\left(g_u\left(\underline{\check{X}}_t\right), g_v\left(\underline{\check{X}}_s\right)\right) \\ & \qquad \pm Cov^*\left(\tilde{g}_u\left(\underline{\check{X}}^*_t\right), \tilde{g}_v\left(\underline{\check{X}}^*_s\right)\right) \pm \; Cov\left(\tilde{g}_u\left(\underline{\check{X}}_t\right), \tilde{g}_v\left(\underline{\check{X}}_s\right)\right)\Big| \\ & = \mathcal{O}_P\left(N(n)^{-1}\right) + o_P(1) \end{aligned}$$

for all $t, s \bmod l \in \tau_n$, and, thus, is applicable to the first sum of $R_{1,4}$. To handle the further sums of $R_{1,4}$, we have to consider the covariances in (40) where k_1 or k_2 is not in τ_n. Hence, consider the case $k_1 \notin \tau_n$ and $k_2 \in \tau_n$ and obtain by using (62)

$$\begin{aligned} & Cov^*\left(g_u\left(\underline{\check{X}}^*_{k_1}\right), g_v\left(\underline{\check{X}}^*_{k_2}\right)\right) - \; Cov\left(g_u\left(\underline{\check{X}}_{k_1}\right), g_v\left(\underline{\check{X}}_{k_2}\right)\right) \qquad (63) \\ & = \; Cov^*\left(g_u\left(\underline{\check{X}}^*_{k_1}\right) - g_u\left(\underline{\check{X}}^*_{l-m+1}\right), g_v\left(\underline{\check{X}}^*_{k_2}\right)\right) \\ & \qquad + Cov^*\left(g_u\left(\underline{\check{X}}^*_{l-m+1}\right), g_v\left(\underline{\check{X}}^*_{k_2}\right)\right) - \; Cov\left(g_u\left(\underline{\check{X}}_{k_1}\right), g_v\left(\underline{\check{X}}_{k_2}\right)\right) \\ & = \; Cov^*\left(g_u\left(\underline{\check{X}}^*_{k_1}\right) - g_u\left(\underline{\check{X}}^*_{l-m+1}\right), g_v\left(\underline{\check{X}}^*_{k_2}\right)\right) \\ & \qquad - Cov\left(g_u\left(\underline{\check{X}}_{k_1}\right) - g_u\left(\underline{\check{X}}_{l-m+1}\right), g_v\left(\underline{\check{X}}_{k_2}\right)\right) + \mathcal{O}_P(N(n)^{-1}). \end{aligned}$$

Then by application of the Taylor expansion and by Cauchy Schwarz inequality one obtains for the first covariance

$$Cov^*\left(g_u\left(\underline{\check{X}}^*_{k_1}\right) - g_u\left(\underline{\check{X}}^*_{l-m+1}\right), g_v\left(\underline{\check{X}}^*_{k_2}\right)\right) \qquad (64)$$

$$\leq \left(\sum_{1\leq|\alpha|\leq h-1} \frac{1}{\alpha!} E^* \left[\left(D^\alpha g_u \left(\underline{\check{X}}^*_{l-m+1} \right) \right)^4 \right]^{1/4} \right.$$
$$\times E^* \left[\left(\underline{\check{X}}^*_{k_1} - \underline{\check{X}}^*_{l-m+1} \right)^{4\alpha} \right]^{1/4}$$
$$\left. + \sum_{|\alpha|=h} \frac{1}{\alpha!} E^* \left[\left(D^\alpha g_u (\underline{\tau}) \right)^4 \right]^{1/4} E^* \left[\left(\underline{\check{X}}^*_{k_1} - \underline{\check{X}}^*_{l-m+1} \right)^{4\alpha} \right]^{1/4} \right)$$
$$\times E^* \left[g_v \left(\underline{\check{X}}^*_{k_2} \right)^2 \right]^{1/2},$$

for some $\underline{\tau} = \underline{\check{X}}^*_{l-m+1} + \lambda \left(\underline{\check{X}}^*_{k_1} - \underline{\check{X}}^*_{l-m+1} \right)$, $\lambda \in (0,1)$. Indeed, this expression is similar to (39) and can be handled analogously. One obtains

$$\begin{aligned} & E^* \left[\left(\underline{\check{X}}^*_{k_1} - \underline{\check{X}}^*_{l-m+1} \right)^{4\alpha} \right] \\ & \leq \prod_{r=1}^{m} E^* \left[\left(\sum_{j=k_1+M(n)}^{l-m+M(n)} \hat{\psi}_j \left(U^*_{k_1+r-1-j} - U^*_{l+r-1-j} \right) \right)^{4\alpha_r} \right] \\ & = \mathcal{O}_P \left(\rho^{(k_1+M(n))\cdot|\alpha|} \right), \end{aligned} \tag{65}$$

