On the Whittle estimator for linear random noise spectral density parameter in continuous-time nonlinear regression models

A. V. Ivanov¹ · N. N. Leonenko² · I. V. Orlovskyi¹

Received: 4 February 2019 / Accepted: 5 September 2019
© The Author(s) 2019

Abstract
A continuous-time nonlinear regression model with Lévy-driven linear noise process is considered. Sufficient conditions of consistency and asymptotic normality of the Whittle estimator for the parameter of spectral density of the noise are obtained in the paper.

Keywords  Nonlinear regression model · Lévy-driven linear noise process · The least squares estimator · Spectral density · Whittle estimator · Consistency · Asymptotic normality · Levitan polynomials

1 Introduction

The paper is focused on such an important aspect of the study of regression models with correlated observations as an estimation of random noise functional characteristics. When considering this problem the regression function unknown parameter becomes nuisance and complicates the analysis of noise. To neutralise its presence, we must estimate the parameter and then build estimators, say, of spectral density parameter of a stationary random noise using residuals, that is the difference between the values of the observed process and fitted regression function.

So, in the first step we employ the least squares estimator (LSE) for unknown parameter of nonlinear regression, because of its relative simplicity. Asymptotic properties of the LSE in nonlinear regression model were studied by many authors. Numerous results on the subject can be found in monograph by Ivanov and Leonenko (1989), Ivanov (1997).

¹ Department of Mathematical Analysis and Probability Theory, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Peremogi Avenue 37, Kiev 03056, Ukraine
² School of Mathematics, Cardiff University, Senghennydd Road, Cardiff CF24 4AG, UK

Published online: 12 September 2019
In the second step we use the residual periodogram to estimate the unknown parameter of the noise spectral density using the Whittle-type contrast process (Whittle 1951, 1953).


In the article by Koul and Surgailis (2000) in the linear regression model the asymptotic properties of the Whittle estimator of strongly dependent random noise spectral density parameters were studied in a discrete-time setting.

In the paper by Ivanov and Prykhod’ko (2016) sufficient conditions on consistency and asymptotic normality of the Whittle estimator of the spectral density parameter of the Gaussian stationary random noise in continuous-time nonlinear regression model were obtained using residual periodogram. The current paper continues this research extending it to the case of the Lévy-driven linear random noise and more general classes of regression functions including trigonometric ones. We use the scheme of the proof in the case of Gaussian noise (Ivanov and Prykhod’ko 2016) and some results of the papers (Avram et al. 2010; Anh et al. 2004). For linear random noise the proofs utilize essentially another types of limits theorems. In comparison with Gaussian case it leads to the use of special conditions on linear Lévy-driven random noise, new consistency and asymptotic normality conditions.

In the present publication continues-time model is considered. However, the results obtained can also be used for discrete time observations using the statements like Theorem 3 of Alodat and Olenko (2017) or Lemma 1 of Leonenko and Taufer (2006).

2 Setting

Consider a regression model

\[ X(t) = g(t, \alpha_0) + \varepsilon(t), \quad t \geq 0, \]

where \( g: (-\gamma, \infty) \times \mathcal{A}_\gamma \rightarrow \mathbb{R} \) is a continuous function, \( \mathcal{A} \subset \mathbb{R}^q \) is an open convex set, \( \mathcal{A}_\gamma = \bigcup \{ \mathcal{A} + \gamma e \} \), \( \gamma \) is some positive number, \( \alpha_0 \in \mathcal{A} \) is a true value of unknown parameter, and \( \varepsilon \) is a random noise described below.

**Remark 1** The assumption about domain \((-\gamma, \infty)\) for function \( g \) in \( t \) is of technical nature and does not effect possible applications. This assumption makes it possible to formulate the condition \( N_2 \), which is used in the proof of Lemma 7.

Throughout the paper \( (\Omega, \mathcal{F}, P) \) denotes a complete probability space.

A Lévy process \( L(t), t \geq 0 \), is a stochastic process, with independent and stationary increments, continuous in probability, with sample-paths which are right-continuous with left limits (càdlàg) and \( L(0) = 0 \). For a general treatment of Lévy processes we refer to Applebaum (2009) and Sato (1999).

Let \( (a, b, \Pi) \) denote a characteristic triplet of the Lévy process \( L(t), t \geq 0 \), that is for all \( t \geq 0 \)

\[ \log \mathbb{E} \exp \{ i z L(t) \} = t \kappa(z) \]
for all $z \in \mathbb{R}$, where
\[ \kappa(z) = iaz - \frac{1}{2}b z^2 + \int_{\mathbb{R}} \left( e^{izu} - 1 - iz \tau(u) \right) \Pi(du), \quad z \in \mathbb{R}, \quad (2) \]
where $a \in \mathbb{R}$, $b \geq 0$, and
\[ \tau(u) = \begin{cases} u, & |u| \leq 1; \\ \frac{u}{|u|}, & |u| > 1. \end{cases} \]
The Lévy measure $\Pi$ in (2) is a Radon measure on $\mathbb{R} \setminus \{0\}$ such that $\Pi(\{0\}) = 0$, and
\[ \int_{\mathbb{R}} \min(1, u^2) \Pi(du) < \infty. \]
It is known that $L(t)$ has finite $p$th moment for $p > 0$ ($\mathbb{E}|L(t)|^p < \infty$) if and only if
\[ \int_{|u| \geq 1} |u|^p \Pi(du) < \infty, \]
and $L(t)$ has finite $p$th exponential moment for $p > 0$ ($\mathbb{E}[e^{pL(t)}] < \infty$) if and only if
\[ \int_{|u| \geq 1} e^{pu} \Pi(du) < \infty, \quad (3) \]
see, i.e., Sato Sato (1999), Theorem 25.3.

If $L(t)$, $t \geq 0$, is a Lévy process with characteristics $(a, b, \Pi)$, then the process $-L(t)$, $t \geq 0$, is also a Lévy process with characteristics $(-a, b, \tilde{\Pi})$, where $\tilde{\Pi}(A) = \Pi(-A)$ for each Borel set $A$, modifying it to be càdlàg (Anh et al. 2002).

We introduce a two-sided Lévy process $L(t)$, $t \in \mathbb{R}$, defined for $t < 0$ to be equal an independent copy of $-L(-t)$.

Let $\hat{a} : \mathbb{R} \to \mathbb{R}_+$ be a measurable function. We consider the Lévy-driven continuous-time linear (or moving average) stochastic process
\[ \varepsilon(t) = \int_{\mathbb{R}} \hat{a}(t - s) dL(s), \quad t \in \mathbb{R}. \quad (4) \]

For causal process (4) $\hat{a}(t) = 0$, $t < 0$.

In the sequel we assume that
\[ \hat{a} \in L_1(\mathbb{R}) \cap L_2(\mathbb{R}) \text{ or } \hat{a} \in L_2(\mathbb{R}) \text{ with } \mathbb{E} L(1) = 0. \quad (5) \]

Under the condition (5) and
\[ \int_{\mathbb{R}} u^2 \Pi(du) < \infty, \]
the stochastic integral in (4) is well-defined in $L_2(\Omega)$ in the sense of stochastic integration introduced in Rajput and Rosinski (1989).

The popular choices for the kernel in (4) are Gamma type kernels:

- $\hat{a}(t) = t^\alpha e^{-t \lambda I_{[0, \infty)}(t)}$, $\lambda > 0, \alpha > -\frac{1}{2}$;
- $\hat{a}(t) = e^{-\lambda t I_{[0, \infty)}(t)}$, $\lambda > 0$ (Ornstein-Uhlenbeck process);
\[ \hat{a}(t) = e^{-\lambda |t|}, \lambda > 0 \text{ (well-balanced Ornstein-Uhlenbeck process)}. \]

\textbf{A}_1. The process \( \varepsilon \) in (1) is a measurable causal linear process of the form (4), where a two-sides Lévy process \( L \) is such that \( E\, L(1) = 0, \hat{a} \in L_1(\mathbb{R}) \cap L_2(\mathbb{R}) \). Moreover the Lévy measure \( \Pi \) of \( L(1) \) satisfies (3) for some \( p > 0 \).

From the condition \( \text{A}_1 \) it follows Anh et al. (2002) for any \( r \geq 1 \)
\[
\log E \exp \left\{ i \sum_{j=1}^{r} z_j \varepsilon(t_j) \right\} = \int_{\mathbb{R}} \kappa \left( \sum_{j=1}^{r} z_j \hat{a}(t_j - s) \right) ds. \tag{6}
\]

In turn from (6) it can be seen that the stochastic process \( \varepsilon \) is stationary in a strict sense.

Denote by
\[
m_r(t_1, \ldots, t_r) = E \varepsilon(t_1) \cdots \varepsilon(t_r),
c_r(t_1, \ldots, t_r) = i^{-r} \frac{\partial^r}{\partial z_1 \cdots \partial z_r} \log E \exp \left\{ i \sum_{j=1}^{r} z_j \varepsilon(t_j) \right\} \bigg|_{z_1=\cdots=z_r=0}
\]

the moment and cumulant functions correspondingly of order \( r, \ r \geq 1 \), of the process \( \varepsilon \).

Thus \( m_2(t_1, t_2) = B(t_1 - t_2), \) where
\[
B(t) = d_2 \int_{\mathbb{R}} \hat{a}(t + s)\hat{a}(s) ds, \ t \in \mathbb{R},
\]
is a covariance function of \( \varepsilon \), and the fourth moment function
\[
m_4(t_1, t_2, t_3, t_4) = c_4(t_1, t_2, t_3, t_4) + m_2(t_1, t_2)m_2(t_3, t_4) + m_2(t_1, t_3)m_2(t_2, t_4) + m_2(t_1, t_4)m_2(t_2, t_3). \tag{7}
\]

The explicit expression for cumulants of the stochastic process \( \varepsilon \) can be obtained from (6) by direct calculations:
\[
c_r(t_1, \ldots, t_r) = d_r \int_{\mathbb{R}} \prod_{j=1}^{r} \hat{a}(t_j - s) \ ds, \tag{8}
\]
where \( d_r \) is the \( r \)th cumulant of the random variable \( L(1) \). In particular,
\[
d_2 = E\, L^2(1) = -\kappa^{(2)}(0), \quad d_4 = E\, L^4(1) - 3 \left( E\, L^2(1) \right)^2.
\]

Under the condition \( \text{A}_1 \), the spectral densities of the stationary process \( \varepsilon \) of all orders exist and can be obtained from (8) as
\[
f_r(\lambda_1, \ldots, \lambda_{r-1}) = (2\pi)^{-r+1} d_r \cdot a \left( -\sum_{j=1}^{r-1} \lambda_j \right) \cdot \prod_{j=1}^{r-1} a(\lambda_j). \tag{9}
\]
where \( a \in L_2(\mathbb{R}), \ a(\lambda) = \int_{\mathbb{R}} \hat{a}(t)e^{-\lambda t} dt, \ \lambda \in \mathbb{R}, \) if complex-valued functions \( f_r \in L_1(\mathbb{R}^{r-1}), \ r > 2, \) see, e.g., Avram et al. (2010) for definitions of the spectral densities of higher order \( f_r, \ r \geq 3 \).

For \( r = 2 \), we denote the spectral density of the second order by
\[
f(\lambda) = f_2(\lambda) = (2\pi)^{-1} d_2 a(\lambda)a(-\lambda) = (2\pi)^{-1} d_2 |a(\lambda)|^2.
\]

\( \copyright \) Springer
A2. (i) Spectral densities (9) of all orders \( f_r \in L_1(\mathbb{R}^{r-1}), r \geq 2; \)
(ii) \( a(\lambda) = a(\lambda, \theta^{(1)}), d_2 = d_2(\theta^{(2)}), \theta = (\theta^{(1)}, \theta^{(2)}) \in \Theta_T, \Theta_T = \bigcup_{\|e\|<1} (\Theta + \tau e), \)
\( \tau > 0 \) is some number, \( \Theta \subset \mathbb{R}^m \) is a bounded open convex set, that is \( f(\lambda) = f(\lambda, \theta), \theta \in \Theta_T, \) and a true value of parameter \( \theta_0 \in \Theta; \)
(iii) \( f(\lambda, \theta) > 0, (\lambda, \theta) \in \mathbb{R} \times \Theta^c. \)

In the condition A2(ii) above \( \theta^{(1)} \) represents parameters of the kernel \( \hat{a} \) in (4), while \( \theta^{(2)} \) represents parameters of Lévy process.

Remark 2 The last part of the condition A1 is fully used in the proof of Lemma 5 and Theorem B.1 in "Appendix B". The condition A2(i) is fully used just in the proof of Lemma 5. When we refer to these conditions in other places of the text we use them partially: see, for example, Lemma 3, where we need in the existence of \( f_4 \) only.

Definition 1 The least squares estimator (LSE) of the parameter \( a_0 \in A \) obtained by observations of the process \( \{X(t), t \in [0, T]\} \) is said to be any random vector \( \hat{a}_T = (\hat{a}_{1T}, \ldots, \hat{a}_{qT}) \in A^c (A^c \text{ is the closure of } A), \) such that
\[
S_T (\hat{a}_T) = \min_{\alpha \in A^c} S_T (\alpha), \quad S_T (\alpha) = \int_0^T (X(t) - g(t, \alpha))^2 dt.
\]

We consider the residual periodogram
\[
I_T (\lambda, \hat{a}_T) = (2\pi T)^{-1} \left| \int_0^T (X(t) - g(t, \hat{a}_T)) e^{-it\lambda} dt \right|^2, \quad \lambda \in \mathbb{R},
\]
and the Whittle contrast field
\[
U_T (\theta, \hat{a}_T) = \int_{\mathbb{R}} \left( \log f(\lambda, \theta) + \frac{I_T (\lambda, \hat{a}_T)}{f(\lambda, \theta)} \right) w(\lambda) d\lambda, \quad \theta \in \Theta^c, \quad (10)
\]
where \( w(\lambda), \lambda \in \mathbb{R}, \) is an even nonnegative bounded Lebesgue measurable function, for which the integral (10) is well-defined. The existence of integral (10) follows from the condition C4 introduced below.

Definition 2 The minimum contrast estimator (MCE) of the unknown parameter \( \theta_0 \in \Theta \) is said to be any random vector \( \hat{\theta}_T = (\hat{\theta}_{1T}, \ldots, \hat{\theta}_{mT}) \) such that
\[
U_T (\hat{\theta}_T, \hat{a}_T) = \min_{\theta \in \Theta^c} U_T (\theta, \hat{a}_T).
\]
The minimum in the Definition 2 is attained due to integral (10) continuity in \( \theta \in \Theta^c \) as follows from the condition C4 introduced below.

3 Consistency of the minimum contrast estimator

Suppose the function \( g(t, \alpha) \) in (1) is continuously differentiable with respect to \( \alpha \in A^c \) for any \( t \geq 0, \) and its derivatives \( g_i(t, \alpha) = \frac{\partial}{\partial \alpha_i} g(t, \alpha), i = 1, q, \) are locally integrable with respect to \( t. \) Let
Let 
\[ d_T(\alpha) = \text{diag}\left(d_{iT}(\alpha), \ i = 1, q\right), \ d_{iT}^2(\alpha) = \int_0^T g_i^2(t, \alpha)dt, \]

and \( \lim inf_{T \to \infty} T^{-\frac{1}{2}}d_{iT}(\alpha) > 0, \ i = 1, q, \alpha \in \mathcal{A}. \)

Set
\[ \Phi_T(\alpha_1, \alpha_2) = \int_0^T (g(t, \alpha_1) - g(t, \alpha_2))^2dt, \alpha_1, \alpha_2 \in \mathcal{A}^c. \]

We assume that the following conditions are satisfied.

C1. The LSE \( \hat{\alpha}_T \) is a weakly consistent estimator of \( \alpha_0 \in \mathcal{A} \) in the sense that
\[ T^{-\frac{1}{2}}d_T(\alpha_0)(\hat{\alpha}_T - \alpha_0) \xrightarrow{P} 0, \ \text{as} \ T \to \infty. \]

C2. There exists a constant \( c_0 < \infty \) such that for any \( \alpha_0 \in \mathcal{A} \) and \( T > T_0 \), where \( c_0 \) and \( T_0 \) may depend on \( \alpha_0 \),
\[ \Phi_T(\alpha, \alpha_0) \leq c_0\|d_T(\alpha_0)(\alpha - \alpha_0)\|^2, \alpha \in \mathcal{A}^c. \]

The fulfillment of the conditions C1 and C2 is discussed in more detail in “Appendix A”. We need also in 3 more conditions.

C3. \( f(\lambda, \theta_1) \neq f(\lambda, \theta_2) \) on a set of positive Lebesgue measure once \( \theta_1 \neq \theta_2, \theta_1, \theta_2 \in \Theta^c. \)

C4. The functions \( w(\lambda) \log f(\lambda, \theta), \frac{w(\lambda)}{f(\lambda, \theta)} \) are continuous with respect to \( \theta \in \Theta^c \) almost everywhere in \( \lambda \in \mathbb{R} \), and
\( \text{(i)} \ \ w(\lambda) |\log f(\lambda, \theta)| \leq Z_1(\lambda), \theta \in \Theta^c, \text{almost everywhere in} \ \lambda \in \mathbb{R}, \text{and} \ Z_1(\cdot) \in L_1(\mathbb{R}); \)
\( \text{(ii)} \ \sup_{\lambda \in \mathbb{R}, \theta \in \Theta^c} \frac{w(\lambda)}{f(\lambda, \theta)} = c_1 < \infty. \)

C5. There exists an even positive Lebesgue measurable function \( v(\lambda), \lambda \in \mathbb{R} \), such that
\( \text{(i)} \ \frac{v(\lambda)}{f(\lambda, \theta)} \text{ is uniformly continuous in} (\lambda, \theta) \in \mathbb{R} \times \Theta^c; \)
\( \text{(ii)} \ \sup_{\lambda \in \mathbb{R}} \frac{w(\lambda)}{v(\lambda)} < \infty. \)

Theorem 1 Under conditions \( A_1, A_2, C_1–C_5 \) \( \hat{\theta}_T \xrightarrow{P} \theta, \ \text{as} \ T \to \infty. \)

To prove the theorem we need some additional assertions.

