PARAMETER ESTIMATION FOR RECIPROCAL GAMMA ORNSTEIN–UHLENBECK TYPE PROCESSES

UDC 519.21

N. LEONENKO, L. SAKHNO, AND N. ŠUVAK

ABSTRACT. We consider parameter estimation for a process of Ornstein–Uhlenbeck type with reciprocal gamma marginal distribution, to be called reciprocal gamma Ornstein–Uhlenbeck (RGOU) process. We derive minimum contrast estimators of unknown parameters based on both the discrete and the continuous observations from the process as well as moments based estimators based on discrete observations. We prove that proposed estimators are consistent and asymptotically normal. The explicit forms of the asymptotic covariance matrices are determined by using the higher order spectral densities and cumulants of the RGOU process.

Анотація. Розглянуто задачу оцінювання параметрів процесу Орнштейна–Уленбека, що має обернений гамма розподіл. Одержано оцінки параметрів методом мінімального контрасту та методом моментів, встановлено їх консистентність та асимптотичну нормальність.

Аннотация. Рассматривается задача оценивания параметров для процесса Орнштейна–Уленбека с обратным гамма распределением. Получены оценки параметров методом минимального контраста и методом моментов, доказана их консистентность и асимптотическая нормальность.

1. INTRODUCTION

Non-Gaussian processes of Ornstein–Uhlenbeck (OU) type have been introduced by Barndorff-Nielsen and Shephard in 2001 [10]. One of the main advantages of this type of processes is their ability to capture important distributional deviations from normality. Therefore, they are widely applied for modeling time series coming from different fields of science, with an accent to financial mathematics (stochastic volatility models). Also, in comparison to other models (e.g. geometric Gaussian and CEV models), these processes are much more analytically tractable and therefore very suitable as practical models.

This paper deals with the Lévy driven stationary OU type process with reciprocal or inverse gamma marginal distributions, a heavy-tailed distribution which belongs to the Pearson family containing Gaussian, gamma, beta, reciprocal gamma, Fisher-Snedecor and Student distributions as six subclasses (see [25]). The detailed study of Student OU type processes can be found in [17], as well a good motivation for application of this OU type process. Also note that reciprocal gamma distribution is a special case of generalized inverse Gaussian (GIG) distribution and that OU-type processes with GIG marginals are already used in practice. However, the application of the OU-type processes with reciprocal gamma marginals needs further investigation.

In this paper we focus on parameter estimation for our process of interest. Estimation of parameters of Lévy driven OU type processes is already a well studied topic. For example, in [19] nonparametric inference was developed, in [30], [29] maximum likelihood estimation procedures were analyzed while in [28] moment based estimators and their asymptotic properties were studied. Estimation of the parameters of this type of processes based on the sampled AR(1) process is treated in [12]. In the present paper parameters of reciprocal gamma OU type process are estimated by both minimum contrast method based on Whittle and Ibragimov functionals and method of moments. Derived estimators are proved to be consistent and asymptotically normal. For moment based estimators the consistency follows from the ergodicity and stationarity of the process while asymptotic normality of the bivariate estimator of parameters of the marginal distribution follows from the functional central limit theorem for α-mixing sequences (see, e.g., [13]) and the functional delta method (see [27]). Methodology for proving these asymptotic properties in the case of minimum contrast estimators is given in the Appendix of the paper.

The main advantage of these two methods of estimation is that, based on higher order spectra and cumulants of the observed process, it is possible to calculate the explicit form of the asymptotic covariances in the asymptotic normality framework. This makes an important problem of constructing asymptotic confidence intervals for unknown parameters operational and significantly increases applicability of this process in practice.

We shall use the following notations for characteristic functions and cumulant transforms for a random variable (r.v.) X:

$$\phi_X(\zeta) = \mathsf{E}\left\{e^{i\zeta X}\right\}, \qquad \kappa_X(\zeta) = \log \mathsf{E}\left\{e^{i\zeta X}\right\}.$$

We shall use the notation $X \stackrel{d}{=} Y$ for equality of distributions of two random variables, or $X_t \stackrel{d}{=} Y_t$ for equality of the finite-dimensional distributions of stochastic processes.

2. The reciprocal gamma distribution and Lévy processes

We present basic results from the theory of reciprocal or inverse gamma distribution.

2.1. **Density and characteristic functions.** If the r.v. $\gamma_{\alpha,\beta}$ has gamma distribution with probability density function (pdf) of the form

$$\mathfrak{g}(x) = \begin{cases} \frac{\alpha^\beta}{\Gamma(\beta)} x^{\beta-1} e^{-\alpha x}, & x > 0, \\ 0, & x \leq 0, \end{cases} \tag{1}$$

where $\alpha > 0$ is scale parameter and $\beta > 0$ is shape parameter, then the r.v. $r_{\alpha,\beta} = \frac{1}{\gamma_{\alpha,\beta}}$ has reciprocal or inverse gamma distribution with pdf

$$\mathfrak{rg}(x) = \begin{cases} \frac{\alpha^\beta}{\Gamma(\beta)} x^{-\beta-1} e^{-\frac{\alpha}{x}}, & x > 0, \\ 0, & x \leq 0, \end{cases} \tag{2}$$

with the same parameters α and β. These distributions will be denoted by $\gamma_{\alpha,\beta} \sim \mathfrak{G}(\alpha,\beta)$, $r_{\alpha,\beta} \sim \mathfrak{RG}(\alpha,\beta)$, respectively.

