Electronic Journal of Statistics
Vol. 12 (2018) 3875–3907
ISSN: 1935-7524
https://doi.org/10.1214/18-EJS1503

Generalized subsampling procedure for non-stationary time series*

Łukasz Lenart

Department of Mathematics
Cracow University of Economics
27 Rakowicka St.
31-510 Cracow, Poland
e-mail: lukasz.lenart@uek.krakow.pl

Abstract: In this paper, we propose a generalization of the subsampling procedure for non-stationary time series. The proposed generalization is simply related to the usual subsampling procedure. We formulate the sufficient conditions for the consistency of such a generalization. These sufficient conditions are a generalization of those presented for the usual subsampling procedure for non-stationary time series. Finally, we demonstrate the consistency of the generalized subsampling procedure for the Fourier coefficient in mean expansion of Almost Periodically Correlated time series.

MSC 2010 subject classifications: 62G09, 62G20, 62M10.
Keywords and phrases: Generalized subsampling procedure, non-stationary time series, subsampling consistency, almost periodically correlated time series.

Received July 2017.

1. Introduction

The subsampling procedure for non-stationary time series is very well investigated (see Politis et al. (1999)). The sufficient conditions for the consistency of subsampling in the non-stationary (and stationary) case were formulated in the general case. However, in recent years, there have been a growing number of new resampling methods for non-stationary time series. Almost all of them are based on the Moving Block Bootstrap (MBB for short) method introduced first by Kunsch (1989) and Liu and Singh (1992). The MBB procedure is an extension of Efron's i.i.d. bootstrap method to the time series case. Based on MBB, several alternative methods dedicated to the stationary case have been investigated (see Lahiri (2003)).

In the case of non-stationary time series, the examples of consistency of resampling methods often concerns time series with periodic or almost periodic structure in mean or autocovariance functions. Recall that the time series $\{X_t : t \in \mathbb{Z}\}$ is called Periodically Correlated (PC for short) if a mean function $\mu(t) = E(X_t)$ and autocovariance function $\mathrm{cov}(X_t, X_{t+\tau}) = B(t, \tau)$ exist and are

*Publication was financed from the funds granted to the Faculty of Finance and Law at Cracow University of Economics, within the framework of the subsidy for the maintenance of research potential.

periodic functions with the same period $T > 1$, where the period T is taken as the smallest integer such that conditions: $\mu(t) = \mu(t+T)$, $B(t,\tau) = B(t+T,\tau)$ hold for any $t, \tau \in \mathbb{Z}$ (see Gladyshev (1961) and Hurd and Miamee (2007)). To recall from literature the next class of time series we start from the definition of the almost periodic function. A real-valued function $f(t) : \mathbb{Z} \longrightarrow \mathbb{R}$ of an integer variable is called almost periodic if for any $\epsilon > 0$ there exists an integer $L_\epsilon > 0$ such that among any L_ϵ consecutive integers there is an integer p_ϵ with the property $\sup_{t\in\mathbb{Z}} |f(t+p_\epsilon) - f(t)| < \epsilon$ (see to Corduneanu (1989) for more details). The time series is called Almost Periodically Correlated (APC for short) if a mean function and autocovariance function exist and are almost periodic in variable t. The class of APC time series is used in many fields including telecommunications, econometrics (recently in Mazur and Pipień (2012)) and many others (see the review in Serpedin et al. (2005) and Gardner et al. (2006)).

The subsampling consistency was proved in many cases for time series with PC and APC structures without any modification of the subsampling scheme. In the paper of Lenart et al. (2008), the subsampling consistency for the magnitude of Fourier coefficients for autocovariance functions in the PC case was proved. The consistency of subsampling for a spectral density function in the APC case was examined in Lenart (2011), while in Lenart (2013) and Lenart and Pipień (2017), the consistency of subsampling for the Fourier coefficient in a Fourier representation of the expectation function of an APC time series was proved in univariate and multivariate case, respectively. In Dehay et al. (2014), the problem of consistency of the subsampling procedure for estimators in continuous time non-stationary stochastic processes with periodic or almost periodic covariance structures was investigated. The application of the subsampling method for time series with PC and APC structures can be found in Lenart and Pipień (2013a), Lenart and Pipień (2013b), and Lenart (2017).

There are several resampling methods based on the MBB procedure dedicated to the PC and APC cases (see Dudek et al. (2014a), Dudek et al. (2014b), Dudek (2015), Dehay and Dudek (2015), Dudek (2016), Dudek et al. (2016), Dehay and Dudek (2017), and Dudek (2018)). The main motivation for investigating in the mentioned above literature a new resampling scheme based on the MBB procedure for time series with a PC or APC structure is to adjust the Moving Block Bootstrap procedure to the periodic or almost periodic structure of the time series. This modification is desirable since the MBB procedure and numerous modifications use the common idea of generating a new pseudo time series. For example, in Dudek et al. (2014a), the so-called *Generalized Seasonal Block Bootstrap* (GSBB) was motivated in the following manner: *"When time-series data contain a periodic/seasonal component, the usual block bootstrap procedures are not directly applicable."* Unfortunately, in each case, the consistency of resampling methods based on the Moving Block Bootstrap method are proved individually because the general sufficient conditions for the consistency has not been formulated.

Recently, Tewes et al. (2017) use the idea of self-convolution to define the new resampling procedure called *convolved subsampling*. They proved the consistency of this procedure for sampling distributions with normal limits. In particular, they proved the consistency of convolved subsampling for the mean of generally α-mixing sequence. Moreover, Tewes et al. (2017) shown the connection of convolved subsampling to block bootstrap procedure in the case of sample mean. They shown that by applying the convolved subsampling technique, some results concerning block bootstrap consistency can be extend by substantially weakening the assumptions.

Note that the subsampling procedure and methods based on the Moving Block Bootstrap procedure are considered rather separately in the literature. The main reason is that the idea of Moving Block Bootstrap is based on generation of a new pseudo time series, while in subsampling, each individual block (or subsample) is taken under consideration to evaluate the estimator based on that subsample separately. In the Moving Block Bootstrap procedure (or modification), we calculate the estimator for a resampled pseudo time series.

The sufficient conditions for subsampling consistency presented for stationary or non-stationary time series seams to be relatively weak in relation to the conditions for consistency of methods based on block bootstrap and also for convolved subsampling. Note that the consistency of convolved subsampling implies the consistency of subsampling, but not necessary inversely. It is not easy to find an example where the procedure based on block bootstrap is consistent but the well defined subsampling procedure is not consistent. This is probably due to the fact that subsampling does not require specific forms of statistics (for example mean-like statistics) or asymptotic distribution (for example Gaussian). This is one of the motivations for us to generalize subsampling procedure.

In this work, we generalize the subsampling procedure proposed by Politis et al. (1999) (Section 2) and formulate the sufficient conditions for the consistency of such a generalization (Section 3). We show that the generalized subsampling is valid under weak assumptions which generalize these presented by Politis et al. (1999) for subsampling (i.e., α-mixing assumption and assumption requiring the existence of a non-degenerate limiting distribution for the sampling distribution being approximated). Our generalization of subsampling procedure brings the subsampling and usual Moving Block Bootstrap procedures together. Finally, we demonstrate the consistency of this procedure for Fourier coefficients in mean expansion for the APC case (Section 4) and show a short simulation study (Section 5). Some preliminary results related to the presented results can be found in Lenart (2016).

2. Generalized subsampling procedure

Let $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ be a sample from a general non-stationary $\mathbb{R}^r$ - valued time series $\{\mathbf{X}_t : t \in \mathbb{Z}\}$. $\hat{\boldsymbol{\theta}}_n \in \mathbb{R}^s$ denotes the estimator (based on the sample $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$) of the parameter of interest, $\boldsymbol{\theta} \in \mathbb{R}^s$. Assume that we are interested in approximating an unknown cumulative distribution function $J_n(\cdot)$ of

$$\tau_n(\hat{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}), \tag{2.1}$$

where τ_n is a normalizing sequence. The limiting cumulative distribution function of (2.1) is denoted by $J(\cdot)$. In the standard subsampling procedure, for a non-stationary time series, the distribution of (2.1) is approximated on the basis of the empirical distribution based on the following sample $\{\hat{\boldsymbol{\theta}}_{n,b,1}, \hat{\boldsymbol{\theta}}_{n,b,2}, \ldots, \hat{\boldsymbol{\theta}}_{n,b,n-b+1}\}$, where $\hat{\boldsymbol{\theta}}_{n,b,t}$ is the estimator $\hat{\boldsymbol{\theta}}_n$ based on the subsample starting at t $(1 \le t \le n-b+1)$ and with length $b = b(n) \to \infty$. Let $J_{b,t}(\cdot)$ be the distribution function of $\tau_b(\hat{\boldsymbol{\theta}}_{n,b,t} - \boldsymbol{\theta})$. One of the sufficient conditions for the consistency of the subsampling distribution is that for any Borel set $A \in \mathbb{R}^s$ whose boundary has zero mass, we have the following convergence:

$$\frac{1}{n-b+1} \sum_{t=1}^{n-b+1} J_{b,t}(A) \to J(A),$$

as $n \to \infty$. The stronger condition assumes that

$$J_{b,t_b}(A) \to J(A), \tag{2.2}$$

as $n \to \infty$, for any sequence t_n of positive integers $(1 \le t_n \le n-b+1)$ and any Borel set A whose boundary has mass zero. Note that (2.2) means that, uniformly at t, the cumulative distribution function of $\tau_b(\hat{\boldsymbol{\theta}}_{n,b,t} - \boldsymbol{\theta})$ trends to the limiting cumulative distribution function of (2.1). This condition plays a central role in proving subsampling consistency. In this paper, we formulate more general assumptions by taking into consideration more general subsamples than in the usual subsampling.

We start from the definition of a so-called *generalized subsample*. To simplify, consider the set of integers $S_{n,b} = \{1, 2, \ldots, q_{n,b}\}$, where $q_{n,b} = n-b+1$. By $\mathcal{B} = \{\mathcal{B}_i = (X_i, X_{i+1}, \ldots, X_{i+b-1}) \colon\ i = 1, 2, \ldots, q_{n,b}\}$, we denote the set of all overlapping blocks. Take any sequence $k = k(n)$ of positive integers and the related vector $\mathbf{T}_k = (t_1, t_2, \ldots, t_k) \in S_{n,b}^k = S_{n,b} \times S_{n,b} \times \ldots \times S_{n,b}$ with length k (the so-called *vector of starting points*). Define the so-called *generalized subsample* with the vector of starting points $\mathbf{T}_k \in S_{n,b}^k$ and total length kb via

$$\begin{aligned} \mathcal{B}_{b,\mathbf{T}_k} =& (\underbrace{X_{t_1}, X_{t_1+1}, \ldots, X_{t_1+b-1}}_{\mathcal{B}_{t_1}}, \underbrace{X_{t_2}, X_{t_2+1}, \ldots, X_{t_2+b-1}}_{\mathcal{B}_{t_2}}, \ldots, \\ & \underbrace{X_{t_k}, X_{t_k+1}, \ldots, X_{t_k+b-1}}_{\mathcal{B}_{t_k}}). \end{aligned} \tag{2.3}$$

Remark 2.1. Note that the above definition of a generalized subsample cannot be identified with a generated pseudo time series in the Moving Block Bootstrap procedure or with a generated pseudo time series by MBB modifications dedicated to the PC or APC case. Note that, if $k = 1$, $\mathcal{B}_{b,\mathbf{T}_1}$ is a typical subsample used in the classical subsampling approach. If $k \approx n/b$, then the length of the generalized subsample is comparable with the length of a pseudo time series in the classical MBB procedure. However, in numerous modifications of the classical MBB procedure, the generating scheme restricts the possible shape of a pseudo time series to retain the structure of the time series (see, for example, the generating scheme in GSBB introduced by Dudek et al. (2014a)). Summing up, not all possible generalized subsamples $\mathcal{B}_{b,\mathbf{T}_k}$ are allowed by the MBB procedure, because in such a case we are restricted to the case $k \approx n/b$. In the case of existing sophisticated modifications of block bootstrap we additionally try to retain the structure of the time series. For example in GSBB procedure the structure of subsample (or sample bootstrap pseudo-observations) is strictly related to the length of the period for considered PC time series (see Step 1-3. in Dudek et al. (2014a)).

Finally, $\hat{\boldsymbol{\theta}}_{n,b,\mathbf{T}_k}$ denotes the estimator of $\boldsymbol{\theta} \in \mathbb{R}^s$ based on such a generalized subsample. In our generalization, the construction of such an estimator (in particular cases) plays a central role. The expected property of the estimator $\hat{\boldsymbol{\theta}}_{n,b,\mathbf{T}_k}$ is the ability to estimate θ (with the appropriate meaning) for any b, k and vector of starting points $\mathbf{T}_k$. For the APC time series case, the appropriate estimators for the usual subsampling methodology were proposed in Lenart (2013) (for Fourier coefficient estimator) and in Lenart (2011) and Lenart (2016) for the case of spectral characteristics. In particular, for the estimator $\hat{\boldsymbol{\theta}}_{n,b,\mathbf{T}_k}$ based on a generalized subsample, the following form can be proposed:

$$\hat{\theta}_{n,b,(t_1,t_2,\ldots,t_k)} = \frac{1}{k}\sum_{j=1}^{k}\hat{\theta}_{n,b,(t_j)}. \tag{2.4}$$

To explain how the estimator $\hat{\boldsymbol{\theta}}_{n,b,\mathbf{T}_k}$ can be constructed, we present some illustrative examples, **A-D**, below. Example **A** corresponds to the stationary case, while in the next examples, non-stationarity is assumed. To simplify, we assume that $r = s = 1$ in the below examples.

A. Mean for the stationary case is $\theta = E(X_t)$. The estimator based on a generalized subsample is

$$\hat{\theta}_{n,b,(t_1,t_2,\ldots,t_k)} = \frac{1}{k}\sum_{j=1}^{k}\hat{\theta}_{n,b,(t_j)} = \frac{1}{kb}\sum_{j=1}^{k}\sum_{t=t_j}^{t_j+b-1} X_t.$$

As an estimator of $|\theta|$, we take $|\hat{\theta}_{n,b,(t_1,t_2,\ldots,t_k)}|$.

B. Fourier coefficients for the mean function in the APC case: $E(X_t) = \sum_{\psi\in\Psi} m(\psi)e^{i\psi t}$ - an almost periodic function, where $\Psi = \{\psi \in [0, 2\pi) : \lim_{n\to\infty}\sum_{j=1}^{n}\frac{1}{n}E(X_j)e^{-i\psi j} \neq 0\}$ and $m(\psi)$ are complex numbers. For $\psi \in [0, 2\pi)$, we define the estimator of $\theta(\psi) = m(\psi)$ based on a generalized subsample via

$$\hat{\theta}_{n,b,(t_1,t_2,\ldots,t_k)}(\psi) = \frac{1}{k}\sum_{j=1}^{k}\hat{\theta}_{n,b,(t_j)} = \frac{1}{kb}\sum_{j=1}^{k}\sum_{t=t_j}^{t_j+b-1} X_t e^{it\psi}.$$

In the same manner, for parameter $|\theta(\psi)|$, we define $|\hat{\theta}_{n,b,(t_1,t_2,\ldots,t_k)}(\psi)|$.

