Spectral density estimation for nonstationary data with nonzero mean function

Anna E. Dudek, Lukasz Lenart

Spectral density estimation for nonstationary data with nonzero mean function

Anna E. Dudek*
Department of Applied Mathematics, AGH University of Science and Technology,
al. Mickiewicza 30, 30-059 Krakow, Poland
and
Łukasz Lenart †
Department of Mathematics, Cracow University of Economics,
ul. Rakowicka 27, 31-510 Cracow, Poland

March 1, 2021

Abstract

We introduce a new approach for nonparametric spectral density estimation based on the subsampling technique, which we apply to the important class of nonstationary time series. These are almost periodically correlated sequences. In contrary to existing methods our technique does not require demeaning of the data. On the simulated data examples we compare our estimator of spectral density function with the classical one. Additionally, we propose a modified estimator, which allows to reduce the leakage effect. Moreover, we provide a simulation study and two real data economic applications.

Keywords: subsampling, almost periodically correlated, covariance matrix estimation, mixing, periodogram-subsampling estimator.

Anna Dudek acknowledges support from the King Abdullah University of Science and Technology (KAUST) Research Grant OSR-2019-CRG8-4057.2.

†*Lukasz Lenart was financed from the funds granted to the Department of Mathematics at Cracow University of Economics, within the framework of the subsidy for the maintenance of research potential.*

1 Introduction

Spectral theory for stationary time series is well developed (see Priestley (1981), Žurbenko (1986), Stoica and Moses (2005), and many others). However, in the real data applications one deals often with nonstationary phenomena. In such a case statistical tools designed for stationary processes should be appropriately modified or generalized. If this is not possible a new approach needs to be proposed.

In this paper we focus on the spectral density estimation problem in the nonstationary setting. Classical approach for spectral density estimation requires removing the mean function from the data. However, as we discuss below, this may be very challenging in the nonstationary case. Thus, it is crucial to propose a new method that provides consistent estimates and does not assume the nullity of the mean function.

In the sequel we introduce and develop a novel methodology for spectral densities estimation for nonstationary mixing sequences with nonzero mean function. Our main results cover the class of almost periodically correlated (APC) time series, which is an important class of nonstationary time series containing among others periodically correlated (PC) and covariance stationary sequences. APC and PC processes are broadly used to model data with cyclic features. Many interesting examples concerning e.g., telecommunication, vibroacoustics, mechanics, economics and climatology can be found in Antoni (2009), Gardner et al. (2006), Hurd and Miamee (2007), Napolitano (2012, 2016). To be precise, a time series $\{X_t, t \in \mathbb{Z}\}$ with finite second moments is called PC if it has periodic mean and covariance functions, while in APC case the mean and covariance functions are almost periodic functions of the time argument. A function $f(t) : \mathbb{Z} \longrightarrow \mathbb{R}$ 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 such that $\sup_{t\in\mathbb{Z}} |f(t + p_\epsilon) - f(t)| < \epsilon$ (see Corduneanu (1989)).

Problem of spectral density estimation for PC and APC sequences is not new (see Nematollahi and Rao (2005), Lii and Rosenblatt (2006), Hurd and Miamee (2007), Napolitano (2012, 2016)), but as we mentioned before all methods assume that $\mathrm{E}X_t \equiv 0$. In the case of stationary and PC time series this is not a restrictive condition, because the data can be demeaned (see e.g., Brillinger (2001) and Hurd and Miamee (2007)). However, for APC time series this problem is much more difficult. In general case it is not easy to estimate an almost periodic mean function. Moreover, in the presence of an almost periodic mean, the spectral densities and cyclic spectra contain Dirac impulses as a consequence of the presence of an almost periodic component in the cyclic autocovariance (see Napolitano (2016)). Estimation of the spectral densities or the cyclic spectrum without removing the mean, results in spikes in correspondence of the ideal Dirac deltas. In such a case, Napolitano

and Spooner (2000) proposed a median filtering technique to remove the spikes. However, for now the theoretical properties of this approach were not studied in the literature.
Our idea for spectral densities estimation is based on the subsampling method, which allows to approximate the asymptotic variance of the rescaled statistics of interest. Let us indicate that the general consistency result concerning subsampling estimator of the asymptotic variance that we present does not require periodicity or almost periodicity of the data. However, we apply it to an APC process to provide at the same time an application that is interesting and important from practical point of view. For that purpose we use results from Lenart (2013). Lenart considered estimation problem for the Fourier coefficients of the mean function of an APC time series. He showed that their rescaled estimators are asymptotically normal and that elements of the asymptotic covariance matrix are linear functions of the spectral densities of the considered process. Applying subsampling to estimate elements of this matrix we get consistent estimators of the spectral densities. Moreover, our subsampling estimator of the spectral density functions turns out to be a generalized type of a periodogram-subsampling estimator and it is associated with a classical spectral density estimator with the Bartlett window.

Paper is organized as follows. In Section 2 the known methods for spectral density estimation in stationary and almost periodic case are summarized. The main results are provided in Section 3. In particular, the subsampling estimator of the spectral density functions of an APC time series with the nonzero mean function is constructed and its consistency is shown. Moreover, to reduce the leakage effect the modification of this estimator is discussed. Finally, its relationship with the periodogram-subsampling estimator and the classical spectral density estimator with the Bartlett window is derived. Section 4 contains results of the performed simulation study. In Section 5 real data example is provided. Finally, all proofs are presented in the supplementary materials.

2 Spectral density estimation for APC time series

Let $\{X_t, t \in \mathbb{Z}\}$ be an APC time series. Its mean function $\mu(t) = \mathrm{E}(X_t)$ and the autocovariance function $B(t,\tau) = \mathrm{cov}(X_t, X_{t+\tau})$ (for any fixed $\tau \in \mathbb{Z}$) can be represented as Fourier series

$$\mu(t) \sim \sum_{\psi\in\Psi} m(\psi)\exp(i\psi t), \quad B(t,\tau) \sim \sum_{\lambda\in\Lambda_\tau} a(\lambda,\tau)\exp(i\lambda t), \tag{1}$$

where the sets $\Psi = \{\psi \in [0, 2\pi) : m(\psi) \neq 0\}, \Lambda_\tau = \{\lambda \in [0, 2\pi) : a(\lambda, \tau) \neq 0\}$ are countable, and

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

Symbol $\sim$ indicates some association between the Fourier series and the almost periodic function. The sup-norm used in the definition of an almost periodic function (see Section 1) results in the uniform convergence of the Fourier series (1) to $\mu(t)$ and $B(t, \tau)$, respectively. For more details we refer the reader to Napolitano (2012).
If the sets of frequencies Ψ and $\Lambda = \bigcup_{\tau\in\mathbb{Z}} \Lambda_\tau$ are finite, then in the expressions (1) we have equalities. This condition is always fulfilled for PC time series, with $\Psi, \Lambda \subset \{2k\pi/d, k = 0, 1, \ldots, d-1\}$ (see e.g., Hurd and Miamee (2007)).
Under some regularity conditions (see e.g., Lenart (2011)) the spectral density function $P(\cdot, \cdot)$ of an APC time series is defined on the bifrequency square $[0, 2\pi)^2$ and has the form

$$P(\nu, \omega) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} a(\nu - \omega, \tau) \exp(-i\nu\tau), \qquad (\nu, \omega) \in [0, 2\pi)^2. \tag{2}$$

Below we shortly summarize the known results concerning estimation of $P(\nu, \omega)$.
In the stationary case i.e., when $\Lambda = \{0\}$, the consistent estimator of the spectral density function is obtained by tapering the covariance estimator or equivalently by smoothing the periodogram (see e.g., Priestley (1981), Brillinger (2001)). Let us recall that when a PC process with known period d is considered, then the spectral density function $P(\nu, \omega)$ can be nonzero only on the $2d - 1$ diagonal lines $\{(\nu, \omega) \in [0, 2\pi)^2 : \omega = \nu - 2k\pi/d, k = -d + 1, \ldots, d - 1\}$ (see Hurd (1991), Dehay and Hurd (1994)). As a result the consistent estimator of the spectral density functions can be obtained by smoothing the shifted periodogram along the support lines or equivalently by weighting the estimators of the Fourier coefficients of the autocovariance function (ch. 10.2 in Hurd and Miamee (2007)). On the other hand, in Nematollahi and Rao (2005) to construct an estimator of the spectral density matrix, the relation between the PC time series and the multivariate stationary time series is used. Nematollahi and Rao proposed to estimate the spectral matrix of the stationary vector time series and then to use it to construct the estimator of the spectral density matrix of the considered PC time series. Finally, in Lii and Rosenblatt (2006) the estimation of the frequencies in almost periodic autocovariance structure of a zero mean Gaussian harmonizable APC time series is considered. In Napolitano (2016) the review of known

results for the APC processes is provided. Two estimators of the spectral density functions are discussed: the frequency-smoothed cyclic periodogram and the time-smoothed cyclic periodogram (see Section 2 in Napolitano (2016)).
Independently on the type of the process (stationary, PC, APC) to construct the aforementioned estimators of the spectral density function, the nullity of the mean function is assumed. If the considered process is not zero mean, before one estimates the spectral density, the data should be demeaned. This can be easily achieved when the mean function is constant or periodic with the known period length. It is also possible when one deals with an APC time series and the set Ψ is finite and known. Then, having sample $(X_1, \ldots, X_n)$, one can apply the estimator of the mean function $\widehat{\mu}_n(t) = \sum_{\psi\in\Psi} \widehat{m}_n(\psi) \exp(i\psi t)$, where $\widehat{m}_n(\psi) = \frac{1}{n}\sum_{j=1}^{n} X_j \exp(-i\psi j)$ is the estimator of the Fourier coefficient $m(\psi)$, $\psi \in [0, 2\pi)$. Finally, the estimator of the spectral density $P(\nu, \omega)$ can be obtained by replacing $\mu(t)$ by its estimate $\widehat{\mu}_n(t)$ (see Lenart (2011, 2016)). According to our knowledge such estimator was not considered in the literature in APC case when the set Ψ is unknown and needs to be estimated. It is worth to note that in the general APC case, when $\mu(t)$ is an almost periodic function with the unknown set Ψ, it is very difficult to remove entirely the mean function from the data. Moreover, even in purely periodic case one may deal with so called hidden periodicities (Yavorskyj et al. (2012)) and in consequence estimation of the mean function can be challenging. In Napolitano (2016) (see (2.9)) it was shown that if the almost periodic mean is not removed from the data, the classical estimates of the spectral densities for zero mean case contain Dirac impulses, which can be observed as additional spikes. The shape of the spikes is related to the length of the frequency-smoothing window if one uses the frequency-smoothing cyclic periodogram method for the estimation. In such a case the estimator of spectral density function is not consistent. To solve this problem in Napolitano and Spooner (2000) a median filtering technique is proposed, but so far the theoretical properties of this approach were not studied in the literature. In a very special case, when the considered process is a sum of periodic mean component and a stationary time series, one can estimate spectrum without removing the periodic component. The overview of possible techniques can be found in Li (2016) (see ch. 7.2). However, these methods are not suitable for a general PC time series (with a periodic autocovariance function) or more complex nonstationary time series like APC processes. For instance, they assume that the spectral density function is nonzero, which in the APC case means that we know the set Λ. This is a restrictive condition that we do not use for our results.
In the next section we introduce a new method for the spectral density function $P(\nu, \omega)$ (see (2)) estimation on the bifrequency square $(\nu, \omega) \in [0, 2\pi)^2$ in a case when the almost periodic mean function is unknown, i.e., when set Ψ is unknown. Our approach does not

require estimation of the mean function and can be applied directly to the raw data.

3 Main results

In the first part of this section we propose the subsampling covariance matrix estimator for general weakly dependent time series and show its mean square consistency. In the next step we use this general result in the construction of the spectral density function estimator for an APC time series in the case when the sets Ψ and Λ are unknown.
From now on by $\mathrm{Re}(z)$ and $\mathrm{Im}(z)$ we denote the real and the imaginary part of a complex number z. Symbol $(\cdot)'$ denotes transpose of a vector. By $\mathcal{N}_s(\mathbf{a}, \mathbf{\Sigma})$ we denote the s-variate Gaussian distribution with mean $\mathbf{a}$ and covariance matrix $\mathbf{\Sigma}$. Finally, for any random variable Z and real number $p > 0$, $\|Z\|_p = (\mathrm{E}|Z|^p)^{1/p}$.

3.1 General subsampling consistency result

The idea of subsampling (see Politis et al. (1999)) is to approximate the sampling distribution of the statistics using subsamples of the data. Let $\{X_t, t \in \mathbb{Z}\}$ be a considered time series and $(X_1, \ldots, X_n)$ be an observed sample. Moreover, let r be any positive integer. By $\boldsymbol{\theta} = \left(\theta^{(1)}, \theta^{(2)}, \ldots, \theta^{(r)}\right)' \in \mathbb{R}^r$ and $\widehat{\boldsymbol{\theta}}_n = \left(\widehat{\theta}_n^{(1)}, \widehat{\theta}_n^{(2)}, \ldots, \widehat{\theta}_n^{(r)}\right)'$ we denote an unknown parameter and its estimator based on the sample $(X_1, X_2, \ldots, X_n)$. We assume that there exists a non-degenerate limiting distribution of $\tau_n(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta})$, where τ_n is a normalizing sequence. Finally, $\widehat{\boldsymbol{\theta}}_{n,b,t}$ is a subsampling version of $\widehat{\boldsymbol{\theta}}_n$ based on the subsample $(X_t, \ldots, X_{t+b-1})$, where $b = b(n) \to \infty$, $b/n \to 0$ as $n \to \infty$, $1 \le t \le n - b + 1$ and

$$L_{n,b}(A) = \frac{1}{n-b+1} \sum_{t=1}^{n-b+1} \mathbf{1}\left\{\tau_b\left(\widehat{\boldsymbol{\theta}}_{n,b,t} - \widehat{\boldsymbol{\theta}}_n\right) \in A\right\} \tag{3}$$

is an r-variate empirical distribution which approximates the distribution of $\tau_n\left(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\right)$, A is any Borel set in $\mathbb{R}^r$ and $\mathbf{1}\{C\}$ is an indicator function of the event C.
The important feature of subsampling (and its recent generalizations) is that it provides valid quantiles for confidence intervals and hypothesis tests in stationary and nonstationary models under very mild regularity conditions (see e.g., Politis et al. (1999), Dehay et al. (2014), Dudek and Lenart (2017), Lenart (2011, 2013, 2018), Tewes et al. (2019)). Moreover, to obtain subsampling consistency one does not need to know the exact form

of the limiting distribution of $\tau_n(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta})$. However, in this paper we are only interested in a subsampling approximation of the elements of the covariance matrix of this limiting distribution. By $\boldsymbol{\Sigma} = [\sigma_{ij}]_{r\times r}$ we denote this limiting covariance matrix.

Recall that for fixed $1 \leq i, j \leq r$ and any sequence $\tilde{b}$ such that $\tilde{b} \to \infty$, $\tilde{b}/n \to 0$ the estimator

$$\widehat{\sigma}^{(i,j)}_{n,\tilde{b}} = \frac{\tau^2_{\tilde{b}}}{n-\tilde{b}+1} \sum_{t=1}^{n-\tilde{b}+1} \left(\widehat{\theta}^{(i)}_{n,\tilde{b},t} - \widehat{\theta}^{(i)}_{n,\tilde{b},\cdot}\right)\left(\widehat{\theta}^{(j)}_{n,\tilde{b},t} - \widehat{\theta}^{(j)}_{n,\tilde{b},\cdot}\right) \tag{4}$$

is a classical subsampling estimator of σ_{ij} based on the block length $\tilde{b}$, where $\widehat{\theta}^{(i)}_{n,\tilde{b},\cdot} = 1/(n-\tilde{b}+1)\sum_{t=1}^{n-\tilde{b}+1}\widehat{\theta}^{(i)}_{n,\tilde{b},t}$ and $\widehat{\theta}^{(i)}_{n,\tilde{b},t}$ is a subsampling estimator of $\theta^{(i)}$ (see Section 3.8.1 and 4.6 in Politis et al. (1999)). However, in this paper we define some more general subsampling estimator of σ_{ij}, which is based on the generalized subsampling procedure introduced by Lenart (2018). For this purpose let us define two sequences of positive integers b_1, b_2 such that $b_1, b_2 \to \infty$ as $n \to \infty$, $b_1 \leq b_2$ and $b_2/n \to 0$ as $n \to \infty$. Next, denote $w = w(b_1, b_2) = b_2 - b_1 + 1$ and $q = q(n, b) = n - b + 1$ for $b_1 \leq b \leq b_2$. Finally, by

$$\widehat{\sigma}^{(i,j)}_{n,b_1,b_2} = \frac{1}{w}\sum_{b=b_1}^{b_2}\frac{\tau^2_b}{q}\sum_{t=1}^{q}\left(\widehat{\theta}^{(i)}_{n,b,t} - \widehat{\theta}^{(i)}_{n,b,\cdot}\right)\left(\widehat{\theta}^{(j)}_{n,b,t} - \widehat{\theta}^{(j)}_{n,b,\cdot}\right), \tag{5}$$

we define the generalized subsampling estimator of the σ_{ij}, which is based on block lengths belonging to the set $\{b_1, b_1+1, \ldots, b_2\}$.