Important property of gamma and reciprocal gamma r.v. is their scaling property. Note that $\gamma_{\alpha,\beta} \stackrel{d}{=} \frac{\gamma_{1,\beta}}{\alpha}$ and $\mathsf{E}\{\gamma_{\alpha,\beta}\} = \frac{\beta}{\alpha}$, $\mathsf{E}\{\gamma_{\alpha,\beta}^2\} = \frac{\beta(\beta+1)}{\alpha^2}$. Also, $r_{\alpha,\beta} \stackrel{d}{=} \alpha \cdot r_{1,\beta}$. In this case only moments less then β exist for $\mathfrak{RG}(\alpha,\beta)$ and if so:

$$\mathsf{E}\{r_{\alpha,\beta}\} = \frac{\alpha}{\beta-1}, \ \beta > 1; \quad \mathrm{Var}\{r_{\alpha,\beta}\} = \frac{\alpha^2}{(\beta-1)^2(\beta-2)}, \ \beta > 2.$$

Both pdf's (1) and (2) have tails of Pareto type and they belong to the class of the generalized inverse Gaussian distributions (GIG) (see, e.g., [10, Section 2.2.]). In particular, when $\alpha = n/2$, $\beta = n/2$ and $n \geq 1$ is an integer, the gamma random variable $\gamma_{\frac{n}{2},\frac{n}{2}} \stackrel{d}{=} \frac{1}{n}\gamma_{\frac{1}{2},\frac{n}{2}} \stackrel{d}{=} \frac{1}{n}\chi_n^2$, where χ_n^2 has chi-square distribution with n degrees of freedom.

The expectation of reciprocal gamma r.v. exists when $\beta > 1$, the variance exists when $\beta > 2$, the n-th moments exists when $\beta > n$ and can be calculated using the formula:

$$\mathsf{E}\{r_{\alpha,\beta}^n\} = \frac{\alpha^n}{\prod_{i=1}^n (\beta - i)} = \alpha^n \frac{\Gamma(\beta - n)}{\Gamma(\beta)}, \qquad \beta > n. \tag{3}$$

According to Witkovsky [31], characteristic function of r.v. $X \sim \mathfrak{RG}(\alpha, \beta)$ is

$$\phi_{\mathfrak{RG}(\zeta)} = \mathsf{E}[e^{i\zeta X}] = \frac{2\alpha^\beta \cdot (-i\zeta\alpha^{-1})^{\beta/2} \cdot K_\beta\left[2\alpha(-i\zeta\alpha^{-1})^{1/2}\right]}{\Gamma(\beta)}, \qquad \zeta \in \mathbb{R}, \tag{4}$$

where $K_\beta(\cdot)$ denotes the modified Bessel function of the third kind (see, e.g., [1]).

2.2. **Infinite divisibility and self-decomposability.** It is known that the generalized hyperbolic distributions (GHD) are infinite divisible (ID) and self-decomposable (SD) (see, e.g., [10] and references therein). Thus, the reciprocal gamma distribution has to be infinitely divisible and self-decomposable.

Recall (see [10]) that a r.v. X is SD if its characteristic function $\phi(\zeta)$ has the property that for every $c \in (0,1)$ there exists a characteristic function $\phi_c(\zeta)$ such that

$$\phi(\zeta) = \phi(c\zeta) \cdot \phi_c(\zeta) \tag{5}$$

for all $\zeta \in \mathbb{R}$. Property (5) means that for any $c \in (0,1)$ the exists a r.v. X_c such that $X \overset{d}{=} cX + X_c$, where random variables X and X_c are independent.

All SD characteristic functions (i.e., SD r.v.) are ID. It means that for every $n \geq 1$ there exists a characteristic function $\phi_n(\zeta)$, such that $\phi(\zeta) = [\phi_n(\zeta)]^n$, $\zeta \in \mathbb{R}$.

A stochastically continuous process $\{L(t), t \geq 0\}$, $L(0) = 0$, with strictly stationary and independent increments is called the (homogeneous) Lévy process. Usually we may choose a version with cádlág paths or with paths (a.s.) from the Skorokhod space of cádlág functions (see [11] or [26]). Then the law of $L(t)$ is determined by the law of $L(1)$ which is ID. The independence and stationarity of the increments of the Lévy process means that the cumulant transforms are

$$\kappa_{L(t)}(\zeta) = t \cdot \kappa_{L(1)}(\zeta), \qquad t \geq 0, \ \zeta \in \mathbb{R}. \tag{6}$$

Familiar special classes of Lévy process are Brownian motion and the compound Poisson processes. All Lévy processes except for Brownian motion have jumps.

We say that $X(t)$ has the scaling property if for each $c \in (0,1)$ there exists a nonrandom function $M(c)$ such that

$$X(ct) \overset{d}{=} M(c) \cdot X(t), \qquad t \geq 0, \tag{7}$$

For every infinitely divisible random variable T there exists a Lévy process such that $L(1) \overset{d}{=} X$. The following theorem is proved by Jurek [20].

Theorem 2.1. *Let $L(t)$ be a Lévy process with the strong Markov property. That is, for any independent r.v. $T \geq 0$ $L(t + T) - L(T)$ and $L(t)$ have the same probability distributions. Assume the scaling property (7) holds. If T is SD and the (nonrandom) scaling function $M(c)$ is a homeomorphism of the unit interval, then $L(T)$ is SD.*

From Theorem 2.1 we obtain the following result.

Theorem 2.2. *The reciprocal gamma distribution $\mathfrak{RG}(\alpha, \beta)$ with probability density function (2) and characteristic function (4) is self-decomposable (and infinitely divisible).*

3. Reciprocal gamma Ornstein–Uhlenbeck type processes

Based on the theory of non-Gaussian Ornstein–Uhlenbeck-based models (see [7], [10]) we can introduce the reciprocal gamma Ornstein–Uhlenbeck type processes.

The key result is known as the random integral representation (see [21, Theorem 3.9.3.] and the bibliographic comments there and [10, formula (12)]).

Theorem 3.1. *The random variable X has self-decomposable distribution if and only if there exists a Lévy process $Y(t)$ such that*

$$\mathsf{E}\left[\log(1 + |Y(1)|)\right] < +\infty \quad and \quad X \overset{d}{=} \int_0^\infty e^{-s}\, dY(s). \tag{8}$$

Then Y is unique in distribution. Moreover, if the cumulant transform $\kappa_X(\zeta)$ is differentiable for $\zeta \neq 0$ and $\zeta \cdot \kappa_X(\zeta) \to 0$ for $0 \neq \zeta \to 0$, then we have

$$\kappa_{Y(1)}(\zeta) = \zeta \cdot \frac{d}{d\zeta}\kappa_X(\zeta). \tag{9}$$

$Y(t)$ is referred to as the background driving Lévy process (BDLP) corresponding to X.