C. Fourier coefficients for the autocovariance function in the APC case: $E(X_t) = 0$, $\mathrm{cov}(X_t, X_{t+\tau}) = E(X_t X_{t+\tau}) = \sum_{\lambda \in \Lambda_\tau} a(\lambda, \tau) e^{i\lambda t}$ - an almost periodic function, where $\Lambda_\tau = \{\lambda \in [0, 2\pi) : \lim_{n\to\infty} \sum_{j=1}^{n} \frac{1}{n} E(X_j X_{j+\tau}) e^{-i\lambda j} \neq 0\}$, $a(\lambda, \tau)$ are complex numbers and $\tau \in \mathbb{Z}$. We define the estimator of $\theta(\lambda, \tau) = a(\lambda, \tau)$ based on a generalized subsample via ($\lambda \in [0, 2\pi)$, $\tau \in \mathbb{Z}$)

$$\begin{aligned} \hat{\theta}_{n,b,(t_1,t_2,\ldots,t_k)}(\lambda, \tau) &= \frac{1}{k} \sum_{j=1}^{k} \hat{\theta}_{n,b,(t_j)} \\ &= \frac{1}{kb} \sum_{j=1}^{k} \sum_{t=t_j}^{t_j+b-1} X_t X_{t+\tau} \mathbf{1}_{\{1 \le t+\tau \le n\}} e^{i\lambda t}. \end{aligned} \tag{2.5}$$

In the same manner, for parameter $|\theta(\lambda, \tau)|$, we define $|\hat{\theta}_{n,b,(t_1,t_2,\ldots,t_k)}(\lambda, \tau)|$.

D. Spectral density matrix estimator in the APC case (see Lenart (2016)).

Note that, in the resampling world the normalizing sequence $\tau_{(\cdot)}$ that corresponds to $\hat{\theta}_{n,b,\mathbf{T}_k}$ may depend on both b and $\mathbf{T}_k$. To explain this let us consider the example **A** with $k = 2$ and assume additionally that $\sqrt{d_n}(\frac{1}{d_n} \sum_{t=c_n+1}^{c_n+d_n} X_t - \theta)$ has non-generate normal limit $N(0, \sigma^2)$, where d_n is any sequence of positive integers tending to infinity, as $n \to \infty$, and c_n is any sequence of nonnegative integers (see for example Theorem 2.1 in Lenart (2013) in a more general case). If $t_1 = t_2 = 1$, then $\mathcal{B}_{t_1} = \mathcal{B}_{t_2}$ and hence $\sqrt{b}(\hat{\theta}_{n,b,(t_1,t_2)} - \theta) = \sqrt{b}(\frac{1}{b} \sum_{t=1}^{b} X_t - \theta)$ has the same normal limit $N(0, \sigma^2)$, as $b \to \infty$. But if we consider $t_1 = 1$ and $t_2 = n - b + 1$ then the normalizing sequence should be $\sqrt{2b}$ to provide convergence of $\sqrt{2b}(\hat{\theta}_{n,b,(t_1,t_2)} - \theta)$ to the same normal limit $N(0, \sigma^2)$, as $n, b \to \infty$, $b/n \to 0$. Therefore, we introduce the notation $\tau_{b,\mathbf{T}_k}$ (rather than τ_b, as in the usual subsampling approach).

In the definition of generalized subsample, three parameters play a crucial role: the number of blocks k, the length of each block b, and finally the vector of starting points for blocks, i.e., $\mathbf{T}_k = (t_1, t_2, \ldots, t_k)$. To introduce the generalized subsampling (GS for short), we need to define the probability space in the resampling world. In this section, we formulate the general assumptions for the probability space in the resampling world, while in the next section, we restrict our attention to some particular cases. In brief, we assume below that the number of blocks k, the length of the blocks b, and finally the vector of starting points can be random. Let $b_1 = b_1(n)$ and $b_2 = b_2(n)$ be two sequences of the length of the blocks such that $b_1 \to \infty$, $b_2 \to \infty$ as $n \to \infty$ and such that $b_1 \le b_2$. Let $k_1 = k_1(n)$ and $k_2 = k_2(n)$ be two sequences of the number of blocks in a generalized subsample such that $k_1 \le k_2$. Finally, (b^*, k^*) denotes the random vector with the support

$$\mathbf{B} \times \mathbf{K} = \{b_1, b_1 + 1, b_1 + 2, \ldots, b_2\} \times \{k_1, k_1 + 1, k_1 + 2, \ldots, k_2\} \tag{2.6}$$

and with probability mass function denoted by $p_{(b,k)} = P^*\big((b^*,k^*) = (b,k)\big)$ for $(b,k) \in \mathbf{B}\times\mathbf{K}$. Next, conditionally at $(b^*,k^*) = (b,k) \in \mathbf{B}\times\mathbf{K}$, let $\mathbf{T}^*_k = (t^*_1, t^*_2, \ldots, t^*_k)$ be a random vector with the support $S^k_{n,b}$ and conditional probabilities denoted by $p_{(t_1,t_2,\ldots,t_k)|(b,k)} = p_{\mathbf{T}_k|(b,k)} = P^*\Big(\mathbf{T}^*_k = \mathbf{T}_k \in S^k_{n,b}\Big|(b^*,k^*) = (b,k) \in \mathbf{B}\times\mathbf{K}\Big)$. Note that the vector $(b^*,k^*,\mathbf{T}^*_{k^*})$ has a random length equal to k^*+2. For such a random element $(b^*,k^*,\mathbf{T}^*_{k^*})$, we denote $p_{(b,k,\mathbf{T}_k)} = P^*\big((b^*,k^*,\mathbf{T}^*_{k^*}) = (b,k,\mathbf{T}_k)\big)$.

In our generalization, we propose to approximate the distribution function $J_n(A)$ of (2.1) for any Borel set $A\in\mathbb{R}^s$ by the distribution function of the form

$$\begin{aligned} L_{n,\mathbf{B},\mathbf{K},p}(A) &= P^*(\tau_{b^*,\mathbf{T}^*_{k^*}}(\hat{\boldsymbol\theta}_{n,b^*,\mathbf{T}^*_{k^*}} - \hat{\boldsymbol\theta}_n) \in A) \\ &= E^*\left(\mathbf{1}\left\{\tau_{b^*,\mathbf{T}^*_{k^*}}(\hat{\boldsymbol\theta}_{n,b^*,\mathbf{T}^*_{k^*}} - \hat{\boldsymbol\theta}_n) \in A\right\}|(X_1,X_2,\ldots,X_n)\right), \end{aligned} \tag{2.7}$$

which is equivalent to

$$L_{n,\mathbf{B},\mathbf{K},p}(A) = \sum_{b_1\le b\le b_2}\sum_{k_1\le k\le k_2}\sum_{\mathbf{T}_k\in S^k_{n,b}} p_{(b,k,\mathbf{T}_k)}\mathbf{1}\left\{\tau_{b,\mathbf{T}_k}(\hat{\boldsymbol\theta}_{n,b,\mathbf{T}_k} - \hat{\boldsymbol\theta}_n)\in A\right\}. \tag{2.8}$$

Remark 2.2. Note that

$$L_{n,\mathbf{B},\mathbf{K},p}(A) = \sum_{b_1\le b\le b_2}\sum_{k_1\le k\le k_2} p_{(b,k)}P^*(\tau_{b,\mathbf{T}^*_k}(\hat{\boldsymbol\theta}_{n,b,\mathbf{T}^*_k} - \hat{\boldsymbol\theta}_n)\in A|(b^*,k^*) = (b,k))$$

$$= \sum_{b_1\le b\le b_2}\sum_{k_1\le k\le k_2} p_{(b,k)}\left(\sum_{\mathbf{T}_k\in S^k_{n,b}} p_{\mathbf{T}_k|(b,k)}\mathbf{1}\left\{\tau_{b,\mathbf{T}_k}(\hat{\boldsymbol\theta}_{n,b,\mathbf{T}_k} - \hat{\boldsymbol\theta}_n)\in A\right\}\right).$$

This means that $L_{n,\mathbf{B},\mathbf{K},p}(A)$ can be interpreted as a mixture of $(b_2-b_1+1)(k_2-k_1+1)$ generalized subsampling distributions with fixed (b,k) and weights $p_{(b,k)}$.

Assume that for any $b\in\mathbf{B}$, $k,k'\in\mathbf{K}$ and for any vectors of starting points $\mathbf{T}_k = (t_1,t_2,\ldots,t_k)\in S^k_{n,b}$, $\mathbf{T}'_{k'} = (t'_1,t'_2,\ldots,t'_{k'})\in S^{k'}_{n,b}$ such that $\bigcup_{j=1}^{k}\{t_j\} = \bigcup_{j=1}^{k'}\{t'_j\}$ we have $\tau_{b,\mathbf{T}_k} = \tau_{b,\mathbf{T}'_{k'}}$ and $\hat{\boldsymbol\theta}_{n,b,\mathbf{T}_k} = \hat{\boldsymbol\theta}_{n,b,\mathbf{T}'_{k'}}$. In such a case, the GS cumulative distribution function can be written as the following:

$$L_{n,\mathbf{B},\mathbf{K},p}(A) = \sum_{b_1\le b\le b_2}\sum_{\mathbf{T}_{k_2}\in S^{k_2}_{n,b}} p'_{(b,\mathbf{T}_{k_2})}\mathbf{1}\left\{\tau_{b,\mathbf{T}_{k_2}}(\hat{\boldsymbol\theta}_{n,b,\mathbf{T}_{k_2}} - \hat{\boldsymbol\theta}_n)\in A\right\}, \tag{2.9}$$

with appropriate $p'_{(b,\mathbf{T}_{k_2})}$ depending on the probabilities $p_{(b,k,\mathbf{T}_k)}$. The equality $\bigcup_{j=1}^{k}\{t_j\} = \bigcup_{j=1}^{k'}\{t'_j\}$ means that vectors $(t_1,t_2,\ldots,t_k)$ and $(t'_1,t'_2,\ldots,t'_{k'})$ conation the same elements but not necessarily in the same order and with the same multiplicity.

As a simple example, let's consider constant sequences $k_1 \equiv 1$ and $k_2 \equiv 2$ with the estimator $\hat{\boldsymbol{\theta}}_{n,b,\mathbf{T}_k}$ from above example A. In such a case for any fixed b ($b_1 \leq b \leq b_2$) and any $l \in S_{n,b}$ we have $\hat{\boldsymbol{\theta}}_{n,b,(l)} = \hat{\boldsymbol{\theta}}_{n,b,(l,l)} = \frac{1}{b}\sum_{t=l}^{l+b-1} X_t$ and hence $\tau_{b,(l)} = \tau_{b,(l,l)}$. Then

$$\begin{aligned} L_{n,\mathbf{B},\mathbf{K},p}(A) &= \sum_{b_1 \leq b \leq b_2} \sum_{k_1 \leq k \leq k_2} \sum_{\mathbf{T}_k \in S^k_{n,b}} p_{(b,k,\mathbf{T}_k)} \mathbf{1}\left\{\tau_{b,\mathbf{T}_k}(\hat{\boldsymbol{\theta}}_{n,b,\mathbf{T}_k} - \hat{\boldsymbol{\theta}}_n) \in A\right\}. \\ &= \sum_{b_1 \leq b \leq b_2} \sum_{l \in S_{n,b}} p_{(b,1,(l))} \mathbf{1}\left\{\tau_{b,(l)}(\hat{\boldsymbol{\theta}}_{n,b,(l)} - \hat{\boldsymbol{\theta}}_n) \in A\right\} \\ &+ \sum_{b_1 \leq b \leq b_2} \sum_{(l,w) \in S^2_{n,b}} p_{(b,2,(l,w))} \mathbf{1}\left\{\tau_{b,(l,w)}(\hat{\boldsymbol{\theta}}_{n,b,(l,w)} - \hat{\boldsymbol{\theta}}_n) \in A\right\} \\ &= \sum_{b_1 \leq b \leq b_2} \sum_{(l,w) \in S^2_{n,b}} p'_{(b,2,(l,w))} \mathbf{1}\left\{\tau_{b,(l,w)}(\hat{\boldsymbol{\theta}}_{n,b,(l,w)} - \hat{\boldsymbol{\theta}}_n) \in A\right\}, \end{aligned}$$

where $p'_{(b,2,(l,w))} = p_{(b,2,(l,w))} + \mathbf{1}\{l = w\}\, p_{(b,1,(l))}$.

Example 2.1. Take $k_1 = k_2 = 1$ and $b_1 = b_2 = b$ and a uniform distribution for t_1^* on $S_{n,b}$. Assume that $\tau_{b,(t)}$ does not depend on t and that $\tau_{b,(t)} = \tau_b$. In such a case, GS simplifies to the usual subsampling procedure examined in Politis et al. (1999),

$$\begin{aligned} L_{n,\mathbf{B},\mathbf{K},p}(A) &= E^*\left(\mathbf{1}\left\{\tau_{b,\mathbf{T}_1}(\hat{\boldsymbol{\theta}}_{n,b,(t_1^*)} - \hat{\boldsymbol{\theta}}_n) \in A\right\} | (X_1, X_2, \ldots, X_n)\right) \\ &= \frac{1}{n-b+1} \sum_{t_1=1}^{n-b+1} \mathbf{1}\left\{\tau_b(\hat{\boldsymbol{\theta}}_{n,b,(t_1)} - \hat{\boldsymbol{\theta}}_n) \in A\right\}. \end{aligned} \tag{2.10}$$

As was mentioned in the example above, if $b_2 = b_1 = b$, $k_1 = k_2 = 1$ and under uniform distribution for t_1^* on $S_{n,b}$ the GS reduces to the usual subsampling procedure. We refer to that case as subsampling of order one (S(1) for short). Generally, for any constant $k \in \mathbb{N}$ and with $k_1 = k_2 = k$ and $b_1 = b_2 = b$ and under uniform distribution for $\mathbf{T}_k^*$ on $S^k_{n,b}$, we denote the GS procedure via S(k) and call it *subsampling of order k (with block length b)*. The distribution function of S(k) procedure will be denoted by $L_{n,b,k}(\cdot)$, i.e.,

$$L_{n,b,k}(A) = \frac{1}{q^k_{n,b}} \sum_{\mathbf{T}_k \in S^k_{n,b}} \mathbf{1}\left\{\tau_{b,\mathbf{T}_k}(\hat{\boldsymbol{\theta}}_{n,b,\mathbf{T}_k} - \hat{\boldsymbol{\theta}}_n) \in A\right\}. \tag{2.11}$$

Example 2.2. Take $k_1 = k_2 = 1$ and a uniform distribution for t_1^* (conditionally on b^*). Assume that $\tau_{b,(t)}$ does not depend on t and that $\tau_{b,(t)} = \tau_b$. To simplify we write p_b in place of $p_{(b,1)}$. In such a case,

$$L_{n,\mathbf{B},\mathbf{K},p}(A) = \sum_{b_1 \le b \le b_2} \frac{p_b}{n-b+1} \sum_{t \in S_{n,b}} \mathbf{1}\left\{\tau_b(\hat{\boldsymbol{\theta}}_{n,b,(t)} - \hat{\boldsymbol{\theta}}_n) \in A\right\} \tag{2.12}$$

is a mixture of $b_2 - b_1 + 1$ subsampling distributions of order one (each with different b) with weights p_b.