Lemma 1 Under condition \( A_1 \)
\[ v_T^* = T^{-1} \int_0^T \varepsilon^2(t)dt \xrightarrow{P} B(0), \ \text{as} \ T \to \infty. \]
**Proof** For any $\rho > 0$ by Chebyshev inequality and (7)

\[
P \left\{ \left| v_T^n - B(0) \right| \geq \rho \right\} \leq \rho^{-2} T^{-2} \int_0^T \int_0^T c_4(t, t, s, s) \, dt \, ds + 2 \rho^{-2} T^{-2} \int_0^T \int_0^T B^2(t - s) \, dt \, ds = I_1 + I_2.
\]

From $A_1$ it follows that $I_2 = O(T^{-1})$. Using expression (8) for cumulants of the process $\varepsilon$ we get

\[
I_1 = d_4 \rho^{-2} T^{-2} \int_0^T \int_0^R a^2(t - u) a^2(s - u) \, du \, ds \, dt \leq d_4 \rho^{-2} \|a\|_2^4 T^{-1},
\]

where $\|a\|_2 = \left( \int_R a^2(u) \, du \right)^{1/2}$, that is $I_1 = O(T^{-1})$ as well. \hfill \Box

Let

\[
F_T^{(k)}(u_1, \ldots, u_k) = \frac{1}{(2\pi)^{(k-1)}T^{k-1}} \int_{[0,T]^k} e^{\frac{i}{T} \sum_{j=1}^k t_j u_j} \, dt_1 \ldots dt_k
\]

with $u_k = -(u_1 + \ldots + u_{k-1})$, $u_j \in \mathbb{R}$, $j = 1, \ldots, k$.

The functions $F_T^{(k)}(u_1, \ldots, u_k)$, $k \geq 3$, are multidimensional analogues of the Fejér kernel, for $k = 2$ we obtain the usual Fejér kernel.

The next statement bases on the results by Bentkus (1972a, b), Bentkus and Rutkauskas (1973).

**Lemma 2** Let function $G(u_1, \ldots, u_k)$, $u_k = -(u_1 + \ldots + u_{k-1})$ be bounded and continuous at the point $(u_1, \ldots, u_{k-1}) = (0, \ldots, 0)$. Then

\[
\lim_{T \to \infty} \int_{\mathbb{R}^{k-1}} F_T^{(k)}(u_1, \ldots, u_{k-1}) G(u_1, \ldots, u_k) \, du_1 \ldots du_{k-1} = G(0, \ldots, 0).
\]

We set

\[
g_T(\lambda, \alpha) = \int_0^T e^{-i\lambda t} g(t, \alpha) \, dt, \quad s_T(\lambda, \alpha) = g_T(\lambda, \alpha_0) - g_T(\lambda, \alpha),
\]

\[
\varepsilon_T(\lambda) = \int_0^T e^{-i\lambda t} \varepsilon(t) \, dt, \quad I_T^2(\lambda) = (2\pi T)^{-1} |\varepsilon_T(\lambda)|^2,
\]
and write the residual periodogram in the form
\[
I_T(\lambda, \hat{\alpha}_T) = I_T^T(\lambda) + (\pi T)^{-1} \text{Re} \left\{ \varepsilon_T(\lambda)s_T(\lambda, \hat{\alpha}_T) \right\} + (2\pi T)^{-1} |s_T(\lambda, \hat{\alpha}_T)|^2.
\]

Let \( \varphi = \varphi(\lambda, \theta), (\lambda, \theta) \in \mathbb{R} \times \Theta^c \), be an even Lebesgue measurable with respect to variable \( \lambda \) for each fixed \( \theta \) weight function. We have
\[
J_T(\varphi, \hat{\alpha}_T) = \int_{\mathbb{R}} I_T(\lambda, \hat{\alpha}_T) \varphi(\lambda, \theta) d\lambda = \int_{\mathbb{R}} I_T^T(\lambda) \varphi(\lambda, \theta) d\lambda
\]
\[
+ (\pi T)^{-1} \int_{\mathbb{R}} \text{Re} \left\{ \varepsilon_T(\lambda)s_T(\lambda, \hat{\alpha}_T) \right\} \varphi(\lambda, \theta) d\lambda
\]
\[
+ (2\pi T)^{-1} \int_{\mathbb{R}} |s_T(\lambda, \hat{\alpha}_T)|^2 \varphi(\lambda, \theta) d\lambda
\]
\[
= J_T^T(\varphi) + J_T^{(1)}(\varphi) + J_T^{(2)}(\varphi).
\]

Suppose
\[
\varphi(\lambda, \theta) \geq 0, \quad \sup_{\lambda \in \mathbb{R}, \theta \in \Theta^c} \varphi(\lambda, \theta) = c(\varphi) < \infty. \quad (11)
\]

Then by the Plancherel identity and condition C2
\[
\left| J_T^{(1)}(\varphi) \right| \leq 2c(\varphi) \left( (2\pi T)^{-1} \int_{\mathbb{R}} |\varepsilon_T(\lambda)|^2 d\lambda \right)^{\frac{1}{2}} \left( (2\pi T)^{-1} \int_{\mathbb{R}} |s_T(\lambda, \hat{\alpha}_T)|^2 d\lambda \right)^{\frac{1}{2}}
\]
\[
= 2c(\varphi) \left( v_T^s \right)^{\frac{1}{2}} T^{-\frac{1}{2}} \left( \Phi_T(\hat{\alpha}_T, \alpha_0) \right)^{\frac{1}{2}} \leq 2c_0 \left( v_T^s \right)^{\frac{1}{2}} T^{-\frac{1}{2}} \|d_T(\alpha_0) (\hat{\alpha}_T - \alpha_0)\|.
\]

Taking into account conditions A1, C1, C2 and the result of Lemma 1 we obtain
\[
\sup_{\theta \in \Theta^c} \left| J_T^{(1)}(\varphi) \right| \xrightarrow{P} 0, \quad \text{as } T \to \infty. \quad (12)
\]

On the other hand
\[
J_T^{(2)}(\varphi) \leq c(\varphi) T^{-1} \Phi_T(\alpha_0, \hat{\alpha}_T) \leq c_0 c(\varphi) \left\| T^{-\frac{1}{2}} d_T(\alpha_0) (\hat{\alpha}_T - \alpha_0) \right\|^2,
\]
and again, thanks to C1, C2,
\[
\sup_{\theta \in \Theta^c} J_T^{(2)}(\varphi) \xrightarrow{P} 0, \quad \text{as } T \to \infty. \quad (13)
\]

**Lemma 3** Suppose conditions A1, A2 are fulfilled and the weight function \( \varphi(\lambda, \theta) \) introduced above satisfies (11). Then, as \( T \to \infty, \)
\[
J_T^\varepsilon(\varphi) \xrightarrow{P} J(\varphi) = \int_{\mathbb{R}} f(\lambda, \theta_0) \varphi(\lambda, \theta) d\lambda, \quad \theta \in \Theta^c.
\]

**Proof** The lemma in fact is an application of Lemma 2 in Anh et al. (2002) and Theorem 1 in Anh et al. (2004) reasoning to linear process (4). It is sufficient to prove
\[
(1) \ E J_T^\varepsilon(\varphi) \to J(\varphi); \quad (2) \ J_T^\varepsilon(\varphi) - E J_T^\varepsilon(\varphi) \xrightarrow{P} 0.
\]
To apply Lemma 2 we have to show that the functions $G_2(u)$, $u \in \mathbb{R}$; $G_4^{(1)}(u)$, $G_4^{(2)}(u)$, $u = (u_1, u_2, u_3) \in \mathbb{R}^3$, are bounded and continuous at origins.

Boundedness of $G_2$ follows from (11). Thanks to (11)

$$\sup_{u \in \mathbb{R}^3} |G_4^{(1)}(u)| \leq c^2(\varphi) \|f\|_2^2 < \infty, \quad \|f\|_2 = \left( \int_{\mathbb{R}} f^2(\lambda, \theta_0) d\lambda \right)^{\frac{1}{2}}.$$

On the other hand, by (9)

$$|G_4^{(2)}(u_1, u_2, u_3)| \leq d_4(2\pi)^{-3} \int_{\mathbb{R}} |a(\lambda + u_1)a(-\lambda + u_2)\varphi(\lambda) d\lambda \times$$

$$\times \int_{\mathbb{R}} |a(\mu + u_3)a(-\mu - u_1 - u_2 - u_3)\varphi(\mu) d\mu$$

$$= d_4 \cdot (2\pi)^{-3} \cdot I_3 \cdot I_4,$$

$$I_3 \leq 2\pi c(\varphi) d_2^{-1} \int_{\mathbb{R}} f(\lambda, \theta_0) d\lambda = 2\pi c(\varphi) d_2^{-1} B(0).$$

Integral $I_4$ admits the same upper bound. So,

$$\sup_{u \in \mathbb{R}^3} |G_4^{(2)}(u)| \leq (2\pi)^{-1} \gamma_2 c^2(\varphi) B^2(0),$$

where $\gamma_2 = \frac{d_4}{d_2^2} > 0$ is the excess of $L(1)$ distribution, and functions $G_2$, $G_4^{(1)}$, $G_4^{(2)}$ are bounded. The continuity at origins of these functions follows from conditions of Lemma 3 as well.

**Corollary 1** If $\varphi(\lambda, \theta) = \frac{w(\lambda)}{f(\lambda, \theta)}$, then under conditions $A_1, A_2, C_1, C_2$ and $C_4$

$$U_T(\theta, \widehat{\alpha}_T) \xrightarrow{p} U(\theta) \xrightarrow{} \int_{\mathbb{R}} \left( \log f(\lambda, \theta) + \frac{f(\lambda, \theta_0)}{f(\lambda, \theta)} \right) w(\lambda) d\lambda, \theta \in \Theta^c.$$
Consider the Whittle contrast function

\[ K(\theta_0, \theta) = U(\theta) - U(\theta_0) = \int_{\mathbb{R}} \left( \frac{f(\lambda, \theta_0)}{f(\lambda, \theta)} - 1 - \log \frac{f(\lambda, \theta_0)}{f(\lambda, \theta)} \right) w(\lambda) d\lambda \geq 0, \]

with \( K(\theta_0, \theta) = 0 \) if and only if \( \theta = \theta_0 \) due to \( C_3 \).

**Lemma 4** If the coditions \( A_1, A_2, C_1, C_2, C_4 \) and \( C_5 \) are satisfied, then

\[ \sup_{\theta \in \Theta} |U_T(\theta, \hat{\alpha}_T) - U(\theta)| \xrightarrow{P} 0, \text{ as } T \to \infty. \]

**Proof** Let \( \{\theta_j, j = 1, N \} \) be a \( \delta \)-net of the set \( \Theta \). Then

\[
\sup_{\theta \in \Theta} |U_T(\theta, \hat{\alpha}_T) - U(\theta)| \leq \\
\leq \sup_{||\theta_1 - \theta_2|| \leq \delta} |U_T(\theta_1, \hat{\alpha}_T) - U(\theta_1) - (U_T(\theta_2, \hat{\alpha}_T) - U(\theta_2))| \\
+ \max_{1 \leq j \leq N} |U_T(\theta_j, \hat{\alpha}_T) - U(\theta_j)|,
\]

and for any \( \rho \geq 0 \)

\[
P_2 = P \left\{ \max_{1 \leq j \leq N} |U_T(\theta_j, \hat{\alpha}_T) - U(\theta_j)| \geq \frac{\rho}{2} \right\} \to 0, \text{ as } T \to \infty.
\]

by Corollary 1. On the other hand,

\[
P_1 = P \left\{ \sup_{||\theta_1 - \theta_2|| \leq \delta} \left| U_T(\theta_1, \hat{\alpha}_T) - U(\theta_1) - (U_T(\theta_2, \hat{\alpha}_T) - U(\theta_2)) \right| \geq \frac{\rho}{2} \right\}
\]

\[
\leq P \left\{ \sup_{||\theta_1 - \theta_2|| \leq \delta} \left| \int_{\mathbb{R}} I_T^1(\lambda) \left( \frac{w(\lambda)}{f(\lambda, \theta_1)} - \frac{w(\lambda)}{f(\lambda, \theta_2)} \right) d\lambda \right| \\
+ \sup_{||\theta_1 - \theta_2|| \leq \delta} \left| \int_{\mathbb{R}} f(\lambda, \theta_0) \left( \frac{w(\lambda)}{f(\lambda, \theta_1)} - \frac{w(\lambda)}{f(\lambda, \theta_2)} \right) d\lambda \right| \\
+ 2 \sup_{\theta \in \Theta} J_T^{(1)} \left( \frac{w}{f} \right) + 2 \sup_{\theta \in \Theta} J_T^{(2)} \left( \frac{w}{f} \right) \geq \frac{\rho}{2} \right\}.
\]

By the condition \( C_5(i) \)

\[
\sup_{||\theta_1 - \theta_2|| \leq \delta} \left| \int_{\mathbb{R}} I_T^1(\lambda) \left( \frac{w(\lambda)}{f(\lambda, \theta_1)} - \frac{w(\lambda)}{f(\lambda, \theta_2)} \right) d\lambda \right| \leq \eta(\delta) \int_{\mathbb{R}} I_T^1(\lambda) \frac{w(\lambda)}{v(\lambda)} d\lambda,
\]

where

\[
\eta(\delta) = \sup_{\lambda \in \mathbb{R}, ||\theta_1 - \theta_2|| \leq \delta} \left| \frac{v(\lambda)}{f(\lambda, \theta_1)} - \frac{v(\lambda)}{f(\lambda, \theta_2)} \right| \to 0, \; \delta \to 0.
\]
Since by Lemma 3 and the condition \( C_5(\text{ii}) \)
\[
\int_{\mathbb{R}} I_T^r(\lambda) \frac{w(\lambda)}{v(\lambda)} d\lambda \xrightarrow{p} \int_{\mathbb{R}} f(\lambda, \theta_0) \frac{w(\lambda)}{v(\lambda)} d\lambda, \quad \text{as } T \to \infty,
\]
and the 2nd term under the probability sign in (14) by choosing \( \delta \) can be made arbitrary small, then \( P_1 \to 0 \), as \( T \to 0 \), taking into account that the 3rd and the 4th terms converge to zero in probability, thanks to (12) and (13), if \( \varphi = \frac{w}{f} \).

**Proof of Theorem 1** By Definition 2 for any \( \rho > 0 \)
\[
P \{ \| \hat{\theta}_T - \theta_0 \| \geq \rho \} = P \{ \| \hat{\theta}_T - \theta_0 \| \geq \rho; U_T(\hat{\theta}_T, \hat{\alpha}_T) \leq U_T(\theta_0, \hat{\alpha}_T) \}
\leq P \left\{ \inf_{\| \theta - \theta_0 \| \geq \rho} (U_T(\theta, \hat{\alpha}_T) - U_T(\theta_0, \hat{\alpha}_T)) \leq 0 \right\}
= P \left\{ \inf_{\| \theta - \theta_0 \| \geq \rho} \left[ U_T(\theta, \hat{\alpha}_T) - U(\theta) - (U_T(\theta_0, \hat{\alpha}_T) - U(\theta_0)) + K(\theta_0, \theta) \right] \leq 0 \right\}
\leq P \left\{ \inf_{\| \theta - \theta_0 \| \geq \rho} \left[ U_T(\theta, \hat{\alpha}_T) - U(\theta) - (U_T(\theta_0, \hat{\alpha}_T) - U(\theta_0)) \right] + \inf_{\| \theta - \theta_0 \| \geq \rho} K(\theta_0, \theta) \leq 0 \right\}
\leq P \left\{ \sup_{\theta \in \mathcal{A}} |U_T(\theta, \hat{\alpha}_T) - U(\theta)| + |U_T(\theta_0, \hat{\alpha}_T) - U(\theta)| \geq \inf_{\| \theta - \theta_0 \| \geq \rho} K(\theta_0, \theta) \right\} \to 0,
\]
when \( T \to \infty \) due to Lemma 4 and the property of the contrast function \( K \).

### 4 Asymptotic normality of minimum contrast estimator

The first three conditions relate to properties of the regression function \( g(t, \alpha) \) and the LSE \( \hat{\alpha}_T \). They are commented in “Appendix B”.

**N1.** The normed LSE \( d_T(a_0) (\hat{\alpha}_T - a_0) \) is asymptotically, as \( T \to \infty \), normal \( N(0, \Sigma_{\text{LSE}}) \), \( \Sigma_{\text{LSE}} = (\Sigma_{\text{LSE}}^{ij})_i, j = 1 \).

Let us
\[
g'(t, \alpha) = \frac{\partial}{\partial t} g(t, \alpha); \quad \Phi_T'(\alpha_1, \alpha_2) = \int_0^T (g'(t, \alpha_1) - g'(t, \alpha_2))^2 dt, \quad \alpha_1, \alpha_2 \in \mathcal{A}^c.
\]

**N2.** The function \( g(t, \alpha) \) is continuously differentiable with respect to \( t \geq 0 \) for any \( \alpha \in \mathcal{A}^c \) and for any \( a_0 \in \mathcal{A} \), and \( T > T_0 \) there exists a constant \( c' \) (\( T_0 \) and \( c' \) may depend on \( a_0 \)) such that
\[
\Phi_T'(\alpha_1, a_0) \leq c' \| d_T(a_0)(\alpha - a_0) \|_2^2, \quad \alpha \in \mathcal{A}^c.
\]

Let
\[
g_{il}(t, \alpha) = \frac{\partial^2}{\partial \alpha_i \partial \alpha_l} g(t, \alpha), \quad d_{il, T}^2(\alpha) = \int_0^T g_{il}^2(t, \alpha) dt, \quad i, l = 1, q,
\]
\[
v(r) = \{ x \in \mathbb{R}^q : \| x \| < r \}, \quad r > 0.
\]
N3. The function $g(t, \alpha)$ is twice continuously differentiable with respect to $\alpha \in \mathcal{A}^c$ for any $t \geq 0$, and for any $R \geq 0$ and all sufficiently large $T$ ($T > T_0(R)$)

(i) $d_{T}^{-1}(\alpha_0) \sup_{t \in [0, T], u \in v^{c}(R)} |g_i \left( t, \alpha_0 + d_{T}^{-1}(\alpha_0)u \right)| \leq c^i(R)T^{-\frac{1}{2}}, i = 1, q$;

(ii) $d_{il,T}(\alpha_0) \sup_{t \in [0, T], u \in v^{c}(R)} |g_{il} \left( t, \alpha_0 + d_{T}^{-1}(\alpha_0)u \right)| \leq c^{il}(R)T^{-\frac{1}{2}}, i, l = 1, q$;

(iii) $d_{T}^{-1}(\alpha_0)d_{il,T}^{-1}(\alpha_0)d_{il,T}(\alpha_0) \leq c^{il}T^{-\frac{1}{2}}, i, l = 1, q$,

with positive constants $c^i, c^{il}, c^{il}$, possibly, depending on $\alpha_0$.