From (4), (9) and properties of $K_\beta(\cdot)$ we obtain the following expression for the cumulant transform of the BDLP $Y(t)$ for the reciprocal gamma r.v. $\mathfrak{RG}(\alpha, \beta)$:

$$\kappa_{Y(1)}(\zeta) = \log \mathsf{E}\exp\{i\zeta Y(1)\} = \frac{\zeta i}{\sqrt{-i\zeta\alpha^{-1}}} + \frac{K_{\beta-1}\left(2\alpha\sqrt{\frac{-i\zeta}{\alpha}}\right)}{K_\beta\left(2\alpha\sqrt{\frac{-i\zeta}{\alpha}}\right)}, \tag{10}$$

$$\zeta \in \mathbb{R}, \qquad \zeta \neq 0,$$

and $\kappa_{Y(1)}(0) = 0$.

A stochastic process $X(t)$ is said to be of OU type if it satisfies a stochastic differential equation (SDE) of the form

$$dX_t = -\theta X_t\, dt + dY(\theta t), \tag{11}$$

where $Y(t)$ is BDLP.

From Theorem 3.1 and [10, Theorem 1] using (10) we arrive to the following statement:

Theorem 3.2. *There exists a strictly stationary stochastic process $\{X_t, t \in \mathbb{R}\}$ which has marginal reciprocal gamma distribution $\mathfrak{RG}(\alpha, \beta)$ with density function (2) and BDLP $Y(t)$ with cumulant transform (10), such that:*

(i) *X_t satisfies the SDE*

$$X_t = e^{-\theta t}X_0 + e^{-\theta t}\int_0^t e^{\theta s}\, dY(\theta s) = \int_{-\infty}^t e^{-\theta(t-s)}\, dY(\theta s), \qquad \theta > 0,\ t \in \mathbb{R}, \tag{12}$$

and

(ii) *if $\beta > 1$ the expectation $\mathsf{E}\{X_t\} = \frac{\alpha}{\beta-1}$ and if $\beta > 2$ the autocorrelation function*

$$\rho(\tau) = \mathrm{Corr}\left(X_{t+\tau}, X_t\right) = e^{-\theta|\tau|}, \qquad \tau \in \mathbb{R}, \tag{13}$$

and the spectral density $f(\lambda)$ of $\{X_t, t \in \mathbb{R}\}$ is of the form

$$f(\lambda) = \frac{\alpha^2}{\pi(\beta-1)^2(\beta-2)}\frac{\theta}{(\theta^2+\lambda^2)}, \qquad \lambda \in \mathbb{R},\ \beta > 2, \tag{14}$$

while the discretized version of the spectral density $f_d(\lambda)$, $\lambda \in [-\pi, \pi)$, obtained through the discretization $X_t^d = X_{th}$, $t \in Z$, $h \in \mathbb{R}$, is of the form

$$f(\lambda) = \sum_{k \in Z}\frac{\alpha^2}{h\pi(\beta-1)^2(\beta-2)}\frac{\theta}{\left(\theta^2 + \left(\frac{\lambda+2k\pi}{h}\right)^2\right)}, \qquad \lambda \in [-\pi, \pi). \tag{15}$$

The stationary process $\{X_t, t \in \mathbb{R}\}$ is referred to as the reciprocal gamma OU (RGOU) type process.

Remark 3.1. *According to the general result [24, Theorem 4.3], RGOU type process is strong or α-mixing with an exponentially decaying mixing coefficient. Furthermore, this α-mixing property together with the stationarity of the process imply that the process is ergodic. For properties of Lévy driven OU processes see also [28, Theorem 2.5].*

Remark 3.2. *The higher-order cumulant functions and spectral densities of the stochastic process (12) can be obtained (see [2] for details). We present here the expressions for the spectral densities. Under the assumption that $\mathsf{E}\,|X_t|^p < \infty$ (that is, $\beta > p$ in the case of RGOU-type process X_t) the spectral density of order $2 \leq m \leq p$ can be represented as*

$$f_m(\lambda_1, \ldots, \lambda_{m-1}) = \frac{\theta m k_{m,X}}{(2\pi)^{m-1}} \frac{1}{(\theta + i\lambda_1) \cdots (\theta + i\lambda_{m-1})(\theta - i(\lambda_1 + \ldots + \lambda_{m-1}))}, \quad (16)$$

where $k_{m,X}$ is the m-th cumulant of the marginal reciprocal gamma distribution of X_t.

In the same manner the stationary autoregressive processes with marginal reciprocal gamma distribution can be constructed based on the SD of the reciprocal gamma distribution. Let $X_0 = X$ be a r.v. with reciprocal gamma distribution and we define an autoregressive sequence $(X_n, n \geq 1)$ of the first order by equality

$$X_{n+1} = cX_n + \varepsilon_{n+1}, \qquad n \geq 1, \ 0 < c < 1, \tag{17}$$

where $\{\varepsilon_n\}_{n=0}^{\infty}$ are i.i.d.r.v. (a so called innovation process) and independent of $\{X_n\}_{n=1}^{\infty}$.
If the expression (17) holds for all $c \in (0, 1)$, then X_0 is a SD r.v.
In other words, the self-decomposable reciprocal gamma distribution $X_0 \sim \mathfrak{RG}(\alpha, \beta)$ can be realized as the marginal distribution of an autoregressive sequence. In fact,

$$\varepsilon_1 \overset{d}{=} \int_0^{-\log c} e^{-s} \, dY(s),$$

with the BDLP $Y(t)$ of X_0, and the cumulant transform of BDLP $Y(t)$ is given by (10).
Thus, there exists a stationary (in strict sense) solution X_n of the autoregressive equation (17) with the following properties: $X_n \sim \mathfrak{RG}(\alpha, \beta)$ and

$$\mathsf{E}\{X_n\} = \frac{\alpha}{\beta - 1}, \qquad \text{Corr}\,(X_n, X_{n+\tau}) = c^{|\tau|}, \quad 0 < c < 1, \ \tau = 0, \pm 1, \ldots, \ \beta > 2.$$

Remark 3.3. *A stationary process with pdf (2) and the correlation function $e^{-\theta|\tau|}$, $\tau \in \mathbb{R}$, can be obtained as a unique Markovian weak solution of the following SDE:*

$$dX_t = -\theta \left(X_t - \frac{\alpha}{\beta - 1} \right) dt + \sqrt{\frac{2\theta}{\beta - 1} X_t^2} \, dB_t, \qquad t \geq 0,$$

where $\{B(t), t \geq 0\}$ is a standard Brownian motion or Wiener process (see [17], [8]).