Example 2.3. (Non-overlapping generalized subsample) Take $k_1 = k_2 = 2$, $b_1 = b_2 = b$. In such a case, by a non-overlapping generalized subsample, we mean that $|t_1 - t_2| > b$ for any $(t_1, t_2) \in S^2_{n,b}$. Elementary calculations show that $\sharp\{(t_1, t_2) \in S^2_{n,b} : |t_1 - t_2| \le b\} = n + b(2n-1) - b^2$. Note that in many cases (for considered non-overlapping generalized subsample) it is natural to take $\tau_{b,\mathbf{T}_2} = \tau_{2b}$, where τ_{2b} is given by (2.1). Under a uniform discrete distribution for $\mathbf{T}^*_2 = (t^*_1, t^*_2)$ on the set $\{(t_1, t_2) \in S^2_{n,b} : |t_1 - t_2| > b\}$ we have $p_{(b,2,\mathbf{T}_2)} = \frac{1}{n^2-n-b(2n-1)+b^2}$ and

$$\begin{aligned} L_{n,\mathbf{B},\mathbf{K},p}(A) =& \frac{1}{n^2 - n - b(2n-1) + b^2} \\ & \sum_{\mathbf{T}_2 \in S^2_{n,b} : |t_1 - t_2| > b} \mathbf{1}\left\{\tau_{2b}(\hat{\boldsymbol{\theta}}_{n,b,\mathbf{T}_2} - \hat{\boldsymbol{\theta}}_n) \in A\right\}. \end{aligned} \tag{2.13}$$

Assume that, for each pair of sets $\{t_1, t_2\}$, $\{t'_1, t'_2\}$, such that $\{t_1, t_2\} \ne \{t'_1, t'_2\}$, the estimator $\hat{\boldsymbol{\theta}}_{n,b,(\cdot)}$ has a different value. Then, the number of point masses in such a case equals $(n^2 - n - b(2n-1) + b^2)/2!$. Recall that, in the usual subsampling methodology, there are $n - b + 1$ possible different point masses.

Remark 2.3. Tewes et al. (2017) introduced an excellent idea of convolved subsampling. To compare this idea with generalized subsampling we unify the notation. We identify k in our work with k_n from Tewes et al. (2017) and q_n with N_n. The convolved subsampling estimator $C_{n,k_n}(x)$ of distribution function proposed by Tewes et al. (2017) can be expressed as

$$\begin{aligned} C_{n,k_n}(x) &= P^*(Z^* \le x) \\ &= P^*\Big(\frac{1}{\sqrt{k_n}} \sum_{j=1}^{k_n} (Y^*_{n,j} - m_{n,SUB}) \le x\Big), \qquad x \in \mathbb{R}. \end{aligned} \tag{2.14}$$

Note that the term $Y^*_{n,j}$ (for $j = 1, 2, \ldots, k$) in convolved subsampling is equivalent to $\tau_b(\hat{\theta}_{n,b,(t^*_j)} - \hat{\theta}_n)$ in generalized subsampling notation, where t^*_j are i.i.d.

from uniform distribution on the set $\{1, 2, \ldots, n-b+1\}$. Hence

$$\begin{aligned} C_{n,k_n}(x) &= P^*\left(\frac{1}{\sqrt{k_n}}\sum_{j=1}^{k_n}[\tau_b(\hat{\theta}_{n,b,(t_j^*)} - \hat{\theta}_n) - m_{n,SUB}] \leq x\right) \\ &= P^*\left(\frac{\tau_b}{\sqrt{k_n}}\sum_{j=1}^{k_n}\left[(\hat{\theta}_{n,b,(t_j^*)} - \hat{\theta}_n) - N_n^{-1}\sum_{i=1}^{N_n}(\hat{\theta}_{n,b,(t_i)} - \hat{\theta}_n)\right] \leq x\right) \\ &= P^*\left(\frac{\tau_b}{\sqrt{k_n}}\sum_{j=1}^{k_n}\left[\hat{\theta}_{n,b,(t_j^*)} - E^*(\hat{\theta}_{n,b,(t_j^*)})\right] \leq x\right) \\ &= P^*\left(\tau_b\sqrt{k_n}\left(\frac{1}{k_n}\sum_{j=1}^{k_n}\hat{\theta}_{n,b,(t_j^*)} - E^*\left(\frac{1}{k_n}\sum_{j=1}^{k_n}\hat{\theta}_{n,b,(t_j^*)}\right)\right) \leq x\right). \end{aligned}$$

The last equality shows the direct relation between generalized subsampling and convolved subsampling. Assuming that (2.4) holds and that t_j^* are i.i.d. from uniform distribution on the set $\{1, 2, \ldots, n-b+1\}$ and that $\tau_{b,\mathbf{T}_k} = \tau_b\sqrt{k}$, we conclude that the shape of generalized subsampling distribution is the same as the shape of convolved subsampling distribution. The difference is only in the location. In the case of generalized subsampling, the primary issue is to construct an estimator $\hat{\boldsymbol{\theta}}_{n,b,\mathbf{T}_k}$ based on generalized subsample with parameters b, k and t_k. Second important issue is to construct normalizing sequence $\tau_{b,\mathbf{T}_k}$, to be able to prove the sufficient conditions for consistency. Finally, the generalized subsampling distribution is built on the basis of $\hat{\boldsymbol{\theta}}_{n,b^*,\mathbf{T}^*_{k^*}}$, where (b^*, k^*) and $\mathbf{T}^*_{k^*}$ (conditionally on $(b^*, k^*) = (b, k) \in \mathbf{B}\times\mathbf{K}$) has some given discrete distribution. In relation to generalized subsampling the distribution of convolved subsampling is based on $\hat{\boldsymbol{\theta}}_{n,b^*,\mathbf{T}^*_{k^*}} = \frac{1}{k}\sum_{j=1}^{k}\hat{\theta}_{n,b,(t_j^*)}$, with one-point distribution for b^* (we have $b^* = b$) and one-point distribution for k^* (we have $k^* = k$), and with t_j^* as i.i.d. from uniform distribution on the set $\{1, 2, \ldots, n-b+1\}$ for $j = 1, 2, \ldots, k$. The consistency of convolved subsampling was proved in Tewes et al. (2017) for many cases, when the limiting distribution of $\tau_n(\hat{\theta}_n - \theta)$ is Gaussian. Hence, the problem may arise for convolved subsampling when the asymptotic distribution is not Gaussian. As a problematic example, let us consider the problem of statistical inference for parameter $|\theta(\psi)|$ from point B above. In such a case the subsampling is consistent (see Theorem 2.3 in Lenart (2013)) but the assumption (2.3) from Tewes et al. (2017) do not hold for $\theta(\psi) = 0$ (see Theorem 2.2 in Lenart (2013)).

Example 2.4. (Illustrative example) Analogously to example **B**, for a time series $\{X_t : t \in \mathbb{Z}\}$, we assume that the expectation function exists and that $\mu(t) = E(X_t) = \sum_{\psi\in\Psi} m(\psi)e^{i\psi t}$ is an almost periodic function with $\Psi = \{\psi \in [0, 2\pi) : |m(\psi)| \neq 0\}$, where $m(\psi) = \lim_{n\to\infty}\sum_{t=1}^{n}\mu(t)e^{-i\psi t}$. Note that $\psi \in [0, 2\pi) \Leftrightarrow |m(\psi)| \neq 0$. Take $\psi \in [0, 2\pi)$ and assume that we are interested in approximating the unknown distribution function of

$$\sqrt{n}(|\hat{m}_n(\psi)| - |m(\psi)|), \tag{2.15}$$

where $\hat{m}_n(\psi)$ is the estimator (based on the entire sample) of the Fourier coefficient $m(\psi)$.

Note that the fundamental step in the GS is the appropriate definition of the estimator of the parameter of interest. As in example **B**, we define the estimator (based on the generalized subsample) of the magnitude of the Fourier coefficient (for any $\psi \in [0, 2\pi)$) in natural way:

$$|\hat{\theta}_{n,b,(t_1,t_2,\dots,t_k)}(\psi)| = \left| \frac{1}{k} \sum_{j=1}^{k} \hat{\theta}_{n,b,(t_j)} \right| = \left| \frac{1}{kb} \sum_{j=1}^{k} \sum_{t=t_j}^{t_j+b-1} X_t e^{it\psi} \right|.$$

In section 4, we show that generalized subsampling is consistent in a case such as the approximating distribution function of (2.15). Here, we construct only an illustrative example. Let us consider a simple case: $X_t = 2\sin(0.15t) + \epsilon_t$, where ϵ_t is a Gaussian white noise with variance of one, $b_2 = b_1 = b$, and $k_2 = k_1 = k$. For any $k = 1, 2, 3, 4$, the cumulative distribution function for subsampling of order k (with block length b) was evaluated (see Figure 1). In each case, we arbitrarily fix $b = [3.5\sqrt{n}]$. We consider two sample sizes $n \in \{80, 150\}$ and two frequencies $\psi \in \{0.15, 0.5\}$.

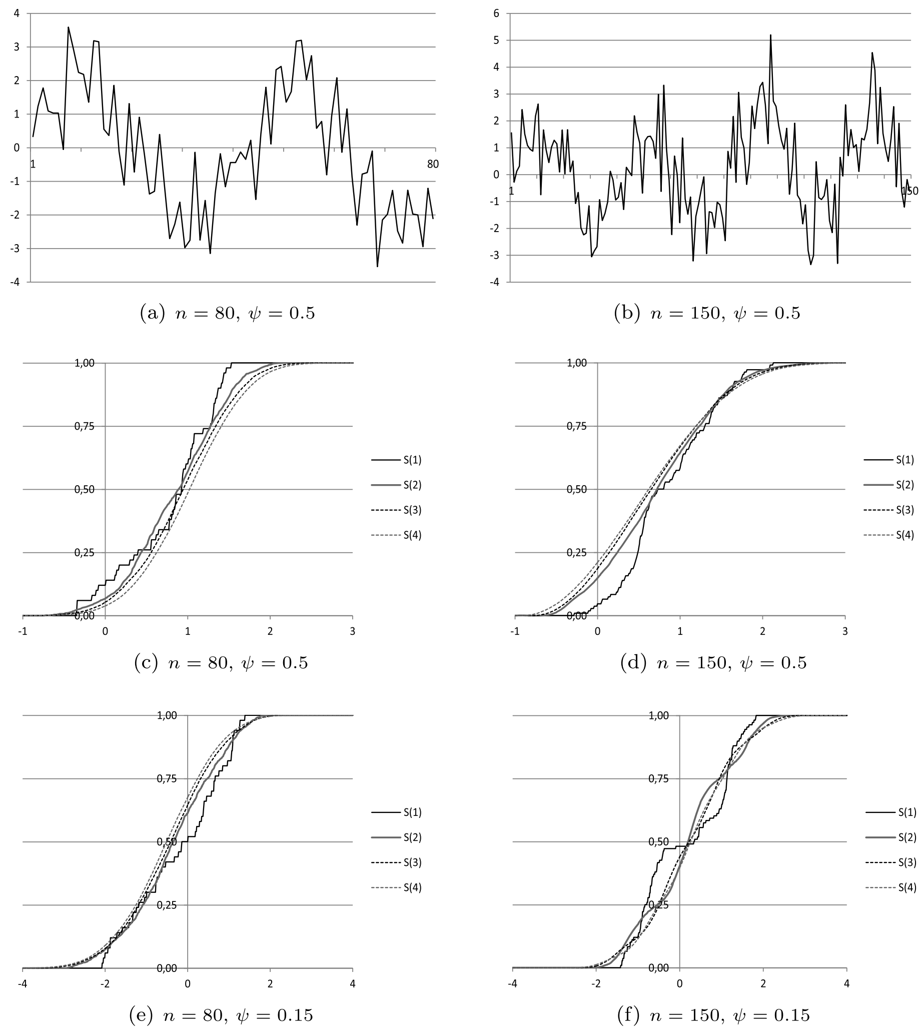
The cumulative distribution function of S(k) for $k > 1$ is 'smoother' than for the usual subsampling ($k = 1$). This is a consequence of the higher number of different point masses for $k > 1$ compared to the usual subsampling procedure. The second reason is the relatively small sample sizes. The shape of the cumulative distribution function for $k = 1$ diverges from subsampling of higher orders. Hence, the above illustrative example shows that the order of subsampling may have serious implications on the obtained results, i.e., the results of testing procedures, quantiles, etc.

Remark 2.4. In some cases, the distribution $L_{n,\mathbf{B},\mathbf{K},p}(A)$ cannot be efficiently calculated due to long computational time. This situation may appear when, for example, k is large. In such a situation, we can approximate the resampling distribution using Monte Carlo (MC) approximation of the expectation value. This can be done by noticing that $L_{n,\mathbf{B},\mathbf{K},p}(A)$ can be interpreted as expectation of $\mathbf{1}\left\{\tau_{b^*,\mathbf{T}^*_{k^*}}(\hat{\boldsymbol{\theta}}_{n,b^*,\mathbf{T}^*_{k^*}} - \hat{\boldsymbol{\theta}}_n) \in A\right\}$ in the resampling world (see (2.7)).

In the next section, we formulate the sufficient conditions for the consistency of the generalized subsampling procedure.

3. Sufficient conditions for consistency of the generalized subsampling procedure

In this section, the sufficient conditions for the consistency of GS are formulated in similar manner to the formulation in Politis et al. (1999). To simplify our consideration, we restrict our attention to the univariate case ($s = 1$) from now on.



(a) $n = 80$, $\psi = 0.5$ (b) $n = 150$, $\psi = 0.5$

(c) $n = 80$, $\psi = 0.5$ (d) $n = 150$, $\psi = 0.5$

(e) $n = 80$, $\psi = 0.15$ (f) $n = 150$, $\psi = 0.15$

FIG 1. *Cumulative distribution functions $L_{n,b,k}(x)$ of S(1) (subsampling), S(2), S(3), and S(4) for different $n \in \{80, 150\}$, $b = [3.5\sqrt{n}]$, and $\psi \in \{0.15, 0.5\}$.*

The generalization to the multivariate case is natural. Furthermore, in this section, we formulate a special assumption concerning probabilities $p_{(b,k,\mathbf{T}_k)}$, which simplifies the proofs of sufficient conditions for consistency of the generalized subsampling procedure. More precisely, we assume that (b^*, k^*) has a discrete uniform distribution and $(t_1^*, t_2^*, \ldots, t_k^*)$ conditionally on $(b^*, k^*) = (b, k)$ has also a discrete uniform distribution (see the assumption below).