We assume also that the function $f(\lambda, \theta)$ is twice differentiable with respect to $\theta \in \Theta^c$ for any $\lambda \in \mathbb{R}$.

Set

$$f_i(\lambda, \theta) = \frac{\partial}{\partial \theta_i} f(\lambda, \theta), \quad f_{ij}(\lambda, \theta) = \frac{\partial^2}{\partial \theta_i \partial \theta_j} f(\lambda, \theta),$$

and introduce the following conditions.

N4. (i) For any $\theta \in \Theta^c$ the functions $\phi_i(\lambda) = \frac{f_i(\lambda, \theta)}{f^2(\lambda, \theta)} w(\lambda), \lambda \in \mathbb{R}, i = 1, m$, possess the following properties:

1. $\phi_i \in L_\infty(\mathbb{R}) \cap L_1(\mathbb{R});$

2. $\sup_{\eta \to 1} \varphi_i < \infty;$

3. $\lim_{\eta \to 1} \sup_{\lambda \in \mathbb{R}} |\phi_i(\eta \lambda) - \phi_i(\lambda)| = 0 ;$

4. $\phi_i$ are differentiable and $\phi_i'$ are uniformly continuous on $\mathbb{R}$.

(ii) $\frac{|f_i(\lambda, \theta)|}{f(\lambda, \theta)} w(\lambda) \leq Z_2(\lambda), \theta \in \Theta, i = 1, m$, almost everywhere in $\lambda \in \mathbb{R}$ and $Z_2(\cdot) \in L_1(\mathbb{R})$.

(iii) The functions $\frac{f_i(\lambda, \theta)f_j(\lambda, \theta)}{f^2(\lambda, \theta)} w(\lambda), \frac{f_{ij}(\lambda, \theta)}{f^2(\lambda, \theta)} w(\lambda)$ are continuous with respect to $\theta \in \Theta^c$ for each $\lambda \in \mathbb{R}$ and

$$\frac{f^2_i(\lambda, \theta)}{f^2(\lambda, \theta)} w(\lambda) + \frac{|f_{ij}(\lambda, \theta)|}{f(\lambda, \theta)} w(\lambda) \leq a_{ij}(\lambda), \lambda \in \mathbb{R}, \theta \in \Theta^c,$$

where $a_{ij}(\cdot) \in L_1(\mathbb{R}), i, j = 1, m$.

N5. (i) $\frac{f^2_i(\lambda, \theta)}{f^3(\lambda, \theta)} w(\lambda), \frac{f_{ij}(\lambda, \theta)}{f^2(\lambda, \theta)} w(\lambda), i, j = 1, m$, are bounded functions in $(\lambda, \theta) \in \mathbb{R} \times \Theta^c$;

(ii) There exists an even positive Lebesgue measurable function $v(\lambda), \lambda \in \mathbb{R}$, such that the functions $\frac{f_i(\lambda, \theta)f_j(\lambda, \theta)}{f^3(\lambda, \theta)} v(\lambda), \frac{f_{ij}(\lambda, \theta)}{f^2(\lambda, \theta)} v(\lambda), i, j = 1, m$, are uniformly continuous in $(\lambda, \theta) \in \mathbb{R} \times \Theta^c$;

(iii) $\sup_{\lambda \in \mathbb{R}} \frac{w(\lambda)}{v(\lambda)} < \infty$.

Conditions N5(iii) and C5(ii) look the same, however the function $v$ in these conditions must satisfy different conditions N5(ii) and C5(i), and therefore, generally speaking, the functions $v$ in these two conditions can be different.
Statistical Inference for Stochastic Processes

The next three matrices appear in the formulation of Theorem 2:

\[ W_1(\theta) = \int_{\mathbb{R}} \nabla_\theta \log f(\lambda, \theta) \nabla'_\theta \log f(\lambda, \theta) w(\lambda) d\lambda, \]

\[ W_2(\theta) = 4\pi \int_{\mathbb{R}} \nabla_\theta \log f(\lambda, \theta) \nabla'_\theta \log f(\lambda, \theta) w^2(\lambda) d\lambda, \]

\[ V(\theta) = \gamma_2 \int_{\mathbb{R}} \nabla_\theta \log f(\lambda, \theta) w(\lambda) d\lambda \int_{\mathbb{R}} \nabla'_\theta \log f(\lambda, \theta) w(\lambda) d\lambda, \]

where \( \gamma_2 = \frac{d_4}{d_2^2} > 0 \) is the excess of the random variable \( L(1) \), \( \nabla_\theta \) is a column vector-gradient, and \( \nabla'_\theta \) is a row vector-gradient.

Theorem 2 Under conditions \( A_1, A_2, C_1–C_5 \) and \( N_1–N_6 \) the normed MCE \( T^{\frac{1}{2}}(\hat{\theta}_T - \theta_0) \) is asymptotically, as \( T \to \infty \), normal with zero mean and covariance matrix

\[ W(\theta) = W_1^{-1}(\theta_0) (W_2(\theta_0) + V(\theta_0)) W_1^{-1}(\theta_0). \] (15)

The proof of the theorem is preceded by several lemmas. The next statement is Theorem 5.1 Avram et al. (2010) formulated in a form convenient to us.

Lemma 5 Let the stochastic process \( \varepsilon \) satisfies \( A_1, A_2 \), spectral density \( f \in L_p(\mathbb{R}) \), a function \( b \in L_q(\mathbb{R}) \cap L_1(\mathbb{R}) \), where \( \frac{1}{p} + \frac{1}{q} = \frac{1}{2} \). Let

\[ \hat{b}(t) = \int_{\mathbb{R}} e^{i\lambda t} b(\lambda) d\lambda. \] (16)

and

\[ Q_T = \int_0^T \int_0^T (\varepsilon(t)\varepsilon(s) - B(t-s)) \hat{b}(t-s) dt ds. \] (17)

Then the central limit theorem holds:

\[ T^{-\frac{1}{2}} Q_T \Rightarrow N(0, \sigma^2), \quad \text{as} \quad T \to \infty, \]

where “\( \Rightarrow \)” means convergence in distributions,

\[ \sigma^2 = 16\pi^3 \int_{\mathbb{R}} b^2(\lambda) f^2(\lambda) d\lambda + \gamma_2 \left( 2\pi \int_{\mathbb{R}} b(\lambda) f(\lambda) d\lambda \right)^2, \] (18)

where \( \gamma_2 = \frac{d_4}{d_2^2} > 0 \) is the excess of the random variable \( L(1) \). In particular, the statement is true for \( p = 2 \) and \( q = \infty \).

Alternative form of Lemma 5 is given in Bai et al. (2016). We formulate their Theorem 2.1 in the form convenient to us.
Lemma 6  Let the stochastic process $\varepsilon$ be such that $\mathbb{E} L(1) = 0$, $\mathbb{E} L^4(1) < \infty$, and $Q_T$ be as in (17). Assume that $\hat{a} \in L_p(\mathbb{R}) \cap L_2(\mathbb{R})$, $\hat{b}$ is of the form (16) with even function $b \in L_1(\mathbb{R})$ and $\tilde{b} \in L_q(\mathbb{R})$ with

$$1 \leq p, q \leq 2, \frac{2}{p} + \frac{1}{q} \geq \frac{5}{2},$$

then

$$T^{-\frac{1}{2}} Q_T \Rightarrow N(0, \sigma^2), \quad \text{as } T \to \infty,$$

where $\sigma^2$ is given in (18).

Remark 3  It is important to note that conditions of Lemma 5 are given in frequency domain, while Lemma 6 employs the time domain conditions.

Theorems similar to Lemmas 5 and 6 can be found in paper by Giraitis et al. (2017), where the case of martingale-differences were considered. Overview of analogous results for different types of processes is given in the paper by Ginovyan et al. (2014).

Set

$$\Delta_T(\varphi) = T^{-\frac{1}{2}} \int_{\mathbb{R}} \varepsilon_T(\lambda) s_T(\lambda, \hat{\alpha}_T) \varphi(\lambda) d\lambda.$$

Lemma 7  Suppose the conditions $A_1, A_2, C_2, N_1-N_3$ are fulfilled, $\varphi(\lambda), \lambda \in \mathbb{R}$, is a bounded differentiable function satisfying the relation 3) of the condition $N_4(i)$, and moreover the derivative $\varphi'(\lambda), \lambda \in \mathbb{R}$, is uniformly continuous on $\mathbb{R}$. Then

$$\Delta_T(\varphi) \xrightarrow{P} 0 \quad \text{as } T \to \infty.$$

Proof  Let $B_\sigma$ be the set of all bounded entire functions on $\mathbb{R}$ of exponential type $0 \leq \sigma < \infty$ (see “Appendix C”), and $\delta > 0$ is an arbitrarily small number. Then there exists a function $\varphi_\sigma \in B_\sigma, \sigma = \sigma(\delta)$, such that

$$\sup_{\lambda \in \mathbb{R}} |\varphi(\lambda) - \varphi_\sigma(\lambda)| < \delta.$$

Let $T_n(\varphi_\sigma; \lambda) = \sum_{j=-n}^{n} c_j^{(n)} e^{j \pi \frac{\lambda}{n}}, \ n \geq 1$, be a sequence of the Levitan polynomials that corresponds to $\varphi_\sigma$. For any $\Lambda > 0$ there exists $n_0 = n_0(\delta, \Lambda)$ such that for $n > n_0$

$$\sup_{\lambda \in [-\Lambda, \Lambda]} |\varphi_\sigma - T_n(\varphi_\sigma; \lambda)| \leq \delta.$$

Write

$$\Delta_T(\varphi) = \Delta_T(\varphi - \varphi_\sigma) + \Delta_T(\varphi_\sigma - T_n) + \Delta_T(T_n),$$

$$|\Delta_T(\varphi - \varphi_\sigma)| \leq \delta T^{-\frac{1}{2}} \int_{\mathbb{R}} \left| \varepsilon_T(\lambda) s_T(\lambda, \hat{\alpha}_T) \right| d\lambda$$

$$\leq \delta T^{-\frac{1}{2}} \left( \int_{\mathbb{R}} |\varepsilon_T(\lambda)|^2 d\lambda \right)^{\frac{1}{2}} \left( \int_{\mathbb{R}} |s_T(\lambda, \hat{\alpha}_T)|^2 d\lambda \right)^{\frac{1}{2}}$$

$$= 2\pi \delta \left( v_T^* \right)^{\frac{1}{2}} \Phi_T^{\frac{1}{2}}(\hat{\alpha}_T, \alpha_0) \leq 2\pi c_0^{\frac{1}{2}} \delta \left( v_T^* \right)^{\frac{1}{2}} \|d_T(\alpha_0) (\hat{\alpha}_T - \alpha_0)\|.$$

© Springer
So, under the condition $C_2$, for any $\rho > 0$

$$P \{|\Delta_T (\varphi - \varphi_\sigma) | \geq \rho \} \leq$$

$$\leq P \left\{ \| d_T (\alpha_0) (\hat{\alpha}_T - \alpha_0) \| \geq \frac{\rho}{2 \pi c_0 \delta (B(0) + 1)^{\frac{1}{2}}} \right\} + P \{ \nu_T^* - B(0) > 1 \} = P_3 + P_4.$$  

The probability $P_4 \to 0$, as $T \to \infty$, and the probability $P_3$ under the condition $N_1$ for sufficiently large $T$ (we will write $T > T_0$) can be made less than a preassigned number by choosing $\delta > 0$ for a fixed $\rho > 0$.

As far as the function $\varphi_\sigma \in B_\sigma$ and the corresponding sequence of Levitan polynomials $T_n$ are bounded by the same constant, we obtain

$$|\Delta (\varphi_\sigma - T_n)| \leq \delta T^{-\frac{1}{2}} \int_{-\Lambda}^{\Lambda} \left| \varepsilon_T (\lambda) s_T (\lambda, \hat{\alpha}_T) \right| d\lambda$$

$$+ 2c(\varphi_\sigma) T^{-\frac{1}{2}} \int_{\mathbb{R} \setminus [-\Lambda, \Lambda]} \left| \varepsilon_T (\lambda) s_T (\lambda, \hat{\alpha}_T) \right| d\lambda = D_1 + D_2.$$  

The integral in the term $D_1$ can be majorized by an integral over $\mathbb{R}$ and bounded as earlier. We have further

$$s_T (\lambda, \hat{\alpha}_T) = (i\lambda)^{-1} \left[ e^{i\lambda T} (g(T, \alpha_0) - g(T, \hat{\alpha}_T)) - (g(0, \alpha_0) - g(0, \hat{\alpha}_T)) - s_T^* (\lambda, \hat{\alpha}_T) \right].$$

where $s_T^* (\lambda, \hat{\alpha}_T) = \int_0^T e^{-i\lambda t} (g'(t, \alpha_0) - g'(t, \hat{\alpha}_T)) dt$.

Under the Lemma conditions

$$T^{-\frac{1}{2}} \int_{\mathbb{R} \setminus [-\Lambda, \Lambda]} \left| \varepsilon_T (\lambda) s_T (\lambda, \hat{\alpha}_T) \right| d\lambda \leq T^{-\frac{1}{2}} \left( \int_{\mathbb{R} \setminus [-\Lambda, \Lambda]} \left| \varepsilon_T (\lambda) \right|^2 d\lambda \right)^{\frac{1}{2}}$$

$$\cdot \left[ 3 \int_{\mathbb{R} \setminus [-\Lambda, \Lambda]} \lambda^{-2} \left( |g(T, \alpha_0) - g(T, \hat{\alpha}_T)|^2 + |g(0, \alpha_0) - g(0, \hat{\alpha}_T)|^2 + |s_T^* (\lambda, \hat{\alpha}_T)|^2 \right) d\lambda \right]^{\frac{1}{2}}$$

$$\leq \sqrt{3} (2\pi \nu_T^*)^{\frac{1}{2}} \left( \sqrt{2} \Lambda^{-\frac{1}{2}} \left( |g(T, \hat{\alpha}_T) - g(T, \alpha_0)| + |g(0, \hat{\alpha}_T) - g(0, \alpha_0)| \right) + \right.$$

$$\left. + (2 \pi c_0)^{\frac{1}{2}} \Lambda^{-1} \| d_T (\alpha_0) (\hat{\alpha}_T - \alpha_0) \| \right).$$

Obviously,

$$g(T, \hat{\alpha}_T) - g(T, \alpha_0) = \sum_{i=1}^{q} g_i (T, \alpha_T^*, (\hat{\alpha}_T - \alpha_i_0),$$

\textcopyright Springer
\[ \alpha^*_T = \alpha_0 + \eta (\hat{\alpha}_T - \alpha_0), \eta \in (0, 1), d_T(\alpha_0) (\alpha^*_T - \alpha_0) = \eta d_T(\alpha_0) (\hat{\alpha}_T - \alpha_0), \] and for any \( \rho > 0 \) and \( i = 1, q \)

\[
P \left\{ \left| g_i(T, \alpha^*_T), (\hat{\alpha}_i T - \alpha_{i0}) \right| \geq \rho \right\} 
\leq P \left\{ \left| g_i(T, \alpha^*_T), (\hat{\alpha}_i T - \alpha_{i0}) \right| \geq \rho, \| d_T(\alpha_0) (\hat{\alpha}_T - \alpha_0) \| \leq R \right\} 
+ P \left\{ \| d_T(\alpha_0) (\hat{\alpha}_T - \alpha_0) \| > R \right\} = P_5 + P_6.
\]

By condition \( N_3(i) \) for any \( R \geq 0 \)

\[
P_5 \leq P \left\{ \left( d_{i, T}^{-1}(\alpha_0) \sup_{t \in [0, T], \| u \| \leq R} \left| g_i \left( t, \alpha_0 + d_{i, T}^{-1}(\alpha_0)u \right) \right| \right) \cdot \left( d_{i, T}^{-1}(\alpha_0) \| \hat{\alpha}_iT - \alpha_{i0} \| \right) \geq \rho \right\} 
\leq P \left\{ \Lambda^{-1} \| d_T(\alpha_0) (\hat{\alpha}_T - \alpha_0) \| \geq \rho \right\} \leq P_6
\]

and the second probability is equal to zero, if \( \Lambda > \frac{\rho}{R} \).

Thus, for any fixed \( \rho > 0 \), similarly to the probability \( P_3 \), the probability \( P_7 = P \{ D_2 \geq \rho \} \) for \( T > T_0 \) can be made less than preassigned number by the choice of the value \( \Lambda \).