where α_r indicates the r-th component of α. The further expectations are finite by Assumptions 1 and 3. For the non-bootstrap covariance in (63) one proceeds exactly in the same way. Altogether this leads to

$$\begin{aligned} & Cov^* \left(g_u \left(\underline{\check{X}}^*_{k_1} \right), g_v \left(\underline{\check{X}}^*_{k_2} \right) \right) - \; Cov \left(g_u \left(\underline{\check{X}}_{k_1} \right), g_v \left(\underline{\check{X}}_{k_2} \right) \right) \\ & = \mathcal{O}_P \left(N(n)^{-1} + \rho^{k_1+M(n)} \right), \end{aligned} \tag{66}$$

for all $k_1 \notin \tau_n$ and $k_2 \in \tau_n$. Of course, if $k_1 \in \tau_n$ and $k_2 \notin \tau_n$ this result is directly adaptable.

Now we have all preliminaries at hand to revisit the remainder term $R_{1,4}$ given in (40). Using the previous considerations and the summability of the autocovariances it holds

$$\begin{aligned} R_{1,4} &= \mathcal{O}_P \left(\frac{l}{N(n)} + \frac{b\, M(n)^2}{n} + M(n) \left(N(n)^{-1} + \rho^{M(n)} \right) \right) \\ &= \mathcal{O}_P \left(\frac{l}{N(n)} + \frac{b\, M(n)^2}{n} + M(n)\rho^{M(n)} \right), \end{aligned} \tag{67}$$

which converges to zero as n increases by the conditions on the rates ($l/N(n) \to 0$ and $\rho < 1$) and by the choice of $M(n)$. The further remainder terms $R_{2,4}$ and $R_{3,4}$ yield analogue results. Altogether this proofs (16), the assertion on the limiting covariance.

Now consider (17), the central limit result. By Lemma 1 it suffices to show the result for

$$\frac{1}{\sqrt{n-m+1}}\sum_{t=1}^{n-m+1}\left(g\left(\underline{\check{X}}_t^*\right)-E^*g\left(\underline{\check{X}}_t^*\right)\right) \tag{68}$$

to conclude the proof. Define, for all $t=1,\dots,n-m+1$,

$$\underline{Z}_{n,t}^* := \frac{1}{\sqrt{n-m+1}}\left(g\left(\underline{\check{X}}_t^*\right)-E^*g\left(\underline{\check{X}}_t^*\right)\right) \tag{69}$$

which is a triangular array with $E^*\underline{Z}_{n,t}^*=0$ for all $t=1,\dots,n-m+1$, and denote by $Z_{n,t,u}^*$ the u-th component of $\underline{Z}_{t,n}^*$, $u=1,\dots,q$. Due to the underlying block structure and the definition of $\check{X}_t^*$, the sequence is $2l$-dependent and does not yield stationarity. To proceed, a central limit result for $2l$-dependent triangular random variables is required where l is allowed to increase to infinity with sample size by some rate. In here we make use of Theorem 2.1 in [41] and need to check their six conditions. Using the notation of [41], we define

$$\begin{aligned}
&B_{n,r,a}^2 \equiv B_{n,r,a}^2(u,v) \qquad (70)\\
&:= Cov^*\left(\sum_{t=a}^{a+r-1} Z_{n,t,u}^*, \sum_{s=a}^{a+r-1} Z_{n,s,v}^*\right)\\
&= \frac{\left\lfloor \frac{r}{l}\right\rfloor}{n-m+1}\sum_{h=-(l-m)}^{l-m}\left((l-m+1)-|h|\right)\ Cov\left(g_u\left(\underline{\check{X}}_0\right), g_v\left(\underline{\check{X}}_h\right)\right)+o_P(1)
\end{aligned}$$

for all $u,v=1,\dots,q$, and note that $\left\lfloor \frac{r}{l}\right\rfloor l=\mathcal{O}(r)$. Further compute

$$\begin{aligned}
&B_n^2 \equiv B_{n,n,1}^2 \qquad (71)\\
&:= Cov^*\left(\sum_{t=1}^{n} Z_{n,t,u}^*, \sum_{s=1}^{n} Z_{n,s,v}^*\right)\\
&= \frac{b}{n-m+1}\sum_{h=-(l-m)}^{l-m}\left((l-m+1)-|h|\right)\ Cov\left(g_u\left(\underline{\check{X}}_0\right), g_v\left(\underline{\check{X}}_h\right)\right)+o_P(1)
\end{aligned}$$

for all $u,v=1,\dots,q$, by the same manner. If one sets $\gamma=0$ (in the notation of [41]), it can be straightforwardly checked that all conditions of the Theorem are fulfilled (also see their Example 3.3 on the moving blocks bootstrap for further details), and thus the central limit theorem is applicable to the present scenario.

