Теорія ймовір. та матем. статист.
Вип. 95, 2016, стор. 75–91

Teor. Ĭmovir. ta Matem. Statyst.
No. 95, 2016, pp. 75–91

ROBUST ESTIMATION FOR CONTINUOUS-TIME LINEAR MODELS WITH MEMORY

UDC 519.21

MAMIKON S. GINOVYAN AND ARTUR A. SAHAKYAN

Abstract. In time series analysis, much of statistical inferences about unknown spectral parameters or spectral functionals are concerned with the discrete-time stationary models, in which case it is assumed that the models are centered, or have constant means. The present paper deals with a question involving robustness of inferences, carried out on Lévy-driven continuous-time linear models, possibly exhibiting long memory, contaminated by a small trend. We show that a smoothed periodogram approach to both parametric and nonparametric estimation is robust to the presence of a small trend in the model.

1. Introduction

In time series analysis, much of statistical inferences about unknown spectral parameters or spectral functionals are concerned with the discrete-time stationary models, in which case it is assumed that the model is centered, or has a constant mean (see, e.g., Beran et al. [7], Dzhaparidze [13], Giraitis et al. [27], Taniguchi and Kakizawa [38], and references therein). In this paper we are concerned with the robustness of inferences, carried out on a continuous-time stationary process, possibly exhibiting long memory, contaminated by a small trend.

Specifically, let $\{Y(t),\, t \in \mathbb{R}\}$ be a centered stationary process possessing a spectral density $f(\lambda)$, $\lambda \in \mathbb{R}$. Assuming that either f is known with the exception of a vector parameter $\theta \in \Theta \subset \mathbb{R}^p$, or f is completely unknown and belongs to a given class $\mathcal{F}$, we want to make inferences about θ or the value $\Phi(f)$ of a given functional $\Phi(\cdot)$ at an unknown point $f \in \mathcal{F}$ in the case where the actual observed data are in the contaminated form:

$$X(t) = Y(t) + M(t), \quad 0 \le t \le T, \tag{1.1}$$

where $M(t)$ is a deterministic trend.

The process $Y(t)$ is what we believe is being observed but in reality the data are in the contaminated form $X(t)$. In this case standard inferences can be carried on the basis of the stationary model $Y(t)$, and we are interested in question whether the conclusions are robust against this kind of departure from the stationarity.

A sufficiently developed inferential theory is now available for a stationary model $Y(t)$. For instance, sufficient conditions ensuring consistency and asymptotic normality of various estimators of an unknown spectral parameter θ, constructed on the basis of the stationary data $\{Y(t),\, 0 \le t \le T\}$ were obtained by a number of authors (see, e.g., Anh et al. [3, 4], Avram et al. [5], Casas and Gao [9], Gao [15], Gao et al. [16, 17], Leonenko and Sakhno [34], and references therein). Asymptotic properties of

---

2000 *Mathematics Subject Classification.* Primary 60F05, 60G22; Secondary 62G05, 62G20.

Key words and phrases. Trend; robust inference; Lévy-driven continuous-time model; memory; smoothed periodogram; parametric and nonparametric estimation.

The research of M. S. Ginovyan was partially supported by National Science Foundation Grant #DMS-1309009 at Boston University.

nonparametric estimators for linear and some nonlinear smooth spectral functionals $\Phi(f)$ for continuous time stationary Gaussian models were studied in Ginovyan [19, 20, 23, 24].

In this paper we first prove the above asymptotic properties of estimators of θ and $\Phi(f)$ for a Lévy-driven continuous-time stationary linear model $Y(t)$, and then show that under some conditions on the process $Y(t)$ and the deterministic trend $M(t)$ these asymptotic properties of estimators of θ and $\Phi(f)$ remain valid for the model $X(t)$ of the form (1.1), that is, both the parametric and nonparametric estimating procedures are robust against replacing the stationary model $Y(t)$ by the non-stationary model (1.1).

Throughout the paper the letters C and c are used to denote positive constants, the values of which can vary from line to line.

The paper is structured as follows. In Section 2 we describe the statistical model. In Sections 3 and 4 we discuss the nonparametric and parametric estimation problems for Lévy-driven continuous time stationary linear models, respectively. Section 5 contains the robustness results. Section 6 is devoted to the proofs of the theorems.

2. The model

Let $\{Y(t), t \in \mathbb{R}\}$ be a Lévy-driven, real-valued, continuous-time stationary linear process defined by

$$Y(t) = \int_{\mathbb{R}} b(t-s)\xi(ds), \tag{2.1}$$

where $b(\cdot)$ is a function from $L^2(\mathbb{R})$, and $\xi(t)$ is a Lévy process satisfying the conditions:

$$\mathbb{E}\xi(t) = 0,\ \mathbb{E}\xi^2(1) = 1 \text{ and } \mathbb{E}\xi^4(1) < \infty.$$

A Lévy process, $\{\xi(t), t \in \mathbb{R}\}$ is a process with independent and stationary increments, continuous in probability, with sample-paths which are right-continuous with left limits (càdlàg) and $\xi(0) = \xi(0-) = 0$. The Wiener process $\{B(t), t \geq 0\}$ and the centered Poisson process $\{N(t) - \mathbb{E}N(t), t \geq 0\}$ are typical examples of centered Lévy processes. In the case where $\xi(t) = B(t)$, $Y(t)$ is a Gaussian process.

Notice that the covariance function of $Y(t)$ is given by

$$r(t) = \mathbb{E}X(t)X(0) = \int_{\mathbb{R}} b(t+x)b(x)dx, \tag{2.2}$$

it possesses the spectral density

$$f(\lambda) = \frac{\sigma^2}{2\pi}|\widehat{b}(\lambda)|^2 = \frac{\sigma^2}{2\pi}\left|\int_{\mathbb{R}} e^{-i\lambda t}b(t)dt\right|^2, \quad \lambda \in \mathbb{R}. \tag{2.3}$$

and $r(t)$ and $f(\lambda)$ are connected by the Fourier integral:

$$f(\lambda) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-i\lambda t}r(t)dt, \quad \lambda \in \mathbb{R}. \tag{2.4}$$

The function $b(\cdot)$ plays the role of a time-invariant filter.

Processes of the form (2.1) appear in many fields of science (economics, finance, physics, etc.), and cover a large class of popular models in continuous-time time series modeling. For instance, the so-called continuous-time autoregressive moving average (CARMA) models, which are the continuous-time analogs of the classical autoregressive moving average (ARMA) models in discrete-time case, are of the form (2.1) and play a central role in the representations of continuous-time stationary time series.

Memory (dependence) structure of the model. There are several possible definitions of the notion of "memory" of a stationary process, and they are not necessarily identical (see, e.g., Beran et al. [7], Gao [15], Giraitis et al. [27], Heyde and Dai [29], Taniguchi and Kakizawa [38], and references therein).

In this paper, we define the memory concept basing on the integrability property of covariance function $r(t)$, and depending on the memory structure we will distinguish the following types of stationary models:

(a) short memory or short-range dependent,

(b) intermediate memory or anti-persistent,

(c) long memory or long-range dependent.

We will say that the process $Y(t)$ displays *short memory (SM)* or *short-range dependence (SRD)* if the covariance function $r(t)$ is integrable: $r \in L^1(\mathbb{R})$ and $\int_{\mathbb{R}} r(t)dt \neq 0$. In this case the spectral density $f(\lambda)$ is bounded away from zero and infinity at frequency $\lambda = 0$, that is, $0 < f(0) < \infty$.

A typical continuous-time short memory model example is the stationary continuous-time autoregressive moving average (CARMA) process whose spectral density is a rational function (see, e.g., Brockwell [8]).

Much of statistical inference is concerned with the short memory stationary models. However, data in many fields of science (economics, finance, hydrology, telecommunications, etc.) is well modeled by a stationary process with *unbounded* or *vanishing* at the origin spectral density (see, e.g., Beran et al. [7], Casas and Gao [9], Gao [15], Tsai and Chan [39], and references therein).

The process $Y(t)$ is said to be *anti-persistent* or exhibits *intermediate memory (IM)* if the covariance function $r(t)$ is integrable: $r \in L^1(\mathbb{R})$ and $\int_{\mathbb{R}} r(t)dt = 0$. In this case the spectral density $f(\lambda)$ vanishes at frequency zero: $f(0) = 0$.

We will say that the process $Y(t)$ displays *long memory (LM)* or *long-range dependence (LRD)* if the covariance function $r(t)$ is not integrable: $r \notin L^1(\mathbb{R})$. In this case the spectral density $f(\lambda)$ has a pole at frequency zero, that is, it is unbounded at the origin.

The memory property of a stationary process can also be characterized by the behavior of spectral density $f(\lambda)$ in the neighborhood of zero, or by the behavior of covariance function $r(t)$ at infinity (see, e.g., Beran et al. [7], Section 1.3.4).