Consider

\[
\Delta_T(T_n) = T^{-\frac{1}{2}} \sum_{j=-n}^{n} c_j^{(n)} \int_{\mathbb{R}} \varepsilon_T(\lambda) \bar{s}_T(\lambda, \hat{\alpha}_T) e^{ij\frac{\sigma}{\pi} \lambda} d\lambda,
\]

\[
\bar{s}_T(\lambda, \hat{\alpha}_T) e^{ij\frac{\sigma}{\pi} \lambda} = \int_{j\frac{\sigma}{\pi}}^{(j+1)\frac{\sigma}{\pi}} e^{i\lambda t} \left( g \left( t - j \frac{\sigma}{n}, \alpha_0 \right) - g \left( t - j \frac{\sigma}{n}, \hat{\alpha}_T \right) \right) dt, j = \frac{-n}{\sigma}, \frac{n}{\sigma}.
\]

It means that

\[
\Delta_T(T_n) = 2\pi \sum_{j=1}^{n} c_j^{(n)} T^{-\frac{1}{2}} \int_{\frac{j\sigma}{n}}^{\frac{(j+1)\sigma}{n}} \varepsilon(t) \left( g \left( t - j \frac{\sigma}{n}, \alpha_0 \right) - g \left( t - j \frac{\sigma}{n}, \hat{\alpha}_T \right) \right) dt 
+ 2\pi \sum_{j=-n}^{0} c_j^{(n)} T^{-\frac{1}{2}} \int_{0}^{\frac{(j+1)\sigma}{n}} \varepsilon(t) \left( g \left( t - j \frac{\sigma}{n}, \alpha_0 \right) - g \left( t - j \frac{\sigma}{n}, \hat{\alpha}_T \right) \right) dt.
\]
For $j > 0$ consider the value

$$T^{-\frac{1}{2}} \int_{\frac{j}{n}}^{T} \varepsilon(t) \left( g \left( t - j \frac{\sigma}{n}, \hat{\alpha}_T \right) - g \left( t - j \frac{\sigma}{n}, \alpha_0 \right) \right) dt$$

$$= \sum_{i=1}^{q} \left( T^{-\frac{1}{2}} d_{iT}^{-1}(\alpha_0) \right) \int_{\frac{j}{n}}^{T} \varepsilon(t) g_i \left( t - j \frac{\sigma}{n}, \alpha_0 \right) dt$$

$$+ \frac{1}{2} \sum_{i,k=1}^{q} \left( T^{-\frac{1}{2}} d_{iT}^{-1}(\alpha_0) \right) \int_{\frac{j}{n}}^{T} \varepsilon(t) g_{ik} \left( t - j \frac{\sigma}{n}, \alpha_0 \right) dt$$

$$= S_{1T} + \frac{1}{2} S_{2T},$$

$$\alpha_T^* = \alpha_0 + \tilde{\eta} (\hat{\alpha}_T - \alpha_0), \tilde{\eta} \in (0, 1).$$

Note that for $i = 1, q$

$$d_{iT}(\alpha_0) (\hat{\alpha}_T - \alpha_0) \Rightarrow N(0, \Sigma_{LSE}^{ii}), \text{ as } T \to \infty,$$

by the condition $N_1$. Moreover,

$$E \left( T^{-\frac{1}{2}} d_{iT}^{-1}(\alpha_0) \int_{\frac{j}{n}}^{T} \varepsilon(t) g_i \left( t - j \frac{\sigma}{n}, \alpha_0 \right) dt \right)^2$$

$$= T^{-1} d_{iT}^{-2}(\alpha_0) \int_{\frac{j}{n}}^{T} \int_{\frac{j}{n}}^{T} B(t-s) g_i \left( t - j \frac{\sigma}{n}, \alpha_0 \right) g_i \left( s - j \frac{\sigma}{n}, \alpha_0 \right) dt ds$$

$$\leq \left( T^{-2} \int_{0}^{T} \int_{0}^{T} B^2(t-s) dt ds \right)^{\frac{1}{2}} = O \left( T^{-\frac{1}{2}} \right),$$

since

$$T^{-1} \int_{0}^{T} \int_{0}^{T} B^2(t-s) dt ds \to 2\pi \|f\|_2^2, \text{ as } T \to \infty.$$
Under condition \(|d_T(\alpha_0) (\widehat{\alpha}_T - \alpha_0)| \leq R\) using assumptions \(N_3(ii)\) and \(N_3(iii)\) we get as in the estimation of the probability \(P_5\)

\[
|S_{2T}^{ik}| \leq \left( T^{-\frac{1}{2}} \int_0^T |\varepsilon(t)| dt \right) \cdot \left( d^{-1}_{i,k,T}(\alpha_0) \sup_{t \in [0,T], u \in \sigma(R)} |g_{ik}(t, \alpha_0 + d^{-1}_{i,k}(\alpha_0)u)| \right) \cdot \left( d^{-1}_{i,T}(\alpha_0) d^{-1}_{k,T}(\alpha_0) d_{ik,T}(\alpha_0) \cdot |d_{i,T}(\alpha_0)(\widehat{\alpha}_{iT} - \alpha_{i0})| \cdot |d_{k,T}(\alpha_0)(\widehat{\alpha}_{kT} - \alpha_{k0})| \right)
\]

\[
\leq \varepsilon^{ik}(R) \widehat{\varepsilon}^{ik} T^{-\frac{3}{2}} \int_0^T |\varepsilon(t)| dt \cdot |d_{i,T}(\alpha_0)(\widehat{\alpha}_{iT} - \alpha_{i0})| \cdot |d_{k,T}(\alpha_0)(\widehat{\alpha}_{kT} - \alpha_{k0})|.
\]

By Lemma 1

\[
T^{-\frac{1}{2}} \int_0^T |\varepsilon(t)| dt \leq \frac{1}{2} T^{-\frac{1}{2}} + \frac{1}{2} T^{-\frac{1}{2}} \int_0^T \varepsilon^2(t) dt \xrightarrow{p} 0, \quad \text{as } T \to \infty.
\]

So, by condition \(N_1\) \(P_8 \to 0\), as \(T \to \infty\), that is \(S_{2T} \xrightarrow{p} 0\), as \(T \to \infty\). For \(j \leq 0\) the reasoning is similar, and

\[
\Delta_T(T_n) \xrightarrow{p} 0, \quad T \to \infty.
\]

\( \square \)

**Lemma 8** Let the function \(\varphi(\lambda, \theta)w(\lambda)\) be continuous in \(\theta \in \Theta^c\) for each fixed \(\lambda \in \mathbb{R}\) with

\[
|\varphi(\lambda, \theta)| \leq \varphi(\lambda), \quad \theta \in \Theta^c, \quad \text{and } \varphi(\cdot)w(\cdot) \in L_1(\mathbb{R}).
\]

If \(\theta_T^\bullet \xrightarrow{p} \theta_0\), then

\[
I(\theta_T^\bullet) = \int_{\mathbb{R}} \varphi(\lambda, \theta_T^\bullet)w(\lambda) d\lambda \xrightarrow{p} \int_{\mathbb{R}} \varphi(\lambda, \theta_0)w(\lambda) d\lambda = I(\theta_0).
\]

**Proof** By a Lebesgue dominated convergence theorem the integral \(I(\theta), \theta \in \Theta^c\), is a continuous function. Further argument is standard. For any \(\rho > 0\) and \(\varepsilon = \frac{\rho}{2}\) we find such a \(\delta > 0\), that \(|I(\theta) - I(\theta_0)| < \varepsilon\) as \(|\theta - \theta_0| < \delta\). Then

\[
P \left\{ |I(\theta_T^\bullet) - I(\theta_0)| \geq \rho \right\} = P_9 + P_{10},
\]

where

\[
P_9 = P \left\{ |I(\theta_T^\bullet) - I(\theta_0)| \geq \frac{\rho}{2}, \quad |\theta_T^\bullet - \theta_0| < \delta \right\} = 0,
\]

due to the choice of \(\varepsilon\), and

\[
P_{10} = P \left\{ |I(\theta_T^\bullet) - I(\theta_0)| \geq \frac{\rho}{2}, \quad |\theta_T^\bullet - \theta_0| \geq \delta \right\} \to 0, \quad \text{as } T \to \infty.
\]

\( \square \)
Lemma 9 If the conditions $A_1, C_2$ are satisfied and \[ \sup_{\lambda \in \mathbb{R}, \theta \in \Theta^c} |\varphi(\lambda, \theta)| = c(\varphi) < \infty, \]
then
\[ T^{-1} \int_{\mathbb{R}} \varphi(\lambda, \theta_T^*) \delta_T(\lambda, \vartheta_T) d\lambda \xrightarrow{P} 0, \text{ as } T \to \infty. \]

Proof These relations are similar to (12), (13), and can be obtained in the same way. □

Lemma 10 Let under conditions $A_1, A_2$ there exists an even positive Lebesgue measurable function $v(\lambda), \lambda \in \mathbb{R}$, and an even Lebesgue measurable in $\lambda$ for any fixed $\theta \in \Theta^c$ function $\varphi(\lambda, \theta) = (\lambda, \theta) \in \mathbb{R} \times \Theta^c$, such that

(i) $\varphi(\lambda, \theta)v(\lambda)$ is uniformly continuous in $(\lambda, \theta) \in \mathbb{R} \times \Theta^c$;
(ii) $\sup_{\lambda \in \mathbb{R}} w(\lambda) < \infty$;
(iii) $\sup_{\lambda \in \mathbb{R}, \theta \in \Theta^c} |\varphi(\lambda, \theta)|w(\lambda) < \infty$.

Suppose also that $\theta_T^* \xrightarrow{P} \theta_0$, then, as $T \to \infty$,
\[ \int_{\mathbb{R}} I_T^c(\lambda) \varphi(\lambda, \theta_T^*) w(\lambda) d\lambda \xrightarrow{P} \int_{\mathbb{R}} f(\lambda, \theta_0) \varphi(\lambda, \theta_0) w(\lambda) d\lambda. \]

Proof We have
\[ \int_{\mathbb{R}} I_T^c(\lambda) \varphi(\lambda, \theta_T^*) w(\lambda) d\lambda = \int_{\mathbb{R}} I_T^c(\lambda) \varphi(\lambda, \theta_T^*) - \varphi(\lambda, \theta_0) \frac{w(\lambda)}{v(\lambda)} d\lambda + \int_{\mathbb{R}} I_T^c(\lambda) \varphi(\lambda, \theta_0) w(\lambda) d\lambda = I_5 + I_6. \]

By Lemma 3 and the condition (iii)
\[ I_6 \xrightarrow{P} \int_{\mathbb{R}} f(\lambda, \theta_0) \varphi(\lambda, \theta_0) w(\lambda) d\lambda, \text{ as } T \to \infty. \]

On the other hand, for any $r > 0$ under the condition (i) there exists $\delta = \delta(r)$ such that for $\|\theta_T^* - \theta_0\| < \delta$
\[ |I_5| \leq r \int_{\mathbb{R}} I_T^c \frac{w(\lambda)}{v(\lambda)} d\lambda, \]
and by the condition (ii)
\[ \int_{\mathbb{R}} I_T^c \frac{w(\lambda)}{v(\lambda)} d\lambda \xrightarrow{P} \int_{\mathbb{R}} f(\lambda, \theta_0) \frac{w(\lambda)}{v(\lambda)} d\lambda. \]

The relations (19)–(21) prove the lemma. □

Proof of Theorem 2 By definition of the MCE $\hat{\theta}_T$, formally using the Taylor formula, we get
\[ 0 = \nabla_\theta U_T(\hat{\theta}_T, \vartheta_T) = \nabla_\theta U_T(\theta_0, \vartheta_T) + \nabla_\vartheta U_T(\theta_0^*, \vartheta_T)(\hat{\theta}_T - \theta_0). \]
Since there is no vector Taylor formula, (22) must be taken coordinatewise, that is each row of vector equality (22) depends on its own random vector $\theta^*_{T,i}$, such that $\|\theta^*_{T,i} - \theta_0\| \leq \|\tilde{\theta}_T - \theta_0\|$. In turn, from (22) we have formally

$$T^{\frac{1}{2}}(\tilde{\theta}_T - \theta_0) = \left(\nabla_\theta \nabla'_\theta U_T(\theta^*_{T,i}, \tilde{\alpha}_T)\right)^{-1} \left(-T^{\frac{1}{2}}\nabla_\theta U_T(\theta_0, \tilde{\alpha}_T)\right).$$

As far as the condition $N_4$ implies the possibility of differentiation under the sign of the integrals in (10), then

$$-T^{\frac{1}{2}}\nabla_\theta U_T(\theta_0, \tilde{\alpha}_T) = -T^{\frac{1}{2}} \int_\mathbb{R} \left(\nabla_\theta \log f(\lambda, \theta_0) + \nabla_\theta \left(\frac{1}{f(\lambda, \theta_0)}\right) I_T(\lambda, \tilde{\alpha}_T)\right) w(\lambda) d\lambda$$

$$= T^{\frac{1}{2}} \int_\mathbb{R} \left(\frac{\nabla_\theta f(\lambda, \theta_0)}{f^2(\lambda, \theta_0)} I_T(\lambda) - \frac{\nabla_\theta f(\lambda, \theta_0)}{f(\lambda, \theta_0)}\right) w(\lambda) d\lambda$$

$$+ (2\pi)^{-\frac{1}{2}} T^{\frac{1}{2}} \int_\mathbb{R} \left(2 \Re \left\{\varepsilon_T(\lambda)\tilde{s}_T(\lambda, \tilde{\alpha}_T)\right\} + |s_T(\lambda, \tilde{\alpha}_T)|^2\right) \frac{\nabla_\theta f(\lambda, \theta_0)}{f^2(\lambda, \theta_0)} w(\lambda) d\lambda$$

$$= A_T^{(1)} + A_T^{(2)} + A_T^{(3)}.$$

(23)

Similarly

$$\nabla_\theta \nabla'_\theta U_T(\theta^*_{T,i}, \tilde{\alpha}_T) = \int_\mathbb{R} \left(\nabla_\theta \nabla'_\theta \log f(\lambda, \theta^*_{T,i}) + \nabla_\theta \nabla'_\theta \left(\frac{1}{f(\lambda, \theta^*_{T,i})}\right) I_T(\lambda, \tilde{\alpha}_T)\right) w(\lambda) d\lambda$$

$$= \int_\mathbb{R} \left\{\left(\frac{\nabla_\theta \nabla'_\theta f(\lambda, \theta^*_{T,i})}{f(\lambda, \theta^*_{T,i})} - \frac{\nabla_\theta f(\lambda, \theta^*_{T,i}) \nabla'_\theta f(\lambda, \theta^*_{T,i})}{f^2(\lambda, \theta^*_{T,i})}\right)\right.$$

$$+ \left(2 \nabla_\theta f(\lambda, \theta^*_{T,i}) \nabla'_\theta f(\lambda, \theta^*_{T,i}) - \frac{\nabla_\theta \nabla'_\theta f(\lambda, \theta^*_{T,i})}{f^2(\lambda, \theta^*_{T,i})}\right) \times$$

$$\times (I_T^x(\lambda) + (\pi T)^{-1} \Re \{\varepsilon_T(\lambda)s_T(\lambda, \tilde{\alpha}_T)\}) + (2\pi T)^{-1} |s_T(\lambda, \tilde{\alpha}_T)|^2\right\} w(\lambda) d\lambda$$

$$= B_T^{(1)} + B_T^{(2)} + B_T^{(3)} + B_T^{(4)},$$

(24)

where the terms $B_T^{(3)}$ and $B_T^{(4)}$ contain values $\Re \{\varepsilon_T(\lambda)s_T(\lambda, \tilde{\alpha}_T)\}$ and $|s_T(\lambda, \tilde{\alpha}_T)|^2$, respectively.

Bearing in mind the 1st part of the condition $N_4(i)$, we take in Lemma 7 the functions

$$\varphi(\lambda) = \varphi_i(\lambda) = \frac{f_i(\lambda, \theta)}{f^2(\lambda, \theta)} w(\lambda), \ i = 1, m.$$

Then in the formula (23) $A_T^{(2)} \xrightarrow{P} 0$, as $T \to \infty$.

Consider the term $A_T^{(3)} = (a_{iT}^{(3)})_{i=1}^m$, in the sum (23)

$$a_{iT}^{(3)} = (2\pi)^{-\frac{1}{2}} T^{-\frac{1}{2}} \int_\mathbb{R} |s_T(\lambda, \tilde{\alpha}_T)|^2 \varphi_i(\lambda) d\lambda.$$
Statistical Inference for Stochastic Processes

where \( \varphi_i(\lambda) \) are as before. Under conditions \( C_1, C_2, N_1 \) and \( (1) \) of \( N_4(i) \) \( A_T^{(3)} \overset{P}{\rightarrow} 0 \), as \( T \to \infty \), because

\[
|a_{iT}^{(3)}| \leq c(\varphi_i)T^{-1/2} \Phi_T(\widehat{\alpha}_T, \alpha_0) \\
\leq c(\varphi_i)c_0 T^{-1/2} d_T(\alpha_0) \| \hat{d}_T(\alpha_0) - \alpha_0 \| \overset{P}{\rightarrow} 0, \text{ as } T \to \infty.
\]

Examine the behaviour of the terms \( B_T^{(1)} - B_T^{(4)} \) in formula (24). Under conditions \( C_1 \) and \( N_4(iii) \) we can use Lemma 8 with functions

\[
\varphi(\lambda, \theta) = \varphi_{ij}(\lambda, \theta) = \frac{f_{ij}(\lambda, \theta)f_j(\lambda, \theta)}{f^2(\lambda, \theta)}, \quad i, j = 1, \ldots, m,
\]

to obtain the convergence

\[
B_T^{(1)} \overset{P}{\rightarrow} \int_R \left( \frac{\nabla_{\theta} \nabla_{\theta}' f(\lambda, \theta_0)}{f(\lambda, \theta_0)} - \frac{\nabla_{\theta} f(\lambda, \theta_0)\nabla_{\theta}' f(\lambda, \theta_0)}{f^2(\lambda, \theta_0)} \right) w(\lambda)d\lambda, \quad \text{as } T \to \infty. \tag{25}
\]

Under the condition \( N_5(i) \) we can use Lemma 9 with functions

\[
\varphi(\lambda, \theta) = \varphi_{ij}(\lambda, \theta) = \frac{f_{ij}(\lambda, \theta)}{f^3(\lambda, \theta)} w(\lambda), \quad i, j = 1, \ldots, m,
\]

to obtain that

\[
B_T^{(3)} \overset{P}{\rightarrow} 0, \quad B_T^{(4)} \overset{P}{\rightarrow} 0, \quad \text{as } T \to \infty.
\]

Under conditions \( C_1 \) and \( N_5 \)

\[
B_T^{(2)} \overset{P}{\rightarrow} \int_R \left( 2\frac{\nabla_{\theta} f(\lambda, \theta_0)\nabla_{\theta}' f(\lambda, \theta_0)}{f^2(\lambda, \theta_0)} - \frac{\nabla_{\theta} \nabla_{\theta}' f(\lambda, \theta_0)}{f(\lambda, \theta_0)} \right) w(\lambda)d\lambda, \tag{26}
\]

if we take in Lemma 10 in conditions (i) and (iii)

\[
\varphi(\lambda, \theta) = \varphi_{ij}(\lambda, \theta) = \frac{f_{ij}(\lambda, \theta)f_j(\lambda, \theta)}{f^3(\lambda, \theta)} w(\lambda), \quad i, j = 1, \ldots, m.
\]

So, under conditions \( C_1, C_2, N_4(iii) \) and \( N_5 \)

\[
\nabla_{\theta} \nabla_{\theta}' U_T(\theta_T, \widehat{\alpha}_T) \overset{P}{\rightarrow} \int_R \frac{\nabla_{\theta} f(\lambda, \theta_0)\nabla_{\theta}' f(\lambda, \theta_0)}{f^2(\lambda, \theta_0)} w(\lambda)d\lambda \\
= \int_R \nabla_{\theta} \log f(\lambda, \theta_0)\nabla_{\theta}' \log f(\lambda, \theta_0)w(\lambda)d\lambda = W_1(\theta_0), \tag{27}
\]

because \( W_1(\theta_0) \) is the sum of the right hand sides of (25) and (26).