The idea behind the construction of the estimator (5) is as follows. Take any sequence $\tilde{b}$ of integers such that $b_1 \leq \tilde{b} \leq b_2$. Let t^* be a random variable with uniform distribution on the set $\{1, 2, \ldots, n-\tilde{b}+1\}$. Note that the elements of the covariance matrix of the distribution $\tau_{\tilde{b}}(\widehat{\boldsymbol{\theta}}_{n,\tilde{b},t^*} - \widehat{\boldsymbol{\theta}}_{\boldsymbol{n}})$ (in the resampling world) are of the form (4) i.e., $\mathrm{cov}^*(\tau_{\tilde{b}}(\widehat{\boldsymbol{\theta}}_{n,\tilde{b},t^*} - \widehat{\boldsymbol{\theta}}_{\boldsymbol{n}})) = \widehat{\boldsymbol{\Sigma}}_{n,\tilde{b}} = [\widehat{\sigma}^{(i,j)}_{n,\tilde{b}}]_{i,j=1,2,\ldots,r}$. Thus, $\sigma^{(i,j)}_{n,b_1,b_2} = \frac{1}{w}\sum_{b=b_1}^{b_2}\widehat{\sigma}^{(i,j)}_{n,b}$ is a mean of the subsampling estimators (4) based on different block lengths.

Remark 3.1. *Note that, if $\tilde{b}^*$ has the uniform distribution on the set $\{b_1, b_1+1, \ldots, b_2\}$ and $t^*|(\tilde{b}^* = b)$ $(b_1 \leq b \leq b_2)$ has uniform distribution on the set $\{1, 2, \ldots, n-b+1\}$, then $cov^*(\tau_{\tilde{b}^*}(\widehat{\boldsymbol{\theta}}_{n,\tilde{b}^*,t^*} - \widehat{\boldsymbol{\theta}}_{\boldsymbol{n}})) = \frac{1}{w}\sum_{b=b_1}^{b_2}\widehat{\boldsymbol{\Delta}}_{n,b} = [\frac{1}{w}\sum_{b=b_1}^{b_2}\widehat{\delta}^{(i,j)}_{n,b}]_{i,j=1,2,\ldots,r}$, where*

$$\widehat{\delta}^{(i,j)}_{n,b} = \frac{1}{n-b+1}\sum_{t=1}^{n-b+1}\left(\tau_b\widehat{\theta}^{(i)}_{n,b,t} - \widehat{\theta}^{(i)}_{n,\cdot,\cdot}\right)\left(\tau_b\widehat{\theta}^{(j)}_{n,b,t} - \widehat{\theta}^{(j)}_{n,\cdot,\cdot}\right)$$

and $\widehat{\theta}^{(i)}_{n,\cdot,\cdot} = \frac{1}{w}\sum_{k=b_1}^{b_2} \frac{\tau_k}{n-k+1}\sum_{t=1}^{n-k+1}\widehat{\theta}^{(i)}_{n,k,t}$ *(see Lenart (2018)). Then* $\widehat{\delta}^{(i,j)}_{n,b}$ *is the natural candidate for the estimator. However, for the sake of simplicity, we decided to consider only* $\widehat{\sigma}^{(i,j)}_{n,b}$ *in this paper.*

Let $\mathbf{A} = [a_{ij}]_{r\times r}$ be a known complex valued matrix and let

$$\sigma = \sum_{i=1}^{r}\sum_{j=1}^{r} a_{ij}\sigma_{ij} \tag{6}$$

be a parameter of interests, which is a linear combination of elements of the matrix $\mathbf{\Sigma}$. Based on (5) we propose the following subsampling estimator of σ

$$\widehat{\sigma}_n = \sum_{i=1}^{r}\sum_{j=1}^{r} a_{ij}\left[\frac{1}{w}\sum_{b=b_1}^{b_2}\frac{\tau_b^2}{q}\sum_{t=1}^{q}\left(\widehat{\theta}^{(i)}_{n,b,t} - \widehat{\theta}^{(i)}_{n,b,\cdot}\right)\left(\widehat{\theta}^{(j)}_{n,b,t} - \widehat{\theta}^{(j)}_{n,b,\cdot}\right)\right]. \tag{7}$$

Note that the proposed estimator is based on a standard subsampling estimator (4), which is averaged over different block lengths (one may also consider only single block length). We decided to provide very general formula for the estimator to make it suitable for variety of the possible future applications. Unfortunately, in contrary to the stationary case for any resampling method (including subsamping) for nonstationary processes (like considered in this paper periodic and almost periodic time series) there is no result concerning the optimal subsample or block length choice. Many such results in the stationary setting are obtained using the mean square error expansions of the bias and the variance of the estimator. However, even in the simplest problem of the overall mean estimation for PC/APC time series, such expansion has not been obtained. There is also no heuristic approach for any standard parameter of interest. A broad simulation study was performed in Dudek and Potorski (2020) to check if known heuristic block/subsample length approaches for stationary time series may be adapted for periodic time series and none of them seemed to work. In stationary case one may also find very interesting results (see e.g., ch. 10 in Politis et al. (1999)) stating that taking a linear combination of subsampling distribution estimators with different subsample sizes allows to construct an improved estimator in the sense of the second-order accuracy. It is not possible to state, if such result may hold in our nonstationary case, since no Edgeworth expansion exists for PC/APC processes.

Below we formulate conditions under which the subsampling estimator $\widehat{\sigma}_n$ is mean square consistent.

Theorem 3.1. *Assume that $\{X_t, t \in \mathbb{Z}\}$ is an α-mixing time series. Moreover, let b_1 and b_2 be subsample lengths such that $0 < b_1 \leq b_2 < n$, $b_1, b_2 \to \infty$ as $n \to \infty$, but $b_2/n \to 0$ as $n \to \infty$. Additionally, assume that*

(i) there exists $\tilde{\sigma}$ such that for any sequence of integers $\tilde{b} = \tilde{b}(n)$ such that $b_1 \leq \tilde{b} \leq b_2$, and any sequence of integers t_n such that $1 \leq t_n \leq n - \tilde{b} + 1$ we have that

$$\mathrm{E}\left[\sum_{i=1}^{r}\sum_{j=1}^{r} a_{ij}\tau_{\tilde{b}}^2 \left(\widehat{\theta}_{n,\tilde{b},t_n}^{(i)} - \mathrm{E}\left(\widehat{\theta}_{n,\tilde{b},\cdot}^{(i)}\right)\right)\left(\widehat{\theta}_{n,\tilde{b},t_n}^{(j)} - \mathrm{E}\left(\widehat{\theta}_{n,\tilde{b},\cdot}^{(j)}\right)\right)\right] \longrightarrow \tilde{\sigma};$$

(i') OR there exists $\tilde{\sigma}$ such that for any sequence of integers $\tilde{b}$ such that $b_1 \leq \tilde{b} \leq b_2$, we have that

$$\frac{1}{n-\tilde{b}+1}\sum_{t=1}^{n-\tilde{b}+1} \mathrm{E}\left[\sum_{i=1}^{r}\sum_{j=1}^{r} a_{ij}\tau_{\tilde{b}}^2 \left(\widehat{\theta}_{n,\tilde{b},t}^{(i)} - \mathrm{E}\left(\widehat{\theta}_{n,\tilde{b},\cdot}^{(i)}\right)\right)\left(\widehat{\theta}_{n,\tilde{b},t}^{(j)} - \mathrm{E}\left(\widehat{\theta}_{n,\tilde{b},\cdot}^{(j)}\right)\right)\right] \longrightarrow \tilde{\sigma};$$

(ii) there exist $\xi > 0$ and $\Delta > 0$ such that

$$\sup_{1\leq i\leq r,\, b_1\leq b\leq b_2,\, 1\leq t\leq n-b+1,\, n\in\mathbb{Z}} \tau_b \left\|\widehat{\theta}_{n,b,t}^{(i)} - \mathrm{E}\left(\widehat{\theta}_{n,b,\cdot}^{(i)}\right)\right\|_{4+2\xi} < \Delta \quad \text{and} \quad \sum_{k=1}^{\infty} \alpha^{\frac{\xi}{2+\xi}}(k) < \infty;$$

(iii) for any $1 \leq i \leq r$ and any sequence $\tilde{b}$ such that $b_1 \leq \tilde{b} \leq b_2$ we have convergence

$$\left\|\frac{\tau_{\tilde{b}}}{n-\tilde{b}+1}\sum_{1\leq t\leq n-\tilde{b}+1}\left(\widehat{\theta}_{n,\tilde{b},t}^{(i)} - \mathrm{E}\left(\widehat{\theta}_{n,\tilde{b},\cdot}^{(i)}\right)\right)\right\|_2 \longrightarrow 0 \qquad \text{as } n \longrightarrow \infty.$$

Then the estimator $\widehat{\sigma}_n$ (given by (7)) is mean square consistent i.e.,

$$\widehat{\sigma}_n \xrightarrow{L_2} \tilde{\sigma} \qquad \text{as } n \longrightarrow \infty.$$

3.2 Spectral density estimation for general APC time series

In this section we show how to apply results from the previous section in the problem of spectral density estimation of an APC time series with an almost periodic mean function, under general assumption that the set Ψ is unknown and finite. At first we formulate the

assumptions and introduce some additional notation. Next, we provide an auxiliary central limit theorem for the Fourier coefficients of the mean function (Section 3.2.1). Additionally, we present the main results concerning consistency of our subsampling estimators (Sections 3.2.2 and 3.2.5). In Section 3.2.3 we discuss the relation of our estimator with the periodogram-subsampling estimator of the spectral density function and with classical spectral density estimator based on the Bartlett window. A short illustrative simulation data examples are contained in Section 3.2.4.

Let $(X_{c_n}, \ldots, X_{c_n+d_n-1})$ be a sample of the size d_n from an APC time series $\{X_t, t \in \mathbb{Z}\}$ with the mean function $\mu(t)$, where $\{c_n\}_{n\in\mathbb{Z}}$ and $\{d_n\}_{n\in\mathbb{Z}}$ are arbitrary sequences of positive integers such that $d_n \to \infty$ as $n \to \infty$. Let us recall that when the set Ψ is finite, the mean function $\mu(t)$ can be expressed as follows $\mu(t) = \sum\limits_{\psi\in\Psi} m(\psi)\exp(i\psi t)$. Moreover, when the set Ψ is known its estimator based on the sample $(X_{c_n}, \ldots, X_{c_n+d_n-1})$ (see also Lenart (2011)) is of the form $\widehat{\mu}_{d,c}(t) = \sum_{\psi\in\Psi} \widehat{m}_{d,c}(\psi)\exp(i\psi t)$, where

$$\widehat{m}_{d,c}(\psi) = \frac{1}{d_n}\sum_{j=c_n}^{c_n+d_n-1} X_j \exp(-i\psi j). \tag{8}$$

In the sequel we assume the following conditions

A1 $\{X_t, t \in \mathbb{Z}\}$ is a real valued APC time series such that the sets Ψ and Λ are finite and unknown;

A2 $\{X_t, t \in \mathbb{Z}\}$ is α-mixing (for definition see e.g., Doukhan (1994));

A3 the complex measure r_0, where $a(0,\tau) = \int\limits_0^{2\pi} e^{i\xi\tau} r_0(\mathrm{d}\xi)$ is absolutely continuous with respect to the Lebesgue measure.

Under the condition **A1** the considered APC time series is of the form $X_t = \mu(t) + Y_t$, where Y_t is a zero mean harmonizable sequence (in the sense of Loève (1963); see Definition 5.4 in Hurd and Miamee (2007)). Harmonizability of Y_t can be obtained using the same arguments as presented in Hurd and Miamee (2007) for PC sequences (see also Gladyshev (1961)). Moreover, harmonizability ensures existence of the measure r_0 in the assumption **A3** (see e.g., Hurd and Miamee (2007)). The condition **A3** ensures that for any $|\lambda| \in \Lambda$ there exists the spectral density function $g_\lambda(\nu) = 1/(2\pi)\sum_{\tau=-\infty}^{\infty} a(\lambda,\tau)\exp(-i\nu\tau)$, $\nu \in [0, 2\pi)$. The set of spectral densities can be simply extended to the bifrequency square $[0, 2\pi)^2$ (see (2)) taking $P(\nu,\omega) = g_\lambda(\nu)$ for $|\nu - \omega| = \lambda$, $(\nu,\omega) \in [0, 2\pi)^2$ and $P(\nu,\omega) = 0$ otherwise (see

Hurd (1991), Dehay and Hurd (1994), Hurd and Miamee (2007), Napolitano (2012), Lenart (2016)). Finally, the existence of complex measure r_0 follows from Lemma 7.3 in Lenart (2016).

3.2.1 Central limit theorem for the Fourier coefficients

Take $\nu, \omega \in [0, 2\pi)$. Below we state the asymptotic distribution of

$$\widehat{\mathbf{m}}_{d,c}(\nu,\omega) = (\mathrm{Re}(\widehat{m}_{d,c}(\nu)),\, \mathrm{Im}(\widehat{m}_{d,c}(\nu)),\, \mathrm{Re}(\widehat{m}_{d,c}(\omega)),\, \mathrm{Im}(\widehat{m}_{d,c}(\omega)))', \tag{9}$$

i.e., the estimator of the parameter

$$\mathbf{m}(\nu,\omega) = (\mathrm{Re}(m(\nu)),\, \mathrm{Im}(m(\nu)),\, \mathrm{Re}(m(\omega)),\, \mathrm{Im}(m(\omega)))'. \tag{10}$$

However, our main goal is to provide the exact form of the asymptotic covariance matrix. It turns out that its elements are linear combinations of the spectral density function. In a consecutive subsection we use this feature to construct the estimator of $P(\nu, \omega)$.

Theorem 3.2. *Assume that conditions* $\mathbf{A1} - \mathbf{A3}$ *hold and that* $\sup_t \|X_t\|_{2+\delta} < \infty$ *and* $\sum_{k=0}^{\infty} \alpha^{\frac{\delta}{2+\delta}}(k) < \infty$*, for some* $\delta > 0$*. Then* $\sqrt{d_n}\,(\widehat{\mathbf{m}}_{d,c}(\nu,\omega) - \mathbf{m}(\nu,\omega)) \xrightarrow{d} \mathcal{N}_4(\mathbf{0}, \mathbf{\Gamma}(\nu,\omega))$*, where* $\mathbf{\Gamma}(\nu,\omega) = [\gamma_{ij}(\nu,\omega)]_{4\times 4}$ *is of the form*

$$\pi \begin{pmatrix} g_0(\nu) + P_R(\nu,-\nu) & P_I(\nu,-\nu) & P_R(\nu,\omega) + P_R(\nu,-\omega) & -P_I(\nu,\omega) + P_I(\nu,-\omega) \\ P_I(\nu,-\nu) & g_0(\nu) - P_R(\nu,-\nu) & -P_I(\omega,\nu) + P_I(\omega,-\nu) & P_R(\nu,\omega) - P_R(\nu,-\omega) \\ P_R(\nu,\omega) + P_R(\nu,-\omega) & -P_I(\omega,\nu) + P_I(\omega,-\nu) & g_0(\omega) + P_R(\omega,-\omega) & P_I(\omega,-\omega) \\ -P_I(\nu,\omega) + P_I(\nu,-\omega) & P_R(\nu,\omega) - P_R(\nu,-\omega) & P_I(\omega,-\omega) & g_0(\omega) - P_R(\omega,-\omega) \end{pmatrix}$$

and $P_R(\cdot,\cdot) = \mathrm{Re}[P(\cdot,\cdot)]$, $P_I(\cdot,\cdot) = \mathrm{Im}[P(\cdot,\cdot)]$.