An example of a continuous-time model that displays the above defined memory structures is the continuous-time autoregressive fractionally integrated moving-average (CARFIMA) process. (see, e.g., Chambers [10], Tsai and Chan [39]).

Remark 2.1. In the continuous context, a basic process which has commonly been used to model LRD is fractional Brownian motion (fBm) $B_H(t)$ with Hurst index H. This is a Gaussian process with stationary increments and spectral density of the form

$$f(\lambda) \sim c\,|\lambda|^{1-2H}, \quad c > 0, \;\; 1/2 < H < 1, \tag{2.5}$$

as $\lambda \to 0$, and covariance function:

$$r(t) \sim c\,t^{2H-2}, \; 1/2 < H < 1, \tag{2.6}$$

as $t \to \infty$, where the symbol " $\sim$ " indicates that the ratio of left- and right-hand sides tends to 1. Notice that the form (2.5) can be understood in a limiting sense, since the fBm B_H is a nonstationary process (see, e.g., Solo [36], Gao et al. [16]). A proper stationary model in lieu of fBm is the fractional Riesz-Bessel motion (fRBm), introduced in Anh et al. [1], and then extensively discussed in a number of papers (see, e.g., Anh et al. [2], Gao et al. [16], Leonenko and Sakhno [34], and references therein). The fRBm is defined to be a continuous-time Gaussian stationary process with spectral density of the form

$$f(\lambda) = \frac{c}{|\lambda|^{2u}(1+\lambda^2)^v}, \quad \lambda \in \mathbb{R}, \, 0 < c < \infty, \, 0 < u < 1/2, \;\; v > 0. \tag{2.7}$$

Observe that the spectral density (2.7) behaves as $O(|\lambda|^{-2u})$ as $|\lambda| \to 0$ and as $O(|\lambda|^{-2(u+v)})$ as $|\lambda| \to \infty$. Thus, under the conditions $0 < u < 1/2$, $v > 0$ and $u+v > 1/2$ the function $f(\lambda)$ in (2.7) is well-defined for both $|\lambda| \to 0$ and $|\lambda| \to \infty$ due to the presence of the component $(1+\lambda^2)^{-v}$, which is the Fourier transform of the Bessel potential.

The exponent u determines the LRD, while the exponent v indicates the second-order intermittency of the fRBm (see, e.g., Anh et al. [2] and Gao et al. [16]). Comparing (2.5) and (2.7), we observe that the spectral density of fBm is the limiting case as $v \to 0$ that of fRBm with Hurst index $H = u + 1/2$. Thus, the form (2.7) means that fRBm may exhibit both LRD and second-order intermittency.

3. Nonparametric Estimation Problem

Suppose we observe a finite realization $\mathbf{Y}_T = \{Y(t), 0 \leq t \leq T\}$ of a zero mean stationary process $Y(t)$ with an *unknown* spectral density $f(\lambda)$, $\lambda \in \mathbb{R}$. We assume that $f(\lambda)$ belongs to a given (infinite-dimensional) class $\mathcal{F} \subset L^p := L^p(\mathbb{R})$ $(p \geq 1)$ of spectral densities possessing some specified smoothness properties. The problem is to estimate the value $\Phi(f)$ of a given functional $\Phi(\cdot)$ at an *unknown* "point" $f \in \mathcal{F}$ on the basis of an observation $\mathbf{Y}_T$, and investigate the asymptotic (as $T \to \infty$) properties of the suggested estimators, depending on the dependence structure of the model $Y(t)$ and smoothness structure of the "parametric" set $\mathcal{F} \subset L^p(\mathbb{R})$ $(p \geq 1)$.

This problem for discrete time stationary Gaussian processes has been considered in a number of papers. We cite merely the papers Dahlhaus and Wefelmeyer [12], Ginovyan [18, 22], Ibragimov and Khas'minskii [28, 30], and Millar [35], where can be found additional references.

For continuous time stationary Gaussian processes the problem was studied in Ginovyan [19, 20, 23, 24], where efficient nonparametric estimators for linear and some nonlinear smooth spectral functionals were constructed and asymptotic bounds for minimax mean square risks of these estimators were obtained.

The objective of this section is construction of consistent and asymptotically normal nonparametric estimators for linear and some nonlinear smooth spectral functionals in the case where the underlying model $Y(t)$ is a Lévy-driven continuous-time stationary linear process defined by (2.1) with possibly unbounded or vanishing spectral density function.

Assume first that the estimand functional $\Phi(f)$ is linear and continuous in $L^p(\mathbb{R})$, $p > 1$. Then $\Phi(f)$ admits the representation

$$\Phi(f) = \int_{\mathbb{R}} f(\lambda) g(\lambda) d\lambda, \tag{3.1}$$

where $g(\lambda) \in L^q(\mathbb{R})$, $1/p + 1/q = 1$, which will be referred as generating function for $\Phi(f)$.

As an estimator for $\Phi(f)$ we consider the averaged periodogram statistic, that is, the *simple "plug-in"* statistic:

$$\widehat{\Phi}_T := \Phi(I_{TY}) = \int_{\mathbb{R}} I_{TY}(\lambda) g(\lambda) d\lambda, \tag{3.2}$$

where $I_{TY}(\lambda)$ is the continuous periodogram of $Y(t)$ defined by

$$I_{TY}(\lambda) = \frac{1}{2\pi T} \left| \int_0^T Y(t) e^{i\lambda t} dt \right|^2. \tag{3.3}$$

Given numbers $p \geq 1$, $0 < \alpha < 1$, $r \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$, we set $\beta = \alpha + r$ and denote by $H_p(\beta)$ the L^p-Hölder class, that is, the class of those functions $\psi(\lambda) \in L^p(\mathbb{R})$, which have r-th derivatives in $L^p(\mathbb{R})$ and with some positive constant C satisfy

$$||\psi^{(r)}(\cdot + h) - \psi^{(r)}(\cdot)||_p \leq C|h|^\alpha.$$

The next result contains sufficient conditions for a simple "plug-in" statistic $\Phi(I_{TY})$ to be a consistent and asymptotically normal estimator for a linear functional $\Phi(f)$.

Theorem 3.1. *Let $\Phi(f)$ and $\widehat{\Phi}_T = \Phi(I_{TY})$ be defined by (3.1) and (3.2), respectively. Let the spectral density $f(\lambda) \in H_p(\beta_1)$, $\beta_1 > 0$, $p \geq 1$ and let the generating function $g(\lambda) \in H_q(\beta_2)$, $\beta_2 > 0$, $q \geq 1$ with $1/p + 1/q = 1$. Assume that one of the conditions a)–d) is fulfilled:*

a) $\beta_1 > 1/p$, $\beta_2 > 1/q$
b) $\beta_1 \leq 1/p$, $\beta_2 \leq 1/q$ and $\beta_1 + \beta_2 > 1/2$
c) $\beta_1 > 1/p$, $1/q - 1/2 < \beta_2 \leq 1/q$
d) $\beta_2 > 1/q$, $1/p - 1/2 < \beta_1 \leq 1/p$.

Then the statistic $\widehat{\Phi}_T := \Phi(I_{TY})$ is a consistent and asymptotically normal estimator for functional $\Phi(f)$. More precisely, we have

$$T^{1/2}\left[\Phi(I_{TY}) - \Phi(f)\right] \xrightarrow{d} \eta \quad \text{as} \quad T \to \infty, \tag{3.4}$$

where η is $N(0, \sigma^2)$ with σ^2 given by

$$\sigma^2 = 16\pi^3 \int_{\mathbb{R}} f^2(\lambda) g^2(\lambda) d\lambda + \kappa_4 \left[2\pi \int_{\mathbb{R}} f(\lambda) g(\lambda) d\lambda\right]^2. \tag{3.5}$$

Here $N(0, \sigma^2)$ denotes the normal law with mean zero and variance σ^2, $\xrightarrow{d}$ stands for convergence in distribution, and κ_4 is the fourth cumulant of $\xi(1)$.

Remark 3.1. Notice that if the underlying process $Y(t)$ is Gaussian, then in (3.5) we have only the first term.

The estimation problem becomes somewhat more complicated for non-linear functionals. In this case the simple "plug-in" statistic $\Phi(I_{TY})$ is not necessary a consistent estimator for the functional $\Phi(f)$, and hence instead of the periodogram $I_{TY}(\lambda)$, we need to use a suitable sequence of consistent estimators $\widehat{f}_{TY}$ of f.

We assume that the spectral density $f \in H_p(\beta_1)$, $\beta_1 > 0$, $p \geq 1$, and as an estimator for unknown spectral density f we take the statistic:

$$\widehat{f}_T(\lambda) = \int_{-\infty}^{\infty} W_T(\lambda - \mu) I_{TY}(\mu)\, d\mu, \tag{3.6}$$

where $I_{TY}(\lambda)$ is the periodogram of $Y(t)$ defined by (3.3). For the kernel $W_T(\lambda)$ we set down the following assumptions.