4. Parameter estimation

4.1. **Minimum contrast estimation.** The minimum contrast estimation theory provided in Appendix A is applicable to the RGOU process, if we suppose $\beta > 4$. Namely, we can use the Whittle functional (38) to estimate parameters α, β, θ, or the Ibragimov functional (20) to estimate the parameter θ and the variance of the reciprocal gamma marginal distribution of the process X_t.

4.1.1. *Whittle contrast.* Basing on the observations $\{X_t, t \in D_T\}$ we consider the Whittle objective function

$$U_T(\alpha, \beta, \theta) = \frac{1}{4\pi} \int_{\mathbb{R}^1} \left(\log f(\lambda; \alpha, \beta, \theta) + \frac{I_T(\lambda)}{f(\lambda; \alpha, \beta, \theta)} \right) w(\lambda) \, d\lambda, \qquad (18)$$

where $I_T(\lambda)$ is the periodogram (39) and we can choose the weight function $w(\lambda) = \frac{1}{1+\lambda^2}$.

Assuming that the true value δ_0 of the parameter vector $\delta = (\alpha, \beta, \theta)$ belongs to a compact set $\Delta \subset (0, \infty) \times (4, \infty) \times (0, \infty)$, we construct the estimate

$$\widehat{\delta}_T = \left(\widehat{\alpha}_T, \widehat{\beta}_T, \widehat{\theta}_T \right) = \underset{\delta \in \Delta}{\arg \min} \, U_T(\alpha, \beta, \theta), \qquad (19)$$

and according to Theorem A.1 this estimate is consistent.

Furthermore, according to Theorem A.2 the estimate (19) is asymptotically normal:

$$T^{1/2} \left(\widehat{\delta}_T - \delta_0 \right) \xrightarrow{\mathcal{D}} \mathcal{N} \left(0, W_1^{-1}(\delta_0) (W_2(\delta_0) + V(\delta_0)) W_1^{-1}(\delta_0) \right), \qquad T \to \infty,$$

where the elements of the matrices $W_1(\delta) = W_1(\alpha, \beta, \theta)$, $W_2(\delta) = W_2(\alpha, \beta, \theta)$ and $V(\delta) = V(\alpha, \beta, \theta)$ participating in the limiting normal law are of the following form:

$$w_{11}^{(1)}(\alpha, \beta, \theta) = \frac{1}{\alpha^2}; \qquad w_{22}^{(1)}(\alpha, \beta, \theta) = \frac{c^2}{4}; \qquad w_{33}^{(1)}(\alpha, \beta, \theta) = \frac{\theta^2 + 1}{4(\theta+1)^2};$$

$$w_{12}^{(1)}(\alpha, \beta, \theta) = w_{21}^{(1)}(\alpha, \beta, \theta) = \frac{c}{2\alpha}; \qquad w_{13}^{(1)}(\alpha, \beta, \theta) = w_{31}^{(1)}(\alpha, \beta, \theta) = \frac{1 - \theta}{2\alpha(1 + \theta)};$$

$$w_{23}^{(1)}(\alpha, \beta, \theta) = w_{32}^{(1)}(\alpha, \beta, \theta) = \frac{c(1 - \theta)}{4(1 + \theta)};$$

$$w_{11}^{(2)}(\alpha, \beta, \theta) = \frac{1}{2\alpha^2}; \qquad w_{22}^{(2)}(\alpha, \beta, \theta) = \frac{c^2}{8}; \qquad w_{12}^{(2)}(\alpha, \beta, \theta) = w_{21}^{(2)}(\alpha, \beta, \theta) = \frac{c}{4\alpha};$$

$$w_{33}^{(2)}(\alpha, \beta, \theta) = \frac{\theta^6 - 7\theta^4 + 12\theta^3 - 9\theta^2 + 4\theta - 1}{8(\theta^2 - 1)^3};$$

$$w_{13}^{(2)}(\alpha, \beta, \theta) = w_{31}^{(2)}(\alpha, \beta, \theta) = \frac{\theta^6 - 7\theta^4 + 12\theta^3 - 9\theta^2 + 4\theta - 1}{4\alpha(\theta^2 - 1)^3};$$

$$w_{23}^{(2)}(\alpha, \beta, \theta) = w_{32}^{(2)}(\alpha, \beta, \theta) = c \cdot \frac{\theta^6 - 7\theta^4 + 12\theta^3 - 9\theta^2 + 4\theta - 1}{8(\theta^2 - 1)^3};$$

$$v_{11}(\alpha, \beta, \theta) = v \frac{4\pi^2}{\alpha^2}; \quad v_{22}(\alpha, \beta, \theta) = v \pi^2 c^2;$$

$$v_{33}(\alpha, \beta, \theta) = v \frac{\pi^2 (1 - \theta)^2}{(1 + \theta)^2}; \qquad v_{12}(\alpha, \beta, \theta) = v_{21}(\alpha, \beta, \theta) = v \frac{\pi c}{4\alpha};$$

$$v_{13}(\alpha, \beta, \theta) = v_{31}(\alpha, \beta, \theta) = v \frac{\pi(1 - \theta)}{4\alpha(1 + \theta)}; \qquad v_{23}(\alpha, \beta, \theta) = v_{32}(\alpha, \beta, \theta) = v \frac{\pi c(1 - \theta)}{8(1 + \theta)}.$$

In the above formulas we have denoted

$$c = \frac{5 - 3\beta}{(\beta - 1)(\beta - 2)}, \qquad v = \frac{k_{4,X}}{\theta(k_{2,X})^2}.$$