Note that the elements of the matrix $\mathbf{\Gamma}(\nu,\omega)$ are functions of $P(\cdot,\cdot)$ i.e., they are functions of the extension of the spectral density matrix function to the bifrequency square $[0, 2\pi)^2$. Moreover, taking linear combinations of elements of $\mathbf{\Gamma}(\nu,\omega)$, one can get $P(\nu,\omega)$. In particular, we have that

$$\gamma_{13}(\nu,\omega) + \gamma_{24}(\nu,\omega) = 2\pi P_R(\nu,\omega), \quad \gamma_{23}(\nu,\omega) - \gamma_{14}(\nu,\omega) = 2\pi P_I(\nu,\omega), \tag{11}$$

because $P_I(\omega,\nu) = -P_I(\nu,\omega)$ and $P_I(\omega,-\nu) = P_I(\nu,-\omega)$.
Our idea to obtain an estimate of $P(\nu,\omega)$ is to apply the subsampling method in order to approximate the covariance matrix of the distribution $\mathcal{N}_4(\mathbf{0}, \mathbf{\Gamma}(\nu,\omega))$. Then, having estimates of elements of the matrix $\mathbf{\Gamma}(\nu,\omega)$ and using the formulas (11), we can easily get the

estimate of $P(\nu,\omega)$. In this way, in our approach we do not need to know the form of the set Ψ i.e., we do not need to know the Fourier frequencies of the mean function.
It is worth to indicate that in the literature dedicated to PC and APC processes, asymptotic normality results corresponding to Theorem 3.2, are treated as the preliminary step for construction of confidence intervals for parameters of interest. Since the asymptotic covariance matrix usually is of complicated form and depends on unknown parameters, in practice it is very difficult to estimate. Thus, in such situations to approximate quantiles of the asymptotic distribution, resampling methods are often applied (see e.g., Dudek et al. (2014); Lenart (2016); Dudek (2018) and references therein). Although, in our problem we need to estimate the asymptotic covariance matrix $\mathbf{\Gamma}(\nu,\omega)$ and not $\mathbf{m}(\cdot,\cdot)$, we also use a resampling method. More precisely, we adopt the idea of the subsampling estimator of the limiting variance examined in the stationary case in Carlstein (1986) and generalized to the nonstationary cases in Fukuchi (1999) and Politis et al. (1999). It is necessary to indicate that the consistency of subsampling is not a sufficient condition to show that the subsampling estimator of the limiting variance (or covariance matrix) converges to the asymptotic variance (or the covariance matrix) and hence the mentioned consistency results are not sufficient for our purpose. Thus, in the next subsection we introduce the consistent subsampling estimators of the linear combinations of the elements of the matrix $\mathbf{\Gamma}(\nu,\omega)$ (see Theorem 3.2). To show their consistency we adopt ideas presented in Fukuchi (1999), Carlstein (1986) and Politis et al. (1999).

3.2.2 Subsampling consistency result

Let $(\nu,\omega)\in[0,2\pi)^2$, $r=4$ and $(X_1,\ldots,X_n)$ be an observed sample. Following the notation introduced in Section 3.1 we set $\boldsymbol{\theta}=\boldsymbol{\theta}(\nu,\omega)=\mathbf{m}(\nu,\omega)$, $\widehat{\boldsymbol{\theta}}_n=\widehat{\boldsymbol{\theta}}_n(\nu,\omega)=\widehat{\mathbf{m}}_{n,1}(\nu,\omega)$ and $\widehat{\boldsymbol{\theta}}_{n,b,t}=\widehat{\boldsymbol{\theta}}_{n,b,t}(\nu,\omega)=\widehat{\mathbf{m}}_{b,t}(\nu,\omega)$. In such a case the estimator (7) is of the form

$$\widehat{\sigma}_n(\nu,\omega)=\sum_{i,j=1}^{4}a_{ij}\left[\frac{1}{w}\sum_{b=b_1}^{b_2}\frac{\tau_b^2}{q}\sum_{t=1}^{q}\left(\widehat{\theta}^{(i)}_{n,b,t}(\nu,\omega)-\widehat{\theta}^{(i)}_{n,b,\cdot}(\nu,\omega)\right)\left(\widehat{\theta}^{(j)}_{n,b,t}(\nu,\omega)-\widehat{\theta}^{(j)}_{n,b,\cdot}(\nu,\omega)\right)\right]. \quad (12)$$

Theorem below states the mean square consistency of $\widehat{\sigma}_n(\nu,\omega)$.

Theorem 3.3. *Assume that the conditions of Theorem 3.2 hold. Additionally, assume that there exists $\xi>0$ such that $\sup_t\|X_t\|_{4+3\xi}<\infty$ and $\sum_{k=1}^{\infty}(k+1)^{c-2}\alpha^{\frac{\xi}{c+\xi}}(k)<\infty$, where c is the smallest even integer such that $c\geq 4+2\xi$. Then the conclusion of Theorem 3.1 holds for any complex valued matrix $\mathbf{A}=[a_{ij}]_{i,j=1\ldots,4}$ and $\tilde{\sigma}=\sigma(\nu,\omega)$ given by (6) with*

$\sigma_{ij}(\nu,\omega)=\gamma_{ij}(\nu,\omega)$, *i.e.*,

$$\widehat{\sigma}_n(\nu,\omega) \xrightarrow{L_2} \sigma(\nu,\omega)=\sum_{i=1}^{4}\sum_{j=1}^{4} a_{ij}\gamma_{ij}(\nu,\omega).$$

Setting $a_{13}=1$, $a_{24}=1$, $a_{14}=-i$, $a_{23}=i$ and $a_{ij}=0$ otherwise, we get the estimate of $2\pi P(\nu,\omega)$ (see (11)). From now on the estimator $\widehat{\sigma}_n(\nu,\omega)$ obtained for this particular choice of coefficients a_{ij} will be denoted by by $\widehat{W}^{\text{sub}}_{n,b_1,b_2}(\nu,\omega)$. Note that $\widehat{W}^{\text{sub}}_{n,b_1,b_2}(\nu,\omega)$ is an average of the estimators $\widehat{W}^{\text{sub}}_{n,b,b}(\nu,\omega)$ i.e., estimators calculated on the basis of one fixed subsample length b. Thus, we have

$$\widehat{W}^{\text{sub}}_{n,b_1,b_2}(\nu,\omega)=\frac{1}{w}\sum_{b=b_1}^{b_2}\widehat{W}^{\text{sub}}_{n,b}(\nu,\omega), \tag{13}$$

where

$$\widehat{W}^{\text{sub}}_{n,b}(\nu,\omega)=\frac{b}{q}\sum_{t=1}^{n-b+1}\left(\widehat{m}_{b,t}(\nu)-\frac{1}{q}\sum_{j=1}^{n-b+1}\widehat{m}_{b,j}(\nu)\right)\left(\widehat{m}_{b,t}(-\omega)-\frac{1}{q}\sum_{j=1}^{n-b+1}\widehat{m}_{b,j}(-\omega)\right). \tag{14}$$

In the next section we provide another interpretation of $\widehat{W}^{\text{sub}}_{n,b}(\nu,\omega)$ and discuss its relationship with the classical spectral density estimator.

3.2.3 Another look on subsampling spectral density estimator $\widehat{W}^{\text{sub}}_{n,b_1,b_2}$

In this section we show that $\widehat{W}^{\text{sub}}_{n,b_1,b_2}(\nu,\omega)$ is in fact an average of generalized type of a periodogram-subsampling estimator for the spectral density function. Moreover, we present its relation to the classical spectral density estimator with the Bartlett window.

For the sake of clarity at first we recall a periodogram-subsampling estimator for the spectral density in a stationary case.
Let $\{X'_t, t\in\mathbb{Z}\}$ be a real-valued stationary time series with constant mean μ and spectral density function $f(\cdot)$ and $(X'_1,\ldots,X'_n)$ be the available sample. Moreover, let $Z_i = X'_i - 1/n\sum_{i=1}^{n}X'_i$, i.e., $\mathrm{E}Z_i=0$, $i=1,\ldots,n$. Then the periodogram based on centered sample $(Z_1,\ldots,Z_n)$ is of the form

$$I_n(\omega)=\frac{1}{2\pi n}d_n(\omega)d_n(-\omega), \tag{15}$$

where $d_n(\omega) = \sum_{j=1}^{n} Z_j \exp(-ij\omega),\ \omega \in [0, 2\pi)$ is the discrete Fourier transform of the sample. The subsampling version of (15) based on the subsample $(Z_t, \ldots, Z_{t+b-1}),\ t = 1, \ldots, n-b+1$ is defined as

$$I_{n,b,t}(\omega) = \frac{1}{2\pi b} d_{t,b,n}(\omega) d_{t,b,n}(-\omega), \tag{16}$$

where $d_{n,b,t}(\omega) = \sum_{j=t}^{t+b-1} Z_j \exp(-ij\omega),\ \omega \in [0, 2\pi)$. Then, the subsampling periodogram estimator of the spectral density f is given by

$$f_{n,b}^{sub}(\omega) = \frac{1}{n-b+1} \sum_{t=1}^{n-b+1} I_{n,b,t}(\omega). \tag{17}$$

In (17) subsample is demeaned using the estimator of the overall mean, however it is also possible to demean each subsample using the subsample average. In such a case at $\omega = 0$, the estimator $2\pi f_{n,b}^{sub}(0)$ is the standard subsampling estimator of the variance of the sample mean or of the scaled spectral density at 0, $2\pi f(0) = \lim_{n\to\infty} n\mathrm{Var}(1/n \sum_{i=1}^{n} X_i')$ (see e.g., Politis et al. (1999), Politis and Romano (1994)). Under appropriate mixing and moment assumptions, the estimator $f_{n,b}^{sub}(\omega)$ can be proven to estimate consistently $f(\omega)$. This will also hold true for some classes of nonstationary time series. In particular one may expect consistency for zero mean periodically and almost periodically correlated time series. The proof of this fact can be obtained e.g., following ideas of Soedjak (2002). Soedjak constructed an estimator using sequence of dependent sample paths of a zero mean strongly harmonizable process. Applying subsampling allows to replace the replicates of the process by the subsamples.

One may also expect that subsampling will work also for the time series with nonstationary mean structures (e.g., periodic, almost periodic) where there is a long-run pattern in the expected value of sample averages. Then subsampling can be used to remove the mean at the level of the subsample averages, as opposed to demeaning individual time series observations, which seems to be the important feature for the method's validity with time series having mean structures with convergent long-run averages. And in fact this feature of subsampling is used by the estimator $\widehat{W}_{n,b_1,b_2}^{\mathrm{sub}}(\nu, \omega)$.

Note that introduced in this paper estimator $\widehat{W}_{n,b}^{\mathrm{sub}}(\nu, \omega)$ for nonzero mean case can be expressed as follows

$$\widehat{W}_{n,b}^{\mathrm{sub}}(\nu, \omega) = \frac{1}{q} \sum_{t=1}^{q} \frac{1}{b} d_{n,b,t}(\nu) d_{n,b,t}(-\omega), \tag{18}$$

where $q = n - b + 1$ and

$$d_{n,b,t}(\nu) = \sum_{j=t}^{t+b-1} X_j \exp(-i\nu j) - \frac{1}{q}\sum_{k=1}^{q}\sum_{j=k}^{k+b-1} X_j \exp(-i\nu j).$$

Finally $I_{n,b,t}(\nu,\omega) = \frac{1}{2\pi b} d_{n,b,t}(\nu) d_{n,b,t}(-\omega)$ is the generalization to nonzero mean case of subsampling version of the dual-frequency periodogram. The dual frequency periodogram was considered in zero mean case e.g., in Gorrostieta et al. (2019), Lenart (2011). According to our knowledge in the literature there are no results for $I_{n,b,t}(\nu,\omega)$ when the set Ψ is unknown.

The subsampling estimator (18) is related with a classical estimator of the spectral density function. More specifically, as we see later $\widehat{W}^{\text{sub}}_{n,b}(\nu,\omega)$ given by (14) (or equivalently by (18)) can be expressed using the classical spectral density estimator with the Bartlett window applied to the centered data $Y_t = X_t - \mu(t)$ i.e.,

$$\widehat{H}_{n,b}(\nu,\omega) = \frac{1}{n}\sum_{s=1}^{n}\sum_{t=1}^{n}\left(1 - \frac{|s-t|}{b}\right)\mathbb{I}_{\{|s-t|\le b\}} Y_s Y_t e^{-is\nu} e^{it\omega} \tag{19}$$

(see formula (8) in Lenart (2011) with $L_n = b$, page 293). Note that if the mean function $\mu(t)$ is unknown, then $\widehat{H}_{n,b}(\nu,\omega)$ cannot be used, while one can apply the estimator $\widehat{W}^{\text{sub}}_{n,b}(\nu,\omega)$. Theorem below states relation between $\widehat{H}_{n,b}(\nu,\omega)$ and $\widehat{W}^{\text{sub}}_{n,b}(\nu,\omega)$.

Theorem 3.4. *Assume that conditions* **A1** $-$ **A3** *hold and that* $\sup_t \|X_t\|_{6+3\delta} < \infty$ *and* $\sum_{k=0}^{\infty}(k+1)^2\alpha^{\frac{\delta}{2+\delta}}(k) < \infty$*, for some* $\delta > 0$*. Then* $\widehat{W}^{sub}_{n}(\nu,\omega)$ *given by (14) can be expressed as*

$$\sqrt{\frac{n}{b}}\widehat{W}^{sub}_{n,b}(\nu,\omega) = \sqrt{\frac{n}{b}}\widehat{H}_{n,b}(\nu,\omega) + \sqrt{\frac{n}{b}}C_{n,b}(\nu,\omega) + o_p(1), \tag{20}$$

where

$$C_{n,b}(\nu,\omega) = \frac{1}{qb}\sum_{(\psi_1,\psi_2)\in\Psi^2\setminus(\nu,\omega)} m(\psi_1)\overline{m(\psi_2)}\frac{e^{i\nu}\left(1-e^{ib(\psi_1-\nu)}\right)}{e^{i\nu}-e^{i\psi_1}}\frac{e^{i\psi_2}\left(1-e^{ib(\omega-\psi_2)}\right)}{e^{i\psi_2}-e^{i\omega}}\sum_{t=1}^{n-b+1} e^{it(\psi_1-\nu+\omega-\psi_2)}$$

is deterministic and $\widehat{H}_{n,b}(\nu,\omega)$ *is given by (19).*

Remark 3.2. *(a) (20) also holds under conditions of Theorem 3.2 with an additional assumption on the subsample length, i.e.* $n = O(b^{2+\kappa})$ *for some* $\kappa > 0$*, as* $n \longrightarrow \infty$*.*

(b) A relation analogous to (20) for the stationary at zero frequency case can be found in Politis et al. (1999) (see Section 3.8.2). It was established for demeaned data, i.e. under assumption that the set Ψ is known. Note also that since we assume that Ψ is unknown, $\widehat{H}_{n,b}(\nu,\omega)$ is not an estimator.

Note that for any fixed $(\nu,\omega)\in[0,2\pi)^2$ it can be shown that $C_{n,b}(\nu,\omega)=O(1/b)$ for any $(\nu,\omega)\in S$ and $C_{n,b}(\nu,\omega)=O(1/(qb))$ otherwise, where $S=\{(x,y)\in[0,2\pi)^2 : |x-y|\in\tilde{\Psi}\}$ and $\tilde{\Psi}=\{(\psi_1-\psi_2) \mod 2\pi : (\psi_1,\psi_2)\in\Psi^2\}$. Also note that the set $\tilde{\Psi}$ depends only of the set Ψ, i.e., of the frequencies of the mean function.
The following corollary illustrates the relation between the limiting distributions of the rescaled $\widehat{W}^{\text{sub}}_{n,b}(\nu,\omega)$ and $\widehat{H}_{n,b}(\nu,\omega)$.