Assumption 3.1. $W_T(\lambda) = M_T W(M_T \lambda)$, where $M_T = O(T^s)$, and $b_T := M_T^{-1}$ is the bandwidth. The choice of the number s $(0 < s < 1)$ will depend on the appriori knowledge about f and Φ.

Assumption 3.2. $W(\lambda)$ is bounded, even, nonnegative function with $W(\lambda) \equiv 0$ for $|\lambda| > 1$ and

$$\int_{-1}^{1} W(\lambda)\, d\lambda = 1, \quad \int_{-1}^{1} \lambda^k W(\lambda)\, d\lambda = 0, \quad k = 1, 2, \ldots, r,$$

where $r = [\beta_1]$ is the integer part of β_1.

We assume the functional $\Phi(\cdot)$ to be Fréchet differentiable in L^2 with derivative $\Phi'(f) := \Phi'(f; \lambda)$ satisfying a Hölder condition: there exist constants $C > 0$ and δ $(0 < \delta \leq 1)$ such that for any $f_1, f_2 \in L^2$,

$$||\Phi'(f_1) - \Phi'(f_2)|| \leq C||f_1 - f_2||_2^\delta. \tag{3.7}$$

The next result contains sufficient conditions for a "plug-in" statistic $(\Phi(\widehat{f}_{TY})$ to be a consistent and asymptotically normal estimator for a nonlinear functional $\Phi(f)$.

Theorem 3.2. *Let the spectral density f and the functional $\Phi(f)$ be such that:*
(i) the pair $(f, g := \Phi'(f))$ satisfies the conditions of Theorem 3.1.
(ii) $\Phi(\cdot)$ satisfies the condition (3.7) with $\delta \geq (2\beta_1 - 1)^{-1}$.
Let the estimator $\widehat{f}_{TY}$ for f be defined by (3.6) with the kernel $W_T(\lambda)$ satisfying Assumptions 3.1 and 3.2 with $\frac{1}{2\beta_1} < s < \frac{\delta}{\delta+1}$.

Then the "plug-in" statistic $(\Phi(\widehat{f}_{TY})$ is consistent and asymptotically normal estimator for functional $\Phi(f)$ with asymptotic variance σ^2 as in (3.5) with $g = \Phi'(f))$, that is,

$$T^{1/2}\left[\Phi(\widehat{f}_{TY}) - \Phi(f)\right] \xrightarrow{d} \eta \quad \text{as} \quad T \to \infty, \tag{3.8}$$

where η is $N(0, \sigma^2)$ with σ^2 given by (3.5) with $g = \Phi'(f))$.

Remark 3.2. In fact, in Theorems 3.1 and 3.2 the estimators $\Phi(I_{TY})$ and $\Phi(\widehat{f}_{TY})$ are also $\sqrt{T}$-consistent in the mean square sense, that is, $E[T^{1/2}\{\Phi(\widehat{I}_{TY}) - \Phi(f)\}] \to 0$ and $E[T^{1/2}\{\Phi(\widehat{f}_{TY}) - \Phi(f)\}] \to 0$ as $T \to \infty$ (see Theorem 9 of Ginovyan [21] and the proofs of Theorems 3.1 and 3.2).

Remark 3.3. In the case where the underlying process $Y(t)$ is Gaussian, the results of Theorems 3.1 and 3.2 were proved in Ginovyan [20] and [24], respectively.

4. Parametric Estimation Problem

Now we assume that the spectral density $f = f(\lambda, \theta)$ is known with the exception of a vector parameter $\theta := (\theta_1, \ldots, \theta_p) \in \Theta \subset \mathbb{R}^p$. The problem of interest is to estimate θ on the basis of a sample $\mathbf{Y}_T := \{Y(t),\ 0 \leq t \leq T\}$, and investigate the asymptotic properties of the suggested estimators, depending on the dependence structure of the model.

There are different methods of estimation: maximum likelihood, Whittle, minimum contrast, etc. Here we focus on Whittle method, which is based on the smoothed periodogram analysis on a frequency domain, involving approximation of the likelihood function and asymptotic distributions of empirical spectral functionals.

The Whittle estimation procedure, originally devised for discrete-time short memory stationary processes (see [40]), has played a major role in the parametric estimation in the frequency domain, and was the focus of interest of many statisticians. Their aim was to weaken the conditions needed to guarantee the validity of the Whittle approximation for short memory models, to find analogues for long and intermediate memory models, and to show that the Whittle estimator is asymptotically equivalent to the exact maximum likelihood estimator (see, e. g., Dahlhaus [11], Dzhaparidze [13], Fox and Taqqu [14], Giraitis and Surgailis [26], Giraitis et al. [27], and references therein). In particular, it was shown that for Gaussian and linear stationary models the Whittle approach leads to consistent and asymptotically normal estimators with the standard rate of convergence under short, intermediate and long memory assumptions.

Continuous versions of Whittle estimation procedure have been considered, for example, in Anh et al. [3, 4], Avram et al. [5], Casas and Gao [9], Gao [15], Gao et al. [16, 17], Leonenko and Sakhno [34].

The Whittle procedure of estimation of a parameter θ involved in the spectral density $f(\lambda, \theta)$ of the model, based on a finite realization $\mathbf{Y}_T := \{Y(t), 0 \leq t \leq T\}$ of the centered stationary process $Y(t)$, is to choose the estimator $\widehat{\theta}_{TY}$ to minimize the weighted Whittle functional:

$$U_{TY}(\theta) := \frac{1}{4\pi}\int_{\mathbb{R}}\left[\log f(\lambda, \theta) + \frac{I_{TY}(\lambda)}{f(\lambda, \theta)}\right] \cdot w(\lambda)\, d\lambda, \tag{4.1}$$

where $I_{TY}(\lambda)$ is the continuous periodogram of $Y(t)$, given by (3.3), and $w(\lambda)$ is an even weight function (that is, $w(-\lambda) = w(\lambda)$, $w(\lambda) \geq 0$, and $w(\lambda) \in L^1(\mathbb{R})$) for which the integral in (4.1) is well defined. The choice of an appropriate weight function depends on the specific form of the spectral density (see, e.g., Anh et al. [4]). An example of common used weight function is $w(\lambda) = 1/(1+\lambda^2)$.

Thus, the Whittle estimator $\widehat{\theta}_{TY}$ with weight function $w(\lambda)$ is defined to be a solution of the following estimating equation

$$\int_{\mathbb{R}} [I_{TY}(\lambda) - f(\lambda,\theta)] \frac{\partial}{\partial\theta} f^{-1}(\lambda,\theta) \cdot w(\lambda) d\lambda = 0, \tag{4.2}$$

obtained by differentiating under the integral sign in (4.1).

The asymptotic properties of the Whittle estimator $\widehat{\theta}_{TY}$ then can be obtained using the standard Taylor expansion methods based on the following smoothed periodogram convergence result, obtained in Theorem 3.1:

$$T^{1/2} \int_{\mathbb{R}} g(\lambda,\theta) [I_{TY}(\lambda) - f(\lambda,\theta)] d\lambda \xrightarrow{d} \eta \sim N(0,\sigma^2) \quad \text{as} \quad T\to\infty \tag{4.3}$$

where $g(\lambda,\theta) = \frac{\partial}{\partial\theta} f^{-1}(\lambda,\theta) w(\lambda)$, and σ^2 is as in (3.5).

To state the next result we first introduce the following assumptions.

Assumption 4.1. The parametric set Θ is a compact set in $\mathbb{R}^p$, the true value of the parameter θ_0 is in the interior of Θ, and $f(\lambda,\theta_1) \neq f(\lambda,\theta_2)$ whenever $\theta_1 \neq \theta_2$ almost everywhere in $\mathbb{R}^1$ with respect to the Lebesgue measure.