4.1.2. *Ibragimov contrast.* To apply the Ibragimov objective function we calculate first

$$\sigma^2(\alpha, \beta, \theta) := \int_{\mathbb{R}^1} f(\lambda; \alpha, \beta, \theta) w(\lambda) \, d\lambda = \frac{1}{\pi} \frac{\alpha^2 \theta}{(\beta - 1)^2 (\beta - 2)} \int_{\mathbb{R}^1} \frac{1}{\theta^2 + \lambda^2} \frac{1}{1 + \lambda^2} \, d\lambda$$

$$= \frac{\alpha^2}{(\beta - 1)^2 (\beta - 2)(1 + \theta)},$$

and define

$$\psi(\lambda, \theta) = \frac{f(\lambda; \alpha, \beta, \theta)}{\sigma^2(\alpha, \beta, \theta)} = \frac{\theta(1 + \theta)}{\pi(\theta^2 + \lambda^2)},$$

so that we have

$$f(\lambda; \alpha, \beta, \theta) = \sigma^2(\alpha, \beta, \theta)\psi(\lambda, \theta).$$

Basing on the observations $\{X_t, t \in D_T\}$ we consider the objective function

$$U_T(\theta) = -\int_{\mathbb{R}} I_T(\lambda)w(\lambda)\log\psi(\lambda; \theta)\, d\lambda, \qquad \theta \in \Theta. \tag{20}$$

where $I_T(\lambda)$ is the periodogram (39) and the weight function $w(\lambda) = \frac{1}{1+\lambda^2}$. (Note that in this case we have that the objective function depends on the parameter θ only.)

Assuming that the true value θ_0 of the parameter θ belongs to a compact set $\Theta \subset (0, \infty)$, we construct the estimate

$$\widehat{\theta}_T = \arg\min_{\theta \in \Theta} U_T(\theta), \tag{21}$$

and according to Theorem A.3 this estimator is a consistent estimator of the parameter θ, as well as the estimator

$$\widehat{\sigma}_T^2 = \int_{\mathbb{R}^1} I_T(\lambda)w(\lambda)\, d\lambda$$

is a consistent estimator of the parameter $\sigma^2(\alpha, \beta, \theta)$.

Note that $\sigma^2(\alpha, \beta, \theta) = \mathrm{Var}(r_{\alpha,\beta}) \cdot \frac{1}{1+\theta}$, therefore, having the estimators $\widehat{\theta}_T$ and $\widehat{\sigma}_T^2$ we can also obtain the estimator for the variance of the underlying reciprocal gamma marginal distribution as $\widehat{v}_T = (1 + \widehat{\theta}_T)\widehat{\sigma}_T^2$, and, as a result, we can obtain the expressions (estimated) for the spectral density and covariance function of our process X_t, as for these we need only to know θ and $Var(r_{\alpha,\beta})$ (see, Theorem 3.2 (ii)).

Furthermore, according to Theorem A.2 the estimate (19) is asymptotically normal:

$$T^{1/2}(\widehat{\theta}_T - \theta_0) \xrightarrow{\mathcal{D}} \mathcal{N}(0, a_0 s_0^{-2}), \qquad T \to \infty,$$

where $a_0 = a(\alpha_0, \beta_0, \theta_0)$ and $s_0 = s(\alpha_0, \beta_0, \theta_0)$ are given by the following formulas:

$$s(\alpha, \beta, \theta) = \sigma^2(\alpha, \beta, \theta)\int_{\mathbb{R}} w(\lambda)\left[\frac{\partial^2}{\partial\theta^2}\psi(\lambda, \theta) - \frac{1}{\psi(\lambda, \theta)}\left(\frac{\partial}{\partial\theta}\psi(\lambda, \theta)\right)^2\right]d\lambda,$$

$$a(\alpha, \beta, \theta) = 2\pi\left(\sigma^2(\alpha, \beta, \theta)\right)^2\left[2\int_{\mathbb{R}} w^2(\lambda)\left(\frac{\partial}{\partial\theta}\psi(\lambda; \theta)\right)^2 d\lambda\right.$$

$$\left. + \frac{k_{4,X}}{\theta(k_{2,X})^2}\left\{\int_{\mathbb{R}} w(\lambda)\frac{\partial}{\partial\theta}\psi(\lambda; \theta)\, d\lambda\right\}^2\right]$$

and can be calculated in the closed form.

Here we have the situation where the spectral density is of the form

$$f(\lambda; \alpha, \beta, \theta) = h(\alpha, \beta)\,\theta\,g(\lambda; \theta) \tag{22}$$

and applying estimation procedure based on the Ibragimov contrast function we are able to estimate the parameter θ only. The approach exists (see, [5]) to deal with such a type of spectral densities, for which some parameters appear only in a multiplicative term (as $h(\alpha, \beta)$ above), and the spectral densities of orders > 2 again have these parameters involved only in multiplicative terms. Within this approach it is assumed that taking into consideration the spectral densities of the 2-d and 3-d orders (in which parameters of interest (α, β) appears in multiplicative term $l(\alpha, \beta)$, say), one can obtain first the estimators for $h(\alpha, \beta)$ and $l(\alpha, \beta)$ and then from these estimators derive the estimators for α and β. However in the case under consideration the functional expressions $h(\alpha, \beta)$ and $l(\alpha, \beta)$ are not quite suitable to proceed with this method.

4.1.3. *Minimum contrast estimation procedure under discretization.* In practical situations observations are often made at discrete intervals of time, even for continuous time processes. Also, it is easier to make all necessary calculations basing on computationally simpler discrete periodogram and corresponding contrast functionals.