Assumption 3.1. *Assume that (b^*, k^*) has a discrete uniform distribution on $\boldsymbol{B} \times \boldsymbol{K} = \{b_1, b_1 + 1, b_1 + 2, \ldots, b_2\} \times \{k_1, k_1 + 1, k_1 + 2, \ldots, k_2\}$ (see (2.6)) and conditionally on $(b^*, k^*) = (b, k)$ the random vector $(t_1^*, t_2^*, \ldots, t_k^*)$ has also a* *discrete uniform distribution on* $S_{n,b}^k = \{1, 2, \ldots, n-b+1\}^k$*, which means that*

$$p_{(b,k)} = \frac{1}{(b_2 - b_1 + 1)(k_2 - k_1 + 1)},$$

$$p_{(b,k,(t_1,t_2,\ldots,t_k))} = \frac{1}{(b_2 - b_1 + 1)(k_2 - k_1 + 1)q_{n,b}^k}.$$

Remark 3.1. Note that under Assumption 3.1 the generalized subsampling distribution is a mixture of $(b_2 - b_1 + 1)(k_2 - k_1 + 1)$ subsampling distributions of order k (with block length b) with equal weights. Hence (in multivariate case)

$$L_{n,\mathbf{B},\mathbf{K},p}(A) = \frac{1}{(b_2 - b_1 + 1)(k_2 - k_1 + 1)} \sum_{b \in \mathbf{B}} \sum_{k \in \mathbf{K}} L_{n,b,k}(A). \tag{3.1}$$

Following Assumptions 4.2.1 and 4.2.2 (and 4.3.1 and 4.3.2, respectively, in the multivariate case) in Politis et al. (1999), we formulate sufficient conditions for the consistency of generalized subsampling. Using the same notations as in Politis et al. (1999), we denote the limiting distribution function (as $n \to \infty$) of $\tau_n(\hat{\theta}_n - \theta)$ by $J(\cdot, P)$ and the law of this distribution by $J(P)$. For any $(b, k) \in \mathbf{B} \times \mathbf{K}$ and $(t_1, t_2, \ldots, t_k) \in S_{n,b}^k$, we define

$$J_{b,k,(t_1,t_2,\ldots,t_k)}(x, P) = P(\tau_{b,\mathbf{T}_k}(\hat{\theta}_{n,b,(t_1,t_2,\ldots,t_k)} - \hat{\theta}_n) \leq x).$$

We reformulate Assumptions 4.2.1 and 4.2.2 from Politis et al. (1999) with appropriate adjustment to GS. Assumption 3.2 corresponds to Assumption 4.2.1 in Politis et al. (1999), and Assumption 3.3 corresponds to Assumption 4.2.2. Additionally, in the case where $k > 1$, we formulate new sufficient conditions (Assumptions 3.4 and 3.5) under which generalized subsampling is consistent.

Assumption 3.2. *Assumption 3.1 holds and there exists a limiting low* $J(P)$ *such that*

(i) $J_n(P)$ *converges weakly to* $J(P)$ *as* $n \to \infty$

(ii) *for every continuity point* x *of* $J(P)$*, we have*

$$\sum_{b_1 \leq b \leq b_2} \sum_{k_1 \leq k \leq k_2} \sum_{\boldsymbol{T}_k \in S_{n,b}^k} p_{(b,k,\boldsymbol{T}_k)} J_{b,k,\boldsymbol{T}_k}(x, P) \to J(x, P) \tag{3.2}$$

The next assumption is stronger than 3.2 and is related to Assumption 4.2.2 of Politis et al. (1999).

Assumption 3.3. *Assumption 3.1 holds and there exists a limiting low* $J(P)$ *such that*

(i) $J_n(P)$ *converges weakly to* $J(P)$ *as* $n \to \infty$

(ii) for every continuity point x of $J(P)$, for any integer sequences $\tilde{b}$ and $\tilde{k}$, such that $b_1 \leq \tilde{b} \leq b_2$ and $k_1 \leq \tilde{k} \leq k_2$, and for any vector of integer sequence $(t_{n,1}, t_{n,2}, \dots, t_{n,\tilde{k}})$ such that $1 \leq t_{n,i} \leq n - \tilde{b} + 1$ for $i = 1, 2, \dots, \tilde{k}$, we have

$$J_{\tilde{b},\tilde{k},(t_{n,1},t_{n,2},\dots,t_{n,\tilde{k}})}(x, P) \to J(x, P) \tag{3.3}$$

Note that, for $k > 1$, it is possible to get a vector of starting points $\mathbf{T}_k = (t_1, t_2, \dots, t_k) \in S^k_{n,b}$ such that the blocks $\mathcal{B}_{t_1}, \mathcal{B}_{t_2}, \dots, \mathcal{B}_{t_k}$ of which the subsample $\mathcal{B}_{b,\mathbf{T}_k}$ is composed are not mutually disjoint. For example, if $t_i = t_j$, for some $i \neq j$, then $\mathcal{B}_{t_i} = \mathcal{B}_{t_j}$. This example is a special case of overlapping blocks in the generalized subsample $\mathcal{B}_{b,\mathbf{T}_k}$ (see also Example 2.1 in the case S(2)). Formally, for $k > 1$ by overlapping blocks in the generalized subsample, we mean that

$$|t_i - t_j| \leq b \quad \text{for some } \ i \neq j, \text{ where } i, j \in \{1, 2, \dots, k\}. \tag{3.4}$$

which is equivalent that there exist $i, j \in \{1, 2, \dots, k\}$, $i \neq j$, such that $\mathcal{B}_{t_i}$ and $\mathcal{B}_{t_j}$ are not disjoint. To clarify the problem of overlapping blocks in the generalized subsample, let us define in the natural manner the following set of vectors of starting points $R_{n,b,k} \subset S^k_{n,b}$:

$$R_{n,b,k} = \{(t_1, t_2, \dots, t_k) \in S^k_{n,b} : \min_{i,j \in \{1,2,\dots,n-b+1\},\ i \neq j} |t_i - t_j| \leq b\}. \tag{3.5}$$

By the above definition, the set $R_{n,b,k}$ contains all generalized subsamples where overlapping blocks appear in the sense of (3.4). Elementary combinatorics show that the number of elements in the set $R_{n,b,k}$ has the following bound:

$$\sharp R_{n,b,k} \leq k(k-1)(n + b(2n-1) - b^2)q^{k-2}_{n,b}.$$

This means that, under Assumption 3.1

$$\sum_{b_1 \leq b \leq b_2} \sum_{k_1 \leq k \leq k_2} \sum_{\mathbf{T}_k \in R_{n,b,k}} p_{(b,k,\mathbf{T}_k)} J_{b,k,\mathbf{T}_k}(x, P) = O(k_2^2 b_2 / n), \tag{3.6}$$

since $J_{b,k,(t_1,t_2,\dots,t_k)}(x, P)$ is nonnegative and bounded by one uniformly at $(b, k, \mathbf{T}_k)$. Hence, under the assumption that $k_2^2 b_2/n \to 0$ as $n \to \infty$, Assumption 3.2 is equivalent to the following assumption:

Assumption 3.4. *Assumption 3.1 holds and there exist a limiting low $J(P)$ such that*

(i) $J_n(P)$ converges weakly to $J(P)$ as $n \to \infty$
(ii) for every continuity point x of $J(P)$, we have

$$\sum_{b_1 \leq b \leq b_2} \sum_{k_1 \leq k \leq k_2} \sum_{\boldsymbol{T}_k \in S^k_{n,b} \setminus R_{n,b,k}} p_{(b,k,\boldsymbol{T}_k)} J_{b,k,\boldsymbol{T}_k}(x, P) \to J(x, P) \tag{3.7}$$

The next assumption is stronger than 3.4 and is related to Assumption 4.2.2 of Politis et al. (1999).

Assumption 3.5. *Assume Assumption 3.1 and there exists a limiting low $J(P)$ such that*

(i) $J_n(P)$ *converges weakly to* $J(P)$ *as* $n \to \infty$

(ii) for every continuity point x of $J(P)$, any integer sequences $\tilde{b}$ and $\tilde{k}$ such that $b_1 \le \tilde{b} \le b_2$ and $k_1 \le \tilde{k} \le k_2$, and any vector of integer sequences $(t_{n,1}, t_{n,2}, \dots, t_{n,\tilde{k}})$ such that:

A) $1 \le t_{n,i} \le n - \tilde{b} + 1$, *for* $i = 1, 2, \dots, \tilde{k}$

B) $\min_{i,j \in \{1,2,\dots,\tilde{k}\},\ i \neq j} |t_{n,i} - t_{n,j}| > \tilde{b}$, *for* $\tilde{k} > 1$,

we have

$$J_{\tilde{b},\tilde{k},(t_{n,1},t_{n,2},\dots,t_{n,\tilde{k}})}(x, P) \to J(x, P). \tag{3.8}$$

Notice that Assumption 3.2 is implied by Assumption 3.3. If we assume that $k_2^2 b_2/n \to 0$, Assumption 3.2 is also implied by Assumption 3.5. In the main theorem of this section (see the theorem below), we assume that the convergence of $k_2^2 b_2/n \to 0$ holds. Analogically to the original Theorem 4.2.1 in Politis et al. (1999), the following theorem concerning the consistency of GS holds. This theorem differs from the original in assumptions concerning the α-mixing sequence and additional sequence k (with relations to n and b). We formulate this theorem with adequate adjustment.

Theorem 3.1. *Assume Assumption 3.2 holds and that*

$$\max_{b,k,\boldsymbol{T}_k} \{\tau_{b,\boldsymbol{T}_k} : b_1 \le b \le b_2, k_1 \le k \le k_2, \boldsymbol{T}_k \in S^k_{n,b}\}/\tau_n \to 0,$$

$k_2^2 b_2/n \to 0$, $b_1 \to \infty$, $n \to \infty$. *Additionally, assume that the considered time series is α-mixing with mixing sequence $\alpha(\cdot)$ such that* $\frac{k_2^2}{n} \sum_{s=1}^{n} \alpha(s) \to 0$. *Then, we have that*

(i) If x is a continuity point of $J(\cdot, P)$, then $L_{n,\boldsymbol{B},\boldsymbol{K},p}(x) \to J(x, P)$ in probability

(ii) If $J(\cdot, P)$ is continuous, then

$$\sup_{x \in \mathbb{R}} |L_{n,\boldsymbol{B},\boldsymbol{K},p}(x) - J(x, P)| \to 0$$

in probability

(iii) For $\alpha \in (0, 1)$, let $c_{n,\boldsymbol{B},\boldsymbol{K},p}(1 - \alpha) = \inf\{x : L_{n,\boldsymbol{B},\boldsymbol{K},p}(x) \geqslant 1 - \alpha\}$. Correspondingly, $c(1 - \alpha, P) = \inf\{x : J(x, P) \geqslant 1 - \alpha\}$. If $J(\cdot, P)$ is continuous at point $c(1 - \alpha, P)$, then

$$P(\tau_n(\hat{\theta}_n - \theta) \leqslant c_{n,\boldsymbol{B},\boldsymbol{K},p}(1 - \alpha)) \to 1 - \alpha.$$

Proof. The proof can be found in the Appendix. □

Remark 3.2. Note that, if $k_1 = k_2 \equiv 1$ and $b_1 = b_2 = b$, then above theorem is equivalent to the original Theorem 4.2.1 presented in Politis et al. (1999). The only difference is that we assume in such a case that the assumption $\frac{k_2^2}{n}\sum_{s=1}^{n}\alpha(s) \to 0$ holds. This assumption simplifies to $\frac{1}{n}\sum_{s=1}^{n}\alpha(s) \to 0$, if $k_1 = k_2 \equiv 1$. However, the last convergence is a simple consequence of α-mixing properties.

Remark 3.3. Note that, in the usual subsampling methodology, the problem of overlapping blocks (in the above sense) does not exist since $k = 1$. Note that, if k_2 is a bounded sequence or constant, then Assumption 3.2 is equivalent to Assumption 3.4 under the condition $b_2/n \to 0$. If we take sequences $k_1 = k_2 = k$ and $b_1 = b_2 = b$ such that $kb \approx n$, then our procedure is equivalent to the usual MBB procedure. However, the formulated sufficient conditions for the consistency of generalized subsampling excludes the case $kb \approx n$, by noting that $k^2b/n \to 0$. The condition $k^2b/n \to 0$ can be probably weaken. However, the sufficient conditions for the consistency of the generalized subsampling procedure in the case that $kb \approx n$ can be much more sophisticated (if they exist) because many examples of bootstrap inconsistency are known in the literature.

4. Consistency for Fourier coefficients in the APC case

In this section, we consider an Almost Periodically Correlated univariate time series $\{X_t : t \in \mathbf{Z}\}$ with an almost periodic mean function $\mu(t) = E(X_t)$, with representation:

$$\mu(t) \sim \sum_{\psi\in\Psi} m(\psi)e^{i\psi t}, \tag{4.1}$$

and almost periodic autocovariance function $B(t,\tau) = \operatorname{cov}(X_t, X_{t+\tau})$, with representation:

$$B(t,\tau) \sim \sum_{\lambda\in\Lambda_\tau} a(\lambda,\tau)e^{i\lambda t}, \tag{4.2}$$

for any $\tau \in \mathbf{Z}$. The Fourier coefficients $m(\psi)$ and $a(\lambda,\tau)$ equal to (see Hurd (1989, 1991); Dehay and Hurd (1994))

$$m(\psi) = \lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}\mu(j)e^{-i\psi j}, \qquad a(\lambda,\tau) = \lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}B(j,\tau)e^{-i\lambda j}.$$

The sets $\Psi = \{\psi \in [0,2\pi) : m(\psi) \neq 0\}$ and $\Lambda_\tau = \{\lambda \in [0,2\pi) : a(\lambda,\tau) \neq 0\}$ are countable (see Corduneanu (1989)). To simplify this section, we assume that the sets Ψ and $\Lambda = \bigcup_{\tau\in\mathbb{Z}}$ are finite. In such a case, the representations (4.1) and (4.2) become equalities, and the spectral density function exists (see Lenart (2013)). By $P(\nu,\omega)$ we denote the extension of the spectral density function to the bifrequency square $(0,2\pi]^2$,

$$P(\nu,\omega) = \frac{1}{2\pi}\sum_{\tau=-\infty}^{\infty} a(\nu-\omega,\tau)e^{-i\ \nu\tau}, \tag{4.3}$$

for any $(\nu, \omega) \in (0, 2\pi]^2$.

Take any $\psi \in [0, 2\pi)$ and let $\theta = |m(\psi)|$ be the parameter of interest. Note that $\psi \in \Psi \Leftrightarrow |m(\psi)| \neq 0$, which means that identification of the frequencies in the set Ψ can be based on identification of the non-zero Fourier coefficients $m(\psi)$.