It is worth to note that the stated condition on the block length, $l^{2+2/\delta}/n \to 0$, corresponds to condition C6 of [41] and is required for their theorem to hold. In our setup, it causes that the higher moments of the time series exist the larger the block length is allowed to be chosen. However, as for any block approach, the rate of the block length is not allowed to exceed $n^{1/2}$.

One finally obtains

$$B_n^{-1} \sum_{t=1}^{n-m+1} \underline{Z}_{n,t}^* \xrightarrow{\mathcal{D}^*} \mathcal{N}(0,1) \tag{72}$$

in probability, and furthermore, since $B_n^2 \to (\Sigma_{q\times q})_{u,v}$, for all $u, v = 1, \ldots, q$, it holds

$$\frac{1}{\sqrt{n-m+1}} \sum_{t=1}^{n-m+1} \left(g\left(\underline{\check{X}}_t^*\right) - \theta^* \right) \xrightarrow{\mathcal{D}^*}_{n\to\infty} \mathcal{N}\left(0_q, \Sigma_{q\times q}\right). \tag{73}$$

By application of the delta technique, e.g. [3], the proof of (17) then is concluded for T_n^* with X_t^* replaced by $\check{X}_t^*$. Lemma 1 then immediately concludes the proof for T_n^* itself. □

References

[1] Akaike, H. (1973): Information theory and an extension of the maximum likelihood principle. In: B. N. Petrov and F. Csaki (eds.) *2nd International Symposium on Information Theory*, pp. 267–281. Akademiai Kiado, Budapest. MR0483125

[2] Akaike, H. (1974): A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, Vol. 19, pp. 716–723. MR0423716

[3] Billingsley, P. (1979): *Probability and Measure*. John Wiley and Sons, New York. MR0534323

[4] Bose, A. (1988): Edgeworth correction by bootstrap in autoregressions. *The Annals of Statistics*, Vol. 16, pp. 1709–1722. MR0964948

[5] Brockwell, P.J. and Davis, R.A. (1991): *Time Series: Theory and Methods (2nd Ed.)*. Springer, New York. MR1093459

[6] Brockwell, P.J., Davis, R.A. and Yang, Y. (2010): Estimation for Non-Negative Lévy-Driven CARMA Processes. *Journal of Business and Economic Statistics*, Vol. 29, pp. 250–259.

[7] Broersen, P.M.T. and de Waele, S. (2004): Finite Sample Properties of ARMA order selection. *IEEE*, Vol. 53, No. 3.

[8] Bühlmann, P. (1994): Blockwise bootstrapped empirical processes for stationary sequences. *The Annals of Statistics*, Vol. 22, pp. 995–1012.

[9] Bühlmann, P. (1997): Sieve Bootstrap for Time Series. *Bernoulli*, Vol. 3, pp. 123–148.

[10] Bühlmann, P. and Künsch, H.R. (1995). The blockwise bootstrap for general parameters of a stationary time series. *Scandinavian Journal of Statistics*, Vol. 22, pp. 35–54.

[11] Bühlmann, P. and Künsch, H.R. (1999). Block length selection in the bootstrap for time series. *Computational Statistics and Data Analysis*, Vol. 31, pp. 295–310.

[12] Carlstein, E. (1986). The use of subseries values for estimating the variance of a general statistic from a stationary time series. *The Annals of Statistics*, Vol. 14, pp. 1171–1179.

[13] Carlstein, E., Do, K.-A., Hall, P. Hesterberg, T. and Künsch, H.R. (1998). Matched-block bootstrap for dependent data. *Bernoulli*, Vol. 4, pp. 305–328.

[14] Davison, A.C. and Hinkley, D.V. (1997): *Bootstrap methods and their application*. Cambridge University Press, Cambridge.

[15] Efron, B. (1979). Bootstrap methods: another look at the jackknife. *The Annals of Statistics*, Vol. 7, pp. 1–26. MR0515681

[16] Efron, B. and Tibshirani, R.J. (1986). Bootstrap methods for standard errors, confidence intervals and other measures of statistical accuracy (with discussion). *Statistical Science*, Vol. 1, pp. 54–77. MR0833275

[17] Fischer, W. and Lieb, I. (1988): *Funktionentheorie (5th Ed.)*. Vieweg, Wiesbaden. MR0554918

[18] Freedman, D. (1984): On bootstrapping two-stage least squares estimates in stationary linear models. *The Annals of Statistics*, Vol. 12, pp. 827–842.