Assumption 4.2. The function $1/f(\lambda,\theta)$ is twice continuously differentiable in a neighborhood of the point θ_0, and the matrices

$$W_1(\theta) = \|w_{ij}^{(1)}(\theta)\|_{i,j=1,\dots,p}, \qquad W_2(\theta) = \|w_{ij}^{(2)}(\theta)\|_{i,j=1,\dots,p}$$

and $V(\theta) = \|v_{ij}(\theta)\|_{i,j=1,\dots,p}$, are positive definite, where

$$w_{ij}^{(1)}(\theta) = \frac{1}{4\pi}\int_{\mathbb{R}} \frac{\partial}{\partial\theta_i} \ln f(\lambda,\theta) \frac{\partial}{\partial\theta_j} \ln f(\lambda,\theta) w(\lambda) d\lambda, \tag{4.4}$$

$$w_{ij}^{(2)}(\theta) = \frac{1}{4\pi}\int_{\mathbb{R}} \frac{\partial}{\partial\theta_i} \ln f(\lambda,\theta) \frac{\partial}{\partial\theta_j} \ln f(\lambda,\theta) w^2(\lambda) d\lambda, \tag{4.5}$$

$$v_{ij}^{(1)}(\theta) = \frac{\kappa_4}{8\pi}\int_{\mathbb{R}} \frac{\partial}{\partial\theta_i} \ln f(\lambda,\theta) w(\lambda) d\lambda \int_{\mathbb{R}} \frac{\partial}{\partial\theta_j} \ln f(\lambda,\theta) w(\lambda) d\lambda, \tag{4.6}$$

and κ_4 is the fourth cumulant of $\xi(1)$.

Theorem 4.1. *Suppose that the Assumptions 4.1 and 4.2, and the conditions of Theorem 3.1 with $f = f(\lambda;\theta)$ and $g = w(\lambda)\frac{\partial}{\partial\theta_i} f^{-1}(\lambda,\theta)$ $(i = 1,\dots,p)$ are fulfilled. Then the Whittle estimator $\widehat{\theta}_{TY}$ of an unknown spectral parameter θ is consistent and asymptotically normal. More precisely, we have*

$$T^{1/2}\left(\widehat{\theta}_{TY} - \theta_0\right) \xrightarrow{d} N_p(0, R(\theta_0)) \quad \text{as} \quad T\to\infty, \tag{4.7}$$

where $N_p(\cdot,\cdot)$ denotes the p-dimensional normal law, $\xrightarrow{d}$ stands for convergence in distribution, and

$$R(\theta_0) = W_1^{-1}(\theta_0)\,(W_2(\theta_0) + V(\theta_0))\,W_1^{-1}(\theta_0), \tag{4.8}$$

where the matrices W_1, W_2 and V are defined in (4.4)-(4.6).

Remark 4.1. The result of Theorem 4.1 under different assumptions was proved in Leonenko and Sakhno [34] (see also, Avram et al. [5]).

Remark 4.2. In our analysis we can use a general even integrable smoothing function $g(\lambda;\theta)$ rather than the specific form $g(\lambda,\theta) = w(\lambda)\frac{\partial}{\partial\theta}f^{-1}(\lambda,\theta)$, which is suggested by the Whittle procedure in (4.2). The estimator $\widehat{\theta}_{GW}(\mathbf{Y}_T)$, called a generalized Whittle estimator of θ, is then obtained as a solution of the estimating equation

$$\int_{\mathbb{R}} [I_{T,Y}(\lambda) - f(\lambda,\theta)]\, g(\lambda,\theta)d\lambda = 0. \tag{4.9}$$

Then the asymptotic properties of the estimator $\widehat{\theta}_{GW}(\mathbf{Y}_T)$ can be obtained using the standard Taylor expansion methods based on the smoothed periodogram convergence results of type (4.3) with general smoothing function $g(\lambda;\theta)$. This topic will be considered in detail elsewhere.

5. Robustness to small trends of estimation

In this section, assuming that the actual observed data are in the contaminated form (1.1):

$$X(t) = Y(t) + M(t), \quad 0 \le t \le T,$$

where $Y(t)$ a centered stationary process and $M(t)$ is a deterministic trend, we are interested in question whether the conclusions obtained in Sections 3 and 4 are robust against this kind of departure from the stationarity.

We show that if the trend $M(t)$ is "small", then the asymptotic properties of estimators of θ and $\Phi(f)$, obtained for stationary model $Y(t)$, remain valid for the contaminated model $X(t)$, that is, both the parametric and nonparametric estimating procedures are robust against replacing the stationary model $Y(t)$ by the non-stationary $X(t)$. To this end, we first establish an asymptotic relation between stationary and contaminated periodograms.

5.1. **An asymptotic relation between stationary and contaminated periodograms.** The next result shows that a small trend of the form $|M(t)| \le C|t|^{-\beta}$, $\beta > 1/4$, does not effect the asymptotic properties of the empirical spectral linear functionals of a periodogram. Note that this result is of general nature, and do not require from the model $Y(t)$ to have the specific form (2.1).

Theorem 5.1. *Let $\{Y(t), t \in \mathbb{R}\}$ be a stationary mean zero process, $\{M(t), t \in \mathbb{R}\}$ be a deterministic trend, $X(t) = Y(t) + M(t)$, and let $I_{TY}(\lambda)$ and $I_{TX}(\lambda)$ be the periodograms of $Y(t)$ and $X(t)$, respectively. Let $g(\lambda)$, $\lambda \in \mathbb{R}$ be an even integrable function. If the trend $M(t)$ and the Fourier transform $a(t) := \widehat{g}(t)$ of $g(\lambda)$ are such that $M(t)$ is locally integrable on $\mathbb{R}$ and*

$$|M(t)| \le C|t|^{-\beta}, \qquad |a(t)| \le C|t|^{-\gamma}, \quad t \in \mathbb{R}, \quad 2\beta + \gamma > \frac{3}{2}, \tag{5.1}$$

with some constants $C > 0$, $\gamma > 0$ and $\beta > 1/4$, then

$$T^{1/2} \int_{\mathbb{R}} g(\lambda)\, [I_{TX}(\lambda) - I_{TY}(\lambda)]\, d\lambda \xrightarrow{P} 0 \quad \text{as} \quad T \to \infty, \tag{5.2}$$

where $\xrightarrow{P}$ stands for convergence in probability, provided that one of the following conditions holds:

(i) *the process $Y(t)$ has short or intermediate memory, that is, the covariance function $r(t)$ of $Y(t)$ satisfies $r \in L^1(\mathbb{R})$, and $\beta + \gamma > 1$,*

(ii) *the process $Y(t)$ has long memory with covariance function $r(t)$ satisfying*

$$|r(t)| \le C|t|^{-\alpha}, \quad t \in \mathbb{R}, \quad \alpha + \gamma \ge \frac{3}{2} \tag{5.3}$$

with some constants $C > 0$, $0 < \alpha \le 1$, and $\alpha + 2\beta > 1$ if $\beta < 1 < \gamma$.

Remark 5.1. It is easy to check that the statement of Theorem 5.1 holds, in particular, if the parameters α, β and γ satisfy the following conditions:
in the case *(i)*: $\beta > 1/2$, $\gamma \geq 1/2$,
in the case *(ii)*: $\alpha \geq 3/4$, $\beta > 3/8$, $\gamma \geq 3/4$.

Remark 5.2. The discrete version of Theorem 5.1 (with additional conditions $\gamma = 1$ in the case *(i)*, and $\gamma > 1$, $\alpha < 1/2$ in the case *(ii)*), was proved by Heyde and Dai [29] (see also Taniguchi and Kakizawa [38], Theorems 6.4.1 and 6.4.2). Using the same arguments applied in the proof of Theorem 5.1 one can prove that the complete discrete analog of Theorem 5.1 is also true.

5.2. **Robustness to small trends of nonparametric estimation.** The next result shows that a small trend of the form $|M(t)| \leq C|t|^{-\beta}$ does not effect the asymptotic properties of the estimator of a linear spectral functional $\Phi(f)$, that is, the nonparametric estimation procedure is robust to the presence of a small trend in the model.

Theorem 5.2. *Suppose that the assumptions of Theorems 3.1 and 5.1 are fulfilled. Then the statistic $\Phi(I_{TX})$ is consistent and asymptotically normal estimator for functional $\Phi(f)$ with asymptotic variance σ^2 as in (3.5), that is, the asymptotic relation (3.4) is satisfied with $I_{TY}(\lambda)$ replaced by the contaminated periodogram $I_{TX}(\lambda)$:*

$$T^{1/2}\left[\Phi(I_{TX}) - \Phi(f)\right] \xrightarrow{d} \eta \quad \text{as} \quad T \to \infty, \tag{5.4}$$

where η is $N(0, \sigma^2)$ with σ^2 given by (3.5).

The next result shows that the nonparametric estimation procedure is robust to the presence of a small trend also in the case where the estimand spectral functional $\Phi(f)$ is nonlinear.

Theorem 5.3. *Suppose that the assumptions of Theorems 3.2 and 5.1 are fulfilled. Then the statistic $\Phi(\widehat{f}_{TX})$, where $\widehat{f}_{TX}$ is the kernel estimator of f constructed on the contaminated periodogram $I_{TX}(\lambda)$, is consistent and asymptotically normal estimator for nonlinear functional $\Phi(f)$ with asymptotic variance σ^2 given by (3.5) with $g = \Phi'(f)$, that is, the asymptotic relation*

$$T^{1/2}\left[\Phi(\widehat{f}_{TX}) - \Phi(f)\right] \xrightarrow{d} \eta \quad \text{as} \quad T \to \infty, \tag{5.5}$$

holds, where η is $N(0, \sigma^2)$ with σ^2 given by (3.5) with $g = \Phi'(f)$).