At the beginning, we formulate general assumptions concerning the block size b, their number k and starting points. Let $k = k_n$ and $b = b_n$ be any sequences of positive integers such that $b \to \infty$ and let $t_{n,i}$, for any $i = 1, 2, \ldots, k$, be a sequence of positive integers such that $t_{n,i} < n - b + 1$. For $k > 1$, we additionally assume thet we consider only non-overlapping blocks in the generalized subsample, i.e.,

$$\min_{i,j \in \{1,2,\ldots,k\},\ i \neq j} |t_{n,i} - t_{n,j}| > b. \tag{4.4}$$

Finally, under the above assumptions we define the estimator $\hat{\theta}_{n,b,(t_{n,1},t_{n,2},\ldots,t_{n,k})} = |\hat{m}_{n,b,(t_{n,1},t_{n,2},\ldots,t_{n,k})}|$ of the parameter θ based on the subsample (with a variable vector of starting points)

$$\begin{aligned}(&X_{t_{n,1}}, X_{t_{n,1}+1}, \ldots, X_{t_{n,1}+b-1}, X_{t_{n,2}}, X_{t_{n,2}+1}, \ldots, X_{t_{n,2}+b-1}, \ldots, \\ &X_{t_{n,k}}, X_{t_{n,k}+1}, \ldots, X_{t_{n,k}+b-1})\end{aligned} \tag{4.5}$$

via

$$\hat{\theta}_{n,b,(t_{n,1},t_{n,2},\ldots,t_{n,k})} = \left| \frac{1}{kb} \sum_{i=1}^{k} \sum_{j=t_{n,i}}^{t_{n,i}+b-1} X_j e^{-\mathrm{i}\psi j} \right|. \tag{4.6}$$

For $k = 1$ (usual subsampling), the above estimator was proposed in Lenart (2013). The following theorem is a natural adjustment of Theorem 2.1 presented in Lenart (2013) to the generalized subsample.

Theorem 4.1. *Let $\{X_t : t \in \mathbb{Z}\}$ be an APC time series with finite sets Ψ and Λ. Assume that there exist constants $\delta > 0$, and $\Delta, K \in \mathbb{R}$ such that*

(i) $\sup_{t \in \mathbb{Z}} \|X_t\|^{2+\delta} < \Delta$,

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

Take any $\psi \in [0, 2\pi)$. Assume that $k/b \to 0$ and $b \to \infty$ as $n \to \infty$. Let $\tilde{\mathbf{T}}_{n,k} = (t_{n,1}, t_{n,2}, \ldots, t_{n,k})$ be any vector of sequences of positive integers such that, for any sequence $\{t_{n,i}\}_{n \in \mathbb{Z}}$, $i = 1, 2, \ldots, k$, we have the inequality $t_{n,i} < n - b + 1$ for $n \in \mathbb{Z}$. Additionally, for $k > 1$, we assume that (4.4) holds (i.e., $\min_{i,j \in \{1,2,\ldots,k\},\ i \neq j} |t_{n,i} - t_{n,j}| > b$). Then we have the convergence

$$\sqrt{bk} \left(\begin{bmatrix} \mathrm{Re}(\hat{m}_{n,b,\tilde{\mathbf{T}}_{n,k}}(\psi)) \\ \mathrm{Im}(\hat{m}_{n,b,\tilde{\mathbf{T}}_{n,k}}(\psi)) \end{bmatrix} - \begin{bmatrix} \mathrm{Re}(m(\psi)) \\ \mathrm{Im}(m(\psi)) \end{bmatrix} \right) \xrightarrow{d} \mathcal{N}_2(0, \Omega(\psi)),$$

where

$$\Omega(\psi) = \pi g_0(\psi) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \pi \begin{pmatrix} \mathrm{Re}[P(\psi, 2\pi - \psi)] & \mathrm{Im}[P(\psi, 2\pi - \psi)] \\ \mathrm{Im}[P(\psi, 2\pi - \psi)] & -\mathrm{Re}[P(\psi, 2\pi - \psi)] \end{pmatrix}.$$

Note that through assumption (4.4) it is possible to consider one normalizing sequence $\tau_{b,\tilde{\boldsymbol{T}}_{n,k}} = \sqrt{bk}$ in the theorem above. The above theorem means that the estimator (4.6) is well defined in the sense, that after appropriate normalizing, is asymptotically normally distributed. Hence, as a consequence of the above theorem, we conclude that $\hat{m}_{n,b,\tilde{\boldsymbol{T}}_{n,k}}(\psi) \xrightarrow{p} m(\psi)$ as $n \to \infty$. The next theorem concerns the asymptotic distribution for the normalized estimator $|\hat{m}_{n,b,\tilde{\boldsymbol{T}}_{n,k}}(\psi)|$, while the last concerns the consistency of generalized subsampling for $|m(\psi)|$. These two theorems are generalizations of Theorems 4.2 and 4.3 in Lenart (2013). We formulate these theorems with appropriate adjustment.

Theorem 4.2. *Assume that all assumptions of Theorem 4.1 hold. Then, for any $\psi \in [0, 2\pi)$, we have the convergence*

$$\sqrt{bk}(|\hat{m}_{n,b,\tilde{\mathbf{T}}_{n,k}}(\psi)|-|m(\psi)|) \xrightarrow{d} J^{\{\psi\}} := \begin{cases} \mathcal{L}(Z), & \text{for } m(\psi)=0, \\ \mathcal{N}_1(0, \mathrm{D}_0\Omega(\psi)\mathrm{D}_0^T), & \text{for } m(\psi)\neq 0, \end{cases}$$

where

$$D_0^T = \begin{bmatrix} \mathrm{Re}(m(\psi)) \\ \mathrm{Im}(m(\psi)) \end{bmatrix} / |m(\psi)|,$$

$Z = \sqrt{B_1^2 + B_2^2}$, and the random vector $(B_1, B_2)^T$ has a two-dimensional normal distribution with zero mean and variance-covariance matrix equals $\Omega(\psi)$. If we assume additionally that $g_0(\xi) > 0$ for any $\xi \in [0, 2\pi)$ and that $\det(\Omega(\psi)) > 0$ for any $m(\psi) \neq 0$ and $\psi \in (0, \pi)$, then the law $J^{\{\psi\}}$ is continuous.

Theorem 4.3. *Let $\{X_t : t \in \mathbb{Z}\}$ be an APC time series with finite sets Ψ and Λ. Assume that there exist constants $\delta > 0$, Δ, and $K \in \mathbb{R}$ such that*

(i) $\sup_{t\in\mathbb{Z}} \|X_t\|^{2+\delta} < \Delta$

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

Additionally, assume that $k_2^2 b_2/n \to 0$, $b_1 \to \infty$, and $k_2/b_1 \to 0$ for $n \to \infty$. Take any $\psi \in [0, 2\pi)$. Then, generalized subsampling is consistent for $\theta = |m(\psi)|$ with $\hat{\theta}_{n,b,(t_1,t_2,\ldots,t_k)}$ given by (4.6) and $\tau_{b,\boldsymbol{T}_k} = \sqrt{bk}$, i.e., Theorem 3.1, holds.

At the end of this section we construct an easy example of APC time series for which Moving Block Bootstrap fails and the above Theorem 4.3 holds, i.e., generalized subsampling is consistent.

Example 4.1. We consider now the problem of consistency of generalized subsampling, convolved subsampling and MBB procedure for the examined in this section parameter $\theta = |m(\psi)|$, where $\psi \in [0, 2\pi)$. We start from the simple construction of the APC time series $\{X_t : t \in \mathbb{Z}\}$ for which the parameter $\theta = |m(\psi)|$ will be considered.

Firstly, let us assume that the zero-mean Gaussian time series $\{\eta_t : t \in \mathbb{Z}\}$ is stationary and that the law of the iterated logarithm:

$$\limsup_{n\to\infty} \frac{\eta_1 + \eta_2 + \ldots + \eta_n}{\sqrt{2n\ln(\ln n)}} = \varsigma \quad \text{w.p. } 1 \tag{4.7}$$

holds for some $0 < \varsigma < \infty$. Additionally, we assume that assumption ii) in Theorem 4.3 (concerning α-mixing) holds. To construct such a time series we can use Corollary 6 in Zhao and Woodroofe (2008), where condition with ρ-mixing was formulated. A very simple example of such time series $\{\eta_t : t \in \mathbb{Z}\}$ is Gaussian white noise or zero-mean stationary Gaussian $m-$dependent time series. Now, define the APC time series $\{X_t : t \in \mathbb{Z}\}$ as

$$X_t = a\cos(\lambda t) + \eta_t, \quad t \in \mathbb{Z},$$

with almost periodic mean function $\mu(t) = E(X_t) = a\cos(\lambda t)$, where $a \in \mathbb{R}$ and $\lambda \in (0, \pi)$ are unknown parameters. Note that for the time series $\{X_t : t \in \mathbb{Z}\}$ the above law of the iterated logarithm also holds, i.e.:

$$\limsup_{n\to\infty} \frac{X_1 + X_2 + \ldots + X_n}{\sqrt{2n\ln(\ln n)}} = \varsigma \quad \text{w.p. } 1. \tag{4.8}$$

To justify this, let's note that $\sum_{j=1}^{n}\cos(j\lambda) = \csc\left(\frac{\lambda}{2}\right)\sin\left(\frac{\lambda n}{2}\right)\cos\left(\frac{1}{2}\lambda(n+1)\right)$ and $\lim_{n\to\infty}\left[\csc\left(\frac{\lambda}{2}\right)\sin\left(\frac{\lambda n}{2}\right)\cos\left(\frac{1}{2}\lambda(n+1)\right)/\sqrt{2n\log(\log(n))}\right] = 0$, for $\lambda \in (0, \pi)$.

The consistency of convolved subsampling can not be justified on the basis of results formulated in Tewes et al. (2017), because the asymptotic distribution from Theorem 4.2 is not Gaussian for $|m(\psi)| = 0$.

Note that under assumptions formulated in Theorem 4.3 the generalized subsampling is consistent for $\theta = |m(\psi)|$, for any $\psi \in [0, 2\pi)$. But the assumption $b_2k_2^2/n \to 0$ (from this theorem) excludes the case of Moving Block Bootstrap procedure (see Remark 3.3).

To show the inconsistency of MBB procedure we use the similar arguments as in the Example 1 of Babu (1984). We take $\psi = 0$, which means that $|m(\psi)| = 0$. In such a case the statistics $\sqrt{n}|\hat{\theta}_n| = \left|\frac{1}{\sqrt{n}}\sum_{t=1}^{n} X_t\right|$ has asymptotic half-normal distribution with cumulative distribution function $F(x) = \operatorname{erf}\left(\frac{\sigma x}{\sqrt{\pi}}\right)$, where $\sigma^2 > 0$ is the limiting variance of $\sqrt{n}\hat{\theta}_n$. To consider the MBB procedure we use generalized subsampling notation (with $b_1 = b_2 = b$, $k_1 = k_2 = k$) based on (4.6), where t_j^* are i.i.d from uniform distribution on the set $\{1, 2, \ldots, n-b+1\}$ for $j = 1, 2, \ldots, k$. Then, by additional assumption $bk = n$, the generalized subsampling reduce to MBB procedure. To show the inconsistency of MBB in such a case, notice that for any positive real number x we have (using the same argument as in Example 1 of Babu (1984)):

$$\begin{aligned}
&P^*(\sqrt{bk}(|\hat{\theta}_{n,b,(t_1^*,t_2^*,\ldots,t_k^*)}| - |\hat{\theta}_n|) \le x)\\
&\ge P^*(0 \ge \sqrt{bk}(\hat{\theta}_{n,b,(t_1^*,t_2^*,\ldots,t_k^*)} - \hat{\theta}_n) \ge -2\sqrt{bk}\hat{\theta}_n)\\
&= P^*(\sqrt{n}(\hat{\theta}_n - E^*(\hat{\theta}_n^*)) \ge \sqrt{n}(\hat{\theta}_n^* - E^*(\hat{\theta}_n^*)) \ge \sqrt{n}(\hat{\theta}_n - E^*(\hat{\theta}_n^*)) - 2\sqrt{n}\hat{\theta}_n),
\end{aligned}$$

where $\hat{\theta}_n^* = \hat{\theta}_{n,b,(t_1^*,t_2^*,\ldots,t_k^*)}$. Note that

$$\sqrt{n}(\hat{\theta}_n - E^*(\hat{\theta}_n^*)) = 0 \quad \text{w.p. } 1, \tag{4.9}$$

which follows from sufficiently fast convergence to zero with probability for the Gaussian term $\sqrt{n}(\hat{\theta}_n - E^*(\hat{\theta}_n^*))$ (we omit the details). Additionally from the law of the iterated logarithm (for considered time series $\{X_t : t \in \mathbb{Z}\}$) we have

$$\limsup_{n\to\infty} \sqrt{n}\hat{\theta}_n = \infty \quad \text{w.p. 1.} \tag{4.10}$$

Hence by consistency of moving block bootstrap procedure for mean $\mu = 0$ (i.e., $\mu = \lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} E(X_t) = \lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} a\cos(\lambda t) = 0$) for APC time series under moment and α-mixing conditions (see Corollary 3.2 in Synowiecki (2007) or Corollary 3 in Tewes et al. (2017) with weakened assumptions) and by (4.9) and (4.10) we obtain that

$$\begin{aligned}
&\limsup_{n\to\infty} P^*(\sqrt{bk}(|\hat{\theta}_{n,b,(t_1^*,t_2^*,\ldots,t_k^*)}| - |\hat{\theta}_n|) \leq x) \\
&\geq \limsup_{n\to\infty} P^*(\sqrt{n}(\hat{\theta}_n - E^*(\hat{\theta}_n^*)) \geq \sqrt{n}(\hat{\theta}_n^* - E^*(\hat{\theta}_n^*)) \geq \sqrt{n}(\hat{\theta}_n - E^*(\hat{\theta}_n^*)) - 2\sqrt{n}\hat{\theta}_n) \\
&= \frac{1}{2} \quad \text{w.p. 1.}
\end{aligned}$$

But $F(x) < \frac{1}{2}$ for $x \in (0, \frac{\sqrt{\pi}\,\mathrm{erf}^{-1}(\frac{1}{2})}{\sigma})$, which means that MBB procedure is not consistent.

5. Choice of generalized subsample size in S(k) procedure - a short simulation study

The fundamental problem in applying any resampling method is how to choose the block size. This problem was considered in relation to the classical subsampling procedure in Politis et al. (1999), where it was shown that the results may be strongly influenced by the choice of the block length (subsample size) b. In this section, by simple simulation experiment we also try to analyze the influence of the parameters b and k on the estimation results in the case of subsampling procedure of order k, i.e., S(k) with block length b. Recall, that in S(k) procedure the generalized subsample size is the product of k and b.

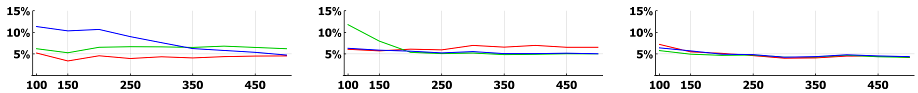
Let us consider the following autoregressive model with almost periodic mean function $\mu(t) = a\cos(\lambda t)$:

$$X_t - \mu(t) = \psi(X_{t-1} - \mu(t-1)) + \epsilon_t, \tag{5.1}$$

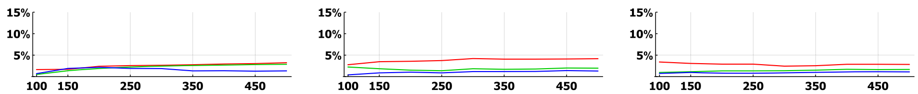
where $|\psi| < 1$ and the innovations $\{\epsilon_t : t \in \mathbb{Z}\}$ are independent and identically distributed (i.i.d.) such that $\mathrm{var}(X_t) = 1$. Note that the time series $\{X_t : t \in \mathbb{Z}\}$ satisfies all the assumptions of Theorem 4.3 since it is α-mixing with geometrically decaying mixing series (see Andrews (1983)).