From the facts obtained, it follows that for the proof of Theorem 2 it is necessary to study an asymptotic behaviour of vector \( A_T^{(1)} \) from (23):

\[
A_T^{(1)} = T^{1/2} \int_R \left( \frac{\nabla_{\theta} f(\lambda, \theta_0)}{f^2(\lambda, \theta_0)} I_T^\tau(\lambda) - \frac{\nabla_{\theta} f(\lambda, \theta_0)}{f(\lambda, \theta_0)} \right) w(\lambda)d\lambda.
\]
We will take
\[ \varphi_i(\lambda) = \frac{f_i(\lambda, \theta_0)}{f^2(\lambda, \theta_0)} w(\lambda), \quad i = 1, m, \]
\[ \Psi(\lambda) = \sum_{i=1}^{m} u_i \varphi_i(\lambda), \quad u = (u_1, \ldots, u_m) \in \mathbb{R}^m, \]
\[ Y_T = \int_{\mathbb{R}} I_T^2(\lambda) \Psi(\lambda) d\lambda, \quad Y = \int_{\mathbb{R}} f(\lambda, \theta_0) \Psi(\lambda) d\lambda, \]
and write
\[ \left\{ A_T^{(1)}, u \right\} = T^{\frac{1}{2}} (Y_T - E Y_T) + T^{\frac{1}{2}} (E Y_T - Y). \]

Under conditions (1) and (2) of \( N_4(i) \) (Bentkus 1972b; Ibragimov 1963) for any \( u \in \mathbb{R}^m \)
\[ T^{\frac{1}{2}} (E Y_T - Y) \longrightarrow 0, \quad \text{as} \quad T \to \infty. \]

On the other hand
\[ T^{\frac{1}{2}} (Y_T - E Y_T) = T^{\frac{1}{2}} \int_{0}^{T} \int_{0}^{T} (\varepsilon(t)\varepsilon(s) - B(t - s)) \hat{b}(t - s) dt ds \]

with
\[ \hat{b}(t) = \int_{\mathbb{R}} e^{i\lambda t} (2\pi)^{-1} \Psi(\lambda) d\lambda. \]

Thus we can apply Lemma 5 taking \( b(\lambda) = (2\pi)^{-1} \Psi(\lambda) \) in the formula (18) to obtain for any \( u \in \mathbb{R}^m \)
\[ T^{\frac{1}{2}} (Y_T - E Y_T) \Rightarrow N(0, \sigma^2), \quad \text{as} \quad T \to \infty, \]
where
\[ \sigma^2 = 4\pi \int_{\mathbb{R}} \Psi^2(\lambda) f^2(\lambda, \theta_0) d\lambda + \gamma_2 \left( \int_{\mathbb{R}} \Psi(\lambda) f(\lambda, \theta_0) d\lambda \right)^2. \]

The relations (28) and (29) are equivalent to the convergence
\[ A_T^{(1)} \Rightarrow N(0, W_2(\theta_0) + V(\theta_0)), \quad \text{as} \quad T \to \infty. \]

From (27) and (30) it follows (15).

Remark 4 From the conditions of Theorem 2 it follows also the fulfillment of Lemma 6 conditions for functions \( \hat{a} \) and \( \hat{b} \). Really by condition \( A_1 \hat{a} \in L_1(\mathbb{R}) \cap L_2(\mathbb{R}) \) and we can take \( p = 1 \) in Lemma 6. On the other hand, if we look at \( \hat{b} = (2\pi)^{-1} \Psi \) as at an original of the Fourier transform, from \( N_4(i)1 \) we have \( b \in L_1(\mathbb{R}) \cap L_2(\mathbb{R}) \). Then according to the Plancherel theorem \( \hat{b} \in L_2(\mathbb{R}) \) and we can take \( q = 2 \) in Lemma 6. Thus
\[ \frac{2}{p} + \frac{1}{q} = \frac{5}{2}, \]
and conclusion of Lemma 6 is true.
5 Example: The motion of a pendulum in a turbulent fluid

First of all we review a number of results discussed in Parzen (1962), Anh et al. (2002), and Leonenko and Papić (2019), see also references therein.

We examine the stationary Lévy-driven continuous-time autoregressive process \( \varepsilon(t), t \in \mathbb{R} \), of the order two (\( CAR(2) \)-process) in the under-damped case (see Leonenko and Papić 2019 for details).

The motion of a pendulum is described by the equation
\[
\ddot{\varepsilon}(t) + 2\alpha \dot{\varepsilon} + \left( \omega^2 + \alpha^2 \right) \varepsilon(t) = \dot{L}(t), \quad t \in \mathbb{R},
\]  
(31)
in which \( \varepsilon(t) \) is the replacement from its rest position, \( \alpha \) is a damping factor, \( \frac{2\pi}{\omega} \) is the damped period of the pendulum (see, i.e., Parzen 1962, pp. 111–113).

We consider the Green function solution of the equation (31), in which \( \dot{L} \) is the Lévy noise, i.e. the derivative of a Lévy process in the distribution sense (see Anh et al. 2002; Leonenko and Papić 2019 for details). The solution can be defined as the linear process
\[
\varepsilon(t) = \int_{\mathbb{R}} \hat{a}(t-s) dL(s), \quad t \in \mathbb{R},
\]
where the Green function
\[
\hat{a}(t) = e^{-\alpha t} \frac{\sin(\omega t)}{\omega} \mathbb{I}_{[0, \infty)}(t), \quad \alpha > 0.
\]  
(32)

Assuming \( EL(1) = 0, d_2 = E L^2(1) < \infty \), we obtain
\[
B(t) = d_2 \int_0^\infty \hat{a}(t+s) \hat{a}(s) ds = \frac{d_2}{4(\alpha^2 + \omega^2)} e^{-\alpha |t|} \left( \frac{\sin(\omega |t|)}{\omega} + \frac{\alpha}{\omega} \cos(\omega |t|) \right).
\]  
(33)

The formula (33) for the covariance function of the process \( \varepsilon \) corresponds to the formula (2.12) in Leonenko and Papić (2019) for the correlation function
\[
\text{Corr} (\varepsilon(t), \varepsilon(0)) = \frac{B(t)}{B(0)} = e^{-\alpha |t|} \left( \cos(\omega |t|) + \frac{\alpha}{\omega} \sin(\omega |t|) \right).
\]

On the other hand for \( \hat{a}(t) \) given by (32)
\[
a(\lambda) = \int_0^\infty e^{-i\lambda t} \hat{a}(t) dt = \frac{1}{\alpha^2 + \omega^2 - \lambda^2 + 2i\alpha\lambda}.
\]

Then the positive spectral density of the stationary process \( \varepsilon \) can be written as (compare with Parzen 1962)
\[
f_2(\lambda) = \frac{d_2}{2\pi} |a(\lambda)|^2 = \frac{d_2}{2\pi} \cdot \frac{1}{(\lambda^2 - \alpha^2 - \omega^2)^2 + 4\alpha^2\lambda^2}, \quad \lambda \in \mathbb{R}.
\]  
(34)

It is convenient to rewrite (34) in the form
\[
f_2(\lambda) = f(\lambda, \theta) = \frac{1}{2\pi} \cdot \frac{\beta}{(\lambda^2 - \alpha^2 - \gamma^2)^2 + 4\alpha^2\lambda^2}, \quad \lambda \in \mathbb{R},
\]  
(35)
where $\alpha = \theta_1$ is a damping factor, $\beta = -\kappa(0) = d_2(\theta_2) = \theta_2$, $\gamma = \omega = \theta_3$ is a damped cyclic frequency of the pendulum oscillations. Suppose that

$$\theta = (\theta_1, \theta_2, \theta_3) = (\alpha, \beta, \gamma) \in \Theta_1 \times \Theta_2 \times \Theta_3$$

where

$$\alpha = \theta_1$$

is a damping factor,

$$\beta = -\kappa(2) = d_2(\theta_2) = \theta_2,$$

$$\gamma = \omega = \theta_3.$$

The condition $C_3$ is fulfilled for spectral density (35).

Assume that

$$\theta = (\theta_1, \theta_2, \theta_3) = (\alpha, \beta, \gamma) \in \Theta_1 \times \Theta_2 \times \Theta_3, \alpha, \beta, \gamma > 0, \alpha, \beta, \gamma < \infty.$$

The condition $C_4$ is fulfilled for spectral density (35).

Assume that

$$w(\lambda) = (1 + \lambda^2)^{-a}, \lambda \in \mathbb{R}, a > 0.$$

More precisely the value of $a$ will be chosen below.

Obviously the functions $w(\lambda) \log f(\lambda, \theta), \frac{w(\lambda)}{f(\lambda, \theta)}$ are continuous on $\mathbb{R} \times \Theta^c$. For any $\Lambda > 0$ the function $|\log f(\lambda, \theta)|$ is bounded on the set $[-\Lambda, \Lambda] \times \Theta^c$. The number $\Lambda$ can be chosen so that for $R \setminus [-\Lambda, \Lambda]$,

$$1 < \frac{8\pi}{\beta} (s^2 - 2\lambda^2) \leq f^{-1}(\lambda, \theta) \leq \frac{2\pi}{\beta} \left(2\left(\lambda^4 + (s^2 + \bar{s}^2)^2\right) + 4s^2\lambda^2\right).$$

Thus the function $Z_1(\lambda)$ in the condition $C_4(i)$ exists.

As for condition $C_4(ii)$, if $a \geq \frac{3}{2}$, then

$$\sup_{\lambda \in \mathbb{R}, \theta \in \Theta^c} \frac{w(\lambda)}{f(\lambda, \theta)} < \infty.$$

As a function $v$ in condition $C_5$ we take

$$v(\lambda) = (1 + \lambda^2)^{-b}, \lambda \in \mathbb{R}, b > 0.$$

Obviously, if $a \geq b$, then $\sup_{\lambda \in \mathbb{R}} \frac{w(\lambda)}{v(\lambda)} < \infty$ (condition $C_5(ii)$), and the function $\frac{v(\lambda)}{f(\lambda, \theta)}$ is uniformly continuous in $(\lambda, \theta) \in \mathbb{R} \times \Theta^c$, if $b \geq 2$ (condition $C_5(i)$).

Further it will be helpful to use the notation $s(\lambda) = (\lambda^2 - \alpha^2 - \gamma^2)^2 + 4\alpha^2\lambda^2$. Then

$$f_\alpha(\lambda, \theta) = \frac{\partial}{\partial \alpha} f(\lambda, \theta) = -\frac{2\alpha\beta}{\pi} (\lambda^2 + \alpha^2 + \gamma^2)^{-2} s^{-2}(\lambda);$$

$$f_\beta(\lambda, \theta) = \frac{\partial}{\partial \beta} f(\lambda, \theta) = (2\pi s(\lambda))^{-2};$$

$$f_\gamma(\lambda, \theta) = \frac{\partial}{\partial \gamma} f(\lambda, \theta) = \frac{2\beta\gamma}{\pi} (\lambda^2 - \alpha^2 - \gamma^2)^{-2} s^{-2}(\lambda).$$

To check the condition $N_4(i)1$ consider the functions

$$\varphi_\alpha(\lambda) = \frac{f_\alpha(\lambda, \theta)}{f^2(\lambda, \theta)} w(\lambda) = -\frac{4\pi \alpha}{\beta} (\lambda^2 + \alpha^2 + \gamma^2)^{-2} w(\lambda);$$

$$\varphi_\beta(\lambda) = \frac{f_\beta(\lambda, \theta)}{f^2(\lambda, \theta)} w(\lambda) = \frac{2\pi}{\beta^2} s(\lambda) w(\lambda);$$

$$\varphi_\gamma(\lambda) = \frac{f_\gamma(\lambda, \theta)}{f^2(\lambda, \theta)} w(\lambda) = \frac{8\pi \gamma}{\beta} (\lambda^2 - \alpha^2 - \gamma^2)^{-2} w(\lambda).$$

Then the condition $N_4(i)1$ is satisfied for $\varphi_\alpha$ and $\varphi_\gamma$ when $a > \frac{3}{2}$, for $\varphi_\beta$ when $a > \frac{5}{2}$. The same values of $a$ are sufficient also to meet the condition $N_4(i)2$.
To verify \textbf{N}_4(\textbf{i})3 fix $\theta \in \Theta^c$ and denote by $\varphi(\lambda), \lambda \in \mathbb{R}$, any of the continuous functions $\varphi_a(\lambda), \varphi_\beta(\lambda), \varphi_\gamma(\lambda), \lambda \in \mathbb{R}$. Suppose $|1 - \eta| < \delta < \frac{1}{2}$. Then

$$
\sup_{\lambda \in \mathbb{R}} |\varphi(\eta \lambda) - \varphi(\lambda)| = \max \left( \sup_{|\lambda| \leq \Lambda} |\varphi(\eta \lambda) - \varphi(\lambda)|, \sup_{|\lambda| > \Lambda} |\varphi(\eta \lambda) - \varphi(\lambda)| \right)
$$

$$
= \max (s_1, s_2),
$$

$$
s_2 \leq \sup_{|\lambda| > \Lambda} |\varphi(\lambda)| + \sup_{|\lambda| > \Lambda} |\varphi(\lambda)| = s_3 + s_4.
$$

By the properties of the functions $\varphi$ under assumption $a > \frac{5}{2}$ for any $\varepsilon > 0$ there exists $\Lambda = \Lambda(\varepsilon) > 0$ such that for $|\lambda| > \frac{2}{3} \Lambda$ $|\varphi(\lambda)| < \frac{\varepsilon}{2}$. So, $s_3 \leq \frac{\varepsilon}{2}$. We have also $s_4 \leq \sup_{|\lambda| > \frac{2}{3} \Lambda} |\varphi(\lambda)| < \frac{\varepsilon}{2}$. On the other hand,

$$
s_1 \leq \sup_{|\lambda| < 2\Lambda} |\varphi(\eta \lambda) - \varphi(\lambda)|, \ |\eta \lambda - \lambda| \leq 2\Lambda \delta = \delta',
$$

and by the proper choice of $\delta$

$$
s_1 \leq \sup_{\lambda_1, \lambda_2 \in \{-2\Lambda, 2\Lambda\}} |\varphi(\lambda_1) - \varphi(\lambda_2)| < \varepsilon,
$$

and condition \textbf{N}_4(\textbf{i})(\textbf{3}) is met.

Using (37) we get for any $\theta \in \Theta^c$, as $\lambda \to \infty$,

$$
\varphi'_a(\lambda) = -\frac{8\pi a}{\beta} \lambda w(\lambda) - \frac{4\pi a}{\beta} \left( \lambda^2 + a^2 + \gamma^2 \right) w'(\lambda) = O \left( \lambda^{-2a+1} \right);
$$

$$
\varphi'_\beta(\lambda) = \frac{2\pi}{\beta^2} \left( s'(\lambda) w(\lambda) + s(\lambda) w'(\lambda) \right) = O \left( \lambda^{-2a+3} \right);
$$

$$
\varphi'_\gamma(\lambda) = \frac{16\pi \gamma}{\beta} \lambda w(\lambda) + \frac{8\pi \gamma}{\beta} \left( \lambda^2 - a^2 - \gamma^2 \right) w'(\lambda) = O \left( \lambda^{-2a+1} \right).
$$

Therefore for $a > \frac{3}{2}$ these derivatives are uniformly continuous on $\mathbb{R}$ (condition \textbf{N}_4(\textbf{i})(\textbf{4})). So, to satisfy condition \textbf{N}_4(\textbf{i}) we can take weight function $w(\lambda)$ with $a > \frac{5}{2}$.

The check of assumption \textbf{N}_4(\textbf{ii}) is similar to the check of \textbf{C}_4(\textbf{i}).

As $\lambda \to \infty$, uniformly in $\theta \in \Theta^c$

$$
\left| \frac{f'_a(\lambda, \theta)}{f(\lambda, \theta)} \right| w(\lambda) = |\varphi'_a(\lambda)| \left| f(\lambda, \theta) w(\lambda) \right| = 2a \left( \lambda^2 + a^2 + \gamma^2 \right) s^{-1}(\lambda) w(\lambda) = O \left( \lambda^{-2a-2} \right);
$$

$$
\left| \frac{f'_\beta(\lambda, \theta)}{f(\lambda, \theta)} \right| w(\lambda) = |\varphi'_\beta(\lambda)| \left| f(\lambda, \theta) w(\lambda) \right| = \beta^{-1} w(\lambda) = O \left( \lambda^{-2a} \right);
$$

$$
\left| \frac{f'_\gamma(\lambda, \theta)}{f(\lambda, \theta)} \right| w(\lambda) = |\varphi'_\gamma(\lambda)| \left| f(\lambda, \theta) w(\lambda) \right| = 4\gamma \left| \lambda^2 - a^2 - \gamma^2 \right| s^{-1}(\lambda) w(\lambda) = O \left( \lambda^{-2a-2} \right).
$$

(38)

On the other hand, for any $\Lambda > 0$ the functions (38) are bounded on the sets $[-\Lambda, \Lambda] \times \Theta^c$. 