[19] Götze, F., and Künsch, H. (1996): *Second-order correctness of the blockwise bootstrap for stationary observations*. The Annals of Statistics, Vol. 24, pp. 1914–1933. MR1421154

[20] Hall, P., Horowitz, J. L., and Jing, B.-Y. (1995): *On blocking rules for the bootstrap with dependent data*. Biometrika, Vol. 82, pp.561–574.

[21] Hannan, E.J. (1980): The estimation of the order of an ARMA process. *The Annals of Statistics*, Vol. 8, pp. 1071–1081.

[22] Kreiss, J.-P. and Franke, J. (1992): Bootstrapping stationary autoregressive moving average models. *Journal of Time Series Analysis*, Vol. 13, pp. 297–319.

[23] Kreiss, J.-P. und Neuhaus, G. (2006): *Einführung in die Zeitreihenanalyse*. Springer, Heidelberg.

[24] Kreiss, J.-P. and Paparoditis, E. (2012): The hybrid wild bootstrap for time series. *Journal of the American Statistical Association*, Vol. 107, No. 499, pp. 1073–1084.

[25] Kreiss, J.-P. and Paparoditis, E. (2003): Autoregressive-aided periodogram bootstrap for time series. *The Annals of Statistics*, Vol. 31, pp. 1923–1955.

[26] Künsch, H.R. (1989). The jackknife and the bootstrap for general stationary observations. *The Annals of Statistics*, Vol. 17, pp. 1217–1241.

[27] Lahiri, S.N. (1993). On the moving block bootstrap under long range dependence. *Statistics and Probability Letters*, Vol. 18, pp. 405–413.

[28] Lahiri, S.N. (1999). Theoretical comparisons of block bootstrap methods. *The Annals of Statistics*, Vol. 27, pp. 386–404.

[29] Lahiri, S.N. (2003). Resampling methods for dependent data. New York: Springer. MR2001447

[30] Lahiri, S.N., Furukawa, K., and Lee, Y.-D. (2007). A nonparametric plug-in method for selecting the optimal block length. *Statistical Methodology*, Vol. 4, pp. 292–321.

[31] Liang, G., Wilkes, D.M. and Cadzow, J.A. (1993). ARMA model order estimation based on the eigenvalues of the covariance matrix. *IEEE transactions on signal processing*, Vol. 41, pp. 3003–3009.

[32] Liu, R.Y. and Singh, K. (1992). Moving blocks Jackknife and Bootstrap

capture weak dependence. In *Exploring the Limits of Bootstrap*, eds. R. lePage and L. Billard. New York, Wiley, pp. 225–248.

[33] Naik-Nimbalkar, U.V. and Rajarshi, M.B. (1994): Validity of blockwise bootstrap for empirical processes with stationary observations. *The Annals of Statistics*, Vol. 22, pp. 980–994.

[34] Niebuhr, T. and Kreiss, J.-P. (2014): Asymptotics for Autocovariances and Integrated Periodograms for Linear Processes Observed at Lower Frequencies. *The International Statistical Review*, Vol. 82, No. 1, pp. 123–140.

[35] Nordman, D. and Lahiri, S. (2014): Convergence rates of empirical block length selectors for block bootstrap. *Bernoulli*, Vol. 20, No. 2, pp. 958–978.

[36] Paparoditis, E. and Politis, D.N. (2001): The tapered block bootstrap. *Biometrika*, Vol. 88, pp. 1105–1119.

[37] Patton, A., Politis, D.N. and White, H. (2009): Correction to 'Automatic block length selection for the dependent bootstrap by D.N. Politis and H. White'. *Econometric Reviews*, Vol. 28, pp. 372–375.

[38] Politis, D.N. and Romano, J.P. (1992): A circular block resampling procedure for stationary data. In *Exploring the Limits of Bootstrap*, eds. R. lePage and L. Billard. New York, Wiley, pp. 263–270.

[39] Politis, D.N. and White, H. (2004): Automatic block length selection for the dependent bootstrap. *Econometric Reviews*, Vol. 23, pp. 53–70. MR2041534

[40] Press, H. and Tukey, J.W. (1956): Power spectral methods of analysis and their application to problems in airplane dynamics. *Bell System Monographs 2606*.

[41] Romano, J.P. and Wolf, M. (2000): A more general central limit theorem for m-dependent random variables with unbounded m. *Statistics and Probability Letters*, Vol. 47, pp. 115–124.

[42] Shao, Q.M. and Yu, H. (1993): Bootstrapping the sample means for stationary mixing sequences. *Stoch. Processes Applic.*, Vol. 48, pp. 175–190.

[43] Shao, X. (2010): The dependent wild bootstrap. *Journal of the American Statistical Association*, Vol. 105, pp. 218–235.