Under the notations from previous sections we consider the following testing problem (for $\{X_t : t \in \mathbb{Z}\}$):

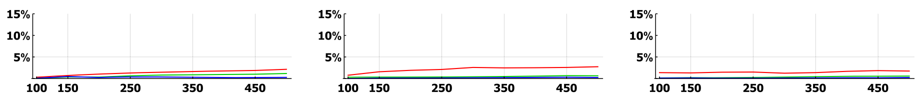
$$H_0: \quad |m(\lambda_0)| = 0, \tag{5.2}$$



(a) S(1) – subsampling



(b) S(2)



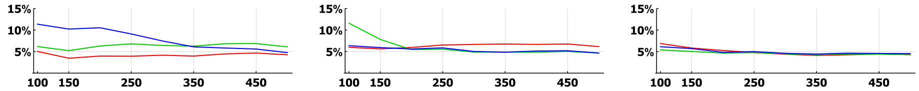
(c) S(3)

FIG 2. *Probability of H_0 rejection under H_0. In columns (respectively): $\lambda = 0.0654$, $\lambda = 0.1496$, $\lambda = 0.2618$. Read line – $b = [\sqrt{n}]$; green line – $b = 2[\sqrt{n}]$; blue line – $b = 3[\sqrt{n}]$. ϵ_t - i.i.d. form Gaussian distribution, $\alpha = 5\%$.*

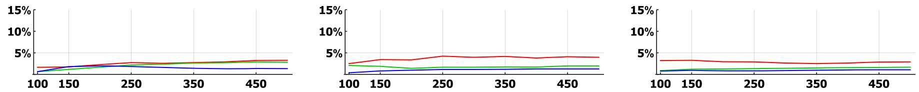
$$H_1 : \quad |m(\lambda_0)| \neq 0,$$

for $\lambda_0 \in [0, \pi]$, with test statistics $\Pi_n(\lambda_0) = \sqrt{n}|\hat{m}_n(\lambda_0)|$ and critical values calculated by subsampling of order k, i.e., procedure S(k). We will consider the problem of probability of type I error for the above testing problem (5.2) with $k \in \{1, 2, 3\}$ and $\psi = 0.7$. This characteristic of the test have been investigated by Monte Carlo simulations. We take $a = 0$ (which is equivalent to $\mu(t) = 0$) and we obtain 50 000 random samples from model (5.1), then we calculate the percentage of rejected hypotheses H_0 for assumed nominal level α.

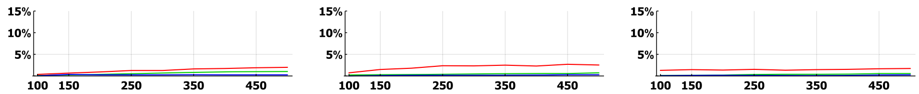
In the first part of our simulation study we consider $b \in \{[\sqrt{n}], [2\sqrt{n}], [3\sqrt{n}]\}$ and $\lambda_0 \in \{\frac{2\pi}{8\times 12}, \frac{2\pi}{3.5\times 12}, \frac{2\pi}{2\times 12}\} \approx \{0.0654, 0.1496, 0.2618\}$. With regard to the monthly data, such a set of frequencies corresponds to the pseudo-period of fluctuations (respectively): eight years, three and a half years and two years. The sample length rises from 100 to 500 ($n \in \{100+(j-1)50 : j = 1, 2, 3, ..., 9\}$). Figures 2-5 presents the estimated nominal level for $\alpha = 5\%$. On each figure we consider different type of innovation distribution (respectively): Gaussian distribution, uniform distribution, Student's t-distribution (with 5 degrees of freedom) and centralized exponential distribution. We can see that our subsampling-based test is very sensitive to b and k. The higher k or b the estimated significance level is more underestimated. Recall, that the problem of the block size in usual subsampling procedure is a delicate issue (see section 9 in Politis et al. (1999)). In generalized subsampling procedure the problem seams to be more complicated since we are dealing with an additional parameter k. Moreover, the estimated significance level depends on the frequency $\lambda_0 \in \{\frac{2\pi}{8\times 12}, \frac{2\pi}{3.5\times 12}, \frac{2\pi}{2\times 12}\}$. This is probably a consequence of the asymptotic distribution (see Theorem 4.1). More precisely, the variance-covariance matrix $\Omega(\lambda)$ of this distribution depends on the frequency λ and is proportional to $\frac{\sigma^2}{\psi^2 - 2\psi\cos(\lambda)+1}\mathbf{I}_2$, where $\mathbf{I}_2$ is 2×2 iden-



(a) S(1) – subsampling

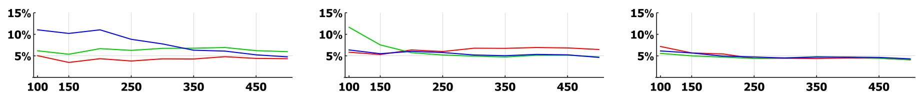


(b) S(2)

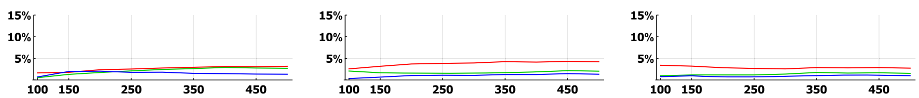


(c) S(3)

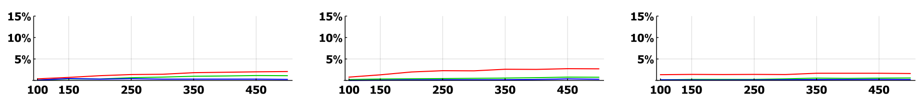
FIG 3. *Probability of H_0 rejection under H_0. In columns (respectively): $\lambda = 0.0654$, $\lambda = 0.1496$, $\lambda = 0.2618$. Read line – $b = [\sqrt{n}]$; green line – $b = 2[\sqrt{n}]$; blue line – $b = 3[\sqrt{n}]$. ϵ_t - i.i.d. form uniform distribution, $\alpha = 5\%$.*



(a) S(1) – subsampling



(b) S(2)



(c) S(3)

FIG 4. *Probability of H_0 rejection under H_0. In columns (respectively): $\lambda = 0.0654$, $\lambda = 0.1496$, $\lambda = 0.2618$. Read line – $b = [\sqrt{n}]$; green line – $b = 2[\sqrt{n}]$; blue line – $b = 3[\sqrt{n}]$. ϵ_t - i.i.d. form Student's t-distribution with 5 degrees of freedom, $\alpha = 5\%$.*

tity matrix, $\psi = 0.7$ and σ meets the condition $\text{var}(X_t) = 1$. Note that the estimated significance level seems to be robust to the distribution of innovations $\{\epsilon_t : t \in \mathbb{Z}\}$. The results are very comparable between different distributions of innovations. Hence, in the next part we restrict our attention only to the Gaussian case. Moreover, for $b \in \{[\sqrt{n}], [2\sqrt{n}], [3\sqrt{n}]\}$ and for $k = 2$ and $k = 3$ the significance level is underestimated (below 5%). This suggests considering lower values of parameter b. Therefore, on the next figure (see Figure 6) we consider $b \in \{[0.4\sqrt{n}], [0.6\sqrt{n}], [0.8\sqrt{n}]\}$ with the same set $\lambda_0 \in \{\frac{2\pi}{8\times 12}, \frac{2\pi}{3.5\times 12}, \frac{2\pi}{2\times 12}\}$ as before. Note that in this case, the estimated significance level is much higher (for $k = 2$ and $k = 3$) than on Figures 2-5, but still clearly depends on parameters

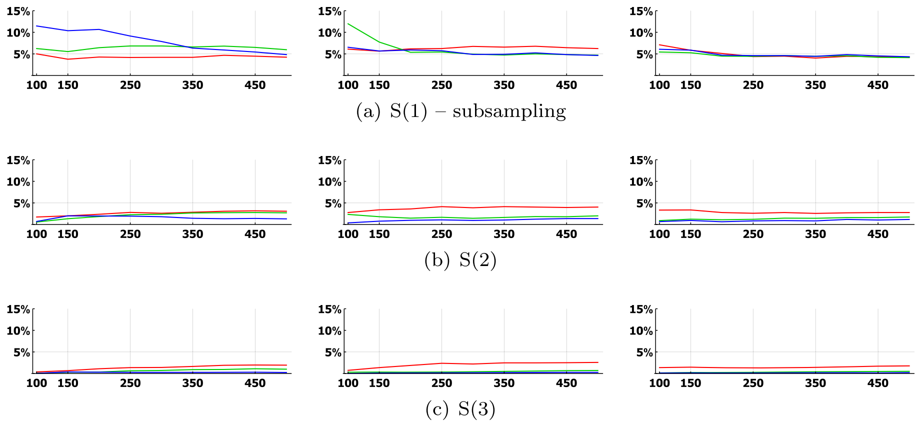


FIG 5. *Probability of H_0 rejection under H_0. In columns (respectively):* $\lambda = 0.0654$, $\lambda = 0.1496$, $\lambda = 0.2618$. *Read line –* $b = [\sqrt{n}]$; *green line –* $b = 2[\sqrt{n}]$; *blue line –* $b = 3[\sqrt{n}]$. ϵ_t *- i.i.d. from centralized exponential distribution,* $\alpha = 5\%$.

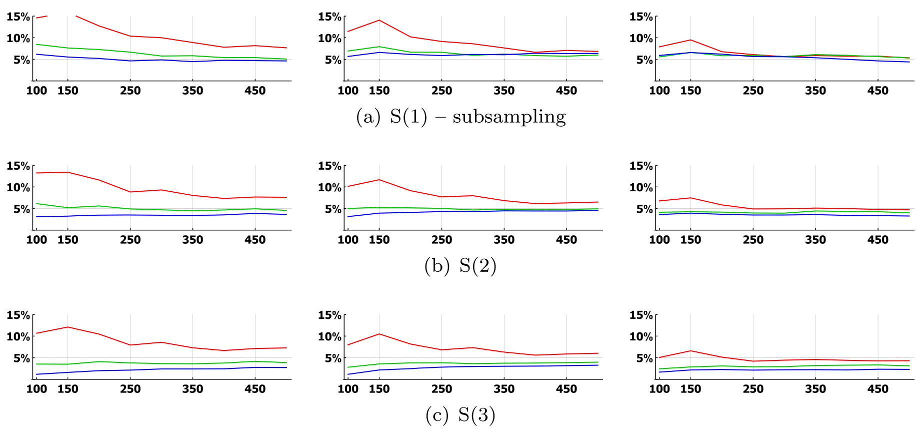
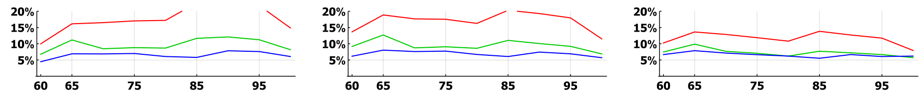


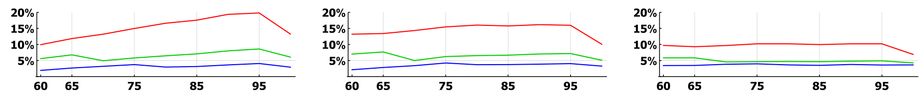
FIG 6. *Probability of H_0 rejection under H_0. In columns (respectively):* $\lambda = 0.0654$, $\lambda = 0.1496$, $\lambda = 0.2618$. *Read line –* $b = [0.4\sqrt{n}]$; *green line –* $b = [0.6\sqrt{n}]$; *blue line –* $b = [0.8\sqrt{n}]$. ϵ_t *- i.i.d. form Gaussian distribution,* $\alpha = 5\%$.

b, k and frequency λ_0. Higher parameter b (or k) give lower estimates of the significance level.

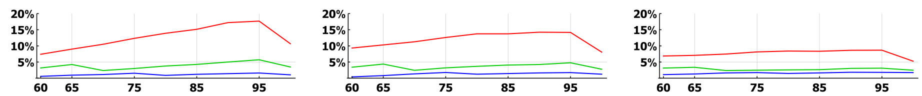
At the end of our simulation study we consider the case where $n \leq 100$. This case seams to be important from practical point of view. In such a case, the low number of different mass atoms of the subsampling distribution of order one (S(1)) may leads to practical problems when determining quantiles from this distribution. Figures (7-8) shows the estimated significance level for sample size $n \leq 100$, where $n \in \{60 + (j-1)5 : j = 1, 2, 3, ..., 9\}$ with $\alpha = 5\%$ (on Figure 7) and $\alpha = 2\%$ (on Figure 8). Note that in the case of subsampling of order one,



(a) S(1) – subsampling

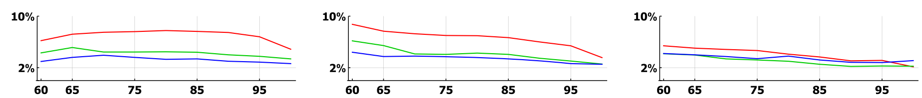


(b) S(2)

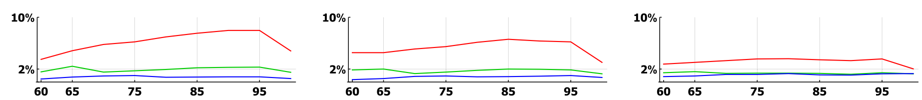


(c) S(3)

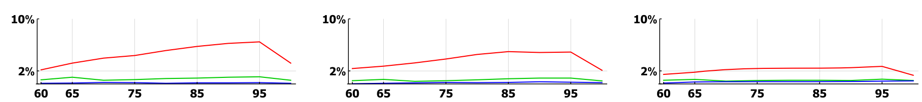
Fig 7. *Probability of H_0 rejection under H_0. In columns (respectively): $\lambda = 0.0654$, $\lambda = 0.1496$, $\lambda = 0.2618$. Read line – $b = [0.4\sqrt{n}]$; green line – $b = [0.6\sqrt{n}]$; blue line – $b = [0.8\sqrt{n}]$. ϵ_t - i.i.d. form Gaussian distribution, $\alpha = 5\%$.*



(a) S(1) – subsampling



(b) S(2)



(c) S(3)

Fig 8. *Probability of H_0 rejection under H_0. In columns (respectively): $\lambda = 0.0654$, $\lambda = 0.1496$, $\lambda = 0.2618$. Read line – $b = [0.4\sqrt{n}]$; green line – $b = [0.6\sqrt{n}]$; blue line – $b = [0.8\sqrt{n}]$. ϵ_t - i.i.d. form Gaussian distribution, $\alpha = 2\%$.*

the estimated significance level is overestimated in each case for $\alpha = 2\%$, which is probably related to the low number of different mass atoms.