Consider the discrete time process

$$X_t^d = X_t, \qquad t \in \mathbb{Z}.$$

The corresponding spectral density can be defined as follows (see Theorem 3.2):

$$\tilde{f}(\lambda; \alpha, \beta, \theta) = \sum_{k=-\infty}^{\infty} f(\lambda + 2k\pi; \alpha, \beta, \theta) = \frac{\alpha^2}{\pi(\beta-1)^2(\beta-2)} \sum_{k=-\infty}^{\infty} \frac{\theta}{\theta^2 + (\lambda + 2k\pi)^2},$$

$\lambda \in [-\pi, \pi)$. Define the discrete version of the Whittle functional

$$U_N(\alpha, \beta, \theta) = \frac{1}{2N} \sum_{k=1}^{N-1} \left(\log \tilde{f}(\lambda_k; \alpha, \beta, \theta) + \frac{I_T(\lambda_k)}{\tilde{f}(\lambda_k; \alpha, \beta, \theta)} \right), \tag{23}$$

where

$$N = [T], \qquad \lambda_k = \frac{2\pi k}{N}, \qquad I_N(\lambda) = \frac{1}{2\pi N} \left| \sum_{t=1}^{N-1} X_t e^{-it\lambda} \right|^2, \quad \lambda \in [-\pi, \pi).$$

Note that in practice we can use as an approximation for $\tilde{f}(\lambda)$ the function

$$\tilde{f}_M(\lambda) = \sum_{k=-M}^{M} f(\lambda + 2k\pi; \alpha, \beta, \theta), \qquad \lambda \in [-\pi, \pi)$$

for large enough M. As $M \to \infty$, $\tilde{f}_M(\lambda)$ converges uniformly to $\tilde{f}(\lambda) = \tilde{f}(\lambda; \alpha, \beta, \theta)$ for $\lambda \neq 0$, $(\alpha, \beta, \theta) \in \Delta$.

The Whittle contrast estimator can be defined as

$$\widehat{\delta}_N = \left(\widehat{\alpha}_N, \widehat{\beta}_N, \widehat{\theta}_N \right) = \arg \min_{\delta \in \Delta} U_N(\alpha, \beta, \theta) \tag{24}$$

with $U_N(\alpha, \beta, \theta)$ given by (23). The estimate $\widehat{\delta}_N$ can be used instead of $\widehat{\delta}_T$ given by (19). It can be shown that $\widehat{\delta}_N$ is consistent and asymptotically normal, namely

$$N^{1/2}\left(\widehat{\delta}_N - \delta_0 \right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \tilde{W}^{-1}(\delta_0) \left(\tilde{W}(\delta_0) + \tilde{V}(\delta_0) \right) \tilde{W}^{-1}(\delta_0) \right), \qquad N \to \infty,$$

where the matrices $\tilde{W}(\delta)$ and $\tilde{V}(\delta)$ are of the following form

$$\tilde{W}(\delta) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \left(\frac{\partial}{\partial \delta} \log \tilde{f}(\lambda; \delta) \right) \left(\frac{\partial}{\partial \delta} \log \tilde{f}(\lambda; \delta) \right)^t d\lambda$$

and

$$\tilde{V}(\delta) = \frac{1}{8\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \frac{\tilde{f}_4(\lambda_1, -\lambda_1, \lambda_2, ; \delta)}{\tilde{f}(\lambda_1; \delta) \tilde{f}(\lambda_2; \delta)} \left(\frac{\partial}{\partial \delta} \log \tilde{f}(\lambda; \delta) \right) \left(\frac{\partial}{\partial \delta} \log \tilde{f}(\lambda; \delta) \right)^t d\lambda_1 \, d\lambda_2.$$

The estimator $\widehat{\delta}_N$ can be obtained in practice as the solution of

$$U_N'\left(\widehat{\delta}_N \right) = 0,$$

where

$$U_N'(\delta) = U_N'(\alpha, \beta, \theta) = \left(\frac{\partial}{\partial \alpha} U_N(\alpha, \beta, \theta), \frac{\partial}{\partial \beta} U_N(\alpha, \beta, \theta), \frac{\partial}{\partial \theta} U_N(\alpha, \beta, \theta) \right)$$

and

$$\frac{\partial}{\partial \alpha} U_N \left(\alpha, \beta, \theta \right) = \frac{1}{2N} \sum_{k=1}^{N-1} \left(1 - \frac{I_T \left(\lambda_k \right)}{\tilde{f} \left(\lambda_k; \alpha, \beta, \theta \right)} \right) \frac{1}{\tilde{f} \left(\lambda_k; \alpha, \beta, \theta \right)} \frac{\partial \tilde{f} \left(\lambda_k; \alpha, \beta, \theta \right)}{\partial \alpha},$$

$$\frac{\partial}{\partial \beta} U_N \left(\alpha, \beta, \theta \right) = \frac{1}{2N} \sum_{k=1}^{N-1} \left(1 - \frac{I_T \left(\lambda_k \right)}{\tilde{f} \left(\lambda_k; \alpha, \beta, \theta \right)} \right) \frac{1}{\tilde{f} \left(\lambda_k; \alpha, \beta, \theta \right)} \frac{\partial \tilde{f} \left(\lambda_k; \alpha, \beta, \theta \right)}{\partial \beta}, \quad (25)$$

$$\frac{\partial}{\partial \theta} U_N \left(\alpha, \beta, \theta \right) = \frac{1}{2N} \sum_{k=1}^{N-1} \left(1 - \frac{I_T \left(\lambda_k \right)}{\tilde{f} \left(\lambda_k; \alpha, \beta, \theta \right)} \right) \frac{1}{\tilde{f} \left(\lambda_k; \alpha, \beta, \theta \right)} \frac{\partial \tilde{f} \left(\lambda_k; \alpha, \beta, \theta \right)}{\partial \theta}.$$

We write down the derivatives used for calculation of the elements of the matrices $\tilde{W} \left(\delta \right)$ and $\tilde{V} \left(\delta \right)$ and the expressions (25):

$$\frac{\partial \tilde{f} \left(\lambda; \alpha, \beta, \theta \right)}{\partial \alpha} = \frac{2}{\alpha} \tilde{f} \left(\lambda; \alpha, \beta, \theta \right), \quad \frac{\partial \tilde{f} \left(\lambda; \alpha, \beta, \theta \right)}{\partial \beta} = \frac{5 - 3\beta}{(\beta - 1)(\beta - 2)} \tilde{f} \left(\lambda; \alpha, \beta, \theta \right),$$

$$\frac{\partial \tilde{f} \left(\lambda; \alpha, \beta, \theta \right)}{\partial \theta} = \sum_{k=-\infty}^{\infty} \frac{\alpha^2}{\pi(\beta - 1)^2(\beta - 2)} \cdot \frac{\theta}{\theta^2 + (\lambda + 2k\pi)^2} \cdot \frac{(\lambda + 2k\pi)^2 - \theta^2}{\theta \left(\theta^2 + (\lambda + 2k\pi)^2 \right)}$$

$$= \sum_{k=-\infty}^{\infty} f(\lambda + 2k\pi; \alpha, \beta, \theta) \frac{(\lambda + 2k\pi)^2 - \theta^2}{\theta \left(\theta^2 + (\lambda + 2k\pi)^2 \right)}.$$

Analogously the discrete version of the Ibragimov contrast can be considered.