5.3. **Robustness to small trends of parametric estimation.** The next result shows that a small trend of the form $|M(t)| \leq C|t|^{-\beta}$, $\beta > 1/4$, does not effect the asymptotic properties of the Whittle estimator of an unknown spectral parameter θ, that is, the Whittle parametric estimation procedure is robust to the presence of a small trend in the model.

Theorem 5.4. *Suppose that the assumptions of 5.1 with $g = f^{-1}(\lambda, \theta)\cdot w(\lambda)$ are fulfilled. Then under the conditions of Theorems 4.1 the Whittle estimator $\widehat{\theta}_{TX}$, constructed on the basis of the contaminated periodogram $I_{T,X}(\lambda)$, is consistent and asymptotically normal estimator for an unknown spectral parameter θ, that is, the asymptotic relation (4.7) is satisfied with $I_{TY}(\lambda)$ replaced by the contaminated periodogram $I_{TX}(\lambda)$:*

$$T^{1/2}\left(\widehat{\theta}_{TX} - \theta_0\right) \xrightarrow{d} N_p\left(0, R(\theta_0)\right) \quad \text{as} \quad T \to \infty, \tag{5.6}$$

where the matrix $R(\theta_0)$ is defined in (4.8).

Remark 5.3. A closely related question to the Whittle estimation procedure discussed in the present paper is estimation of spectral parameters in linear or nonlinear regression models with long or short memory errors. For discrete-time linear regression models with non-Gaussian long memory moving average errors this question was studied by Koul and Surgailis [33], where the asymptotic normality of the Whittle estimator of unknown spectral parameters was established. In continuous-time setting, the problem was considered by Ivanov and Prikhod'ko [31, 32], where the consistency and asymptotic normality of the Whittle and Ibragimov's minimum contrast estimators of unknown spectral parameters in nonlinear regression models with a stationary Gaussian noise was proved.

6. Proofs

Proof of Theorem 3.1. It is enough to show that

$$T^{1/2}\left[E\Phi(I_{TY})-\Phi(f)\right]\to 0 \quad \text{as} \quad T\to\infty \tag{6.1}$$

and

$$T^{1/2}\left[\Phi(I_{TY})-E\Phi(I_{TY})\right]\xrightarrow{d}\eta \quad \text{as} \quad T\to\infty, \tag{6.2}$$

where η is as in the theorem. To this end, observe first that under the conditions of the theorem, by Lemma 4 of Ginovyan [20] we have

$$|E\Phi(I_{TY})-\Phi(f)| \le \begin{cases} C\,T^{-(\beta_1+\beta_2)}, & \text{for} \quad \beta_1+\beta_2<1 \\ C\,T^{-1}\ln T & \text{for} \quad \beta_1+\beta_2=1 \\ C\,T^{-1} & \text{for} \quad \beta_1+\beta_2>1. \end{cases} \tag{6.3}$$

Hence (6.1) follows from (6.3), since under each of the conditions a)-d) of the theorem we have $\beta_1+\beta_2>1/2$.

Next, to prove (6.2) observe that under the conditions of the theorem, there exist numbers p_1 ($p_1>p$) and q_1 ($q_1>q$), such that $H_p(\beta_1)\subset L_{p_1}$, $H_q(\beta_2)\subset L_{q_1}$ and $1/p_1+1/q_1\le 1/2$ (see Ginovyan [20], p. 401-402). Hence (6.2) follows from Theorem 4.1 of Avram et al. [5] (see also Theorem 2.1 of Bai et al. [6], and Theorem 3 of Ginovyan and Sahakyan [25]). □

Proof of Theorem 3.2. First, using the arguments of the proof of Lemma 4.4 of Ginovyan [24], it can be shown that under the conditions of the theorem, we have

$$|\xi_T-\eta_T| = o_P(1) \quad \text{as} \quad T\to\infty, \tag{6.4}$$

where

$$\eta_T := T^{1/2}\int_{-\infty}^{\infty}\Phi'(f;\lambda)\left[\widehat{f}_{TY}(\lambda)-f(\lambda)\right]d\lambda \tag{6.5}$$

and

$$\xi_T := T^{1/2}\int_{-\infty}^{\infty}\Phi'(f;\lambda)\left[I_{TY}(\lambda)-f(\lambda)\right]d\lambda. \tag{6.6}$$

Next, by Lemma 4.5 of Ginovyan [24], under the conditions of the theorem, we have

$$T^{1/2}||\widehat{f}_T-f||_2^{1+\delta} = o_P\,(1) \qquad \text{as} \quad T\to\infty. \tag{6.7}$$

Therefore in view of (6.4) – (6.7) we can write

$$\begin{aligned} T^{1/2}[\Phi(\widehat{f}_T)-\Phi(f)] &= T^{1/2}\int_{-\infty}^{\infty}\Phi'(f;\lambda)\left[\widehat{f}_{TY}(\lambda)-f(\lambda)\right]d\lambda \\ &\quad + O_P\left(T^{1/2}||\widehat{f}_{TY}-f||^{1+\delta}\right) \\ &= T^{1/2}\int_{-\infty}^{\infty}\Phi'(f;\lambda)\left[I_{TY}(\lambda)-f(\lambda)\right]d\lambda + o_P(1). \end{aligned} \tag{6.8}$$

Hence the assertion of the theorem follows from Theorem 3.1 with $g(\lambda)=\Phi'(f;\lambda)$. □

Proof of Theorem 4.1. The proof is similar to that of Theorem 4.3 of Avram et al. [5] (see also, Theorem 2 of Leonenko and Sakhno [34]), and so is omitted. It should only be noted that in these theorems, along other conditions it assumed that

$$T^{1/2}\int_{\mathbb{R}}[EI_{TY}(\lambda)-f(\lambda,\theta)]\,\frac{\partial}{\partial\theta}f^{-1}(\lambda,\theta)\cdot w(\lambda)d\lambda\to 0 \quad \text{as} \quad T\to\infty, \tag{6.9}$$

while in our case (6.9) follows from assumptions of the theorem (see proof of Theorem 3.1). □

Proof of Theorem 5.1. In view of (1.1) and (3.3) we can write

$$\begin{aligned} I_{T,X}(\lambda)-I_{T,Y}(\lambda) &= \frac{1}{2\pi T}\left(\left|\int_0^T e^{i\lambda t}X(t)\,dt\right|^2-\left|\int_0^T e^{i\lambda t}Y(t)\,dt\right|^2\right)\\ &= \frac{1}{2\pi T}\left(\left|\int_0^T e^{i\lambda t}[Y(t)+M(t)]\,dt\right|^2-\left|\int_0^T e^{i\lambda t}Y(t)\,dt\right|^2\right)\\ &= \frac{1}{2\pi T}\int_0^T\int_0^T e^{i\lambda(t-s)}\,[Y(t)M(s)+Y(s)M(t)+M(t)M(s)]\,dtds \end{aligned}$$

and

$$\begin{aligned} &\int_{-\infty}^{+\infty} g(\lambda,\theta)\,[I_{T,X}(\lambda)-I_{T,Y}(\lambda)]\,d\lambda\\ &=\frac{1}{T}\int_0^T\int_0^T [Y(t)M(s)+Y(s)M(t)+M(t)M(s)]\,a(t-s)\,dtds. \end{aligned}\tag{6.10}$$

Thus, to complete the proof it is enough to prove that under the conditions of the theorem we have

$$T^{-1/2}\int_0^T\int_0^T M(t)M(s)a(t-s)\,dtds\to 0 \quad \text{as} \quad T\to\infty \tag{6.11}$$

and

$$T^{-1/2}\int_0^T\int_0^T Y(t)M(s)a(t-s)\,dtds \xrightarrow{P} 0 \quad \text{as} \quad T\to\infty. \tag{6.12}$$

Proof of (6.11). For $T>2$ we set

$$\begin{aligned} I(T) &:= \int_0^T\int_0^T |M(t)M(s)a(t-s)|\,dtds \qquad (6.13)\\ &= \int_0^1\int_0^2+\int_0^1\int_2^T+\int_1^T\int_0^{1/2}+\int_1^T\int_{1/2}^T =: I_1(T)+I_2(T)+I_3(T)+I_4(T), \end{aligned}$$

and estimate the integrals $I_i(T)$, $i=1,2,3,4$, separately.