6. Conclusions

In this paper, the generalized subsampling procedure was investigated. Based on the conditions for the usual subsampling procedure for non-stationary time series, the sufficient conditions for the consistency of the generalized subsampling procedure have been formulated. As was shown, the generalized subsampling is consistent under weak assumptions being a generalization of those formulated for subsampling of order one in Politis et al. (1999). As an example, the consistency was proved for the magnitude of Fourier coefficients in Fourier expansion of the mean function of an APC time series. The simulation experiment was shown, where the influence of the block length to the estimation results was investigated. There are many open problems and future research problems concerning this new procedure. First, the case with constant k may be examined in detail. The condition $k^2b/n \to 0$ probably can be weakened. Optimal sequences for b and k are not known. The next problem is attempting to show the advantage (with appropriate meaning) of generalized subsampling over the usual subsampling procedure.

Appendix A: Proofs

Proof of Theorem 3.1. In this proof, we use the same technique as in the proof of Theorem 4.2.1 in Politis et al. (1999), and therefore we concentrate only on steps with significant differences. Using the analogical steps as in the proof of Theorem 4.2.1 in Politis et al. (1999), we define

$$U_n(x) = \sum_{b_1 \le b \le b_2} \sum_{k_1 \le k \le k_2} \sum_{\mathbf{T}_k \in S^k_{n,b}} p_{(b,k,\mathbf{T}_k)} \mathbf{1}\left\{\tau_{b,\mathbf{T}_k}(\hat{\theta}_{n,b,\mathbf{T}_k} - \theta) \le x\right\}. \quad \text{(A.1)}$$

In the first step, we show that, for any $\epsilon > 0$, we have

$$U_n(x-\epsilon)\mathbf{1}\{E_n\} \le L_{n,b,k}(x)\mathbf{1}\{E_n\} \le U_n(x+\epsilon), \quad \text{(A.2)}$$

where a $\mathbf{1}\{E_n\}$ is the indicator of the event $E_n = \{\max_{b,k,\mathbf{T}_k}\{\tau_{b,\mathbf{T}_k} : b_1 \le b \le b_2, k_1 \le k \le k_2, \mathbf{T}_k \in S^k_{n,b}\}|\hat{\theta}_n - \theta| \le \epsilon\}$. To prove this, consider the following two cases:

1^o If $\mathbf{1}\{E_n\} = 0$, then (A.2) holds

2^o If $\mathbf{1}\{E_n\} = 1$, then $-\epsilon \le \tau_{b,\mathbf{T}_k}(\theta - \hat{\theta}_n) \le \epsilon$ uniformly at $b_1 \le b \le b_2, k_1 \le k \le k_2$, $\mathbf{T}_k \in S^k_{n,b}$. This means that, for any $\mathbf{T}_k = (t_1, t_2, \ldots, t_k) \in S^k_{n,b}$ and b and k such that $b_1 \le b \le b_2$ and $k_1 \le k \le k_2$, we have that

$$\mathbf{1}\left\{\tau_{b,\mathbf{T}_k}(\hat{\theta}_{n,b,\mathbf{T}_k} - \theta) \le x - \epsilon\right\} \le \mathbf{1}\left\{\tau_{b,\mathbf{T}_k}(\hat{\theta}_{n,b,\mathbf{T}_k} - \theta) + \tau_{bk}(\theta - \hat{\theta}_n) \le x\right\}$$

and

$$\mathbf{1}\left\{\tau_{b,\mathbf{T}_k}(\hat{\theta}_{n,b,\mathbf{T}_k} - \theta) + \tau_{bk}(\theta - \hat{\theta}_n) \le x\right\} \le \mathbf{1}\left\{\tau_{b,\mathbf{T}_k}(\hat{\theta}_{n,b,\mathbf{T}_k} - \theta) \le x + \epsilon\right\}.$$

Summing over b, k and $(t_1, t_2, \ldots, t_k) \in S^k_{n,b}$ (with probabilities $p_{(b,k,\mathbf{T}_k)}$) for last two inequalities, we get (A.2). Hence, using the same arguments as in Politis et al. (1999), it is sufficient to show that $E(U_n(x)) \to J(x,P)$ and $\mathrm{var}(U_n(x)) \to 0$, where $J(\cdot,P)$ is continuous at x. Under Assumption 3.2, we get $E(U_n(x)) \to J(x,P)$. Hence, we concentrate on the variance of $U_n(x)$.

Take $(b,k) \in \mathbf{B} \times \mathbf{K}$, $(b',k') \in \mathbf{B} \times \mathbf{K}$ and $\mathbf{T}_k = (t_1, t_2, \ldots, t_k) \in S^k_{n,b}$, $\mathbf{T}'_{k'} = (t'_1, t'_2, \ldots, t'_{k'}) \in S^{k'}_{n,b'}$. Denote by $W_{n,b,\mathbf{T}_k}$ the set of all time indices occurring in the generalized subsample $\mathcal{B}_{b,\mathbf{T}_k}$ (see (2.3)), i.e.,

$$W_{n,b,\mathbf{T}_k} = \bigcup_{i=1}^{k} \{t_i, t_i + 1, \ldots, t_i + b - 1\}.$$

Define the distance $\mathrm{dist}(\cdot,\cdot)$ between sets $W_{n,b,\mathbf{T}_k}$ and $W_{n,b',\mathbf{T}'_{k'}}$ via

$$\mathrm{dist}(W_{n,b,\mathbf{T}_k}, W_{n,b',\mathbf{T}'_{k'}}) = \min\{|l - w| : (l,w) \in W_{n,b,\mathbf{T}_k} \times W_{n,b',\mathbf{T}'_{k'}}\}. \quad \text{(A.3)}$$

Finally, define the sequence of sets L_s for $s = 0, 1, 2, \ldots, n - b - b' + 1$ via

$$L_s = \{(\mathbf{T}_k, \mathbf{T}'_{k'}) \in S^k_{n,b} \times S^{k'}_{n,b'} : \mathrm{dist}(W_{n,b,\mathbf{T}_k}, W_{n,b',\mathbf{T}'_{k'}}) = s\}. \quad \text{(A.4)}$$

Elementary combinatorics show that

$$\mathrm{card}(L_s) \leq \begin{cases} kk'(b + b' - 1) q^k_{n,b} q^{k'-1}_{n,b'} & \text{for } s = 0 \\ 2kk' q^k_{n,b} q^{k'-1}_{n,b'} & \text{for } s = 1, 2, \ldots, n - b - b' + 1. \end{cases} \quad \text{(A.5)}$$

Then, under the notation

$$I_{n,b,(t_1,t_2,\ldots,t_k)} = \mathbf{1}\left\{\tau_{b,\mathbf{T}_k}(\hat{\theta}_{n,b,(t_1,t_2,\ldots,t_k)} - \theta) \leq x\right\},$$

$$p = p_{(b,k,(t_1,t_2,\ldots,t_k))} = \frac{1}{(b_2 - b_1 + 1)(k_2 - k_1 + 1) q^k_{n,b}},$$

$$p' = p_{(b',k',(t'_1,t'_2,\ldots,t'_{k'}))} = \frac{1}{(b_2 - b_1 + 1)(k_2 - k_1 + 1) q^{k'}_{n,b'}}$$

we have that

$$\begin{aligned}
&\mathrm{var}(U_n(x)) \\
&= \sum_{b_1 \leq b \leq b_2} \sum_{k_1 \leq k \leq k_2} \sum_{b_1 \leq b' \leq b_2} \sum_{k_1 \leq k' \leq k_2} \\
&\quad \sum_{(t_1,t_2,\ldots,t_k) \in S^k_{n,b}} \sum_{(t'_1,t'_2,\ldots,t'_k) \in S^{k'}_{n,b'}} pp' \mathrm{cov}(I_{n,b,(t_1,t_2,\ldots,t_k)}, I_{n,b',(t'_1,t'_2,\ldots,t'_{k'})}) \\
&\leq \sum_{b_1 \leq b \leq b_2} \sum_{k_1 \leq k \leq k_2} \sum_{b_1 \leq b' \leq b_2} \sum_{k_1 \leq k' \leq k_2} \sum_{s=0}^{n-b-b'} \\
&\quad \sum_{(t_1,t_2,\ldots,t_k,t'_1,t'_2,\ldots,t'_{k'}) \in L_s} pp' |\mathrm{cov}(I_{n,b,(t_1,t_2,\ldots,t_k)}, I_{n,b',(t'_1,t'_2,\ldots,t'_{k'})})| \qquad \text{(A.6)} \\
&= \sum_{b_1 \leq b \leq b_2} \sum_{k_1 \leq k \leq k_2} \sum_{b_1 \leq b' \leq b_2} \sum_{k_1 \leq k' \leq k_2} pp'
\end{aligned}$$

$$\sum_{(t_1,t_2,\dots,t_k,t'_1,t'_2,\dots,t'_{k'})\in L_0} |\mathrm{cov}(I_{n,b,(t_1,t_2,\dots,t_k)}, I_{n,b',(t'_1,t'_2,\dots,t'_{k'})})|$$

$$+ \sum_{b_1\le b\le b_2}\sum_{k_1\le k\le k_2}\sum_{b_1\le b'\le b_2}\sum_{k_1\le k'\le k_2} pp' \sum_{s=1}^{n-b-b'}$$

$$\sum_{(t_1,t_2,\dots,t_k,t'_1,t'_2,\dots,t'_{k'})\in L_s} |\mathrm{cov}(I_{n,b,(t_1,t_2,\dots,t_k)}, I_{n,b',(t'_1,t'_2,\dots,t'_{k'})})|$$

Hence, by (A.5), we have that

$$\begin{aligned}
&\mathrm{var}(U_n(x)) \\
&\le \sum_{b_1\le b\le b_2}\sum_{k_1\le k\le k_2}\sum_{b_1\le b'\le b_2}\sum_{k_1\le k'\le k_2} pp'kk'(b+b'-1)q_{n,b}^k q_{n,b'}^{k'-1} \\
&+ \sum_{b_1\le b\le b_2}\sum_{k_1\le k\le k_2}\sum_{b_1\le b'\le b_2}\sum_{k_1\le k'\le k_2} pp' \sum_{s=1}^{n-b-b'} 2kk' q_{n,b}^k q_{n,b'}^{k'-1}\alpha(s) \\
&\le \frac{O(1)}{n}\frac{1}{(b_2-b_1+1)^2(k_2-k_1+1)^2} \\
&\sum_{b_1\le b\le b_2}\sum_{k_1\le k\le k_2}\sum_{b_1\le b'\le b_2}\sum_{k_1\le k'\le k_2} kk'(b+b'-1) \\
&+ \frac{O(1)}{n}\frac{1}{(b_2-b_1+1)^2(k_2-k_1+1)^2} \\
&\sum_{b_1\le b\le b_2}\sum_{k_1\le k\le k_2}\sum_{b_1\le b'\le b_2}\sum_{k_1\le k'\le k_2} 2kk' \sum_{s=1}^{n}\alpha(s) \\
&= \frac{O(1)}{n}\frac{1}{(b_2-b_1+1)^2(k_2-k_1+1)^2} \\
&(b_2-b_1+1)(b_2^2-b_1^2+2b_1-1)(k_2-k_1+1)^2(k_1+k_2)^2 \\
&+ \frac{O(1)}{n}\frac{1}{(b_2-b_1+1)^2(k_2-k_1+1)^2} \\
&(b_2-b_1+1)^2(k_2-k_1+1)^2(k_1+k_2)^2\sum_{s=1}^{n}\alpha(s) \\
&= O(k_2^2 b_2/n) + O(k_2^2/n)\sum_{s=1}^{n}\alpha(s) \to 0.
\end{aligned} \tag{A.7}$$

This finishes the proof of (i). The proofs of (ii) and (iii) are analogical, as in Theorem 4.2.1 of Politis et al. (1999). This finishes the proof. □

Proof of Theorem 4.1. We start from the auxiliary lemma.

Lemma A.1. *For any $(\nu,\omega)\in(0,2\pi]^2$, we have*

$$\lim_{n\to\infty} E\left(\frac{1}{2\pi kb}\sum_{i_1=1}^{k}\sum_{i_2=1}^{k}\sum_{s=t_{n,i_1}}^{t_{n,i_1}+b-1}\sum_{t=t_{n,i_2}}^{t_{n,i_2}+b-1}(X_s-\mu(s))(X_t-\mu(t))e^{-i(\nu s-\omega t)}\right)$$
$$= P(\nu,\omega),$$

where $P(\cdot,\cdot)$ is given by (4.3).

Proof of Lemma. The main steps of this proof are analogical to the steps in the proof of Lemma A.5 in Lenart (2011). Take any $(\nu,\omega)\in(0,2\pi]^2$ and notice that by substitution $j=t$ and $\tau=s-t$ we have

$$E\left(\frac{1}{2\pi kb}\sum_{i_1=1}^{k}\sum_{i_2=1}^{k}\sum_{s=t_{n,i_1}}^{t_{n,i_1}+b-1}\sum_{t=t_{n,i_2}}^{t_{n,i_2}+b-1}(X_s-\mu(s))(X_t-\mu(t))e^{-i(\nu s-\omega t)}\right)$$
$$=\frac{1}{2\pi kb}\sum_{i_1=1}^{k}\sum_{i_2=1}^{k}\sum_{j=t_{n,i_2}}^{t_{n,i_2}+b-1}\sum_{\tau=t_{n,i_1}-j}^{t_{n,i_1}+b-1-j}E((X_j-\mu(j))(X_{\tau+j}-\mu(\tau+j)))$$
$$e^{-i(\nu-\omega)j}e^{-i\nu\tau}$$
$$=\frac{1}{2\pi kb}\sum_{i_1=1}^{k}\sum_{i_2=1}^{k}\sum_{j=t_{n,i_2}}^{t_{n,i_2}+b-1}\sum_{\tau=t_{n,i_1}-j}^{t_{n,i_1}+b-1-j}B(j,\tau)e^{-i(\nu-\omega)j}e^{-i\nu\tau}$$
$$=\underbrace{\frac{1}{2\pi kb}\sum_{i=1}^{k}\sum_{j=t_{n,i}}^{t_{n,i}+b-1}\sum_{\tau=t_{n,i}-j}^{t_{n,i}+b-1-j}B(j,\tau)e^{-i(\nu-\omega)j}e^{-i\nu\tau}}_{w_n}$$
$$+\underbrace{\frac{1}{2\pi kb}\sum_{i_1\neq i_2}\sum_{j=t_{n,i_2}}^{t_{n,i_2}+b-1}\sum_{\tau=t_{n,i_1}-j}^{t_{n,i_1}+b-1-j}B(j,\tau)e^{-i(\nu-\omega)j}e^{-i\nu\tau}}_{r_n}.$$