Corollary 3.1. *Assume that conditions of Theorem 3.4 hold. Denote*

$$T_{1n}=\sqrt{\frac{n}{b}}\left(\widehat{W}^{sub}_{n,b}(\nu,\omega)-P(\nu,\omega)\right) \quad \textit{and} \quad T_{2n}=\sqrt{\frac{n}{b}}\left(\widehat{H}_{n,b}(\nu,\omega)-P(\nu,\omega)\right).$$

Then for any fixed $(\nu,\omega)\in[0,2\pi)^2$

i) if $n/b^3\longrightarrow 0$ as $n\longrightarrow\infty$, then T_{1n} and T_{2n} have the same limiting distributions (assuming that T_{2n} is convergent in distribution);

ii) if $n/b^3\to C_1$ as $n\longrightarrow\infty$, where C_1 is some positive constant or infinity, then

- *for any $(\nu,\omega)\notin S$, T_{1n} and T_{2n} have the same limiting distributions (assuming that T_{2n} is convergent in distribution);*
- *for any $(\nu,\omega)\in S$ the deterministic term $\sqrt{n/b}\,C_{n,b}(\nu,\omega)$ is directly related to the mean of the distribution of T_{1n}. Since $\sqrt{n/b}\,C_{n,b}(\nu,\omega)$ does not have limit as $n\to\infty$, the limiting distribution of T_{1n} does not exist.*

Above, we discussed the behavior of $C_{n,b}(\nu,\omega)$ when $(\nu,\omega)\in[0,2\pi)^2$ is fixed. It is also interesting to study the behavior of $C_{n,b}(\nu,\omega)$ as a function of (ν,ω). We have

$$\lim_{(\nu,\omega)\to(\nu_0,\omega_0)} C_{n,b}(\nu,\omega)=\begin{cases} O(b) & \text{for} \quad \nu_0\in\Psi,\omega_0\in\Psi,(\nu_0,\omega_0)\in S \\ O(1) & \text{for} \quad \nu_0\in\Psi,\omega_0\notin\Psi,(\nu_0,\omega_0)\in S \\ O(1) & \text{for} \quad \nu_0\notin\Psi,\omega_0\in\Psi,(\nu_0,\omega_0)\in S \\ O\left(\frac{1}{b}\right) & \text{for} \quad \nu_0\notin\Psi,\omega_0\notin\Psi,(\nu_0,\omega_0)\in S \\ O\left(\frac{1}{nb}\right) & \text{for} \quad \nu_0\notin\Psi,\omega_0\notin\Psi,(\nu_0,\omega_0)\notin S \\ O\left(\frac{1}{n}\right) & \text{for} \quad \nu_0\in\Psi,\omega_0\notin\Psi,(\nu_0,\omega_0)\notin S \\ O\left(\frac{1}{n}\right) & \text{for} \quad \nu_0\notin\Psi,\omega_0\in\Psi,(\nu_0,\omega_0)\notin S \\ O\left(\frac{b}{n}\right) & \text{for} \quad \nu_0\in\Psi,\omega_0\in\Psi,(\nu_0,\omega_0)\notin S \end{cases}.$$

3.2.4 Simulation data examples

Example below illustrates performance of our estimator $\widehat{W}^{\mathrm{sub}}_{n,b_1,b_2}(\nu,\omega)$ and at the same time problem of leakage, which inspired us to propose in the next subsection a modification of $\widehat{W}^{\mathrm{sub}}_{n,b_1,b_2}(\nu,\omega)$.

Example 3.1. *We consider a PC series* $X_t = 5\sin(2\pi t/30) + 2\left(\sin(2\pi t/12) + 21/20\right)\epsilon_t$, *where* ϵ_t *is a normalized Gaussian white noise. Thus,* $\{X_t\}$ *is composition of the PC white noise* $2\left(\sin\left(\frac{2\pi t}{12}\right) + \frac{21}{20}\right)\varepsilon_t$ *and the periodic deterministic component* $5\sin(2\pi t/30)$*), with* $\Lambda \subset \{\lambda_k = 2k\pi/12 : k = 0,\ldots,11\}$. *It is easy to show that the spectral densities (denoted here by* $g_{\lambda_k}(\cdot) = g_k(\cdot)$*) are constant on the interval* $[0, 2\pi)$ *and* $g_0 = 6.41$, $g_1 = g_{11} = -4.2i$, $g_2 = g_{10} = -1$, $g_3 = \cdots = g_9 = 0$ *(see p. 170 in Hurd and Miamee (2007)). To construct our estimator of spectral density function we generated two samples of sizes* $n = 1000$ *and* $n = 10000$*, respectively. We set* $b_1 = b_2 = b$ *and* $b \in \{\lfloor\sqrt{n}\rfloor, \lfloor 2\sqrt{n}\rfloor, \lfloor 3\sqrt{n}\rfloor\}$, *where* $\lfloor x\rfloor$ *is the largest integer that is less than or equal to* x. *The simulation results are presented in Figures 1-2. Since the main conclusions do not depend of the chosen values of parameters, we decided to restrict the number of figures only to those that are necessary to illustrate performance of our estimator.*
In Figure 1 one may observe the estimated rescaled spectral density $2\pi g_0$ *for different values of* b. *Let us recall that the main advantage of our approach is the ability to estimate the spectral densities without need of removing the mean function from the data. In this example the mean function is periodic and its true significant frequency* ψ_0 *is equal to* $2\pi/30$. *First of all, one may notice that independently of the subsample length* b *our estimator provides an accurate estimate of the rescaled spectral density value at* ψ_0. *However, in the neighborhood of* ψ_0 *there are always two side-lobes representing the spectral leakage in the frequency domain. The size of lobes depends on the value of* b *i.e., on the subsample length. In Figure 2 we present the estimate of the spectral density functions on the bifrequency square. One may note that the real part of the estimate is concentrated on the main diagonal, while the imaginary part on two lines that are parallel to diagonal. Moreover, for the real and the imaginary parts, we observe leakage effect around vertical lines* $\nu = \psi_0$, $\nu = 2\pi - \psi_0$ *and horizontal lines* $\omega = \psi_0$, $\omega = 2\pi - \psi_0$.

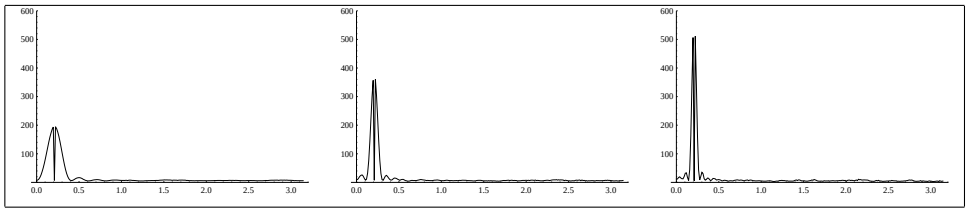

Figure 1: Estimated spectral density $2\pi g_0$ on interval $[0, \pi)$. Sample size $n = 1000$. From left results for subsample length b equal to $\lfloor\sqrt{n}\rfloor, \lfloor 2\sqrt{n}\rfloor, \lfloor 3\sqrt{n}\rfloor$, respectively.

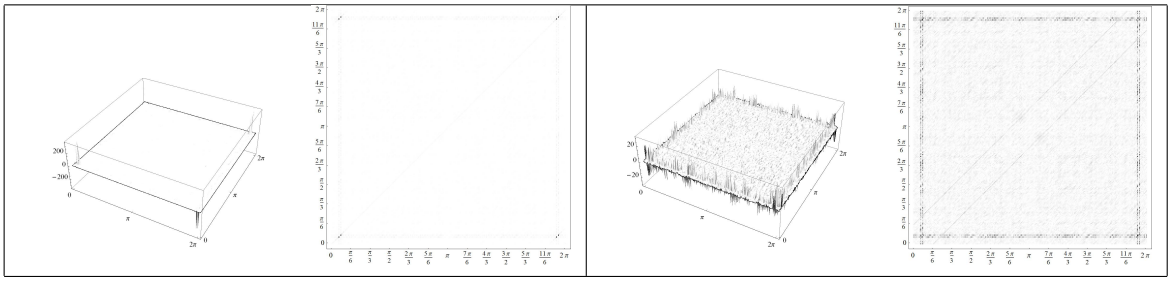

Figure 2: Estimates of $2\pi Re(P(\nu,\omega))$ (left-hand side) and $2\pi Im(P(\nu,\omega))$ (right-hand side) for $(\nu,\omega) \in [0, 2\pi)^2$ together with their projections on the xy-plane. Sample size $n = 1000$, $b = \lfloor 3\sqrt{n}\rfloor$.

As we mentioned above the observed leakage is a consequence of applying the rectangular window of the length b to the data. In the result we do not calculate just the Discrete Fourier Transform (DFT) of $\{X_t\}$ *but of the product of our signal* $\{X_t\}$ *and the rectangular window function. The Fourier transform of the product is equal to the convolution of the Fourier transforms. Moreover, the Fourier transform of the rectangular window function is the Dirichlet kernel function (see e.g., formula (3.3.5) and Table 3.3.1 in Brillinger (2001)). Finally, the convolution of the DFT of the signal with the Dirichlet kernel produces the observed side-lobes.*

The leakage phenomenon is a typical problem in signal processing. Taking sample of the finite size for Fourier transform is equivalent to application of a rectangular window on a data stream. In the literature one may find many methods designed to reduce the leakage effect. The most popular are based on non-rectangular windows. However, reducing the spectral leakage results in decreasing the spectral resolution. More details on that topic can be found for instance in the book Proakis and Manolakis (2009).

Example 3.2. *This example shows some advantages of our estimator over the classical estimators of the spectral density function. In contrary to Example 3.1 we considered time*

series with nonconstant spectral density functions. These are two AR(2) signals of the following form:

M1 $X_t = 0.85X_{t-1} - 0.7X_{t-2} + \epsilon_t,$

M2 $X_t - \mu(t) = 0.85(X_{t-1} - \mu(t-1)) - 0.7(X_{t-2} - \mu(t-2)) + \epsilon_t,$

where ϵ_t *is normalized Gaussian white noise and* $\mu(t) = 1.6\sin(\pi t/3)$. *Model* **M1** *is the example of an AR(2) type stationary time series. Model* **M2** *is having periodic mean function. Both models have the same spectral density function which concentrates the spectral mass around frequency* $\pi/3$. *We chose those two simple series to illustrate problems of classical estimator of spectral density function when the mean function is nonzero and to show that our method provides comparable results to the classical approach in the standard stationary series.*

We estimated the spectral density function using our estimator $\widehat{W}^{sub}_{n,b_1,b_2}(\nu,\omega)$ *(see Section 3.2.2) and the classical estimator based on smoothing the covariance function using the Bartlett window. We considered two sample sizes* $n = 500$ *and* $n = 1000$ *and block lengths* $b_1 = \lfloor 5\sqrt[3]{n} \rfloor$, $b_2 = \lfloor 3\sqrt{n} \rfloor$. *The length of smoothing window* c *was set* $c = \lfloor (b_1 + b_2)/2 \rfloor$. *Since the conclusions do not depend of the sample size we decided to present results only for* $n = 1000$ *(see Figure 3). In panels on the left hand-side one may observe the generated realizations for models* **M1** *and* **M2**. *Panels on the right-hand side contain the theoretical spectral density function (dashed line) together with its estimates obtained using our non-modified subsampling estimator (solid line) and the classical estimator with Bartlett window of length* c *(shadowed area).*

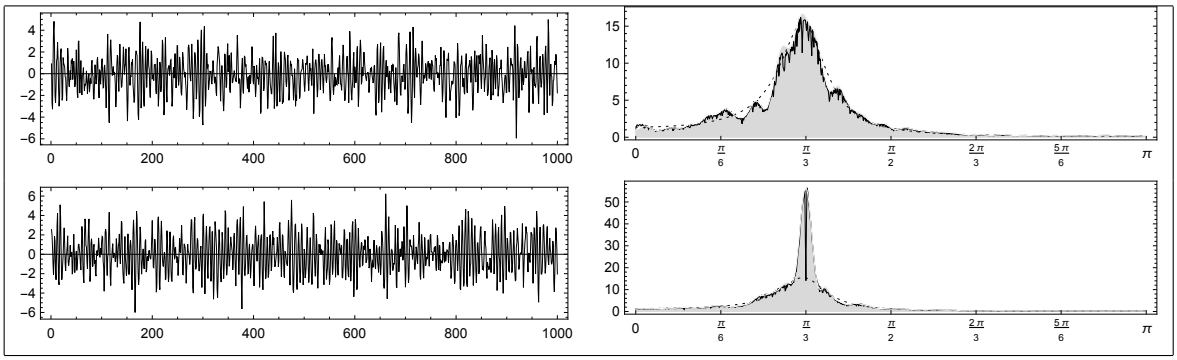

Figure 3: From top results for models **M1**-**M2**, respectively. Left panels: generated realizations; right panels: classical spectral density estimates obtained using Bartlett window of length c (shadowed area), subsampling estimates (solid line), theoretical spectral density functions (dashed line).

In the case of **M1** *the mean function is equal to zero and hence assumption of the classical approach is fulfilled. As a result performances of the classical estimator and the subsampling estimator are comparable and obtained estimates are close to the true values. Model* **M2** *is having periodic mean function with the true significant frequency equal to* $\pi/3$*. Note that the subsampling estimator provides estimate close to the true value of the spectral density function at the frequency* $\pi/3$*. At the same time the classical estimator fails at this frequency. Moreover, for* **M2** *we observe the spectral leakage around* $\pi/3$*.*

Examples above show that our estimator performs correctly and sometimes outperforms the classical one. In contrary to the classical approach we do not need to remove the mean function from the data to be able to estimate the spectral density functions. That is very important in the real data applications since the period length of the mean function may not be known prior to the performed analysis on the data or the mean function may be almost periodic. In such cases removing the first-order component from the data may be very challenging.
It should be emphasized that the leakage phenomenon that we observed appears naturally when DFT is applied to the data. We are aware that there exist variety of approaches trying to deal with the leakage problem. The aim of this paper is not to provide the new solution for the leakage problem, however in the next section we propose the modified subsampling estimator that allows for significant reduction of the leakage effect.

3.2.5 Modified subsampling estimator

Below we introduce the modified subsampling estimator that significantly reduces the leakage problem. Instead of using $\widehat{W}^{\text{sub}}_{n,b_1,b_2}(\nu,\omega)$ we apply $\widehat{W}^{\text{sub}}_{n,b_1,b_2}(\nu+v_n,\omega+v_n)$, where v_n is a deterministic sequence of real numbers such that $v_n \to 0$ as $n \to \infty$. In other words, we consider the small neighborhood of (ν,ω). At first we show mean square convergence of $\widehat{W}^{\text{sub}}_{n,b_1,b_2}(\nu+v_n,\omega+v_n)$ to $\widehat{W}^{\text{sub}}_{n,b_1,b_2}(\nu,\omega)$. In the second step to reduce leakage effect we use a finite number of estimators $\widehat{W}^{\text{sub}}_{n,b_1,b_2}(\nu+v_n,\omega+v_n)$ (with different sequences v_n) to define the new estimator as their median.