Observe first that the Fourier transform $a(t):=\widehat{g}(t)$ is a bounded function on $\mathbb{R}$, since g is integrable on $\mathbb{R}$. Hence, taking into account that by assumption the trend $M(t)$ is locally integrable on $\mathbb{R}$, for $I_1(T)$ we obtain the estimate

$$I_1(T)\le C\|a\|_\infty\int_0^1|M(s)|\,ds\int_0^2|M(t)|\,dt\le C<\infty, \qquad T>2. \tag{6.14}$$

Next, in view of (5.1), for $0<s<1$ and $t>2$ we have $|a(t-s)|\le C(t-s)^{-\gamma}\le Ct^{-\gamma}$, and hence, taking into account that $\beta+\gamma>1$, $I_2(T)$ can be estimated as follows

$$I_2(T) \le C\int_0^1|M(s)|\,ds\int_2^T\frac{1}{t^{\beta+\gamma}}\,dt\le C<\infty, \qquad T>2. \tag{6.15}$$

Similarly, for $I_3(T)$ we have

$$I_3(T) \le C < \infty, \qquad T > 2. \tag{6.16}$$

To estimate $I_4(T)$ observe first that, in view of (5.1), for $1 < s < T$ we can write

$$\begin{aligned}
h(s) &:= \int_{1/2}^{T} |M(t)a(t-s)|\,dt \\
&\le C\left[\int_{(s-1)}^{s+1} \frac{|a(t-s)|}{t^\beta}dt + \int_{s+1}^{2s} \frac{1}{t^\beta(t-s)^\gamma}dt \right.\\
&\qquad \left. + \int_{2s}^{T} \frac{1}{t^\beta(t-s)^\gamma}dt + \int_{1/2}^{s/2} \frac{1}{t^\beta(s-t)^\gamma}dt + \int_{s/2}^{s-1} \frac{1}{t^\beta(s-t)^\gamma}dt\right] \\
&\le C\left[\|a\|_\infty \cdot s^{-\beta} + s^{-\beta}\int_1^s \frac{1}{\tau^\gamma}d\tau + \int_{2s}^{T} \frac{1}{t^{\beta+\gamma}}dt + s^{-\gamma}\int_{1/2}^{s/2} \frac{1}{t^\beta}dt + s^{-\beta}\int_1^{s/2} \frac{1}{\tau^\gamma}d\tau\right] \\
&\le C\left[s^{-\beta} + L(\gamma,T)s^{1-\beta-\gamma} + L(\beta+\gamma,T)\left(T^{1-\beta-\gamma} + s^{1-\beta-\gamma}\right)\right. \\
&\qquad \left. + L(\beta,T)\left(s^{1-\beta-\gamma} + s^{-\gamma}\right) + L(\gamma,T)s^{1-\beta-\gamma}\right] \\
&\le C\log T \cdot \left(T^{1-\beta-\gamma} + s^{1-\beta-\gamma} + s^{-\beta} + s^{-\gamma}\right),
\end{aligned} \tag{6.17}$$

where the function $L(u,T)$ is defined by

$$L(u,T) = \begin{cases} \log T & \text{if} \quad u = 1, \\ 1 & \text{otherwise.} \end{cases}$$

Taking into account that $\beta + \gamma > 1$, from (6.17) we get

$$h(s) \le C\log T \cdot \left(s^{1-\beta-\gamma} + s^{-\beta} + s^{-\gamma}\right), \qquad 1 < s < T, \tag{6.18}$$

and hence for $T > 2$, $I_4(T)$ can be estimated as follows

$$\begin{aligned}
I_4(T) &= \int_1^T |M(s)h(s)|\,ds \\
&\le C\log T\left[\int_1^T \frac{1}{s^{2\beta+\gamma-1}}\,ds + \int_1^T \frac{1}{s^{2\beta}}\,ds + \int_1^T \frac{1}{s^{\beta+\gamma}}\,ds\right] \\
&\le C\log T\left[1 + L(2\beta+\gamma-1,T)\,T^{2-2\beta-\gamma} + L(2\beta,T)\,T^{1-2\beta} + T^{1-\beta-\gamma}\right] \\
&\le C\log^2 T\left(1 + T^{2-2\beta-\gamma} + T^{1-2\beta}\right).
\end{aligned} \tag{6.19}$$

Finally, taking into account that by assumption $2\beta+\gamma > 3/2$ and $\beta > 1/4$, from (6.13)–(6.16) and (6.19) we obtain

$$T^{-1/2} \cdot I(T) \le C\log^2 T\left(T^{-1/2} + T^{3/2-2\beta-\gamma} + T^{1/2-2\beta}\right) \to 0 \quad \text{as} \quad T \to \infty,$$

which implies (6.11).

Proof of (6.12). Observe first that the inequality

$$\beta + \gamma > 1 \tag{6.20}$$

holds also in the case *(ii)*, since by (5.1) and (5.3), we have $2\beta + 2\gamma \ge 2\beta + \gamma + 3/2 - \alpha > 3/2 + 1/2 = 2$.

Denote

$$\nu(s) = \nu(T,s) := \int_0^T M(t)a(t-s)\,dt, \quad 0 < s < T,$$

and observe that

$$\int_0^{1/2} |M(t)a(t-s)|\,dt \le C\int_0^{1/2} |M(t)|\frac{1}{(s-1/2)^\gamma}\,dt \le C\cdot s^{-\gamma}, \qquad 1<s<T,$$

and by (6.18),

$$|\nu(s)| \le C\log T\cdot\left(s^{1-\beta-\gamma}+s^{-\beta}+s^{-\gamma}\right), \qquad 1<s<T. \tag{6.21}$$

On the other hand, by (5.1), for $T>2$ and $0<s<1$ we have

$$|\nu(s)| \le C\left[\|a\|_\infty\int_0^2 |M(t)|\,dt+\int_2^T \frac{1}{t^\beta(t-1)^\gamma}dt\right] \le C\log T. \tag{6.22}$$

Observe that from (5.1) and (6.21) it follows that (6.22) holds for $0<s<T$.

Now, we denote

$$Q(T) := T^{-1/2}\int_0^T\int_0^T Y(s)M(t)a(t-s)\,dtds = T^{-1/2}\int_0^T Y(s)\nu(s)ds,$$

and observe that

$$\begin{aligned} E\{Q^2(T)\} &= T^{-1}\int_0^T\int_0^T E\{Y(s)Y(\tau)\}\nu(s)\nu(\tau)dsd\tau \\ &= T^{-1}\int_0^T\int_0^T \nu(s)\nu(\tau)r(s-\tau)dsd\tau. \end{aligned}$$

Hence, to prove (6.12) it is enough show that

$$J(T) := \int_0^T\int_0^T |\nu(s)\nu(\tau)r(s-\tau)|\,dsd\tau = o(T) \quad \text{as} \quad T\to\infty. \tag{6.23}$$

In the case *(i)*, when the process $Y(t)$ has SM or IM, and hence $r\in L^1(\mathbb{R})$, from (6.22) for $T>2$ we get

$$|J(T)| \le C\log T\int_0^T |\nu(s)|\int_0^T |r(s-\tau)|\,d\tau ds \le C\log T\int_0^T |\nu(s)|\,ds. \tag{6.24}$$

In view of (6.21), the last integral in (6.24) can be estimated as follows:

$$\begin{aligned} \int_0^T |\nu(s)|\,ds &\le C\log^2 T\left[\int_0^1 ds+\int_1^T\left(s^{1-\beta-\gamma}+s^{-\beta}+s^{-\gamma}\right)ds\right] \\ &\le C\log^2 T\left[1+L(\beta+\gamma-1,T)\,T^{2-\beta-\gamma}+L(\beta,T)\,T^{1-\beta}+L(\gamma,T)\,T^{1-\gamma}\right] \\ &\le C\log^3 T\left(1+T^{1-\beta}+T^{1-\gamma}+T^{2-\beta-\gamma}\right). \end{aligned} \tag{6.25}$$

Hence, taking into account that $\beta+\gamma>1$, from (6.24) and (6.25) we obtain

$$J(T)=o(T) \quad \text{as} \quad T\to\infty.$$

In the case *(ii)*, when the process $Y(t)$ has LM, using (5.3), (6.20), (6.21) and (6.22), for $1<\tau<T$ we obtain

$$\begin{aligned} q(\tau) &:= \int_0^T |\nu(s)r(s-\tau)|ds \le C\log T\int_0^{1/2}\left(\tau-\frac{1}{2}\right)^{-\alpha}ds \\ &+ C\log T\left[\int_{1/2}^T \frac{|r(s-\tau)|}{s^{\beta+\gamma-1}}ds+\int_{1/2}^T \frac{|r(s-\tau)|}{s^\beta}ds+\int_{1/2}^T \frac{|r(s-\tau)|}{s^\gamma}ds\right]. \end{aligned} \tag{6.26}$$