To finish the proof, it is sufficient to show that

$$w_n\to P(\nu,\omega) \tag{A.8}$$

$$r_n\to 0. \tag{A.9}$$

To show (A.8), notice that

$$w_n=\frac{1}{2\pi kb}\sum_{i=1}^{k}\sum_{j=t_{n,i}}^{t_{n,i}+b-1}\sum_{\tau=t_{n,i}-j}^{t_{n,i}+b-1-j}\sum_{\lambda\in\Lambda_\tau}a(\lambda,\tau)e^{i(\lambda-(\nu-\omega))j}e^{-i\nu\tau}$$
$$=\frac{1}{2\pi kb}\sum_{i=1}^{k}\sum_{j=t_{n,i}}^{t_{n,i}+b-1}\sum_{\tau=t_{n,i}-j}^{t_{n,i}+b-1-j}a(\nu-\omega,\tau)e^{-i\nu\tau}$$
$$+\frac{1}{2\pi kb}\sum_{i=1}^{k}\sum_{j=t_{n,i}}^{t_{n,i}+b-1}\sum_{\tau=t_{n,i}-j}^{t_{n,i}+b-1-j}\sum_{\lambda\in\Lambda_\tau\setminus\{\nu-\omega\}}a(\lambda,\tau)e^{i(\lambda-(\nu-\omega))j}e^{-i\nu\tau}$$

$$= \frac{1}{2\pi} \sum_{|\tau|<b} \left(1 - \frac{|\tau|}{b}\right) a(\nu - \omega, \tau) e^{-i\nu\tau}$$
$$+ \frac{1}{2\pi kb} \sum_{i=1}^{k} \sum_{j=t_{n,i}}^{t_{n,i}+b-1} \sum_{\tau=t_{n,i}-j}^{t_{n,i}+b-1-j} \sum_{\lambda \in \Lambda_\tau \setminus \{\nu-\omega\}} a(\lambda, \tau) e^{i(\lambda-(\nu-\omega))j} e^{-i\nu\tau}.$$

Analogically as in the proof of Lemma A.5 in Lenart (2011), denote the first and second terms of the last equality by $\epsilon_{1,n}$ and $\epsilon_{2,n}$, respectively. Using the same arguments as in the proof of Lemma A.5 in Lenart (2011), $\epsilon_{1,n}$ goes to $P(\nu, \omega)$. Following the same steps for the term $\epsilon_{2,n}$, as in proof of Lemma A.5 in Lenart (2011), we have

$$|\epsilon_{2,n}| \leq \sum_{i=1}^{k} \frac{1}{2\pi bk} \left| \left(\sum_{\tau=-b+1}^{0} \sum_{j=t_{n,i}-\tau}^{t_{n,i}+b-1} + \sum_{\tau=1}^{b-1} \sum_{j=t_{n,i}}^{t_{n,i}+b-1-\tau} \right) \right.$$
$$\left. \sum_{\lambda \in \Lambda_\tau \setminus \{\nu-\omega\}} a(\lambda, \tau) e^{i(\lambda-(\nu-\omega))j} e^{-i\nu\tau} \right|$$
$$\leq \frac{1}{2\pi b} \sum_{\tau=-b+1}^{b-1} |a(\lambda, \tau)| \sum_{\lambda \in \Lambda_\tau \setminus \{\nu-\omega\}} |\csc((\lambda - (\nu - \omega))/2)|$$
$$\leq \frac{O(1)}{2\pi b} \sum_{\tau=-b+1}^{b-1} |a(\lambda, \tau)|,$$

which means that $\epsilon_{2,n} \to 0$ (by (ii) and Lemma A.1. in Lenart (2011)). This completes the convergence of (A.8).

Notice that, by the inequality $|B(j, \tau)| \leq 8\Delta^2 \alpha^{\frac{\delta}{2+\delta}}(|\tau|)$, $j, \tau \in \mathbb{Z}$ (see Lemma A.1. in Lenart (2011)) and condition (4.4), we have

$$|r_n| \leq \frac{1}{2\pi kb} \sum_{i_1 \neq i_2} \sum_{j=t_{n,i_2}}^{t_{n,i_2}+b-1} \sum_{\tau=t_{n,i_1}-j}^{t_{n,i_1}+b-1-j} |B(j, \tau)|$$
$$\leq \frac{1}{2\pi kb} \sum_{i_1 \neq i_2} \sum_{j=t_{n,i_2}}^{t_{n,i_2}+b-1} \sum_{\tau=t_{n,i_1}-j}^{t_{n,i_1}+b-1-j} 8\Delta^2 \alpha^{\frac{\delta}{2+\delta}}(|\tau|) \tag{A.10}$$
$$\leq \frac{O(1)}{2\pi kb} \sum_{\tau=-n}^{n} |\tau| k^2 \alpha^{\frac{\delta}{2+\delta}}(|\tau|) = \frac{O(1)k}{b} \to 0.$$

This completes the proof of the lemma. □

Consider the same decomposition as in the proof of Theorem 2.1 in Lenart (2013)

$$\sqrt{kb}(\hat{m}_{n,b,\tilde{\boldsymbol{T}}_{n,k}}(\psi) - m(\psi)) = \underbrace{\sqrt{kb}(\hat{m}_{n,b,\tilde{\boldsymbol{T}}_{n,k}}(\psi) - E(\hat{m}_{n,b,\tilde{\boldsymbol{T}}_{n,k}}(\psi))}_{k_1(n)} + \underbrace{\sqrt{kb}(E(\hat{m}_{n,b,\tilde{\boldsymbol{T}}_{n,k}}(\psi)) - m(\psi))}_{k_2(n)}.$$

Using the same arguments, it is enough to show that

$$\begin{bmatrix} \text{Re }(k_1(n)) \\ \text{Im }(k_1(n)) \end{bmatrix} \xrightarrow{d} \mathcal{N}_2(0, \Omega(\psi)) \tag{A.11}$$

and

$$k_2(n) \to 0. \tag{A.12}$$

To show (A.11), we use Cramér-Wold device and Lemma A.1 in the same manner. Therefore, we omit this step. In proving (A.12), notice that

$$\begin{aligned}
|k_2(n)| &= \sqrt{kb}\left|E(\hat{m}_{n,b,\tilde{\boldsymbol{T}}_{n,k}}(\psi)) - m(\psi)\right| \\
&= \sqrt{kb}\left|\frac{1}{kb}\sum_{i=1}^{k}\sum_{j=t_{n,i}}^{t_{n,i}+b-1}\mu(j)e^{-\mathrm{i}\psi j} - m(\psi)\right| \\
&= \sqrt{kb}\left|\frac{1}{kb}\sum_{i=1}^{k}\sum_{j=t_{n,i}}^{t_{n,i}+b-1}\sum_{\phi\in\Psi} m(\phi)e^{i\phi j}e^{-i\psi j} - m(\psi)\right| \\
&= \frac{1}{\sqrt{kb}}\left|\sum_{i=1}^{k}\sum_{j=t_{n,i}}^{t_{n,i}+b-1}\sum_{\phi\in\Psi\setminus\{\psi\}} m(\phi)e^{-i(\psi-\phi)j}\right| \\
&\le \frac{O(1)k}{\sqrt{kb}}\sum_{\phi\in\Psi\setminus\{\psi\}}\left|m(\phi)\csc\left(\frac{\phi-\psi}{2}\right)\right| \le \sqrt{\frac{k}{b}}O(1) \to 0.
\end{aligned}$$

Above, the inequality $\left|\sum_{j=t_{n,i}}^{t_{n,i}+b-1} e^{-i(\psi-\phi)j}\right| = \left|\frac{e^{ib\phi}-e^{ib\psi}}{e^{i\psi}-e^{i\phi}}\right| \le \frac{2}{\sqrt{2}}\csc(\frac{\phi-\psi}{2})$ was used. This completes the proof. □

Proof of Theorem 4.2. This proof is analogical to the proof of Theorem 2.2 in Lenart (2013). Therefore, it is omitted. □

Proof of Theorem 4.3. By noting that $k_2^2 b_2/n \to 0$, it is sufficient to show that Assumption 3.5 is fulfilled. Note that, condition (i) in Assumption 3.5 is fulfilled by the same reasons as in the proof of Theorem 2.3 in Lenart (2013). To show that condition (ii) in Assumption 3.5 is fulfilled, take any $\tilde{b}$, $\tilde{k}$ and $(t_{n,1}, t_{n,2}, \ldots, t_{n,\tilde{k}})$ such that A) and B) in Assumption 3.5 (ii) are fulfilled. Then, $\tilde{k}/\tilde{b} \to 0$ by noting that $k_2/b_1 \to 0$, and $\tilde{b} \to \infty$ by noting that $b_1 \to \infty$. Hence, for such $\tilde{b}$, $\tilde{k}$ and $(t_{n,1}, t_{n,2}, \ldots, t_{n,\tilde{k}})$, the asymptotic convergence in Theorem 4.2 holds. The next steps are analogical to the proof of Theorem 2.3 in Lenart (2013) and therefore are omitted. □

References

D. Andrews. First order autoregressive processes and strong mixing. *Cowles Foundation Discussion Papers*, 664, 1983.

G.J. Babu. Bootstrapping statistics with linear combinations of chi-squares as weak limit. *Sankhya: The Indian Journal of Statistics, Series A*, 46(1): 85–93, 1984. MR0768919

C. Corduneanu. *Almost Periodic Functions*. Chelsea, New York, 1989. MR0481915

D. Dehay and A.E. Dudek. Bootstrap method for poisson sampling almost periodic process. *J. Time Ser. Anal.*, 36(3): 327–351, 2015. MR3343005

D. Dehay and A.E. Dudek. Bootstrap for the second-order analysis of poisson-sampled almost periodic processes. *Electronic Journal of Statistics*, 11: 99–147, 2017. MR3599801

D. Dehay and H. Hurd. Representation and estimation for periodically and almost periodically correlated random processes. *In: W.A. Gardner (Ed.), Cyclostationarity in Communications and Signal Processing, IEEE Press*, pages 295–329, 1994.

D. Dehay, A.E. Dudek, and J. Leśkow. Subsampling for continuous-time almost periodically correlated processes. *Journal of Statistical Planning and Inference*, 150: 142–158, 2014. MR3206725

A.E. Dudek. Circular block bootstrap for coefficients of autocovariance function of almost periodically correlated time series. *Metrika*, 78(3): 313–335, 2015. MR3320900

A.E. Dudek. First and second order analysis for periodic random arrays using block bootstrap methods. *Electronic Journal of Statistics*, 10: 2561–2583, 2016. MR3545469

A.E. Dudek. Block bootstrap for periodic characteristics of periodically correlated time series. *Journal of Nonparametric Statistics*, 30(1): 87–124, 2018. MR3756234

A.E. Dudek, J. Leśkow, D. Politis, and E. Paparoditis. A generalized block bootstrap for seasonal time series. *J. Time Ser. Anal.*, 35: 89–114, 2014a. MR3166348

A.E. Dudek, S. Maiz, and M. Elbadaoui. Generalized seasonal block bootstrap in frequency analysis of cyclostationary signals. *Signal Process.*, 104C: 358–368, 2014b.

A.E. Dudek, E. Paparoditis, and D.N. Politis. Generalized seasonal tapered block bootstrap. *Statistics and Probability Letters*, 115: 27–35, 2016. MR3498365

W.A. Gardner, A. Napolitano, and L. Paura. Cyclostationarity: Half a century of research. *Signal Processing*, 86: 639–6, 2006.

E.G. Gladyshev. Periodically correlated random sequance. *Sov. Math.*, 2: 385–388, 1961.

H. Hurd. Nonparametric time series analysis for periodically correlated processes. *IEEE Trans. Inf. Theory*, 35(2): 350–359, 1989. MR0999650

H. Hurd. Correlation theory of almost periodically correlated processes. *J. Multivariate Anal.*, 37: 24–45, 1991. MR1097303

H.L. Hurd and A.G. Miamee. *Periodically Correlated Random Sequences: Spectral Theory and Practice*. Wiley, Hoboken, New Jersey, 2007. MR2348769

H.R. Kunsch. The jacknife and bootstrap for general stationary observation. *Ann. Statist.*, 17(3): 1217–1241, 1989. MR1015147

S.N. Lahiri. *Resampling Methods for Dependent Data*. Springer–Verlag, New York, 2003. MR2001447

Ł. Lenart. Asymptotic distributions and subsampling in spectral analysis for almost periodically correlated time series. *Bernoulli*, 17(1): 290–319, 2011. MR2797993

Ł. Lenart. Non–parametric frequency identification and estimation in mean function for almost periodically correlated time series. *Journal of Multivariate Analysis*, 115: 252–269, 2013. MR3004557

Ł. Lenart. Generalized resampling scheme with application to spectral density matrix in almost periodically correlated class of time series. *J. Time Ser. Anal.*, 37(3): 369–404, 2016. MR3512963

Ł. Lenart. Examination of Seasonal Volatility in HICP for Baltic Region Countries: Non-Parametric Test versus Forecasting Experiment. *Central European Journal of Economic Modelling and Econometrics*, 9(1): 29–67, 2017.

Ł. Lenart and M. Pipień. Almost Periodically Correlated Time Series in Business Fluctuations Analysis. *Acta Phisica Polonica A*, 123: 567–583, 2013a.

Ł. Lenart and M. Pipień. Seasonality Revisited - statistical Testing for Almost Periodically Correlated Stochastic Processes. *Central European Journal of Economic Modelling and Econometrics*, 5(2): 85–102, 2013b.

Ł. Lenart and M. Pipień. Non-Parametric Test for the Existence of the Common Deterministic Cycle: The Case of the Selected European Countries. *Central European Journal of Economic Modelling and Econometrics*, 9(3): 201–241, 2017.

Ł. Lenart, J. Leśkow, and R. Synowiecki. Subsampling in testing autocovariance for periodically correlated time series. *J. Time Ser. Anal.*, 29(6): 995–1018, 2008. MR2464950

R.Y. Liu and K. Singh. Moving blocks jacknife and bootstrap capture weak dependence. *in R. LePage and L. Billard. eds., Exploring the Limits of Bootstrap, Willey, New York*, pages 225–248, 1992. MR1197787

B. Mazur and M. Pipień. On the Empirical Importance of Periodicity in the Volatility of Financial Returns - Time Varying GARCH as a Second Order APC(2) Process. *Central european Journal of Economic modelling and Econometrics*, 4: 95–116, 2012.

D. Politis, J. Romano, and M. Wolf. *Subsampling*. Springer-Verlag, New York, 1999. MR1707286

E. Serpedin, F. Panduru, I. Sari, and G.B. Gian. Bibliography on cyclostationarity. *Signal Processing*, 85: 2233–2303, 2005.

R. Synowiecki. Consistency and application of moving block bootstrap for nonstationary time series with periodic and almost periodic structure. *Bernoulli*, 13(4): 1151–1178, 2007. MR2364230

J. Tewes, D.N. Politis, and D.J. Nordman. Convolved subsampling estimation with applications to block bootstrap. https://arxiv.org/abs/1706.07237, preprint, 2017.

O. Zhao and M. Woodroofe. Law of the iterated logarithm for stationary processes. *The Annals of Probability*, 36(1): 127–142, 2008. MR2370600