\hfill \copyright \ Springer
To check $N_4(\text{iii})$ note first of all that the functions uniformly in $\theta \in \Theta^c$, as $\lambda \to \infty$,

$$
\frac{f_\alpha^2(\lambda, \theta)}{f^2(\lambda, \theta)} w(\lambda) = \varphi_\alpha(\lambda) f(\lambda, \theta) = 8\alpha^2 \left(\lambda^2 + \alpha^2 + \gamma^2\right) s^{-2}(\lambda)w(\lambda) = O \left(\lambda^{-2a-4}\right);
$$

$$
\frac{f_\beta^2(\lambda, \theta)}{f^2(\lambda, \theta)} w(\lambda) = \varphi_\beta(\lambda) f(\lambda, \theta) = \beta^{-2}w(\lambda) = O \left(\lambda^{-2a}\right);
$$

$$
\frac{f_\gamma^2(\lambda, \theta)}{f^2(\lambda, \theta)} w(\lambda) = \varphi_\gamma(\lambda) f(\lambda, \theta) = 16\gamma^2 \left(\lambda^2 - \alpha^2 - \gamma^2\right) s^{-2}(\lambda)w(\lambda) = O \left(\lambda^{-2a-4}\right).
$$

(39)

These functions are continuous on $\mathbb{R} \times \Theta^c$, as well as the functions

$$
\frac{f_\alpha(\lambda, \theta) f_\beta(\lambda, \theta)}{f^2(\lambda, \theta)} w(\lambda) = \varphi_\alpha(\lambda) f_\beta(\lambda, \theta) = -\frac{2\alpha}{\beta} \left(\lambda^2 + \alpha^2 + \gamma^2\right) s^{-1}(\lambda)w(\lambda);
$$

$$
\frac{f_\alpha(\lambda, \theta) f_\gamma(\lambda, \theta)}{f^2(\lambda, \theta)} w(\lambda) = \varphi_\alpha(\lambda) f_\gamma(\lambda, \theta) = -8\alpha\gamma \left(\lambda^4 - (\alpha^2 + \gamma^2)^2\right) s^{-2}(\lambda)w(\lambda);
$$

$$
\frac{f_\beta(\lambda, \theta) f_\gamma(\lambda, \theta)}{f^2(\lambda, \theta)} w(\lambda) = \varphi_\beta(\lambda) f_\gamma(\lambda, \theta) = \frac{4\gamma}{\beta} \left(\lambda^2 - \alpha^2 - \gamma^2\right) s^{-1}(\lambda)w(\lambda).
$$

(40)

Moreover, uniformly in $\theta \in \Theta^c$, as $\lambda \to \infty$,

$$
\frac{f_{\alpha\alpha}(\lambda, \theta)}{f(\lambda, \theta)} w(\lambda) = -4 \left(\lambda^2 + 3\alpha^2 + \gamma^2\right) s^{-1}(\lambda)w(\lambda) + 8\alpha \left(\lambda^2 + \alpha^2 + \gamma^2\right) s^{-2}(\lambda)s'_\alpha(\lambda)w(\lambda)
$$

$= O \left(\lambda^{-2a-2}\right);
$$

$$
\frac{f_{\beta\beta}(\lambda, \theta)}{f(\lambda, \theta)} w(\lambda) = 0;
$$

$$
\frac{f_{\gamma\gamma}(\lambda, \theta)}{f(\lambda, \theta)} w(\lambda) = 4 \left(\lambda^2 - \alpha^2 - 3\gamma^2\right) s^{-1}(\lambda)w(\lambda) - 8\gamma \left(\lambda^2 - \alpha^2 - \gamma^2\right) s^{-2}(\lambda)s'_\gamma(\lambda)w(\lambda)
$$

$= O \left(\lambda^{-2a-2}\right);
$$

$$
\frac{f_{\alpha\beta}(\lambda, \theta)}{f(\lambda, \theta)} w(\lambda) = -\frac{4\alpha}{\beta} \left(\lambda^2 + \alpha^2 + \gamma^2\right) s^{-1}(\lambda)w(\lambda) = O \left(\lambda^{-2a-2}\right);
$$

$$
\frac{f_{\alpha\gamma}(\lambda, \theta)}{f(\lambda, \theta)} w(\lambda) = -8\alpha\gamma s^{-1}(\lambda)w(\lambda) + 16\alpha\gamma \left(\lambda^4 - (\alpha^2 + \gamma^2)^2\right) s^{-2}(\lambda)w(\lambda) = O \left(\lambda^{-2a-4}\right);
$$

$$
\frac{f_{\beta\gamma}(\lambda, \theta)}{f(\lambda, \theta)} w(\lambda) = \frac{4\gamma}{\beta} \left(\lambda^2 - \alpha^2 - \gamma^2\right) s^{-1}(\lambda)w(\lambda) = O \left(\lambda^{-2a-2}\right).
$$

(41)

Note that the functions (41) are continuous on $\mathbb{R} \times \Theta^c$ as well as functions (39) and (40). Therefore the condition $N_4(\text{iii})$ is fulfilled.

Let us verify the condition $N_5(\text{I})$. According to equation (39), uniformly in $\theta \in \Theta^c$, as $\lambda \to \infty$,

$$
\frac{f_\alpha^2(\lambda, \theta)}{f^3(\lambda, \theta)} w(\lambda) = \frac{16\pi\alpha^2}{\beta} \left(\lambda^2 + \alpha^2 + \gamma^2\right) s^{-1}(\lambda)w(\lambda) = O \left(\lambda^{-2a}\right);
$$

$$
\frac{f_\beta^2(\lambda, \theta)}{f^3(\lambda, \theta)} w(\lambda) = \frac{2\pi}{\beta^3} s(\lambda)w(\lambda) = O \left(\lambda^{-2a+4}\right);
$$

$$
\frac{f_\gamma^2(\lambda, \theta)}{f^3(\lambda, \theta)} w(\lambda) = \frac{32\pi\gamma^2}{\beta} \left(\lambda^2 - \alpha^2 - \gamma^2\right) s^{-1}(\lambda)w(\lambda) = O \left(\lambda^{-2a}\right).
$$

(42)
Therefore the continuous in \((\lambda, \theta) \in \mathbb{R} \times \Theta^c\) functions (42) are bounded in \((\lambda, \theta) \in \mathbb{R} \times \Theta^c\), if \(a \geq 1\).

Using equations (40) and (41) we obtain uniformly in \(\theta \in \Theta^c\), as \(\lambda \to \infty\),

\[
f_{aa}(\lambda, \theta) f^2(\lambda, \theta) w(\lambda) = -\frac{8\pi}{\beta} (\lambda^2 + 3a^2 + \gamma^2) w(\lambda) + \frac{16\pi \alpha}{\beta} \left(\lambda^2 + a^2 + \gamma^2\right) s^{-1}(\lambda) s'_a(\lambda) w(\lambda) = O \left(\lambda^{-2a+2}\right);
\]

\[
f_{bb}(\lambda, \theta) f^2(\lambda, \theta) w(\lambda) = 0;
\]

\[
f_{yy}(\lambda, \theta) f^2(\lambda, \theta) w(\lambda) = 8\pi \frac{\alpha}{\beta} (\lambda^2 - a^2 - 3\gamma^2) w(\lambda) - \frac{16\pi \gamma}{\beta} \left(\lambda^2 - a^2 - \gamma^2\right) s^{-1}(\lambda) s'_\gamma(\lambda) w(\lambda) = O \left(\lambda^{-2a+2}\right);
\]

\[
f_{ab}(\lambda, \theta) f^2(\lambda, \theta) w(\lambda) = -\frac{8\pi \alpha}{\beta^2} (\lambda^2 + a^2 + \gamma^2) w(\lambda) = O \left(\lambda^{-2a+2}\right);
\]

\[
f_{ay}(\lambda, \theta) f^2(\lambda, \theta) w(\lambda) = -\frac{16\pi \gamma}{\beta} w(\lambda) + \frac{32\pi \alpha \gamma}{\beta} \left(\lambda^4 - (a^2 + \gamma^2)^2\right) s^{-1}(\lambda) w(\lambda) = O \left(\lambda^{-2a}\right);
\]

\[
f_{by}(\lambda, \theta) f^2(\lambda, \theta) w(\lambda) = \frac{8\pi \gamma}{\beta^2} (\lambda^2 - a^2 - \gamma^2) w(\lambda) = O \left(\lambda^{-2a+2}\right).
\]

(43)

So, continuous on \(\mathbb{R} \times \Theta^c\) functions (43) are bounded in \((\lambda, \theta) \in \mathbb{R} \times \Theta^c\), if \(a \geq 1\).

To check \(N_5(\text{ii})\) consider the weight function

\[v(\lambda) = (1 + \lambda^2)^{-b}, \quad \lambda \in \mathbb{R}, \quad b > 0.\]

If \(a \geq b\), then function \(\frac{w(\lambda)}{v(\lambda)}\) is bounded on \(\mathbb{R}\) (condition \(N_5(\text{iii})\)). Using (42) we obtain uniformly in \(\theta \in \Theta^c\), as \(\lambda \to \infty\),

\[
f_a(\lambda, \theta) f^3(\lambda, \theta) v(\lambda) = \frac{16\pi \alpha^2}{\beta} \left(\lambda^2 + a^2 + \gamma^2\right)^2 s^{-1}(\lambda) v(\lambda) = O \left(\lambda^{-2b}\right);
\]

\[
f_b(\lambda, \theta) f^3(\lambda, \theta) v(\lambda) = \frac{2\pi}{\beta^3} s(\lambda) v(\lambda) = O \left(\lambda^{-2b+4}\right);
\]

\[
f_y(\lambda, \theta) f^3(\lambda, \theta) v(\lambda) = \frac{32\pi \gamma^2}{\beta} \left(\lambda^2 - a^2 - \gamma^2\right)^2 s^{-1}(\lambda) v(\lambda) = O \left(\lambda^{-2b}\right).
\]

(44)

In turn, similarly to (40) it follows uniformly in \(\theta \in \Theta^c\), as \(\lambda \to \infty\),

\[
f_a(\lambda, \theta) f_b(\lambda, \theta) f^3(\lambda, \theta) v(\lambda) = -\frac{4\pi \alpha}{\beta^2} \left(\lambda^2 + a^2 + \gamma^2\right) v(\lambda) = O \left(\lambda^{-2b+2}\right);
\]

\[
f_a(\lambda, \theta) f_y(\lambda, \theta) f^3(\lambda, \theta) v(\lambda) = -\frac{16\pi \alpha \gamma}{\beta} \left(\lambda^4 - (a^2 + \gamma^2)^2\right) s^{-1}(\lambda) v(\lambda) = O \left(\lambda^{-2b}\right);
\]

\[
f_b(\lambda, \theta) f_y(\lambda, \theta) f^3(\lambda, \theta) v(\lambda) = \frac{8\pi \gamma}{\beta^2} \left(\lambda^2 - a^2 - \gamma^2\right) v(\lambda) = O \left(\lambda^{-2b+2}\right).
\]

(45)

The functions (44) and (45) will be uniformly continuous in \((\lambda, \theta) \in \mathbb{R} \times \Theta^c\), if they converge to zero, as \(\lambda \to \infty\), uniformly in \(\theta \in \Theta^c\), that is if \(b > 2\).
Similarly to (43) uniformly in $\theta \in \Theta^c$, as $\lambda \to \infty$,
\[
\frac{f_{a\alpha}(\lambda, \theta)}{f^2(\lambda, \theta)} v(\lambda) = O\left(\lambda^{-2b+2}\right); \quad \frac{f_{b\beta}(\lambda, \theta)}{f^2(\lambda, \theta)} v(\lambda) = 0; \quad \frac{f_{\gamma\gamma}(\lambda, \theta)}{f^2(\lambda, \theta)} v(\lambda) = O\left(\lambda^{-2b+2}\right);
\]
\[
\frac{f_{a\alpha}(\lambda, \theta)}{f^2(\lambda, \theta)} v(\lambda) = O\left(\lambda^{-2b+2}\right); \quad \frac{f_{\alpha\gamma}(\lambda, \theta)}{f^2(\lambda, \theta)} v(\lambda) = O\left(\lambda^{2b}\right);
\]
\[
\frac{f_{\beta\gamma}(\lambda, \theta)}{f^2(\lambda, \theta)} v(\lambda) = O\left(\lambda^{-2b+2}\right).
\]
Thus the functions (44)–(46) are uniformly continuous in $(\lambda, \theta) \in \mathbb{R} \times \Theta^c$, if $b > 2$.

Proceeding to the verification of condition $N_6$, we note that for any $x = (x_\alpha, x_\beta, x_\gamma) \neq 0$
\[
\langle W_1(\theta) x, x \rangle = \int_{\mathbb{R}} \left(x_\alpha f_{\alpha}(\lambda, \theta) + x_\beta f_{\beta}(\lambda, \theta) + x_\gamma f_{\gamma}(\lambda, \theta)\right) \frac{w(\lambda)}{f^2(\lambda, \theta)} d\lambda.
\]
From equation (36) it is seen that the positive definiteness of the matrix $W_1(\lambda)$ follows from linear independence of the functions $\lambda^2 + \alpha^2 + \gamma^2$, $s(\lambda)$, $\lambda^2 - \alpha^2 - \gamma^2$. Positive definiteness of the matrix $W_2(\theta)$ is established similarly.

In our example to satisfy the consistency conditions $C_4$ and $C_5$ the weight functions $w(\lambda)$ and $v(\lambda)$ should be chosen so that $a \geq b > 2$. On the other hand to satisfy the asymptotic normality conditions $N_4$ and $N_5$ the functions $w(\lambda)$ and $v(\lambda)$ should be such that $a > \frac{5}{2}$ and $a \geq b > 2$.

The spectral density (35) has no singularity at zero, so that the functions $v(\lambda)$ in the conditions $C_5(i)$ and $N_5(ii)$ could be chosen to be equal to $w(\lambda)$, for example, $a = b = 3$. However we prefer to keep in the text the function $v(\lambda)$, since it is needed when the spectral density could have a singularity at zero or elsewhere, see, e.g., Example 1 (Leonenko and Sakhno 2006), where linear process driven by the Brownian motion and regression function $g(t, \alpha) \equiv 0$ have been studied. Specifically in the case of Riesz-Bessel spectral density
\[
f(\lambda, \theta) = \frac{\beta}{2\pi |\lambda|^{2\alpha}(1 + \lambda^2)^\gamma}, \quad \lambda \in \mathbb{R}, \quad (47)
\]
where $\theta = (\theta_1, \theta_2, \theta_3) = (\alpha, \beta, \gamma) \in \Theta = (\alpha, \alpha) \times (\beta, \bar{\beta}) \times (\gamma, \bar{\gamma})$, $\alpha > 0$, $\bar{\alpha} < \frac{1}{2}$, $\bar{\beta} > 0$, $\bar{\beta} < \infty$, $\bar{\gamma} > \frac{1}{2}$, $\bar{\gamma} < \infty$, and the parameter $\alpha$ signifies the long range dependence, while the parameter $\gamma$ indicates the second-order intermittency (Anh et al. 2004; Gao et al. 2001; Lim and Teo 2008), the weight functions have been chosen in the form
\[
w(\lambda) = \frac{\lambda^{2b}}{(1 + \lambda^2)^a}, \quad a > b > 0; \quad v(\lambda) = \frac{\lambda^{2b'}}{(1 + \lambda^2)^{a'}}, \quad a' > b' > 0, \quad \lambda \in \mathbb{R}.
\]

Unfortunately, our conditions do not cover so far the case of the general non-linear regression function and Lévy driven continuous-time strongly dependent linear random noise such as Riesz-Bessel motion.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.
Appendix A: LSE consistency

Some results on consistency of the LSE $\hat{\alpha}_T$ in the observation model of the type (1) with stationary noise $\varepsilon(t)$, $t \in \mathbb{R}$, were obtained, for example, in Ivanov and Leonenko (1989, 2004, 2007, 2008), Ivanov (1980, 2010), Ivanov et al. (2015) to mention several of the relevant works. In this section we formulate a generalization of Malinvaud theorem (Malinvaud 2004, 2007, 2008), Ivanov (1980, 2010), Ivanov et al. (2015) to mention several of the relevant works. In this section we formulate a generalization of Malinvaud theorem (Malinvaud 1970) on $\hat{\alpha}_T$ consistency for linear stochastic process (4) and consider an example of nonlinear regression function $g(t, \alpha)$ satisfying the conditions of this theorem and conditions $C_1, C_2$. Then we consider another possibilities of $C_1$ and $C_2$ fulfillment.

Set

$$w_T(\alpha_1, \alpha_2) = \int_0^T \varepsilon(t) (g(t, \alpha_1) - g(t, \alpha_2)) \, dt, \quad \alpha_1, \alpha_2 \in \mathcal{A}^c,$$

$$\Psi_T(u_1, u_2) = \Phi_T \left( \alpha_0 + T^2 d_T^{-1}(\alpha_0) u_1, \alpha_0 + T^2 d_T^{-1}(\alpha_0) u_1 \right).$$

For any fixed $\alpha_0 \in \mathcal{A}$, the function $\Psi_T(u_1, u_2)$ is defined on the set $U_T(\alpha_0) \times U_T(\alpha_0)$, $U_T(\alpha_0) = T^{-\frac{1}{2}} d_T(\alpha_0) (\mathcal{A}^c - \alpha_0)$.

Assume the following.