Theorem 3.5. *Assume that the conditions of Theorem 3.3 hold. Additionally, assume that* $v_n = Db_1^{-\rho}$*, where* $0 < \rho < 1/4$ *and* $D \in \mathbb{R}$*. Then, as* $n \longrightarrow \infty$ *we have*

a) $\widehat{W}^{sub}_{n,b_1,b_2}(\nu+v_n,\omega+v_n) \xrightarrow{L_2} \sigma(\nu,\omega) = 2\pi P_R(\nu,\omega)$ *for* $A = [a_{ij}]_{4\times 4}$ *with* $a_{13} = 1$, $a_{24} = 1$ *and* $a_{ij} = 0$ *otherwise;*

b) $\widehat{W}^{sub}_{n,b_1,b_2}(\nu+v_n,\omega+v_n) \xrightarrow{L_2} \sigma(\nu,\omega) = 2\pi P_I(\nu,\omega)$ *for* $A=[a_{ij}]_{4\times 4}$ *with* $a_{14}=-1$, $a_{23}=1$ *and* $a_{ij}=0$ *otherwise.*

Let $k>0$ be a fixed odd integer and $\nu,\omega\in[0,2\pi)$ be fixed frequencies. Moreover, let sequences $v_{n,1},\ldots,v_{n,k}$ be of the form $v_{n,i}=D_i b_1^{-\rho_i}$, where $0<\rho_i<1/4$ and $D_i\in\mathbb{R}$, for $i=1,2,\ldots,k$. To reduce the leakage effect we will use the following estimator

$$\widehat{M}^{\text{sub}}_{n,b_1,b_2}(\nu,\omega)=\text{med}\left\{\widehat{W}^{\text{sub}}_{n,b_1,b_2}(\nu+v_{n,1},\omega+v_{n,1}),\ldots,\widehat{W}^{\text{sub}}_{n,b_1,b_2}(\nu+v_{n,k},\omega+v_{n,k})\right\}, \quad (21)$$

where the symbol 'med' denotes the median. Applying $\widehat{M}^{\text{sub}}_{n,b_1,b_2}(\nu,\omega)$ to the points (ν,ω) that are in the area of the side-lobes, especially in the case when the side-lobes are steep, allows to reduce size of the lobes. On the other hand, when (ν,ω) is in the area, in which values of the variance estimator are not changing much, then using $\widehat{M}^{\text{sub}}_{n,b_1,b_2}(\nu,\omega)$ provides values very similar to those obtained with $\widehat{W}^{\text{sub}}_{n,b_1,b_2}(\nu,\omega)$. Thus, the main idea behind $\widehat{M}^{\text{sub}}_{n,b_1,b_2}(\nu,\omega)$ is to reduce the leakage effect by decreasing the size of the side-lobes. Below we state consistency of $\widehat{M}^{\text{sub}}_{n,b_1,b_2}(\nu,\omega)$ and we discuss the possible choice of the sequences $v_{n,i}$, $i=1,2,\ldots,k$.

Theorem 3.6. *Under assumptions of Theorem 3.5 the conclusions of Theorem 3.5 hold for* $\widehat{M}^{sub}_{n,b_1,b_2}(\nu,\omega)$ *with any* $v_{n,i}=D_i b_1^{-\rho_i}$, *where* $0<\rho_i<1/4$ *and* $D_i\in\mathbb{R}$, $i=1,2,\ldots,k$.

Remark 3.3. *Let* $v_n=Db_1^{-\rho}$, *where* $0<\rho<1/4$ *and* $D\in\mathbb{R}$ *and take any* $\tilde{k}\in\mathbb{N}$. *Then each sequence of the form* $v_{n,i}=(i-\tilde{k}-1)v_n/\tilde{k}$, *where* $i=1,\ldots,2\tilde{k}+1$ *can be equivalently expressed as* $v_{n,i}=D_i b_1^{-\rho}$ *with* $D_i=D(i-\tilde{k}-1)/\tilde{k}$, $i=1,2,\ldots,2\tilde{k}+1$ *and hence Theorem 3.6 holds for the estimator* $\widehat{M}^{sub}_{n,b_1,b_2}(\nu,\omega)$ *with the form*

$$med\left\{\widehat{W}^{sub}_{n,b_1,b_2}\left(\nu+(i-\tilde{k}-1)v_n/\tilde{k},\omega+(i-\tilde{k}-1)v_n/\tilde{k}\right): i=1,\ldots,2\tilde{k}+1\right\}. \quad (22)$$

Note that the median in (22) is calculated using $2\tilde{k}-1$ equally spaced points on the line segment joining $(\nu-v_n,\omega-v_n)$ and $(\nu+v_n,\omega+v_n)$ and contained in $[0,2\pi)^2$.

Example 3.3. *(continuation of Example 3.1) We calculated the value of the estimator* $\widehat{M}^{sub}_{n,b_1,b_2}$ *using the same model and the same values of the parameters as were set in Section 3.2.2. To get* $\widehat{M}^{sub}_{n,b_1,b_2}$ *we took sequences* $v_{n,1}=b^{-0.19}$, $v_{n,i}=b^{-0.19+0.01(i-1)}$, $i=2,\ldots,6$ *and* $v_{n,i}=-v_{n,i-5}$, $i=7,\ldots,11$. *Results are presented in Figures 4-5. As before the obtained conclusions are independent of the values of the parameters and hence we do not provide*

all figures.
In Figure 4 one may observe the estimated rescaled spectral density $2\pi g_0$ for different values of b. Definitely, using the estimator $\widehat{M}^{sub}_{n,b_1,b_2}$ allowed for significant reduction of the leakage effect. The side-lobes are no longer present around $\lambda_0 = 2\pi/30$. The same conclusion can be drawn looking on the estimates on the bifrequency square (see Figure 5).

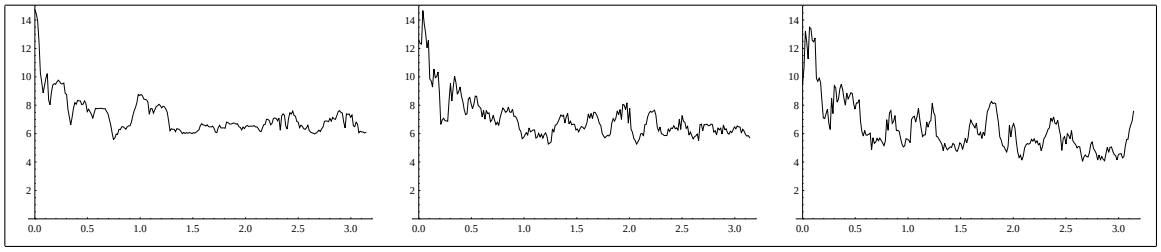

Figure 4: Estimated spectral density $2\pi g_0$ on interval $[0, 2\pi)$. Sample size $n = 1000$. From left results for subsample length b equal to $\lfloor\sqrt{n}\rfloor, \lfloor 2\sqrt{n}\rfloor, \lfloor 3\sqrt{n}\rfloor$, respectively.

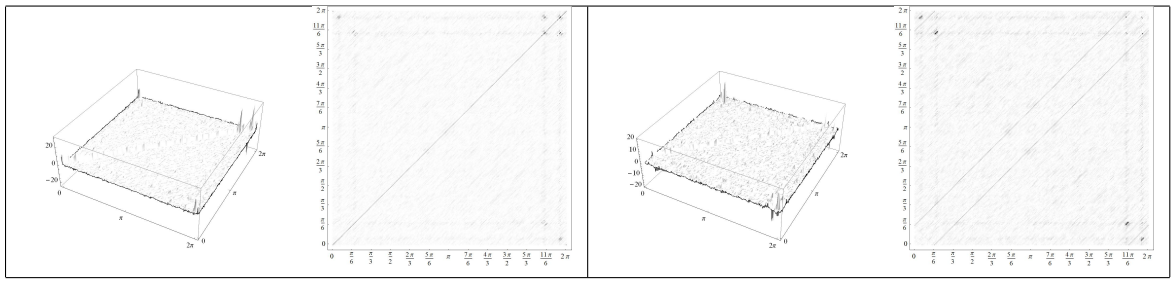

Figure 5: Estimates of $2\pi Re(P(\nu,\omega))$ (left-hand side) and $2\pi Im(P(\nu,\omega))$ (right-hand side) for $(\nu,\omega) \in [0, 2\pi)^2$ together with their projections on the xy-plane. Sample size $n = 1000$, $b = \lfloor 3\sqrt{n}\rfloor$.

4 Simulation study

In this section we present the results of the performed simulation study. Our aim was to investigate the performance of our subsampling estimators of the spectral density function over different APC processes. For that purpose we defined a random measure, which is an average distance between the estimator and the true parameter (spectral density function)

i.e.,

$$w_n(\lambda) = \frac{1}{2\pi} \int_0^{2\pi} \left| \widetilde{P}_n(\nu, \omega - \lambda) - P(\nu, \nu - \lambda) \right| \mathrm{d}\nu, \quad \lambda \in [0, \pi], \tag{23}$$

where $\widetilde{P}_n(\nu,\omega)$ is $\widehat{W}^{\mathrm{sub}}_{n,b_1,b_2}(\nu,\omega)$, $\widehat{M}^{\mathrm{sub}}_{n,b_1,b_2}(\nu,\omega)$ or the classical spectral density estimator with the Bartlett window. Recall that the spectral density of an APC process can be nonzero only on the main diagonal and lines that are parallel to the main diagonal in the bifrequency square $[0,2\pi)^2$. Thus, $w_n(\lambda)$ is measuring the distance between $\widetilde{P}_n(\nu,\omega-\lambda)$ and $P(\nu,\nu-\lambda)$ only along those lines. Moreover, note that the generalized spectral density function $P(\nu,\omega)$ is not continuous function on the square $[0,2\pi)^2$, but for any fixed $\lambda \in [0,2\pi)$, $P(\nu,\nu-\lambda)$ is continuous and hence $w_n(\cdot)$ is well defined. Finally, since $w_n(\pi-\lambda) = w_n(\pi+\lambda)$, we took only $\lambda \in [0,\pi]$. In order to be able to interpret the results of our simulations, we set $b_1 = b_2$ as this allows us to compare performance of our estimator with the classical estimator based on the Bartlett window.

In our study we considered time series of the form

$$X_t = \mu(t) + \left(1 + \frac{9}{10} \sin\left(\frac{t\pi}{2}\right)\right) \zeta_t, \tag{24}$$

where ζ_t is a zero mean stationary process with unit variance and $\mu(t) = a\left(\frac{1}{2} + \sin t\right)$ with $a \in \{0,1\}$. If $a = 0$ then $\{X_t\}$ is a zero mean PC time series with period $T = 4$ (i.e., $\Psi = \emptyset$ and $\Lambda \subset \{0, \pi/2, \pi, 3\pi/2\}$). For $a = 1$, the mean function $\mu(t)$, $t \in \mathbb{Z}$ is almost periodic (while it is periodic function for $t \in \mathbb{R}$). Moreover, the autocovariance function of $\{X_t\}$ is periodic and in consequence $\{X_t\}$ is APC with $\Psi = \{0, 1, 2\pi - 1\}$ and $\Lambda \subset \{0, \pi/2, \pi, 3\pi/2\}$. Note that if the spectral density function of ζ_t exists, then the generalized spectral density function $P(\nu,\omega)$ of the process $\{X_t - \mu(t)\}$ also exists and its explicit form can be found e.g., in Hurd and Miamee (2007) (pp. 171-172). For our study we chose a few examples of zero mean stationary time series ζ_t with unit variance that are listed below.

S1 AR(1) process with coefficient 0.8;

S2 AR(2) process with coefficients $1.9, -0.95$;

S3 MA(2) process with coefficients $-0.2, 1.2$;

S4 white noise;

S5 FARIMA(1,-0.4,1) with autoregression coefficient 0.7 and moving average coefficient 1.0;

S6 GARCH(1,1) process with coefficients $0.6, 0.1$.

To illustrate the complexity of the estimation task for **S1**-**S6** in Figure 6 we present the magnitude of $P(\nu, \nu-\lambda)$ for $\nu \in [\lambda, 2\pi+\lambda]$, where $\lambda \in \{0, \pi/2, \pi, 3\pi/2\}$ ($|P(\nu,\nu-\lambda)| = 0$ for $\lambda \notin \{0, \pi/2, \pi, 3\pi/2\}$). This function is constant for **S4** and **S6** and it has particulary complex shape in the case of **S2** and **S5**.

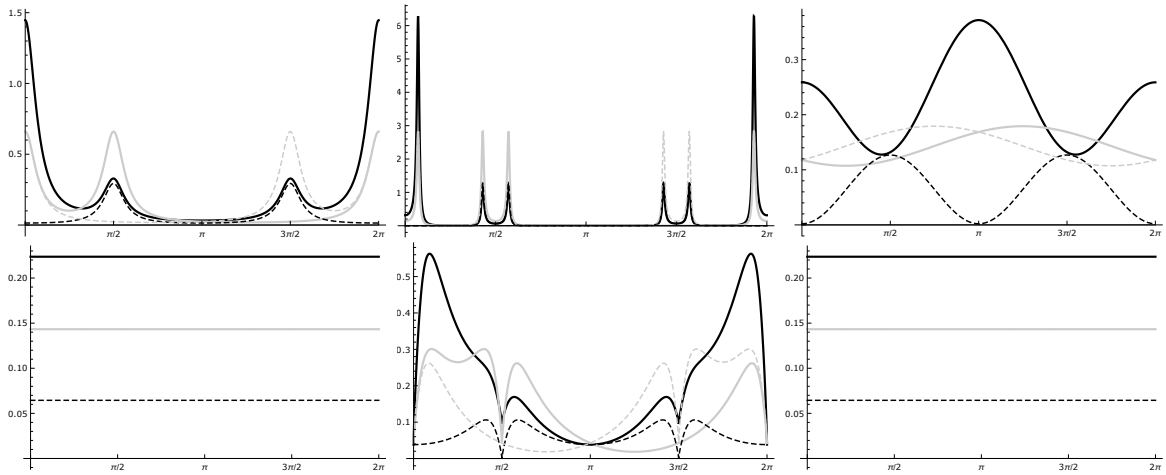

Figure 6: First row: models **S1-S3** (from left). Second row: models **S4-S6** (from left). $|P(\nu,\nu-\lambda)|$ for $\nu \in [\lambda, 2\pi+\lambda]$ and for $\lambda = 0$ (solid black line); $\lambda = \pi/2$ (solid gray line); $\lambda = \pi$ (dashed black line); $\lambda = 3\pi/2$ (dashed gray line).

For the sake of clarity and comparability of the results between different models, in all cases (**S1**-**S6**) the variance of white noise was set in a way to ensure the unity of the (unconditional) variance of ζ_t (i.e., $\mathrm{E}(\zeta_t^2) = 1$). Finally, in all cases we calculated the estimates of the generalized spectral density function $P(\nu,\omega)$ using $\widehat{W}^{\mathrm{sub}}_{n,b_1,b_2}(\nu,\omega)$ and $\widehat{M}^{\mathrm{sub}}_{n,b_1,b_2}(\nu,\omega)$. Moreover, to compare their performance with the classical estimator designed for the zero mean case, we considered the estimator of the form (19) applied for raw data $Y_t = X_t$. All calculations were performed on the set of frequencies $(\nu,\omega) \in K \times K$, where $K = \{k_j = \pi(j-1)/600 : j = 1, 2, \ldots, 1200\}$. For our simulation we took the sample size $n \in \{400, 800, 1200, 2400\}$ and the subsample length $b_1 = b_2 = b = \lfloor 5\sqrt{n} \rfloor$. Additionally, for $\widehat{M}^{\mathrm{sub}}_{n,b}(\nu,\omega)$ we set $\tilde{k} \in \{2, 4, 8, 16, 32\}$ and $v_{n,i} = \frac{(i-\tilde{k}-1)2\pi}{1200} \in K$, $i = 1, 2, \ldots, 2\tilde{k}+1$

(see Remark 3.3). Thus, $\widehat{M}^{\mathrm{sub}}_{n,b}(\nu,\omega)$ was obtained as median of the estimators calculated over the set of frequencies $\{(\nu + 2\pi j/1200, \omega + 2\pi j/1200) : j = -\tilde{k}, \ldots, 0, \ldots, \tilde{k}\}$.

To verify the performance of our subsampling estimators $\widehat{W}^{\mathrm{sub}}_{n,b}(\nu,\omega)$, $\widehat{M}^{\mathrm{sub}}_{n,b}(\nu,\omega)$ and additionally the classical estimator given by (19), in all considered cases we calculated the approximated value of $\mathrm{E}(w_n(\lambda))$ using Monte Carlo (MC) simulations. For that purpose, for each given sample

1. we computed the spectral density estimators in each point $(\nu,\omega) \in K \times K$, where $K = \{k_j = \pi(j-1)/600 : j = 1, 2, \ldots, 1200\}$;

2. for each $\lambda \in K$ we computed the approximated value of the integrate $w_n(\lambda)$.