Taking into account that r is bounded ($|r(t)| \leq r(0) = E|Y(t)|^2 < \infty$, $t \in \mathbb{R}$), and using similar arguments as in (6.17), from (5.3) we obtain that for any $\eta > 0$

$$\begin{aligned}\int_{1/2}^{T} \frac{|r(t-\tau)|}{t^{\eta}}\, dt &\leq C \log T \left(T^{1-\alpha-\eta} + \tau^{1-\alpha-\eta} + \tau^{-\alpha} + \tau^{-\eta}\right) \\ &\leq C \log T \left(1 + T^{1-\alpha-\eta}\right), \qquad 1 < \tau < T.\end{aligned}$$

Applying this inequality for $\eta = \beta + \gamma - 1$, $\eta = \beta$ and $\eta = \gamma$, from (6.26) we obtain

$$q(\tau) \leq C \log^2 T \left(1 + T^{2-\alpha-\beta-\gamma} + T^{1-\alpha-\beta}\right), \qquad 1 < \tau < T, \tag{6.27}$$

since $\alpha + \gamma > 1$. On the other hand, by (6.21) and (6.22) for $T > 2$ and $0 < \tau < 1$, we have

$$\begin{aligned}q(\tau) &\leq C \left[\log T \int_0^2 |r(s-\tau)|\, ds + \int_2^T \frac{|\nu(s)|}{(s-1)^{\alpha}} ds\right] \qquad (6.28) \\ &\leq C \log T \left(1 + T^{2-\alpha-\beta-\gamma} + T^{1-\alpha-\beta}\right) \leq C \log T \left(1 + T^{1-\beta}\right), \qquad 0 < \tau < 1,\end{aligned}$$

since $\alpha + \gamma > 1$ and $\alpha > 0$.

Next, we denote

$$J(T) = \int_0^T |\nu(\tau)|\, q(\tau)\, d\tau = \int_0^1 + \int_1^T =: J_1(T) + J_2(T), \tag{6.29}$$

and estimate $J_1(T)$ and $J_2(T)$. By (6.22) and (6.28), for $J_1(T)$ we have

$$J_1(T) \leq C \log^2 T \left(1 + T^{1-\beta}\right) = o(T) \quad \text{as} \quad T \to \infty, \tag{6.30}$$

since $\beta > 0$.

To estimate $J_2(T)$ we consider three cases, and use conditions (5.1), (5.3), (6.20) and inequalities (6.21), (6.27).

Case 1. If $\beta \geq 1$, then we have

$$|\nu(\tau)| \leq C \log T \left(\tau^{-\beta} + \tau^{-\gamma}\right), \quad q(\tau) \leq C \log^2 T, \qquad 1 < \tau < T,$$

and hence

$$J_2(T) \leq C \log^3 T \left(1 + T^{1-\beta} + T^{1-\gamma}\right) = o(T) \quad \text{as} \quad T \to \infty. \tag{6.31}$$

Case 2. If $\beta < 1 < \gamma$, then we have

$$|\nu(\tau)| \leq C \log T \cdot \tau^{-\beta}, \qquad q(\tau) \leq C \log^2 T \left(1 + T^{1-\alpha-\beta}\right) \qquad 1 < \tau < T,$$

and hence

$$\begin{aligned}J_2(T) &\leq C \log^3 T \left(1 + T^{1-\beta}\right)\left(1 + T^{1-\alpha-\beta}\right) \\ &\leq C \log^3 T \left(1 + T^{1-\alpha-\beta} + T^{1-\beta} + T^{2-\alpha-2\beta}\right) = o(T) \quad \text{as} \quad T \to \infty \qquad (6.32)\end{aligned}$$

since in this case by assumption $\alpha + 2\beta > 1$.

Case 3. If $\beta < 1$ and $\gamma \leq 1$, then we have

$$|\nu(\tau)| \leq C \log T \cdot \tau^{1-\beta-\gamma}, \qquad q(\tau) \leq C \log^2 T \left(1 + T^{2-\alpha-\beta-\gamma}\right), \qquad 1 < \tau < T,$$

and hence

$$\begin{aligned}J_2(T) &\leq C \log^3 T \left(1 + T^{2-\beta-\gamma}\right)\left(1 + T^{2-\alpha-\beta-\gamma}\right) \qquad (6.33) \\ &\leq C \log^3 T \left(1 + T^{2-\alpha-\beta-\gamma} + T^{2-\beta-\gamma} + T^{4-\alpha-2\beta-2\gamma}\right) = o(T) \quad \text{as} \quad T \to \infty,\end{aligned}$$

since $\beta + \gamma > 1$ and $\alpha + 2\beta + 2\gamma = (2\beta + \gamma) + (\alpha + \gamma) > 3$.

Finally, from (6.29)–(6.33) we obtain

$$J(T) = o(T) \quad \text{as} \quad T \to \infty.$$

Thus, the relation (6.23) and hence (6.12) are proved.

Theorem 5.1 is proved. □

Proof of Theorem 5.2. In view of (3.1) and (3.2) we can write

$$T^{1/2}\left[\Phi(I_{TX})-\Phi(f)\right]=T^{1/2}\left[\int_{\mathbb{R}}I_{TX}(\lambda)g(\lambda)d\lambda-\int_{\mathbb{R}}f(\lambda)g(\lambda)d\lambda\right]$$
$$=T^{1/2}\left[\int_{\mathbb{R}}I_{TX}(\lambda)g(\lambda)d\lambda-\int_{\mathbb{R}}I_{TY}(\lambda)g(\lambda)d\lambda\right]$$
$$+T^{1/2}\left[\int_{\mathbb{R}}I_{TY}(\lambda)g(\lambda)d\lambda-\int_{\mathbb{R}}f(\lambda)g(\lambda)d\lambda\right]$$
$$=T^{1/2}\int_{\mathbb{R}}g(\lambda)\left[I_{TX}(\lambda)-I_{TY}(\lambda)\right]d\lambda+T^{1/2}\left[\Phi(I_{TY})-\Phi(f)\right]. \tag{6.34}$$

By Theorem 5.1, the first term on the right-hand side of (6.34) goes to zero in probability as $T\to\infty$, while by Theorem 3.1, the second term on the right-hand side of (6.34) goes in distribution to η, and the result follows. □

Proof of Theorem 5.3. Write

$$T^{1/2}\left[\Phi(\widehat{f}_{TX})-\Phi(f)\right]=T^{1/2}\left[\Phi(\widehat{f}_{TY})-\Phi(f)\right]+T^{1/2}\left[\Phi(\widehat{f}_{TX})-\Phi(\widehat{f}_{TY})\right], \tag{6.35}$$

and observe that by Theorem 3.2, the first term on the right-hand side of (6.35) goes in distribution to η as $T\to\infty$.

Next, it follows from the proof of Theorem 3.1 of Ginovyan [24] that

$$T^{1/2}[\Phi(\widehat{f}_{TY})-\Phi(f)]=T^{1/2}\int_{\mathbb{R}}\Phi'(f;\lambda)\left[I_{TY}(\lambda)-f(\lambda)\right]d\lambda+o_P(1),$$
$$T^{1/2}[\Phi(\widehat{f}_{TX})-\Phi(f)]=T^{1/2}\int_{\mathbb{R}}\Phi'(f;\lambda)\left[I_{TX}(\lambda)-f(\lambda)\right]d\lambda+o_P(1).$$

Therefore we have

$$T^{1/2}\left[\Phi(\widehat{f}_{TX})-\Phi(\widehat{f}_{TY})\right]=T^{1/2}\int_{\mathbb{R}}\Phi'(f;\lambda)\left[I_{TX}(\lambda)-I_{TY}(\lambda)\right]d\lambda+o_P(1). \tag{6.36}$$

Now we can apply Theorem 5.1 with $g=\Phi'(f))$ to conclude that the second term on the right-hand side of (6.35) goes to zero in probability as $T\to\infty$, and hence the result follows. □

Proof of Theorem 5.4. By Taylor's formula for $\frac{\partial}{\partial\theta}U_{TY}\left(\widehat{\theta}_{TY}\right)$, where $U_{TY}(\cdot)$ is the Whittle functional defined by (4.1) and $\widehat{\theta}_{TY}$ is the Whittle estimator constructed on the basis of observation $\mathbf{Y}_T=\{Y(t),\,0\le t\le T\}$, for $|\widehat{\theta}^*_T-\theta_0|<|\widehat{\theta}_{TY}-\theta_0|$ and for sufficiently large T, we can write

$$T^{1/2}\left[\widehat{\theta}_{TY}-\theta_0\right]=-T^{1/2}\left[\frac{\partial^2}{\partial\theta\partial\theta'}U_{TY}(\theta^*_T)\right]^{-1}\left[\frac{\partial}{\partial\theta}U_{TY}(\theta_0)\right]+o_P(1). \tag{6.37}$$

Next, by Theorem 5.1, we have

$$U_{TX}(\theta_T)=U_{TY}(\theta_T)+o_P(1). \tag{6.38}$$

Again using Taylor's formula for $\frac{\partial}{\partial\theta}U_{TX}\left(\widehat{\theta}_{TX}\right)$, where now $U_{TX}(\cdot)$ and $\widehat{\theta}_{TX}$ are respectively the Whittle functional and the Whittle estimator, constructed on the basis of the contaminated observation $\mathbf{X}_T=\{X(t),0\le t\le T\}$, and taking into account the relations (6.37) and (6.38), we can infer that

$$T^{1/2}\left[\widehat{\theta}_{TX}-\theta_0\right]=T^{1/2}\left[\widehat{\theta}_{TY}-\theta_0\right]+o_P(1),$$

showing that the estimator $\widehat{\theta}_{TX}$ possesses the same asymptotic properties as $\widehat{\theta}_{TY}$.