4.2. Method of moments estimation.

4.2.1. *Estimation of parameter θ.* Parameter $\theta > 0$ is estimated under the assumption that parameters α and β of the RGOU process are known. Existence of the first two moments of the marginal reciprocal gamma distribution is ensured by assuming $\beta > 2$.

Let us consider the sample $X_1, \ldots, X_T$ from the stationary RGOU process and the corresponding sample of paired observations $(X_1, X_{t+1}), (X_2, X_{t+2}), \ldots, (X_{T-t}, X_T)$, where $t < T$. The consistent estimator of the autocorrelation function (13) is given by the Pearson's sample correlation coefficient

$$\widehat{\rho}_T(t) = \frac{\frac{1}{T-t} \sum_{i=1}^{T-t} X_i X_{t+i} - \frac{1}{T-t} \sum_{i=1}^{T-t} X_i \cdot \frac{1}{T-t} \sum_{i=1}^{T-t} X_{t+i}}{\sqrt{\frac{1}{T-t} \sum_{i=1}^{T-t} X_i^2 - \left(\frac{1}{T-t} \sum_{i=1}^{T-t} X_i \right)^2} \sqrt{\frac{1}{T-t} \sum_{i=1}^{T-t} X_{t+i}^2 - \left(\frac{1}{T-t} \sum_{i=1}^{T-t} X_{t+i} \right)^2}}.$$
(26)

From the expression (13) we see that the autocorrelation function $\rho(t)$ of the RGOU process takes only positive values. However, according to [15, Theorem 2.1.] the sum of the sample autocorrelation function at lag $t \geq 1$ is always $-1/2$ for any stationary time series with arbitrary length $T \geq 2$, i.e.

$$S_{acf} = \sum_{t=1}^{T-1} \widehat{\rho}_T(t) = -\frac{1}{2}.$$

Therefore, instead of the Pearson's sample correlation coefficient (26) we observe here its absolute value $|\widehat{\rho}_T(t)|$. Since $|\widehat{\rho}_T(t)|$ is the continuous transformation of the consistent estimator (26), according to the continuity mapping theorem it is the consistent estimator of the autocorrelation function (13) for any fixed $t > 0$, i.e.

$$|\widehat{\rho}_T(t)| \xrightarrow{\mathrm{P}} \rho(t), \qquad T \to \infty, \quad \text{for any fixed } t > 0. \tag{27}$$

Therefore, here we observe the equation $|\widehat{\rho}_T(t)| = \rho(t)$, where $\rho(t) = e^{-\theta t}$. Solving this estimation equation in terms of the unknown parameter θ results in the estimator $\widehat{\theta}$ of

the autocorrelation parameter θ, i.e.

$$\widehat{\theta} = -\frac{1}{t} \log |\widehat{\rho}_T(t)|. \tag{28}$$

The continuity of the logarithmic function and the relation (27) imply that $\widehat{\theta}$ is the consistent estimator of the autocorrelation parameter θ, i.e. $\widehat{\theta} \xrightarrow{\text{P}} \theta$, $T \to \infty$.

4.2.2. *Estimation of the parameters α and β.* In this section we present method of moments estimators of parameters α and β of marginal distribution of stationary RGOU process and analyze their asymptotic properties. We will suppose that the parameter θ is known. Existence of the second moment of the marginal distribution and assumptions of the central limit theorem for α-mixing processes is ensured by assuming $\beta > 4$.

Equating first two theoretical moments, $\mathsf{E}[X_t]$ and $\mathsf{E}[X_t^2]$, with corresponding empirical counterparts

$$\overline{m}_1 = \frac{1}{n} \sum_{t=1}^{n} X_t, \qquad \overline{m}_2 = \frac{1}{n} \sum_{t=1}^{n} X_t^2. \tag{29}$$

and solving the resulting system of equations in terms of unknown parameters α and β results in method of moments estimators of parameters α and β:

$$\widehat{\alpha} = \frac{\overline{m}_1 \cdot \overline{m}_2}{\overline{m}_2 - \overline{m}_1^2} = g_1(\overline{m}_1, \overline{m}_2), \qquad \widehat{\beta} = 1 + \frac{\overline{m}_2}{\overline{m}_2 - \overline{m}_1^2} = g_2(\overline{m}_1, \overline{m}_2). \tag{30}$$

From expressions (30) we see that estimators $\widehat{\alpha}$ and $\widehat{\beta}$ are continuous transformations of the empirical counterparts of the first and the second moment given by (29). In particular, $\widehat{\alpha} = g_1(\overline{m}_1, \overline{m}_2)$ and $\widehat{\beta} = g_2(\overline{m}_1, \overline{m}_2)$. Therefore, asymptotic properties of the estimator $(\widehat{\alpha}, \widehat{\beta})$ are implied by asymptotic properties of the estimator $(\overline{m}_1, \overline{m}_2)$:

- RGOU process is strictly stationary and, according to [24, Theorem 4.3], α-mixing with exponential decaying rate. These two properties imply ergodicity of the RGOU process. Therefore, according to the ergodic theorem for stationary sequences (see [22]), $(\overline{m}_1, \overline{m}_2)$ is the consistent estimator of the $(\mathsf{E}[X_t], \mathsf{E}[X_t^2])$
- According to the functional central limit theorem for α-mixing sequences (see [14]; for multidimensional version [13]), $(\overline{m}_1, \overline{m}_2)$ is asymptotically normal, i.e.

$$\boldsymbol{\Sigma}^{-1/2} \sqrt{n} \, (\overline{m}_1 - \mathsf{E}[\overline{m}_1], \overline{m}_2 - \mathsf{E}[\overline{m}_2]) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{I}), \tag{31}$$

where $\mathsf{E}[\overline{m}_1] = \mathsf{E}[X_t] = \frac{\alpha}{\beta-1}$, $\mathsf{E}[\overline{m}_2] = \mathsf{E}[X_t^2] = \frac{\alpha^2}{(\beta-1)(\beta-2)}$.