Finally, for each considered model we generated 500 samples and we approximated $\mathrm{E}(w_n(\lambda))$ by the mean of the values obtained in the steps 1-2.
Note that our choice of the set K is not accidental. Since $\Lambda \subset K$, the measure $w_n(\lambda)$ is computed for all $\lambda \in \Lambda$. Moreover, despite the fact that the set $\tilde{\Psi} = \{(\psi_1 - \psi_2) \mod 2\pi : (\psi_1, \psi_2) \in \Psi^2\} = \{0, 1, 2, 2\pi - 2, 2\pi - 1\} \not\subset K$, the set K contains elements that are very close to frequencies from the set $\tilde{\Psi}$ ($k_{192} = 1.00007 \approx 1$, $k_{383} = 2.00015 \approx 2$).
In Corollary 3.1 we discussed the influence of the term $C_n(\nu,\omega)$ on the limiting behavior of the rescaled estimator $\widehat{W}^{\mathrm{sub}}_{n,b}(\nu,\omega)$. Let us recall that we set $b = \lfloor 5\sqrt{n} \rfloor$ and hence the limiting distribution of $T_{1n} = \sqrt{n/b}\,(\widehat{W}^{\mathrm{sub}}_{n,b}(\nu,\omega) - P(\nu,\omega))$ exists if the limiting distribution of T_{2n} exists. However, $C_n(\nu,\omega)$ still may affect the performance of our estimator by slowing down the convergence of T_{1n}. Thus, in Figure 7 we present $\log_{10}|C_n(\nu,\omega)|, (\nu,\omega) \in K^2$ for different sample sizes. The largest values (displayed as lighter colors) can be observed in the neighborhood of lines $\{(\nu,\omega) \in [0, 2\pi^2) : |\nu - \omega| \in \tilde{\Psi}$ or $(\nu,\omega) \in ([0, 2\pi) \times \Psi) \cup (\Psi \times [0, 2\pi))\}$ i.e., precisely in the neighborhood of the discontinuity set of $|C_n(\nu,\omega)|$.

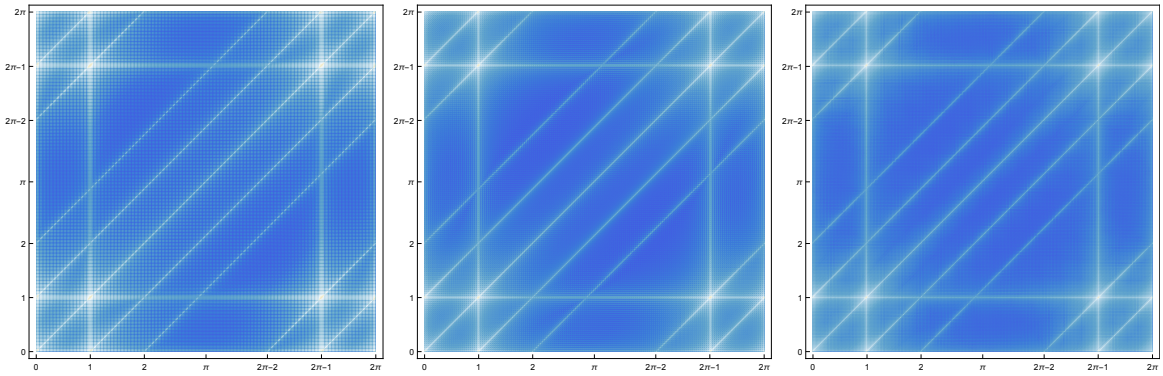

Figure 7: $\log_{10}|C_n(\nu,\omega)|$, $(\nu,\omega)\in K^2$ for $n=400,1200,2400$. Larger values are displayed as lighter colors.

Results below are split into two parts. At first we discuss the APC models **S1**-**S6** with Gaussian errors. The second part is dedicated to **S4** in non-Gaussian case.

4.1 Gaussian case

At first we discuss results for $a=1$ i.e., for $\{X_t\}$ with nonzero almost periodic mean function. Let us recall that in this case $\Psi=\{0,1,2\pi-1\}$, $\tilde{\Psi}=\{0,1,2,2\pi-2,2\pi-1\}$ and $\Lambda\subset\{0,\pi/2,\pi,3\pi/2\}$. Due to the large number of figures, we decided to restrict presentation only to those that are necessary to illustrate the conclusions.
In Figure 8 we present the approximated values of $\mathrm{E}(w_n(\lambda)),\lambda\in K\cap[0,\pi]$ for the models **S1** and **S2** and the sample sizes $n=400$ and $n=2400$. One can notice that for **S1** the function $\mathrm{E}(w_n(\lambda))$ is approximately constant, except for a few frequencies where spikes occur. In fact, exactly the same phenomenon is observed for **S3**-**S6**. In a case of **S2** spikes occur at the same frequencies, but $\mathrm{E}(w_n(\lambda)),\lambda\in K\cap[0,\pi]$ is oscillating on the whole considered interval. This is probably a consequence of a much more complicated shape (compared to other models) of the true generalized spectral density function (see Figure 6).
Note that independently of the considered model and the estimator of the generalized spectral density function, spikes occur precisely at frequencies related to the first and the second order frequencies of $\{X_t\}$ i.e., at those frequencies from the interval $[0,\pi]$ that are elements of the sets Ψ, $\tilde{\Psi}$ and Λ. Also, the order of the curves is always the same except for some spikes. The curve for the classical estimator with the Bartlett window is above all others. Below it we have respectively curves for: $\widehat{W}_{n,b}^{\mathrm{sub}}$ and $\widehat{M}_{n,b}^{\mathrm{sub}}$ with $\tilde{k}=2,4,8,16,32$. Thus, the

larger is the value of $\tilde{k}$, the smaller is $\mathrm{E}(w_n(\lambda))$. As this order is sometimes different only in the area of some spikes, in Figures 9-10 we zoom in on some of these regions. Again, since conclusions for **S1** and **S3-S6** are the same, we restrict presentation of Figures only to those for **S1** and **S2**. For **S1** curves sometimes intersect. It happens in the neighborhood of the frequencies $\lambda = 0$ and $\lambda = \pi/2$. In those two cases $\mathrm{E}(w_n(\lambda))$ for the classical estimator is sometimes smaller than the corresponding value for the subsampling estimator $\widehat{W}_{n,b}^{\mathrm{sub}}$. For **S2** curves intersect additionally in the neighborhood of $\lambda = \pi$. In that case for all sample sizes the smallest values of $\mathrm{E}(w_n(\lambda))$ around $\lambda = \pi$ are obtained for $\widehat{M}_{n,b}^{\mathrm{sub}}$ with $\tilde{k} = 4$. The worst results are observed for $\widehat{M}_{n,b}^{\mathrm{sub}}$ with $\tilde{k} = 32$. Let us recall that this estimator is usually providing the smallest values of the measure for all frequencies that are not in the neighborhood of the first and the second order frequencies of $\{X_t\}$. In the neighborhood of $\lambda = \pi/2$ (same as for $\lambda = \pi$) independently of the sample size the greatest values are for $\widehat{M}_{n,b}^{\mathrm{sub}}$ with $\tilde{k} = 32$ and the smallest for $\widehat{M}_{n,b}^{\mathrm{sub}}$ with $\tilde{k} = 4$. Finally, around $\lambda = 0$ curves for $\widehat{M}_{n,b}^{\mathrm{sub}}$ with $\tilde{k} = 16, 32$ and the one for the classical estimator intersect the remaining curves. As with **S1**, the obtained values of $\mathrm{E}(w_n(\lambda))$ for the classical estimator are sometimes smaller than the corresponding ones for $\widehat{W}_{n,b}^{\mathrm{sub}}$. The smallest $\mathrm{E}(w_n(\lambda))$ around $\lambda = 0$ are observed for $\widehat{M}_{n,b}^{\mathrm{sub}}$ with $\tilde{k} = 16$ ($n = 400, 800$) and for $\widehat{M}_{n,b}^{\mathrm{sub}}$ with $\tilde{k} = 8$ ($n = 1200, 2400$).

For $a = 0$, $\{X_t\}$ is a PC time series with mean zero and $\Lambda \subset \{0, \pi/2, \pi, 3\pi/2\}$. As before the only case, when the results are slightly different than for other models, is **S2**. The order of curves is exactly the same as for $a = 1$ for all models and for all frequencies not belonging to the neighborhood of $0, \pi/2, \pi$. For **S1** (see Figure 11) intersections of curves can be observed only for $\lambda = 0$ and $\lambda = \pi/2$. For the smallest sample $n = 400$, in the very small neighborhoods of those two frequencies the smallest values of $\mathrm{E}(w_n(\lambda))$ are obtained for the classical estimator and $\widehat{M}_{n,b}^{\mathrm{sub}}$ with $\tilde{k} = 4$. However, the differences between all estimators are slight. For $n > 400$ for both frequencies the best is the estimator $\widehat{M}_{n,b}^{\mathrm{sub}}$ with $\tilde{k} = 32$. Finally, for **S2** (see Figure 12) intersections of curves appear around $0, \pi/2, \pi$. The worst performance around those three frequencies is observed for $\widehat{M}_{n,b}^{\mathrm{sub}}$ with $\tilde{k} = 32$, while in most cases $\widehat{M}_{n,b}^{\mathrm{sub}}$ with $\tilde{k} = 4$ outperforms other estimators.
Summarizing, for $a = 0$ and $a = 1$ except the neighborhood of some first and second order frequencies, the smallest values of $\mathrm{E}(w_n(\lambda))$ are obtained for $\widehat{M}_{n,b}^{\mathrm{sub}}$ with $\tilde{k} = 32$. Moreover, independently of the model the values of $\mathrm{E}(w_n(\lambda))$ decrease when the sample size is growing. At the same time for fix n and fixed frequency λ they are very similar for all models.

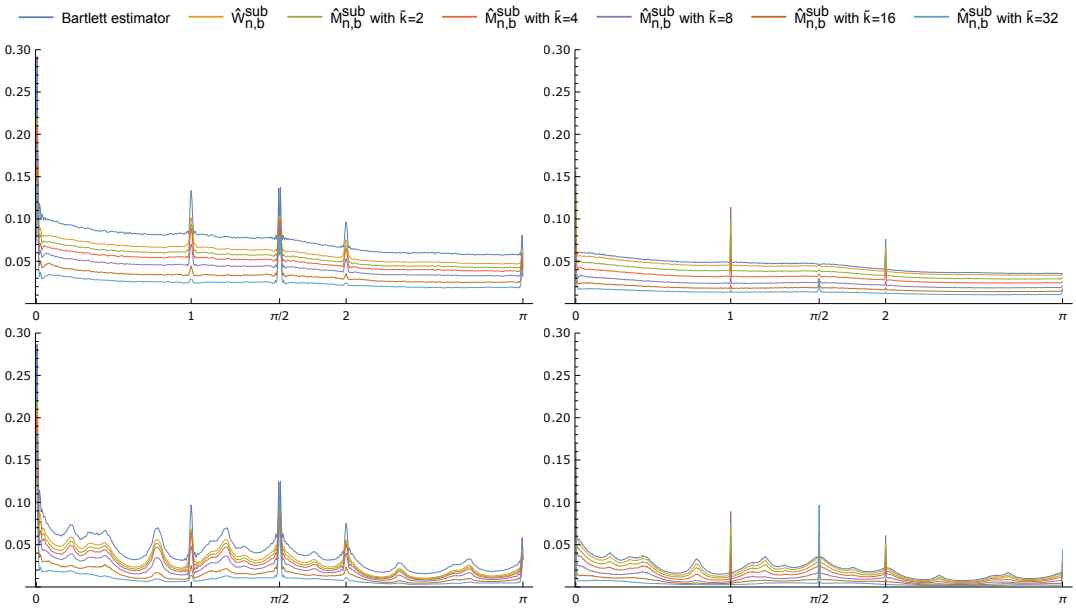

Figure 8: Approximated values of $\mathrm{E}(w_n(\lambda)), \lambda \in K \cap [0, \pi]$ for **S1** (first row) and **S2** (second row) models. Sample size $n = 400$ (first column) and $n = 2400$ (second column).

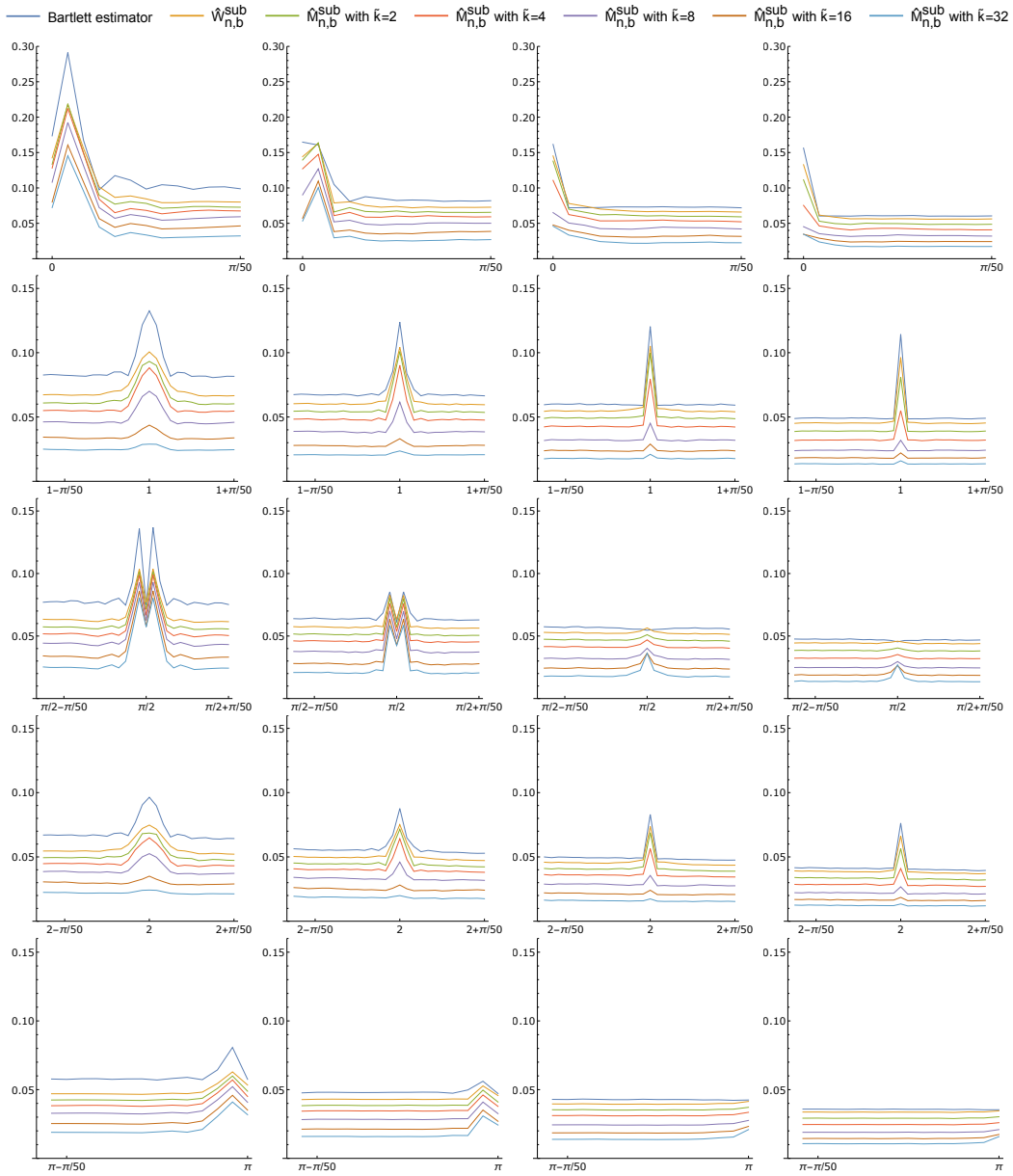

Figure 9: Model **S1**, non-zero mean case ($a = 1$). In columns (from left) $n = 400, 800, 1200, 2400$. In rows (from top) $\mathrm{E}(w_n(\lambda))$ in neighborhood of $\lambda = 0, 1, \pi/2, 2, \pi$, respectively.