(1) For any $\varepsilon > 0$ and $R > 0$ there exists $\delta = \delta(\varepsilon, R)$ such that

$$\sup_{u_1, u_2 \in U_T(\alpha_0) \cap \psi(R)} T^{-1} \Psi_T(u_1, u_2) \leq \varepsilon. \quad (48)$$

(2) For some $R_0 > 0$ and any $\rho \in (0, R_0)$ there exist numbers $a = a(R_0) > 0$ and $b = b(\rho, R_0)$ such that

$$\inf_{u \in U_T(\alpha_0) \cap \psi(R_0)} T^{-1} \Psi(u, 0) \geq b; \quad (49)$$

$$\inf_{u \in U_T(\alpha_0) \cap \psi(R_0)} T^{-1} \Psi(u, 0) \geq 4B(0) + a. \quad (50)$$

It was proven in Lemma 1 that under condition $A_1$

$$E \left( \psi_T^2 - B(0) \right) = O \left( T^{-1} \right). \quad (51)$$

Lemma A.1 Under condition $A_1$,

$$E \, w_T^4(\alpha_1, \alpha_2) \leq c \Phi_T^2(\alpha_1, \alpha_2), \quad \alpha_1, \alpha_2 \in \mathcal{A}^c. \quad (52)$$

Proof By formula (7)

$$E \, w_T^4(\alpha_1, \alpha_2) = \int_{[0,T]^4} c_4(t_1, t_2, t_3, t_4) \prod_{i=1}^4 (g(t_i, \alpha_1) - g(t_i, \alpha_2)) \, dt_1 \, dt_2 \, dt_3 \, dt_4$$

$$+ 3 \left( \int_0^T \int_0^T B(t_1 - t_2) (g(t_1, \alpha_1) - g(t_1, \alpha_2)) (g(t_2, \alpha_1) - g(t_2, \alpha_2)) \, dt_1 \, dt_2 \right)^2$$

$$= I_7 + 3 I_8^2.$$
By condition A₁ and Fubini-Tonelli theorem

\[ |I₈| \leq \frac{1}{2} \int_0^T \int_0^T |B(t₁ - t₂)| \left[ (g(t₁, \alpha₁) - g(t₁, \alpha₂))^2 + (g(t₂, \alpha₁) - g(t₂, \alpha₂))^2 \right] dt₁dt₂ \]

\[ \leq d₂ \| \hat{a} \|₁^2; \]

\[ \| \hat{a} \|₁ = \int \| \hat{a}(t) \| dt. \]

On the other hand by formula (8)

\[ |I₇| \leq d₄ \int \int \int_{[0,T]^4} |\hat{a}(t_i - s)(g(t_i, \alpha₁) - g(t_i, \alpha₂))| dt₁dt₂dt₃dt₄ \]

\[ \leq \frac{1}{2} d₄ \int \int \int_{[0,T]^4} \sum_{i=1}^{4} |\hat{a}(t_i - s)| \left[ (g(t₁, \alpha₁) - g(t₁, \alpha₂))^2 (g(t₂, \alpha₁) - g(t₂, \alpha₂))^2 + \right. \]

\[ \left. + (g(t₃, \alpha₁) - g(t₃, \alpha₂))^2 (g(t₄, \alpha₁) - g(t₄, \alpha₂))^2 \right] dt₁dt₂dt₃dt₄ = I₇⁽¹⁾ + I₇⁽²⁾; \]

\[ I₇⁽¹⁾ = \frac{1}{2} d₄ \int \int \int \int_{[0,T]^4} |\hat{a}(t₁ - s)\hat{a}(t₂ - s)(g(t₁, \alpha₁) - g(t₁, \alpha₂))|^2 \]

\[ \times (g(t₂, \alpha₁) - g(t₂, \alpha₂))^2 dt₁dt₂ \]

\[ \cdot \int \int |\hat{a}(t₃ - s)\hat{a}(t₄ - s)| dt₃dt₄ \]

\[ \leq \frac{1}{4} d₄ \| \hat{a} \|₁^2 \int \int (g(t₁, \alpha₁) - g(t₁, \alpha₂))^2 (g(t₂, \alpha₁) - g(t₂, \alpha₂))^2 \]

\[ \cdot \left( \int \int [\hat{a}^2(t₁ - s) + \hat{a}^2(t₂ - s)] ds \right) dt₁dt₂ \]

\[ \leq \frac{1}{2} d₄ \| \hat{a} \|₁^2 \Phi_T^2(\alpha₁, \alpha₂). \]

For integral I₇⁽²⁾ we get the same bound. So, we obtain inequality (52) with

\[ c = d₄ \| \hat{a} \|₁^2 \| \hat{a} \|₂^2 + 3d₂^2 \| \hat{a} \|₁^4. \]

\[ \square \]

**Theorem A.1** If assumptions (I), (2), and A₁ are valid then for any \( \rho > 0 \)

\[ P \left\{ \left\| T^{-\frac{1}{2}} d_T(\alpha₀) (\hat{a}_T - \alpha₀) \right\| \geq \rho \right\} = O(T⁻¹), \text{ as } T \to \infty. \]

**Proof** The proof of this Malinvaud theorem generalization is similar to the proof of Theorem 3.2.1. in Ivanov and Leonenko (1989) and uses the relations (51) and (52).
Instead of $C_2$ consider the stronger condition. 

$C'_2$. There exist positive constants $c_0$, $c_1 < \infty$ such that for any $\alpha \in \mathcal{A}^c$ and $T > T_0$

$$c_1 \left\| d_T (\alpha) (\alpha_1 - \alpha_2) \right\|^2 \leq \Phi_T (\alpha_1, \alpha_2) \leq c_0 \left\| d_T (\alpha) (\alpha_1 - \alpha_2) \right\|^2, \, \alpha_1, \alpha_2 \in \mathcal{A}^c. \quad (53)$$

Point out a sufficient condition for $C'_2$ fulfillment. Introduce a diagonal matrix

$$s_T = \text{diag} \left( s_{iT}, \, i = \overline{1, q} \right), \, s_{iT} \to \infty, \, \text{as} \, T \to \infty, \, i = \overline{1, q}.$$

$C''_2$. (i) There exist positive constants $c_j$, $\overline{c}_i$, $i = \overline{1, q}$, such that for $T > T_0$ uniformly in $\alpha \in \mathcal{A}$

$$c_j < s_{iT}^{-1} d_{iT} (\alpha) < \overline{c}_i, \, i = \overline{1, q}. \quad (54)$$

(ii) For some numbers $c_0^*$, $c_1^*$ and $T > T_0$,

$$c_0^* \left\| s_T (\alpha_1 - \alpha_2) \right\|^2 \leq \Phi_T (\alpha_1, \alpha_2) \leq c_0^* \left\| s_T (\alpha_1 - \alpha_2) \right\|^2, \, \alpha_1, \alpha_2 \in \mathcal{A}^c.$$

Under condition $C''_2$ as it is easily seen one can take in $C'_2$

$$c_0 = c_0^* \left( \min_{1 \leq i \leq q} c_j \right)^{-1}, \quad c_1 = c_1^* \left( \max_{1 \leq i \leq q} \overline{c}_i \right)^{-1}.$$

The next example demonstrates the fulfillment of the condition $C'_2$ [compare with Ivanov and Orlovskyi (2018)].

**Example A.1** Let

$$g(t, \alpha) = \exp \{ \langle \alpha, \, y(t) \rangle \},$$

with $\langle \alpha, \, y(t) \rangle = \sum_{i=1}^{q} \alpha_i y_i(t)$, regressors $y(t) = \left( y_1(t), \ldots, y_q(t) \right)'$; $t \geq 0$, take values in a compact set $Y \subset \mathbb{R}^q$. Suppose

$$J_T = \left( T^{-1} \int_{0}^{T} y_i(t) y_j(t) dt \right)_{i,j=1}^{q} \rightarrow J = (J_{ij})_{i,j=1}^{q}, \quad \text{as} \, T \to \infty,$$

where $J$ is a positive definite matrix, and the set $\mathcal{A}$ in the model (1) is bounded. Set

$$M = \max_{\alpha \in \mathcal{A}^c, \, y \in Y} \exp \{ \langle \alpha, \, y \rangle \}, \quad L = \min_{\alpha \in \mathcal{A}^c, \, y \in Y} \exp \{ \langle \alpha, \, y \rangle \}.$$

Then for any $\delta > 0$ and $T > T_0$

$$L^2 (J_{ii} - \delta) < T^{-1} d_T^2 (\alpha) < M^2 (J_{ii} + \delta), \, i = \overline{1, q},$$

and condition $C'_2$(i) is fulfilled with matrix $s_T = T^{1/2} \mathbb{I}_q$, $\mathbb{I}_q$ is identity matrix of order $q$, and

$$c_j = L^2 (J_{ii} - \delta), \quad \overline{c}_i = M^2 (J_{ii} + \delta), \, i = \overline{1, q}.$$

Let us check the condition $C''_2$(ii). We have

$$e^{\langle \alpha_1, \, y(t) \rangle} - e^{\langle \alpha_2, \, y(t) \rangle} = e^{\langle \alpha_2, \, y(t) \rangle} \left( e^{\langle \alpha_1 - \alpha_2, \, y(t) \rangle} - 1 \right).$$

As far as $(e^x - 1)^2 \geq x^2, \, x \geq 0$, and $(e^x - 1)^2 \geq e^{2x} x^2, \, x < 0$, then

$$\left( e^{\langle \alpha_1 - \alpha_2, \, y(t) \rangle} - 1 \right)^2 \geq \Delta (\alpha_1 - \alpha_2, \, y(t))^2, \quad \Delta = \min \left\{ 1, \, e^{2\langle \alpha_1 - \alpha_2, \, y(t) \rangle} \right\}.$$
Thus
\[ e^{2(\alpha, y(t))} - e^{2(\alpha_1, y(t))} \geq e^{2(\alpha_2, y(t))} \Delta (\alpha_1 - \alpha_2, y(t))^2 \geq L^2 (\alpha_1 - \alpha_2, y(t))^2, \]
and for any \( \delta > 0 \) and \( T > T_0 \)
\[ \Phi_T(\alpha_1, \alpha_2) \geq L_2 \sum_{i,j=1}^q J_{ij,T} \left( T \frac{1}{z}(\alpha_{i1} - \alpha_{i2}) \right) \left( T \frac{1}{z}(\alpha_{j1} - \alpha_{j2}) \right) \]
\[ \geq L^2 (\lambda_{\min}(J) - \delta) \left\| T \frac{1}{z}(\alpha_1 - \alpha_2) \right\|^2, \]
where \( \lambda_{\min}(J) \) is the least eigenvalue of the matrix \( J \).

On the other hand,
\[ \left| e^{(\alpha_1, y(t))} - e^{(\alpha_2, y(t))} \right| = \left| \sum_{i=1}^q y_i(t) e^{(y(t), \alpha_1 + \eta(t)(\alpha_2 - \alpha_1))} (\alpha_{i1} - \alpha_{i2}) \right| \]
\[ \leq M \left| \sum_{i=1}^q y_i(t)(\alpha_{i1} - \alpha_{i2}) \right|, \]
\( \eta(t) \in (0, 1) \), and
\[ \Phi_T(\alpha_1, \alpha_2) \leq M^2 \int_0^T \left( \sum_{i=1}^q y_i(t)(\alpha_{i1} - \alpha_{i2}) \right)^2 dt \leq M^2 (\lambda_{\max}(J) + \delta) \left\| T \frac{1}{z}(\alpha_1 - \alpha_2) \right\|^2, \]
where \( \lambda_{\max}(J) \) is the maximal eigenvalue of the matrix \( J \). It means that condition \( C_2''(ii) \) is valid for matrix \( s_T = T \frac{1}{z} I_q \).

So the condition \( C_2' \) is valid as well and in (53) one can choose for \( T > T_0 \) some numbers
\[ c_0 > \frac{M^2 \lambda_{\max}(J)}{L^2 \min_{1 \leq i \leq q} J_{ii}}, \quad c_1 < \frac{L^2 \lambda_{\min}(J)}{M^2 \max_{1 \leq i \leq q} J_{ii}}. \]

Inequalities (53) can be rewritten in the equivalent form
\[ c_1 \left\| u - v \right\|^2 \leq T^{-1} \Psi_T(u, v) \leq c_0 \left\| u - v \right\|^2, \quad u, v \in U_T(\alpha), \ \alpha \in A. \] (55)
From the right hand side of (55) it follows (48). Similarly, from the left hand side of (55) taking \( v = 0 \) we obtain (49) for any \( R_0 > 0 \) and it is possible to choose \( R_0 > 0 \) satisfying (50).

In our example \( A_1 \) due to inequalities (54) with \( s_{iT} = T \frac{1}{z}, \ i = \bar{1,q} \), the set \( U_T(\alpha) \) is bounded uniformly in \( T \) and it is not necessary to use condition (50). However in Malinvaud theorem we can not ignore the condition (50) of parameters distinguishability in the cases when the sets \( U_T(\alpha) \) expands to infinity as \( T \to \infty \) or the set \( A \) is unbounded.

It goes without saying not all the interesting classes of nonlinear regression functions satisfy consistency conditions of Malinvaud or, say, Jennrich (1969) types. The important example of such a class is given by the trigonometric regression functions.

**Example A.2** Let
\[ g(t, \alpha) = \sum_{i=1}^N (A_i \cos \varphi_it + B_i \sin \varphi_it), \] (56)
\( \alpha = (\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_{3N-2}, \alpha_{3N-1}, \alpha_{3N}) = (A_1, B_1, \varphi_1, \ldots, A_N, B_N, \varphi_N), 0 < \varphi < \varphi_1 < \ldots < \varphi_N < \varphi < \infty. \)

Under some conditions on angular frequencies \( \varphi = (\varphi_1, \ldots, \varphi_N) \) distinguishability (see Walker Walker 1973; Ivanov 1980; Ivanov et al. 2015) it is possible to prove that at least

\[
T^{-1} \Phi_T(\alpha_0) \xrightarrow{P} 0, \text{ as } T \to \infty, \tag{57}
\]

\( \widehat{\alpha}_T = (A_{1T}, B_{1T}, \varphi_{1T}, \ldots, A_{NT}, B_{NT}, \varphi_{NT}), \alpha_0 = (A^0_1, B^0_1, \varphi^0_1, \ldots, A^0_N, B^0_N, \varphi^0_N), \)

\[
(C_k^0)^2 = (A_k^0)^2 + (B_k^0)^2 > 0, k = 1, N.
\]

The convergence in (57) can be a.s. In turn, form (57) it follows (see cited papers)

\[
A_{iT} \xrightarrow{P} A^0_i, B_{iT} \xrightarrow{P} B^0_i, T (\varphi_{iT} - \varphi^0_i) \xrightarrow{P} 0, \text{ as } T \to \infty. \tag{58}
\]

Note that

\[
T^{-1/2} d_{3k-2,T}(\alpha_0), T^{-1/2} d_{3k-1,T}(\alpha_0) \to \frac{1}{2}, T^{-3/2} d_{3k,T}(\alpha_0) \to \frac{1}{6} \left( (A_k^0)^2 + (B_k^0)^2 \right),
\]

as \( T \to \infty, \tag{59} \)

\( k = 1, N. \)

From (58) and (59) we obtain the relation of condition \( C_1 \) for trigonometric regression:

\[
T^{-1/2} d_T(\alpha_0)(\widehat{\alpha}_T - \alpha_0) \xrightarrow{P} 0, \text{ as } T \to \infty.
\]

To check the fulfillment of the condition \( C_2 \) for regression function (56) we get

\[
|A_i \cos \varphi_i t + B_i \sin \varphi_i t - A^0_i \cos \varphi^0_i t - B^0_i \sin \varphi^0_i t| \leq |A_i - A^0_i| + |B_i - B^0_i| + (|A^0_i| + |B^0_i|) t |\varphi_i - \varphi^0_i|,
\]

\( k = 1, N, \) and therefore

\[
\Phi_T(\alpha_0) \leq 3N \sum_{i=1}^N \left( T (A_i - A^0_i)^2 + T (B_i - B^0_i)^2 + \frac{1}{3} (|A^0_i| + |B^0_i|)^2 T^3 (\varphi_i - \varphi^0_i)^2 \right).
\]

Using again the relation (59) we arrive at the inequality of the condition \( C_2. \)

\[
\Phi_T(\alpha, \alpha_0) \leq c_0 \| d_T(\alpha_0)(\alpha - \alpha_0) \|^2, \alpha \in A^c. \tag{61}
\]

with constant \( c_0 \) depending on \( A^0_i, B^0_i, i = 1, N. \)

The next lemma is the main part of the convergence (57) proof.

**Lemma 11** Under condition \( A_1 \)

\[
\xi(T) = \sup_{\lambda \in \mathbb{R}} \left| T^{-1} \int_0^T e^{-i\lambda t} \varepsilon(t) dt \right| \xrightarrow{P} 0, \text{ as } T \to \infty. \tag{62}
\]

**Proof** Since

\[
\left( \int_0^T e^{-i\lambda t} \varepsilon(t) dt \right)^2 = \int_{-T}^T e^{-i\lambda u} \int_0^{T-|u|} \varepsilon(t+|u|) \varepsilon(t) dt du = 2 \int_0^T \int_0^{T-u} \cos \lambda u \varepsilon(t+u) \varepsilon(t) dt du,
\]

\( \square \) Springer
then
\[ E\xi^2(T) \leq 2T^{-2} \int_0^T \left| \int_0^{T-u} \varepsilon(t+u)\varepsilon(t) \, dt \right| \, du \leq 2T^{-2} \int_0^T K^\frac{1}{2}(u) \, du. \]

By formula (7)
\[ K(u) = \int_0^T \int_0^{T-u} E\varepsilon(t+u)\varepsilon(s+u)\varepsilon(t)\varepsilon(s) \, dt \, ds = \int_0^T \int_0^{T-u} c_4(t+u, s+u, t, s) \, dt \, ds \]
\[ + (T-u)^2 B^2(u) + \int_0^T \int_0^{T-u} B^2(t-s) \, dt \, ds \]
\[ + \int_0^T \int_0^{T-u} B(t-s+u)B(t-s-u) \, dt \, ds \]
\[ \leq K_1(u) + K_2(u) + K_3(u) + |K_4(u)|, \]
and
\[ E\xi^2(T) \leq 2T^{-2} \int_0^T \left( K_1^\frac{1}{2}(u) + K_2^\frac{1}{2}(u) + K_3^\frac{1}{2}(u) + |K_4(u)|^\frac{1}{2} \right) \, du. \quad (63) \]

By formula (8)
\[ K_1(u) = d_4 \int_\mathbb{R} \left( \int_0^{T-u} \hat{a}(t+u-r)\hat{a}(t-r) \right)^2 \, dr \]
\[ \leq d_4 \int_\mathbb{R} \left( \int_0^{T-u} \hat{a}^2(t+u-r) \, dt \int_0^{T-u} \hat{a}^2(t-r) \, dt \right) \, dr \]
\[ \leq d_4 \|\hat{a}\|^4_2 \int_0^{T-u} \hat{a}^2(t+u-r) \, dr \leq d_4 \|\hat{a}\|^4_2 (T-u), \]
that is
\[ T^{-2} \int_0^T K_1^\frac{1}{2}(u) \, du \leq d_4 \|\hat{a}\|^2_2 T^{-2} \int_0^T \sqrt{T-u} \, du = \frac{2}{3} d_4 \|\hat{a}\|^2_2 T^{-\frac{3}{2}}. \quad (64) \]

Obviously,
\[ T^{-2} \int_0^T K_2^\frac{1}{2}(u) \, du = T^{-2} \int_0^T (T-u)|B(u)| \, du \leq 3^{-\frac{1}{2}} \|B\|_2 T^{-\frac{1}{2}}, \quad (65) \]
\[ T^{-2} \int_0^T K_3^\frac{1}{2}(u) \, du \leq \frac{2}{3} \|B\|_2 T^{-\frac{1}{2}}, \quad (66) \]

\( \Box \) Springer
\[ T^{-2} \int_0^T K_4^1(u)du = T^{-2} \int_0^T \left( \frac{1}{2} \int_0^{T-u} \int_0^{T-u} (B^2(t-s+u) + B^2(t-s-u)) \, dt \, ds \right)^{\frac{1}{2}} \, du \]
\[ \leq \frac{2}{3} \|B\|_2 T^{-\frac{1}{2}}. \quad (67) \]

From inequalities (63)–(67) it follows
\[ E \xi^2(T) = O \left( T^{-\frac{1}{2}} \right), \quad \text{as} \ T \to \infty. \]

The result of the lemma can be strengthened to a.s. convergence in (62). Note also that in the proof we did not use the condition \( \hat{a} \in L_1(\mathbb{R}) \).