Hence the result follows from Theorems 4.1. □

Acknowledgment. The authors thank the Associate Editor and two referees for constructive remarks and comments.

References

1. V. V. Anh, J. M. Angulo, and M. D. Ruiz-Medina, *Possible long-range dependence in fractional random fields,* J. Statist. Planning Inference **80** (1999), 95–110.
2. V. V. Anh, N. N. Leonenko, and R. McVinish, *Models for fractional Riesz-Bessel motion and related processes,* Fractals **9** (2001), 329–346.
3. V. V. Anh, N. N. Leonenko, and L. Sakhno, *On a class of minimum contrast estimators for fractional stochastic processes and fields,* J. Statist. Planning Inference **123** (2004), 161–185.
4. V. V. Anh, N. N. Leonenko, and L. Sakhno, *Minimum contrast estimation of random processes based on information of second and third orders,* J. Statist. Planning Inference **137** (2007), 1302–1331.
5. F. Avram, N. N. Leonenko, and L. Sakhno, *On a Szegö type limit theorem, the Hölder-Young-Brascamp-Lieb inequality, and the asymptotic theory of integrals and quadratic forms of stationary fields,* ESAIM: Probability and Statistics **14** (2010), 210–255.
6. S. Bai, M. S. Ginovyan, and M. S. Taqqu, *Limit theorems for quadratic forms of Levy-driven continuous-time linear processes.* Stochast. Process. Appl. **126** (2016), 1036–1065.
7. J. Beran, Y. Feng, S. Ghosh, and R. Kulik, *Long-Memory Processes Probabilistic Properties and Statistical Methods,* Springer, New York, 2013.
8. P. J. Brockwell, *Recent results in the theory and applications of CARMA processes,* Annals of the Institute of Statistical Mathematics **66**, (2014), No. 4, 647–685.
9. I. Casas and J. Gao, *Econometric estimation in long-range dependent volatility models: Theory and practice,* Journal of Econometrics **147** (2008), 72–83.
10. M. J. Chambers, *The Estimation of Continuous Parameter Long-Memory Time Series Models,* Econometric Theory **12** (1996), No. 2, 374–390.
11. R. Dahlhaus, *Efficient parameter estimation for self-similar processes,* Ann. Statist. **17** (1989), 1749–1766.
12. R. Dahlhaus and W. Wefelmeyer, *Asymptotically optimal estimation in misspecified time series models,* Ann. Statist. **24** (1996), 952–974.
13. K. Dzhaparidze, *Parameter Estimation and Hypothesis Testing in Spectral Analysis of Stationary Time Series,* Springer, New York, 1986.
14. R. Fox and M. S. Taqqu, *Large-sample properties of parameter estimation for strongly dependent stationary Gaussian time series,* Ann. Statist. **14** (1986), 517–532.
15. J. Gao, *Modelling long-range dependent Gaussian processes with application in continuous-time financial models,* J. Appl.Probab. **41** (2004), 467–482.
16. J. Gao, V. V. Anh, C. Heyde, and Q. Tieng, *Parameter Estimation of Stochastic Processes with Long-range Dependence and Intermittency,* J. Time Ser. Anal. **22** (2001), 517–535.
17. J. Gao, V. V. Anh, and C. Heyde, *Statistical estimation of nonstationary Gaussian process with long-range dependence and intermittency.* Stochast. Process. Appl. **99** (2002), 295—321.
18. M. S. Ginovyan, *Asymptotically efficient nonparametric estimation of functionals of a spectral density having zeros,* Theory Probab. Appl. **33** (1988), No. 2, 296–303.
19. M. S. Ginovyan, *On estimating the value of a linear functional of the spectral density of a Gaussian stationary process,* Theory Probab. Appl. **33** (1988), No. 4, 722–726.
20. M. S. Ginovyan, *On Toeplitz type quadratic functionals in Gaussian stationary process,* Probab. Theory Relat. Fields **100** (1994), 395–406.
21. M. S. Ginovyan, *Asymptotic properties of spectrum estimate of stationary Gaussian processes,* J. Cont. Math. Anal. **30** (1995), No. 1, 1–16.
22. M. S. Ginovyan, *Asymptotically efficient nonparametric estimation of nonlinear spectral functionals,* Acta Appl. Math. **78** (2003), 145–154.
23. M. S. Ginovyan, *Efficient Estimation of Spectral Functionals for Gaussian Stationary Models,* Comm. Stochast. Anal. **5** (2011), No. 1, 211–232.
24. M. S. Ginovyan, *Efficient Estimation of Spectral Functionals for Continuous-time Stationary Models,* Acta Appl. Math. **115** (2011), No. 2, 233–254.
25. M. S. Ginovyan and A. A. Sahakyan, *Limit Theorems for Toeplitz quadratic functionals of continuous-time stationary process,* Probab. Theory Relat. Fields **138** (2007), 551–579.
26. L. Giraitis and D. Surgailis, *A central limit theorem for quadratic forms in strongly dependent linear variables and its application to asymptotical normality of Whittle's estimate,* Probab. Theory Relat. Fields **86** (1990), 87–104.

27. L. Giraitis, H. L. Koul, and D. Surgailis, *Large Sample Inference for Long Memory Processes,* Imperial College Press, London, 2012.
28. R. Z. Has'minskii and I. A. Ibragimov, *Asymptotically efficient nonparametric estimation of functionals of a spectral density function,* Probab. Theory Related Fields **73** (1986), 447–461.
29. C. Heyde and W. Dai, *On the robustness to small trends of estimation based on the smoothed periodogram,* J. Time Ser. Anal. **17** (1996), No. 2, 141–150.
30. I. A. Ibragimov and R. Z. Khasminskii, *Asymptotically normal families of distributions and efficient estimation,* Ann. Statist., **19** (1991), 1681–1724.
31. A. V. Ivanov and V. V. Prikhod'ko, *On the Whittle Estimator of the Parameters of Spectral Density of Random Noise in the Nonlinear Regression Model,* Ukrainian Math.J. **67** (2016), No. 8, 1183–1203.
32. A. V. Ivanov and V. V. Prikhod'ko, *Asymptotic properties of Ibragimov's estimator for a parameter of the spectral density of the random noise in a nonlinear regression model,* Teor. Imovir. ta Matem. Statyst. No. 93 (2015), 50–66.
33. H. L. Koul and D. Surgailis, *Asymptotic normality of the Whittle estimator in linear regression model with long memory errors,* Statist. Inference for Stochast. Processes **3** (2000), No. 1, 129–147.
34. N. N. Leonenko and L. Sakhno, *On the Whittle estimators for some classes of continuous-parameter random processes and fields,* Stat & Probab. Letters **76** (2006), 781–795.
35. P. W. Millar, *Non-parametric applications of an infinite dimensional convolution theorem,* Z. Wahrsch. verw. Gebiete **68** (1985), 545–556.
36. V. Solo, *Intrinsic random functions and the paradox of $1/f$ noise,* SIAM J. Appl. Math. **52** (1992), 270—291.
37. M. Taniguchi, *Minimum contrast estimation for spectral densities of stationary processes,* J. Roy. Statist. Soc., Ser. B **49** (1987), No. 3, 315–325.
38. M. Taniguchi and Y. Kakizawa, *Asymptotic Theory of Statistical Inference for Time Series,* Springer, New York, 2000.
39. H. Tsai and K. S. Chan, *Quasi-maximum likelihood estimation for a class of continuous-time long memory processes,* J. Time Ser. Anal. **26** (2005), No. 5, 691–713.
40. P. Whittle, *Hypothesis Testing in Time Series,* Hafner, New York, 1951.

Mamikon S. Ginovyan, Department of Mathematics and Statistics,, Boston University, 111 Cummington Mall, Boston, MA, 02215, USA

E-mail address: ginovyan@math.bu.edu

Artur A. Sahakyan, Department of Mathematics and Mechanics,, Yerevan State University, 1 Alex Manoogian, Yerevan, 0025, Armenia

E-mail address: sart@ysu.am

Received 10/17/2016