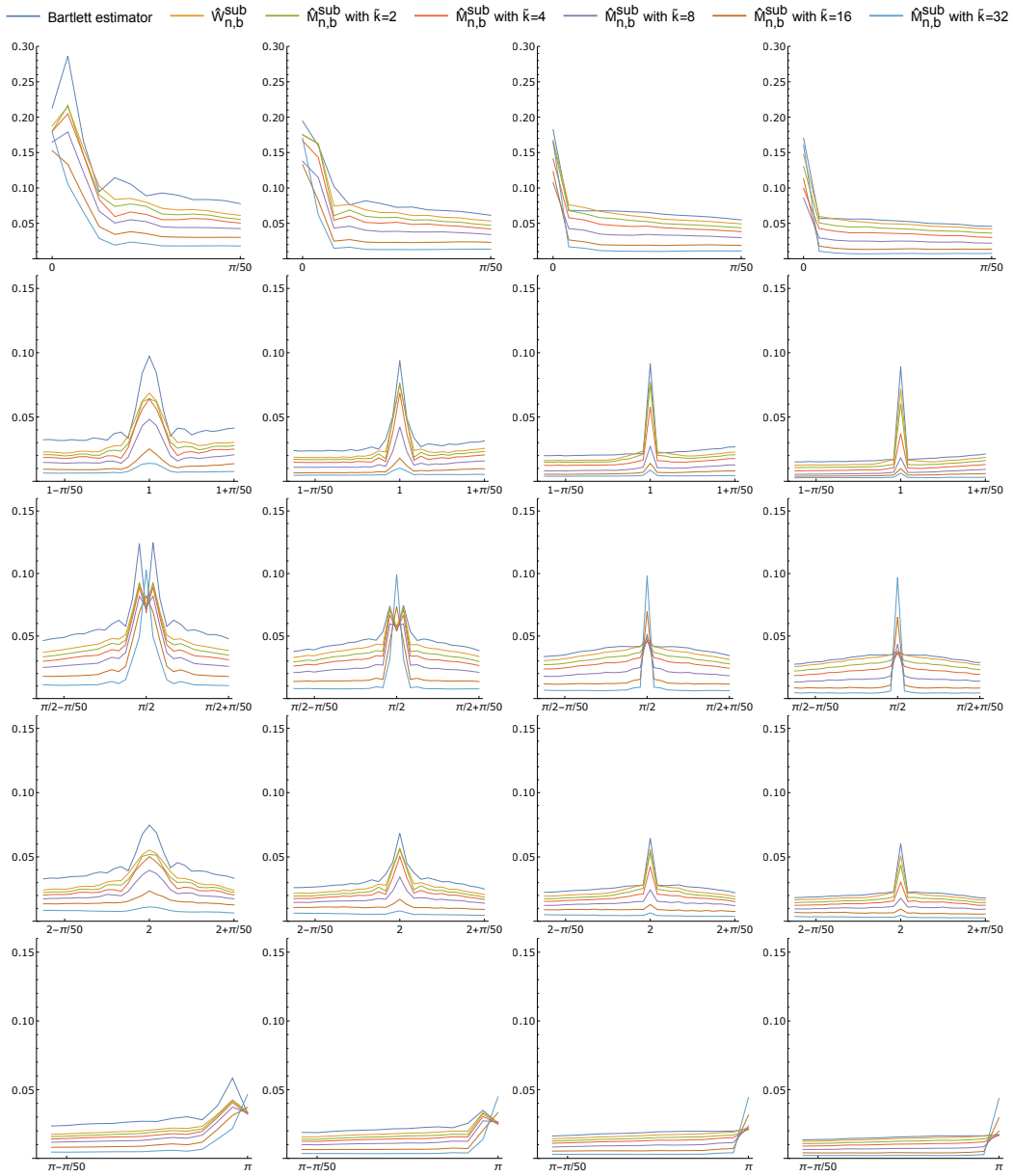

Figure 10: Model **S2**, non-zero mean case ($a = 1$). In columns (from left) $n = 400, 800, 1200, 2400$. In rows (from top) $\mathrm{E}(w_n(\lambda))$ in neighborhood of $\lambda = 0, 1, \pi/2, 2, \pi$, respectively.

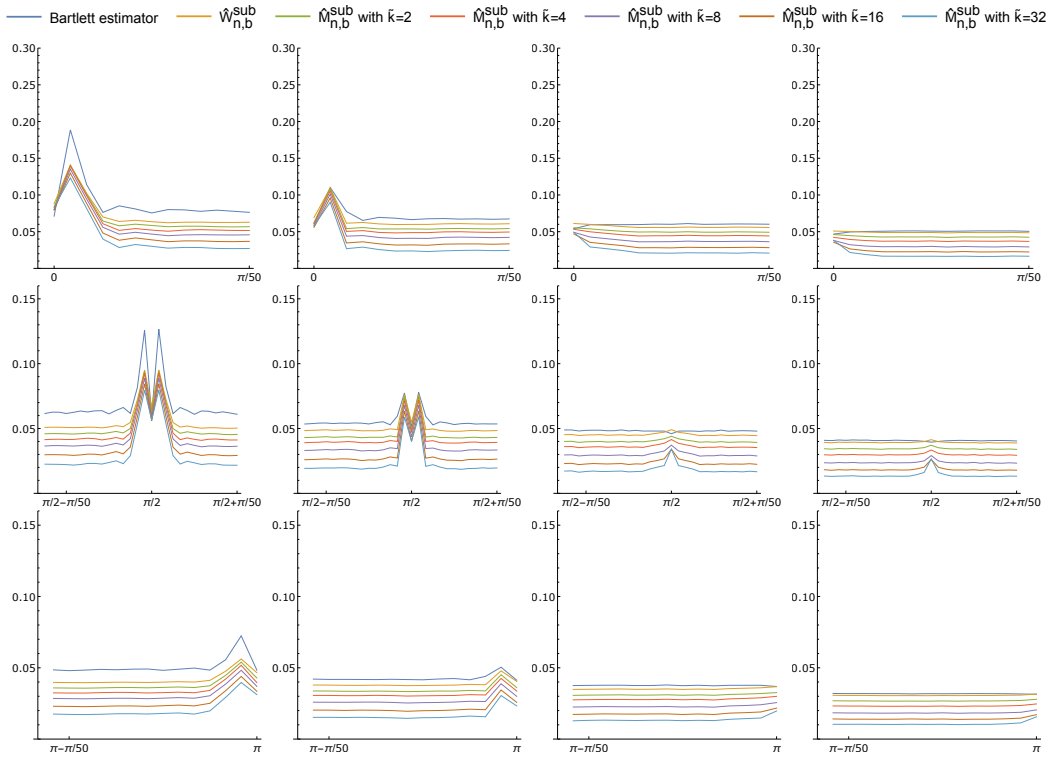

Figure 11: Model **S1**, zero mean case ($a = 0$). In columns (from left) $n = 400, 800, 1200, 2400$. In rows (from top) $\mathrm{E}(w_n(\lambda))$ in neighborhood of $\lambda = 0, \pi/2, \pi$, respectively.

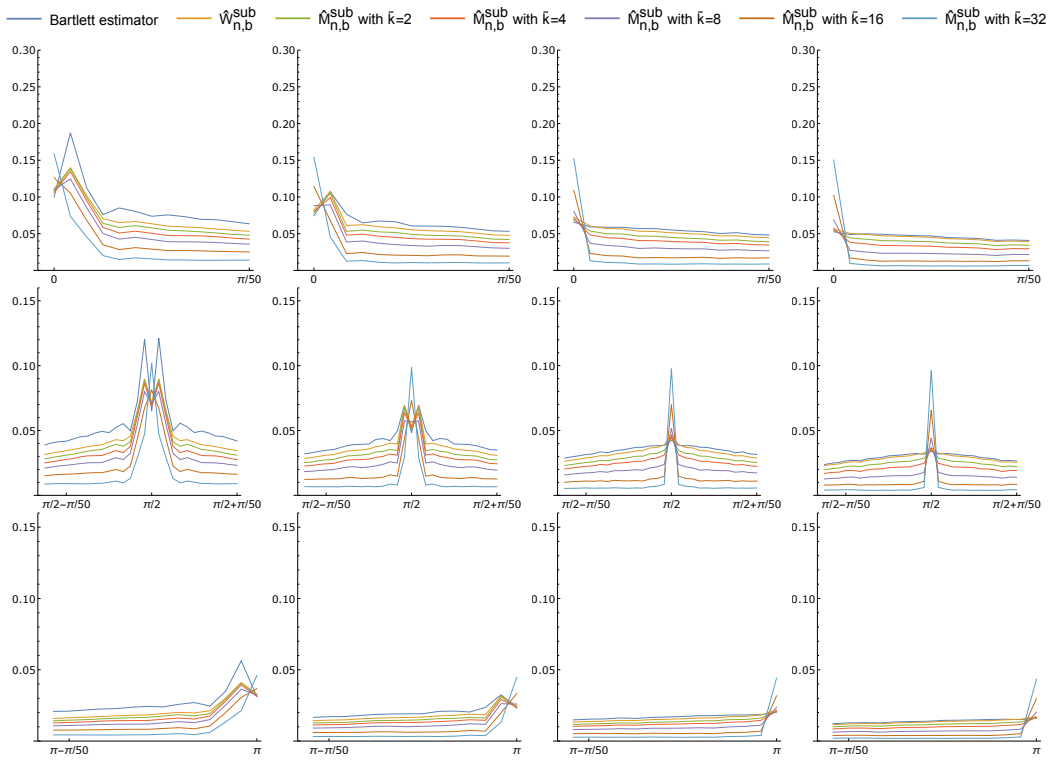

Figure 12: Model **S2**, zero mean case ($a = 0$). In columns (from left) $n = 400, 800, 1200, 2400$. In rows (from top) $\mathrm{E}(w_n(\lambda))$ in neighborhood of $\lambda = 0, \pi/2, \pi$, respectively.

4.2 non-Gaussian case

In this section we consider $\{X_t\}$ of the form **S4** with a non-zero mean function (i.e., the case $a = 1$) and non-Gaussian residuals ζ_t. Let us recall that we set parameters of distributions below in a way to ensure that $\mathrm{var}(\zeta_t) = 1$. Therefore, we do not give parameters of distributions if they are uniquely determined.
The following models for ζ_t were chosen.

W1 Student-t distribution with 3 degrees of freedom;

W2 Student-t distribution with 7 degrees of freedom;

W3 Laplace double-exponential distribution;

W4 skew-normal distribution with shape parameter $\alpha = 5$ (see Azzalini (2005));

W5 uniform distribution;

W6 mixture (with equal weights) of two Gaussian distributions with means respectively $-9/10$ and $9/10$ and standard deviations equal to $\sqrt{19}/10$. (Note that in such a case we get a bimodal distribution).

All of the general conclusions obtained in the previous section for the Gaussian case also hold for the models **W1**-**W6**. Additionally, we decided to investigate if non-Gaussianity significantly affects the values of $\mathrm{E}(w_n(\lambda))$. At first for each considered sample size n we computed the differences between values of $\mathrm{E}(w_n(\lambda)), \lambda \in K$ for **W1**-**W6** and for Gaussian white noise i.e., **S4**. In all cases the obtained values were very similar. The obtained curves (as a function of λ) were quite flat, with very low volatility and values mostly above or below zero. Drawn together in the graph they were virtually indistinguishable. Thus, for each case we computed the average (over the set K) difference. The obtained results are presented in Figure 13. The largest differences equal to 0.026 are observed for **W1**. For other models they belong to the interval $(-0.0022, 0.0032)$ for $n = 400$ and (-0.001,0.002) for $n = 2400$. Independently of the considered non-Gaussian model, the largest absolute differences are obtained for the estimator $\widehat{M}_{n,b}^{\mathrm{sub}}$ with $\tilde{k} = 32$ and the smallest for $\widehat{W}_{n,b}^{\mathrm{sub}}$. The worst performance of all estimators for **W1** is not surprising, because in that case $\{X_t\}$ does not have the sufficient number of finite moments to fulfill conditions of our theorems. In general, for **W1**, **W3** and **W4** for all considered estimators and sample sizes the obtained averages are positive, while for **W5**-**W6** they are negative. Finally, for **W2** they are negative only for $\widehat{W}_{n,b}^{\mathrm{sub}}$ and $\widehat{M}_{n,b}^{\mathrm{sub}}$ with $\tilde{k} = 2$, when $n = 400$. Note that the negative value of the observed average for **W5-W6** means that the Gaussian case can be outperformed (in terms of the measure $\mathrm{E}(w_n(\lambda)), \lambda \in K$). However, the main conclusion is that the observed difference is negligible. Thus, lack of Gaussianity does not seem to be a feature that impacts strongly the obtained results. Definitely, for Gaussian and non-Gaussian distributions independently of the considered estimator, the average differences decrease (the only exception is the heavy-tailed distribution **W1**, for which a third moment does not exist) when the sample sizes increase, which confirms the validity of our approach.

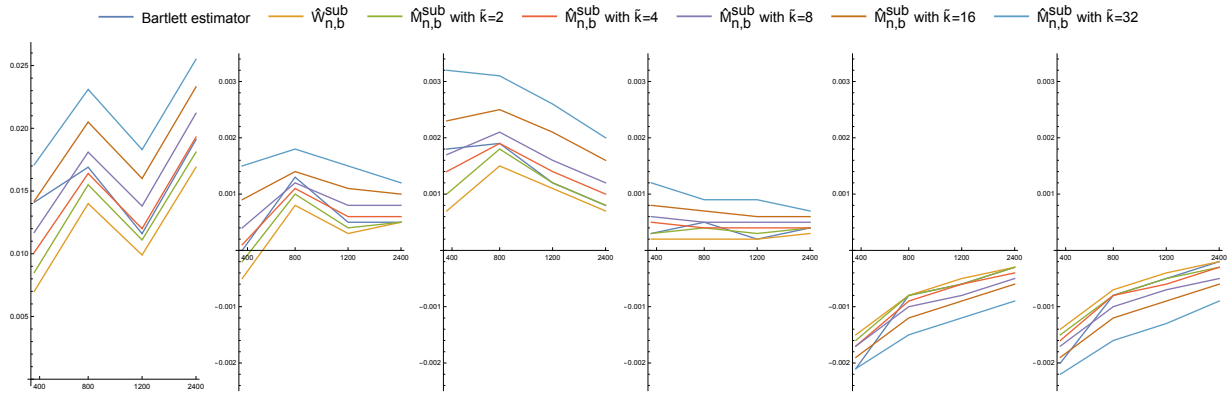

Figure 13: Results for models **W1**-**W6** (from left): average of differences between values of $\mathrm{E}(w_n(\lambda)), \lambda \in K$ for non-Gaussian model and for Gaussian white noise **S4**. X-axis: sample size $n = 400, 800, 1200, 2400$.

5 Real data applications

In this section we present two motivating examples. We focus on economic datasets, however we would like to indicate that our methodology can be applied to any data containing periodic or almost periodic structure. Thus, it can be also particularly useful for instance in telecommunication, mechanics, vibroacoustics, climatology and many others.
Below we show two very different applications of our results. At first, we consider an important problem of distinction between types of cyclical fluctuations, i.e., we discuss the type of cyclical fluctuations (deterministic or stochastic). For that purpose we consider data characterized by different types of cyclical fluctuations, i.e., business fluctuations and fluctuations related to trading-day effect. In the second part we show how to detect the second-order periodic structure in the considered data without need of estimation the almost periodic mean function.

5.1 Cyclicality of industrial production

In this subsection we consider the growth cycle for monthly industrial production (percentage change compared to same period in previous year) for three very well developed European economies: Germany, France and Italy. We use unadjusted data (i.e. neither seasonally adjusted nor calendar adjusted data) available on Eurostat webpage

(`https://ec.europa.eu/eurostat/data/database`). The analyzed sequences cover period of 28 years in the case of France and Italy (January 1991 - December 2018, number of observations $n = 336$) and 27 years in the case of Germany (January 1992 - December 2018, $n = 324$). For Germany the year 1991 is not available on Eurostat.
Usually growth cycles for industrial production are considered to be covariance stationary and hence we assume that the spectral density is 0 beyond the main diagonal line on the bifrequency square, i.e., $\Lambda = \{0\}$. Moreover, we assume that their mean functions are almost periodic and the corresponding sets Ψ are unknown. Our main aim in this subsection is to recognize which of the observed cyclical fluctuations can be classified as deterministic i.e., related to some frequency in the set Ψ and which can be classified rather as stochastic. Let us recall that the set Ψ contains all frequencies for which the corresponding Fourier coefficients of the mean function are nonzero. In the considered real data example these frequencies can be related to well known business cycle or so called calendar effects, which are caused by time-varying numbers of working days in consecutive months. Note that the typical (examined in the literature) business fluctuations are having length from 1.5 to 12 years and are related (in the case of monthly data) to frequencies from interval $[\frac{2\pi}{12\times12}, \frac{2\pi}{1.5\times12}] \approx [0.044, 0.349]$. On the other hand, the calendar effects are related to trading-day frequencies. The first ten predominant theoretical trading-day frequencies are $2.18733, 2.71093, 2.52506, 1.66374, 1.47786, 3.04865, 2.31823, 1.90852, 1.38492, 2.18603$ (see Table 11.3 in Ladiray (2012)). Note that the tenth predominant frequency is almost the same as the first one and hence they are indistinguishable. Therefore in our analysis we consider only the first nine frequencies.
In Figure 14 (left panel, black line) the periodogram for each of considered sequences is presented. In all cases on the upper part of the frame the first nine theoretical trading-day frequencies are marked. Independently of the chosen country one may observe periodogram peaks at some of those theoretical trading-day frequencies and some peaks in the interval $[0.044, 0.349]$, which contains frequencies related to the business fluctuations. To verify which frequencies belong to the set Ψ, we considered the following testing problem: $H_0 : \psi \notin \Psi$ versus $H_1 : \psi \in \Psi$, where $\psi \in [0, \pi]$. We applied a subsampling test introduced by Lenart (2013) and we set the significance level $\alpha = 5\%$ and $b = \lfloor 1.5\sqrt{n} \rfloor$. The obtained critical values of test statistic are presented in Figure 14 (left panel, gray line). Note that in all cases we reject the null hypothesis for frequencies that are mainly (but not limited to) related to trading day frequencies, i.e. the fluctuations associated with calendar effects were considered deterministic in this test. No frequencies were detected in the interval related to business fluctuations, which means that in this case the fluctuations are not related (by our test results) to any fixed frequency in Ψ.