**Appendix B: LSE asymptotic normality**

Cumbersome sets of conditions on the behavior of the nonlinear regression function are used in the proofs of the LSE asymptotic normality of the model parameter can be found, for example, in Ivanov and Leonenko (1989); Ivanov (1997); Ivanov et al. (2015), and it does not make sense to write here all of them. We will comment only on the conditions associated with the proof of the CLT for one weighted integral of the linear process \( \varepsilon \) in the observation model (1).

Consider the family of the matrix-valued measures \( \mu_T(dx; \alpha) = \left( \mu_{jl}^T(dx; \alpha) \right)_{j,l=1}^q \), \( T > T_0, \alpha \in A \), with densities
\[ \mu_{jl}^T(x; \alpha) = g_j^T(x, \alpha) \left( \int_{\mathbb{R}} \left| g_j^T(x, \alpha) \right|^2 dx \int_{\mathbb{R}} \left| g_l^T(x, \alpha) \right|^2 dx \right)^{-\frac{1}{2}}, \quad x \in \mathbb{R}, \quad (68) \]

where
\[ g_j^T(x, \theta) = \int_0^T e^{ixt} g_j(t, \theta) \, dt, \quad j = 1, q. \]

(1) Suppose that the weak convergence \( \mu_T \Rightarrow \mu \) as \( T \to \infty \) holds, where \( \mu_T \) is defined by (68) and \( \mu \) is a positive definite matrix measure.

This condition means that the element \( \mu_{jl}^T \) of the matrix-valued measure \( \mu \) are complex measures of bounded variation, and the matrix \( \mu(A) \) is non-negative definite for any set \( A \in \mathcal{Z} \), with \( \mathcal{Z} \) denoting the \( \sigma \)-algebra of Lebesgue measurable subsets of \( \mathbb{R} \), and \( \mu(\mathbb{R}) \) is positive definite matrix, (see, for example, Ibragimov and Rozanov 1978).

The following definition can be found in Grenander (1954), Grenander and Rosenblatt (1984), Ibragimov and Rozanov (1978) and Ivanov and Leonenko (1989).

**Definition B.1** The positive-definite matrix-valued measure \( \mu(dx; \alpha) = \left( \mu_{jl}^{ij}(dx; \alpha) \right)_{j,l=1}^q \) is said to be the spectral measure of regression function \( g(t, \alpha) \).
Practically the components \( \mu_{jl}(x; \alpha) \) are determined from the relations
\[
R_{jl}(h; \alpha) = \lim_{T \to \infty} d^{-1}_j(\alpha) d^{-1}_l(\alpha) \int_0^T g_j(t + h, \alpha) g_l(t, \alpha) dx = \int_{\mathbb{R}} e^{i\lambda h} \mu_{jl}(d\lambda; \alpha), \quad j, l = 1, q, \tag{69}
\]
where it is supposed that the matrix function \( (R_{jl}(h; \alpha)) \) is continuous at \( h = 0 \).

Continuing Example A.2 with the trigonometric regression function \((56)\) from “Appendix A”, we can state using \((69)\) that the function \( g(t, \alpha) \) has a block-diagonal spectral measure \( \mu(d\lambda; \alpha) \) (see e.g., Ivanov et al. 2015) with blocks
\[
\begin{pmatrix}
\varphi_k & i \rho_k \overline{B}_k \\
-i \rho_k \overline{A}_k & \varphi_k \\
\beta_k & \gamma_k & \varphi_k \\
\end{pmatrix}, \quad k = 1, N, \tag{70}
\]
where
\[
\beta_k = \frac{\sqrt{3}}{2C_k} (B_k \varphi_k + iA_k \rho_k), \quad \gamma_k = \frac{\sqrt{3}}{2C_k} (-A_k \varphi_k + iB_k \rho_k), \quad C_k = \sqrt{A_k^2 + B_k^2}, \quad k = 1, N.
\]
In \((70)\) the measure \( \varphi_k = \varphi_k(d\lambda) \) and the signed measure \( \rho_k = \rho_k(d\lambda) \) are concentrated at the points \( \pm \varphi_k \), and \( \varphi_k(\{\pm \varphi_k\}) = \frac{1}{2}, \rho_k(\{\pm \varphi_k\}) = \pm \frac{1}{2} \).

Returning to the general case let the parameter \( \alpha \in A \) of regression function \( g(t, \alpha) \) be fixed. We will use the notation \( d^{-1}_i(\alpha) g_i(t, \alpha) = b_i T(t, \alpha) \) and condition
\[
(2) \quad \sup_{t \in [0, T]} |b_i T(t, \alpha)| \leq c_i T^{-\frac{1}{2}}, \quad i = 1, q.
\]
The next CLT is an important part of the proof of LSE \( \hat{\alpha}_T \) asymptotic normality in the model \((1)\) and fully uses condition \( A_1 \).

**Theorem B.1** Under conditions \( A_1, 1) \) and \( (2) \) the vector
\[
\zeta_T = d^{-1}(\alpha) \int_0^T \epsilon(t) \nabla g(t, \alpha) dt = \left( \int_0^T \epsilon(t) b_i T(t, \alpha) dt \right)_i \quad \tag{71}
\]
is asymptotically, as \( T \to \infty \), normal \( N(0, \Sigma) \),
\[
\Sigma = 2\pi \int_{\mathbb{R}} f(\lambda) \mu(d\lambda; \alpha) = d_2 \int_{-\infty}^{\infty} |a(\lambda)|^2 \mu(d\lambda; \alpha).
\]

**Proof** For any \( z = (z_1, \ldots, z_q) \in \mathbb{R}^q \) set
\[
\eta_T = \langle \zeta_T, z \rangle = \int_0^T \epsilon(t) S_T(t) dt, \quad S_T(t) = \sum_{i=1}^q b_i T(t, \alpha) z_i.
\]
By condition \((1)\)
\[
\sigma^2(z) = \lim_{T \to \infty} \operatorname{E} \eta_T^2 = 2\pi \int_{\mathbb{R}} f(\lambda) \mu_z(d\lambda; \alpha),
\]
\begin{цвет}[цвет3]
\end{цвет}
\[ \mu_z(d\lambda; \alpha) = \sum_{i,j=1}^{q} \mu^{ij}(d\lambda; \alpha)z_i z_j. \]

To prove the theorem it is sufficient to show for any \( z \in \mathbb{R} \) and \( \nu \geq 1 \), that

\[
\lim_{T \to \infty} E \eta^n_T = E \eta^n = \begin{cases} (n - 1)! \sigma^n(z), & n = 2\nu, \\ 0, & n = 2\nu + 1. \end{cases} \tag{72} \]

Use the Leonov–Shiryaev formula (see, e.g., Ivanov and Leonenko 1989). Let

\[ I = \{1, 2, \ldots, n\}, \quad I_p = \{i_1, \ldots, i_{l_p}\} \subset I, \quad c(I_p) = c_{l_p}(t_{i_1}, \ldots, t_{i_{l_p}}). \]

Then using formula \( (8) \), we have

\[
m(I) = m_n(t_1, \ldots, t_n) = \sum_{A_r} \prod_{p=1}^{r} c(I_p), \tag{73} \]

where \( \sum \) denotes summation over all unordered partitions \( A_r = \left\{ \bigcup_{p=1}^{r} I_p \right\} \) of the set \( I \) into sets \( I_1, \ldots, I_r \) such that \( I = \bigcup_{p=1}^{r} I_p, I_i \cap I_j = \emptyset, i \neq j \).

Since

\[
E \eta^n_T = \int_{[0, T]^n} m_n(t_1, \ldots, t_n) \prod_{k=1}^{n} R_T(t_k) dt_1 \ldots dt_n, \tag{74} \]

then the application of formula \( (73) \) to \( (74) \) shows that to obtain \( (72) \) it is sufficient to prove

\[
I(l) = \int_{[0, T]^l} c_i(t_1, \ldots, t_l) \prod_{k=1}^{l} R_T(t_k) dt_1 \ldots dt_l \to 0, \quad \text{as} \ T \to \infty. \tag{75} \]

for all \( l = 3, n \). Taking into account the equality \( E \varepsilon(t) = 0 \), from \( (75) \) will follow that in \( (72) \) all the odd moments \( E \eta^{2\nu+1} = 0 \). On the other hand, for even moments \( E \eta^{2\nu} \) we shall find that in \( (74) \) thanks to \( (73) \) only those terms correspond to the partitions of the set \( I = \{1, 2, \ldots, 2\nu\} \) into pairs of indices will remain nonzero, i.e. “Gaussian part” : all \( l_p = 2 \). In \( (73) \) it will be \( (2\nu - 1)!! \) of such terms and each of them will be equal to \( \sigma^{2\nu}(z) \).

Let us prove \( (75) \). We note that condition \( (2) \) implies

\[
\sup_{t \in [0, T]} |R_T(t)| \leq \|c\| \|z\| T^{-\frac{1}{2}}, \quad c = (c_1, \ldots, c_q), \quad z = (z_1, \ldots, z_q). \]

Then using formula \( (8) \) we have

\[
|I(l)| = \left| \int_{[0, T]^l} c_i(t_1 - t_i, \ldots, t_{l-1} - t_1, 0) \prod_{k=1}^{l} R_T(t_k) dt_1 \ldots dt_l \right| \leq |d_l| \int_{\mathbb{R}} ds \int_{[0, T]^l} \left| \prod_{i=1}^{l-1} \hat{a} (t_i - t_i - s) \right| |\hat{a}(-s)| \prod_{k=1}^{l} R_T(t_k) dt_1 \ldots dt_l \leq |d_l| \int_{\mathbb{R}} |\hat{a}(-s)| \int_{0}^{T} \left( \int_{0}^{T} \left| \hat{a} (t - t_i - s) R_T(t) \right| dt \right)^{l-1} |R_T(t_i)| dt ds.
\]
\[
\leq |d_t| \left( \|c\|^{l-1} \|z\|^{l-1} \|\hat{a}\|_1 \right) T^{-\frac{1}{T}} \left( \|c\| \|z\| T^\frac{1}{2} \right) = \\
= |d_t| \left( \|c\| \|z\| \|\hat{a}\|_1 \right) T^{-\left(\frac{1}{2}-1\right)} \to 0, \text{ as } T \to \infty, \ l \geq 3.
\]

To obtain (76) we have used \( \hat{a} \in L_1(\mathbb{R}) \) only.

Using the theorem, just as in the works cited above (for definiteness, we turn our attention to Ivanov et al. 2015), it can be proved that, if a number of additional conditions on the regression function are satisfied, the normalized LSE \( d_T(\alpha_0)(\hat{\alpha}_T - \alpha_0) \) is asymptotically normal \( N(0, \Sigma_{LSE}) \), with

\[
\Sigma_{LSE} = 2\pi \left( \int_{\mathbb{R}} \mu(d\lambda; \alpha_0) \right)^{-1} \left( \int_{\mathbb{R}} f(\lambda)\mu(d\lambda; \alpha_0) \left( \int_{\mathbb{R}} \mu(d\lambda; \alpha_0) \right)^{-1} \right) d_2 \left( \int_{\mathbb{R}} \mu(d\lambda; \alpha_0) \right)^{-1} \left( \int_{\mathbb{R}} |a(\lambda)|^2\mu(d\lambda; \alpha_0) \left( \int_{\mathbb{R}} \mu(d\lambda; \alpha_0) \right)^{-1} \right).
\]

Note that, firstly, our conditions \( N_3, (1), (2) \) are included in the conditions for the LSE asymptotic normality of Ivanov et al. (2015), and, secondly, the trigonometric regression function (56) satisfies the conditions of Ivanov et al. (2015). Moreover, using (70) and (59) we conclude that for the trigonometric model the normalized LSE

\[
T^{\frac{1}{2}} (A_{1T} - A_1^0), T^{\frac{1}{2}} (B_{1T} - B_1^0), T^{\frac{1}{2}} (\varphi_{1T} - \varphi_1^0), \ldots, \\
T^{\frac{1}{2}} (A_{NT} - A_N^0), T^{\frac{1}{2}} (B_{NT} - B_N^0), T^{\frac{1}{2}} (\varphi_{NT} - \varphi_N^0)
\]

is asymptotically normal \( N(0, \Sigma_{TRIG}) \), where \( \Sigma_{TRIG} \) is a block diagonal matrix with blocks

\[
\begin{pmatrix}
(A_0^0)^2 + 4(B_0^0)^2 & -3A_0^0B_0^0 & -6B_0^0 \\
-3A_0^0B_0^0 & (B_0^0)^2 + 4(A_0^0)^2 & 6A_0^0 \\
-6B_0^0 & 6A_0^0 & 12
\end{pmatrix}.
\]

The matrix \( \Sigma_{TRIG} \) is positive definite, if \( f(\varphi_k^0) > 0, k = 1, N \). However it follows from our condition \( A_2(iii) \).

Note also that condition \( N_2 \) is satisfied, for example, for the trigonometric regression function (56). Indeed, in this case

\[
g'(t, \alpha) = \sum_{i=1}^{N} (-\varphi_i A_i \sin \varphi_i t + \varphi_i B_i \cos \varphi_i t),
\]

and similarly to (60)

\[
|-\varphi_i A_i \sin \varphi_i t + \varphi_i B_i \cos \varphi_i t + \varphi_i^0 A_i^0 \sin \varphi_i^0 t - \varphi_i^0 B_i^0 \cos \varphi_i^0 t| \\
\leq \varphi (|A_i - A_i^0| + |B_i - B_i^0|) + (|A_i^0| + |B_i^0|) (1 + \varphi_t) |\varphi_i - \varphi_i^0|, \ i = 1, N
\]

which leads to the inequality of condition \( N_2 \) similar to (61), but with a different constant \( c_0 \).
Appendix C: Levitan polynomials

Some necessary facts of approximation theory adapted to needs of this article are represented in this Appendix. All the definitions and results are taken from the book (Akhiezer 1965).

In complex analysis entire function of exponential type is said to be such a function \( F(z) \) that for any complex \( z \) the inequality
\[
F(z) \leq Ae^{B|z|}
\] (77)
holds true, where the numbers \( A \) and \( B \) do not depend on \( z \). Infimum \( \sigma \) of the constant \( B \) values for which inequality (77) takes place is called the exponential type of function \( F(z) \) and can be determined by formula
\[
\sigma = \limsup_{|z| \to \infty} \frac{\ln |F(z)|}{|z|}.
\]

Denote by \( \mathcal{B}_\sigma \) the totality of all the entire functions \( F(z) \) of exponential type \( \leq \sigma \) with property \( \sup_{\lambda \in \mathbb{R}} |F(\lambda)| < \infty \).

Let \( \mathcal{C} \) be linear normed space of bounded continuous functions \( \varphi(\lambda), \lambda \in \mathbb{R} \), with norm \( \| \varphi \| = \sup_{\lambda \in \mathbb{R}} |\varphi(\lambda)| < \infty \). Consider further some set of functions \( \mathcal{M} \subset \mathcal{C} \). For the function of interest \( \varphi \in \mathcal{M} \) suppose that
\[
\lim_{\eta \to 1} \sup_{\lambda \in \mathbb{R}} |\varphi(\eta \lambda) - \varphi(\lambda)| = 0,
\] (78)
and write
\[
\mathcal{A}_\sigma[\varphi] = \inf_{F \in \mathcal{B}_\sigma} \| \varphi - F \|.
\]

Let \( h(\lambda), \lambda \in \mathbb{R} \), be uniformly continuous function. Denote by
\[
\omega(\delta) = \omega(\delta; h) = \sup_{|\lambda_1 - \lambda_2| \leq \delta} |h(\lambda_1) - h(\lambda_2)|, \quad \delta > 0,
\]
the modulus of continuity of the function \( h \). Obviously \( \omega(\delta), \delta > 0 \), is nondecreasing continuous function tending to zero, as \( \delta \to 0 \).

Let the set \( \mathcal{M} \) introduced above consists of differentiable functions such that for \( \varphi \in \mathcal{M} \) the derivatives \( \varphi'(\lambda) = h(\lambda) \), \( \lambda \in \mathbb{R} \), are uniformly continuous on \( \mathbb{R} \). Then for function \( \varphi \) satisfying the property (78) there exists a function \( F_\sigma \subset \mathcal{B}_\sigma \) such that (see Akhiezer 1965, p. 252)
\[
\mathcal{A}_\sigma[\varphi] = \| \varphi - F_\sigma \| \leq \frac{3}{\sigma} \omega \left( \frac{1}{\sigma}; h \right).
\] (79)

The inequality (79) means that for the described function \( \varphi \) and any \( \delta > 0 \) there exists a number \( \sigma = \sigma(\delta) \) and a function \( F_\sigma \in \mathcal{B}_\sigma \) such that
\[
\| \varphi - F_\sigma \| < \delta.
\]

As it has been proved in the 40s of the 20th century by B.M. Levitan for any function \( F \in \mathcal{B}_\sigma \) it is possible to build a sequence of trigonometric sums \( T_n(F; z), n \geq 1 \), bounded on \( \mathbb{R} \) by the same constant as the function \( F \), that converges to \( F(z) \) uniformly in any bounded part of the complex plane. In particular, for any compact set \( K \subset \mathbb{R} \)
\[
\lim_{n \to \infty} \sup_{\lambda \in K} |F(\lambda) - T_n(F; \lambda)| = 0.
\]
Put $s = \frac{\sigma}{n}$, $n \in \mathbb{N}$; $c^{(n)}_j = sE_s(js)$, $j \in \mathbb{Z}$;

$$E_s(x) = (2\pi)^{-1} \int_{\mathbb{R}} e^{-ixu} \left( \frac{2 \sin \frac{su}{2}}{su} \right)^2 F(u) du, \ x \in \mathbb{R}.$$ 

Then the sequence of the Levitan polynomials that corresponds to $F$ can be written as

$$T_n(F; z) = \sum_{j=-n}^{n} c^{(n)}_j e^{ijsz}.$$

References

Akhiezer NI (1965) Lections on approximation theory. Nauka, Moscow (in Russian)
Bentkus R (1972) Asymptotic normality of an estimate of the spectral function. Liet Mat Rink 3(12):5–18
Bentkus R (1972) On the error of the estimate of the spectral function of a stationary process. Liet Mat Rink 1(12):55–71

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.