To investigate how our spectral density estimator (13) performs in this problem we calculated its values and additionally estimates obtained using the classical spectral density estimator based on the Bartlett window (Figure 14 - right panel). As expected the subsampling estimates are less smooth than those obtained for the classical estimator. Moreover, it has local minima with side-lobes due to the leakage effect at those trading-day frequencies in which there are peaks in the periodogram (Figure 14 - left panel). This suggests that the fluctuations associated with these frequencies were considered deterministic by our spectral density estimator (i.e. related to a certain frequency in the set Ψ), which is consistent with the obtained test results. Furthermore, as discussed in Section 3, the classical estimator fails in this case, i.e. it provides incorrect estimates at deterministic frequencies (see Figure 14). Also note that in the interval $[0.044, 0.349]$ (corresponding to business fluctuations) there is no leakage effect. For each country under consideration, both estimates of the spectral density function (classical and subsampling one) have similar shapes in this frequency range, which suggests that our spectral density estimator considered associated fluctuations as stochastic. It is also consistent with the obtained results of statistical test. This confirms the classical paradigm that time-varying amplitude and phase of business cycle fluctuations cannot be described just by an almost periodic mean function and that their presence is reflected in the higher-order moments.

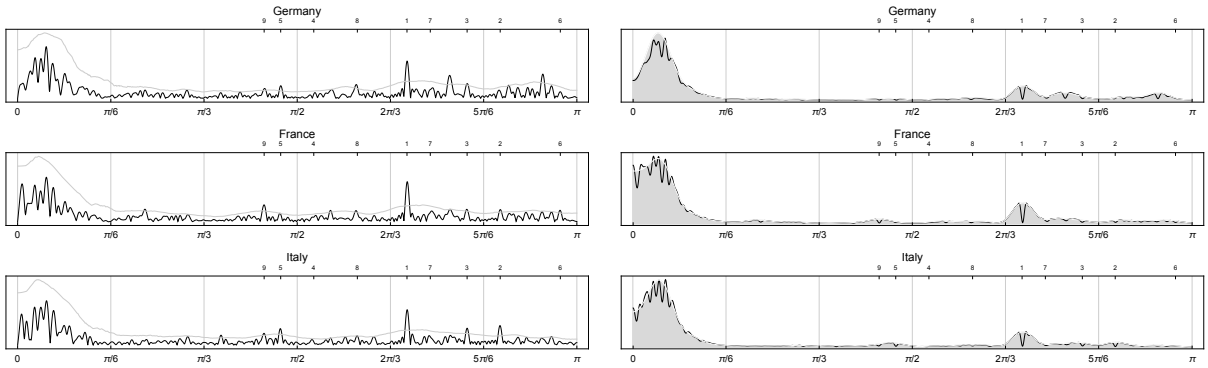

Figure 14: Left-hand side: rescaled periodogram (black line) together with critical values of subsampling test (gray line). Right-hand side: subsampling estimator of spectral density function (13) with $b_1 = \lfloor 5\sqrt[3]{n} \rfloor$ and $b_2 = \lfloor 3\sqrt{n} \rfloor$ (solid line) together with classical estimator with Bartlett window and $c = \lfloor (b_1 + b_2)/2 \rfloor$ (shadowed area).

5.2 Seasonal volatility of electricity spot prices

Energy market data are recognized to have periodic first and second moments (Broszkiewicz-Suwaj et al. (2004), Dudek et al. (2015)). However, so far detection of periodicity in the second moments was possible only when the mean function was removed from the data. In this subsection we study a typical energy market dataset. These are hourly observations from electricity spot prices in EUR/MWh of energy for Lithuania traded on the Nord Pool Spot exchange from 2013 to 2018 (52 584 observations) (`https://www.nordpoolgroup.com/historical-market-data/`). Considered data are having two types of periodicity in the mean function: daily and weakly (see e.g., Weron (2006), Nowotarski and Weron (2015)). Thus, for the mean function it is natural to set the period length $24 \times 7 = 168$ (hours). In consequence to estimate the periodic mean function we need to deal with 168 unknown parameters (Fourier coefficients). However, for our subsampling approach demeaning data is no longer necessary. We apply it to show that we are able to detect periodicity of the autocovariance function (with period length equal to 24 hours).

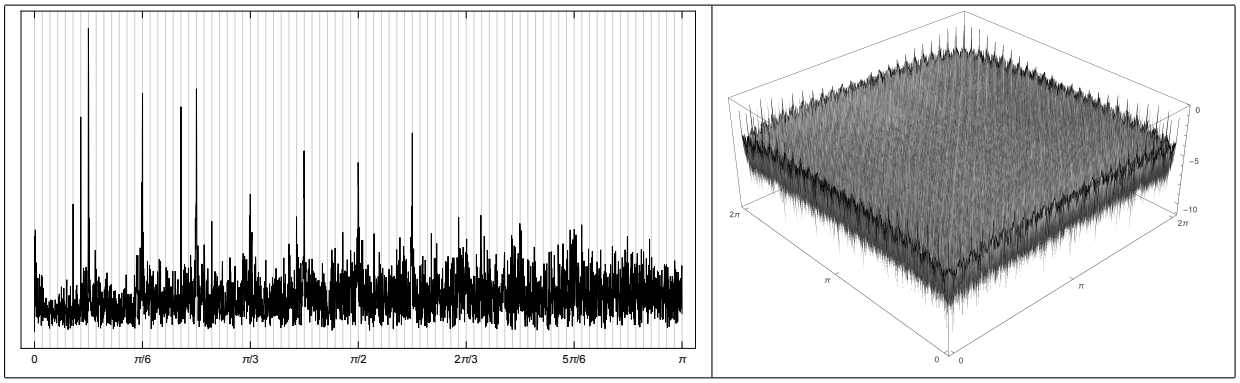

Figure 15: Left-hand side: periodogram for electricity spot prices. Right-hand side: subsampling estimator of spectral density in the log scale with $b_1 = b_2 = \lfloor 20\sqrt[3]{n} \rfloor$.

In Figure 15 we present the periodogram and the subsampling spectral density estimator (13) with $b_1 = b_2 = \lfloor 20\sqrt[3]{n} \rfloor$ computed on the set $K \times K$, where $K = \{k_j = 2\pi(j-1)/960 : j = 1, 2, \ldots, 960\}$. or the sake of clarity in the latter case we decided to use the log scale. However, due to the large number of peaks (on the bifrequency square) it is still difficult to state where the spectral density is nonzero. Thus, we use the smoothed estimator i.e., our modified subsampling estimator $\widehat{M}_{n,b}^{sub}$ (see (22)) with $b_1 = b_2$. We calculate the value of the estimators on the set $K \times K$, with $v_{n,i} = \frac{(i-\tilde{k}-1)2\pi}{960} \in K$, $i = 1, 2, \ldots, 2\tilde{k}+1$, and $\tilde{k} \in \{10, 20\}$

(see Figure 16). One may notice that independently of the chosen value of parameters, the spectral density is nonzero on lines that are parallel to the main diagonal. Moreover, the distance between those lines allows us to state that the autocovariance function is periodic with period length equal to 24 hours.

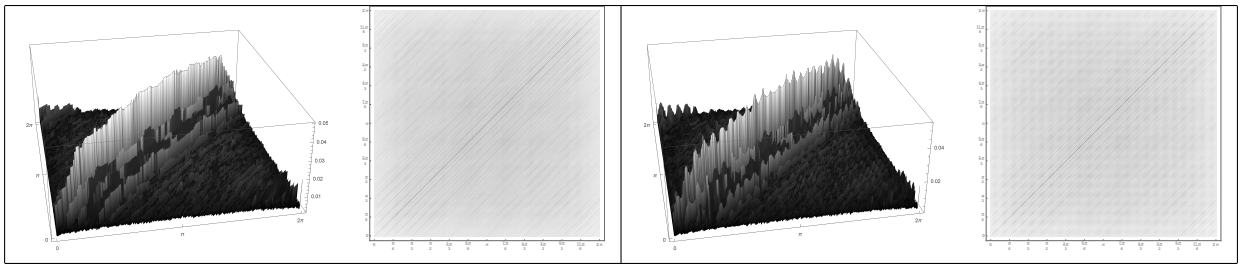

Figure 16: Modified subsampling estimators $\widehat{M}_{n,b}^{sub}$ with $\tilde{k} = 20$ (left panel) and with $\tilde{k} = 10$ (right panel) and their projections on the xy-plane.

6 Conclusions

In this paper we proposed a new nonparametric approach for spectral density estimation, which does not require demeaning of the data and is based on the subsampling method. Our estimator was constructed for almost periodically correlated sequences, which are an important subclass of nonstationary time series. Additionally, we introduced the modified estimator, which allows to reduce the leakage effect. We showed the ability of our approach to analyze the cyclical fluctuations and to determine the period length on the examples of typical economic datasets. Our method may be used to construct testing tools to recognize the type of cyclical fluctuations i.e., to decide if they are described by the periodic mean function or by the the higher-order moments.

Presented results are providing new open problems concerning subsampling estimation for the higher-order moments. Since the classical approach for the higher-order spectral estimation also requires removing the mean function from the data, we believe that ideas presented in this paper can be adapted in higher-order spectral estimation problem under nonzero mean assumption. This and other related issues are the subject of the current research of the authors.

Acknowledgement

The authors express their sincere gratitude to the anonymous referees whose comments allowed for a significant improvement of this paper.

SUPPLEMENTARY MATERIAL

Companion Document: contains all proofs of the paper. (pdf file)

References

Antoni, J. (2009). Cyclostationarity by examples. *Mech. Syst Sig Process. 23*(4), 987–1036.

Azzalini, A. (2005). The skew-normal distribution and related multivariate families. *Scandinavian Journal of Statistics 32*(2), 987–1036.

Brillinger, D. R. (2001). *Time Series: Data Analysis and Theory*. SIAM.

Broszkiewicz-Suwaj, E., A. Makagon, R. Weron, and A. Wyłomańska (2004). On detecting and modeling periodic correlation in financial data. *Physica A 336*, 196–205.

Carlstein, E. (1986). The use of subseries methods for estimating the variance of a general statistic from a stationary time series. *Annals of Statist. 14*, 1171–1179.

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

Dehay, D., A. Dudek, and J. Leśkow (2014). Subsampling for continuous-time nonstationary stochastic processes. *J. Stat. Plan. Inf. 150*, 142–158.

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

Doukhan, P. (1994). *Mixing. Properies and Examples*. Springer-Verlag, New York.

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

Dudek, A., H. Hurd, and W. Wójtowicz (2015). Parma models with applications in r. In F. e. a. Chaari (Ed.), *Cyclostationarity: Theory and Methods - II*, pp. 131–153. Springer International Publishing Switzerland.

Dudek, A. and Ł. Lenart (2017). Subsampling for nonstationary time series with non-zero mean function. *Statistics and Probability Letters 129*, 252–259.

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

Dudek, A. and P. Potorski (2020). Bootstrapping the autocovariance of pc time series - a simulation study. In F. Chaari, J. Leskow, R. Zimroz, A. Wylomanska, and A. Dudek (Eds.), *Cyclostationarity: Theory and Methods - IV; Contributions to the 10th Workshop on Cyclostationary Systems and Their Applications, February 2017, Grodek, Poland*, pp. 41–55. Springer, Cham.

Fukuchi, J. (1999). Subsampling and model selection in time series analysis. *Biometrika 86*(3), 591–604.

Gardner, W., A. Napolitano, and L. Paura (2006). Cyclostationarity: half a century of research. *Signal Processing 86*, 639–697.

Gladyshev, E. (1961). Periodically correlated random sequences. *Soviet mathematics 2*, 385–388.

Gorrostieta, C., H. Ombao, and R. von Sachs (2019). Time-dependent dual-frequency coherence in multivariate non-stationary time series. *Journal of Time Series Analysis 40*(1), 3–22.

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

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

Ladiray, D. (2012). Theoretical and real trading-day frequencies. In W. Bell, S. Holan, and T. McErloy (Eds.), *Economic time Series: Modeling and Seasonality*, pp. 255–279. Chapman and Hall.

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

Lenart, Ł. (2013). Non-parametric frequency identification and estimation in mean function for almost periodically correlated time series. *J. Multivariate Anal. 115*, 252–269.

Lenart, Ł. (2016). Generalized resampling scheme with application to spectral density matrix in almost periodically correlated class of time series. *Journal of Time Series Analysis 37*, 369–404.

Lenart, Ł. (2018). Generalized subsampling procedure for non-stationary time series. *Electronic Journal of Statistics 12(2)*, 3875–3907.

Li, T.-H. (2016). *Time series with mixed spectra*. Chapman and Hall/CRC.

Lii, K. and M. Rosenblatt (2006). Estimation for almost periodic processes. *Ann. Statist. 34*(3), 1115–1139.

Loève, M. (1963). *Probability Theory*. Van Nostrand, Princeton, NJ.

Napolitano, A. (2012). *Generalizations of Cyclostationary Signal Processing: Spectral Analysis and Applications*. Wiley-IEEE Press.

Napolitano, A. (2016). Cyclostationarity: New trends and applications. *Signal Processing 120*, 385–408.

Napolitano, A. and C. Spooner (2000). Median-based cyclic polyspectrum estimation. *IEEE Transactions on Signal Processing 48*(5), 1462–1466.

Nematollahi, A. and T. Rao (2005). On the spectral density estimation of periodically correlated (cyclostationary) time series. *Sankhya 67*(3), 568–589.

Nowotarski, J. and R. Weron (2015). Computing electricity spot price prediction intervals using quantile regression and forecast averaging. *Computational Statistics 30*, 791–803.

Politis, D. and J. Romano (1994). Large sample confidence regions based on subsamples under minimal assumptions. *Ann. Statist. 22*(4), 2031–2050.

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

Priestley, M. B. (1981). *Spectral Analysis and Time Series*. Academic Press, London.

Proakis, J. and D. Manolakis (2009). *Digital Signal Processing*. Pearson Education.

Soedjak, H. (2002). Consistent estimation of the bispectral density function ofa harmonizable process. *J. Statist. Planning Inference 100*, 159–170.

Stoica, P. and R. Moses (2005). *Spectral Analysis of Signals*. Pearson Prentice-Hall.

Tewes, J., D. Politis, and D. Nordman (2019). Convolved subsampling estimation with applications to block bootstrap. *Annals of Statistics 47(1)*, 468–496.

Žurbenko, J. G. (1986). *The Spectral Analysis of Time Series*. North-Holland, Amsterdam.

Weron, R. (2006). *Modeling and Forecasting Electricity Loads and Prices. A Statistical Approach*. John Wiley & Sons, West Susse.

Yavorskyj, I., R. Yuzefovych, I. Kravets, and I. Matsko (2012). Properties of characteristics estimators of periodically correlated random processes in preliminary determination of the period of correlation. *Radioelectronics and Communications Systems 55*(8), 335–348.