Moreover, asymptotic covariance matrix $\boldsymbol{\Sigma}$ is explicitly calculated by using the cumulants $c_2(s,t)$, $c_3(s,t,u)$ and $c_4(s,t,u,v)$ of the RGOU process defined in terms of the higher order spectral densities. In particular, due to the stationarity of the process, expectations

$$\mathsf{E}[X_{s+t}X_s], \quad \mathsf{E}[X_{s+t}X_s^2] \quad \text{and} \quad \mathsf{E}[X_{s+t}^2 X_s^2] \tag{32}$$

needed for calculation of the asymptotic covariance matrix Σ are given by

$$\mathsf{E}\left[X_s X_t\right] = c_2(t-s, 0) + \mu^2,$$
$$\mathsf{E}\left[X_s^2 X_t\right] = c_3(t-s, t-s, 0) + 2\mu c_2(t-s, 0) - \mu \, \mathsf{E}[X_t^2],$$
$$\mathsf{E}\left[X_s^2 X_t^2\right] = c_4(t-s, t-s, 0, 0) + 2\mu \left(\mathsf{E}[X_s X_t^2] + \mathsf{E}[X_s^2 X_t]\right) - 8\mu^2 \, \mathsf{E}[X_s X_t]$$
$$+ 2 \left(\mathsf{E}[X_s X_t]\right)^2 + 6\mu^4 - 4\mu^2 \, \mathsf{E}[X_t^2] + \left(\mathsf{E}[X_t^2]\right)^2,$$

where $\mu = \mathsf{E}[X_t] = \alpha/(\beta - 1)$ and for $t > s$

$$c_2(t - s, 0) = \int_{-\pi}^{\pi} e^{i(t-s)\lambda} f_2(\lambda)\, d\lambda,$$

$$c_3(t - s, t - s, 0) = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} e^{i(t-s)(\lambda_1+\lambda_2)} f_3(\lambda_1, \lambda_2)\, d\lambda_1\, d\lambda_2,$$

$$c_4(t - s, t - s, 0, 0) = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} e^{i(t-s)(\lambda_1+\lambda_2)} f_4(\lambda_1, \lambda_2, \lambda_3)\, d\lambda_1\, d\lambda_2\, d\lambda_3.$$

Therefore

$$\mathsf{E}\left[X_s X_t\right] = \int_{-\pi}^{\pi} \cos\left(\lambda(t - s)\right) f_2(\lambda)\, d\lambda + \frac{\alpha^2}{(\beta - 1)^2},$$

$$\mathsf{E}\left[X_s^2 X_t\right] = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} e^{i(t-s)(\lambda_1+\lambda_2)} f_3(\lambda_1, \lambda_2)\, d\lambda_1\, d\lambda_2 + \frac{2\alpha}{\beta - 1} \int_{-\pi}^{\pi} \cos\left(\lambda(t - s)\right) f_2(\lambda)\, d\lambda$$

$$- \frac{\alpha^3}{(\beta - 1)^2(\beta - 2)}$$

$$= \mathsf{E}\left[X_s X_t^2\right],$$

$$\mathsf{E}\left[X_s^2 X_t^2\right] = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} e^{i(t-s)(\lambda_1+\lambda_2)} f_4(\lambda_1, \lambda_2, \lambda_3)\, d\lambda_1\, d\lambda_2\, d\lambda_3$$

$$+ \frac{2\alpha}{\beta - 1} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \left(e^{i(t-s)(\lambda_1+\lambda_2)} + e^{i(s-t)(\lambda_1+\lambda_2)}\right) f_3(\lambda_1, \lambda_2)\, d\lambda_1\, d\lambda_2$$

$$+ \frac{4\alpha^2}{(\beta - 1)^2} \int_{-\pi}^{\pi} e^{i(s-t)\lambda} f_2(\lambda)\, d\lambda + 2 \left(\int_{-\pi}^{\pi} e^{i(s-t)\lambda} f_2(\lambda)\, d\lambda\right)^2$$

$$+ \frac{\alpha^4}{(\beta - 1)^2(\beta - 2)^2}.$$

Now we calculate the following covariances:

$$\mathrm{Cov}\left(\sqrt{n}\overline{m}_1, \sqrt{n}\overline{m}_1\right) = 2\pi \int_{-\pi}^{\pi} \Phi_n(\lambda) f_2(\lambda)\, d\lambda,$$

$$\mathrm{Cov}\left(\sqrt{n}\overline{m}_1, \sqrt{n}\overline{m}_2\right) = (2\pi)^2 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \Psi_n^{(2)}(\lambda_1, \lambda_2) f_3(\lambda_1, \lambda_2)\, d\lambda_1\, d\lambda_2$$

$$+ \frac{4\alpha\pi}{\beta - 1} \int_{-\pi}^{\pi} \Phi_n(\lambda) f_2(\lambda)\, d\lambda,$$

$$\mathrm{Cov}\left(\sqrt{n}\overline{m}_2, \sqrt{n}\overline{m}_2\right) = (2\pi)^3 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \Psi_n^{(3)}(\lambda_1, \lambda_2, \lambda_3) f_4(\lambda_1, \lambda_2, \lambda_3)\, d\lambda_1\, d\lambda_2\, d\lambda_3$$

$$+ \frac{16\alpha\pi^2}{\beta - 1} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \Psi_n^{(2)}(\lambda_1, \lambda_2) f_3(\lambda_1, \lambda_2)\, d\lambda_1\, d\lambda_2$$

$$+ \frac{8\alpha^2\pi}{(\beta - 1)^2} \int_{-\pi}^{\pi} \Phi_n(\lambda) f_2(\lambda)\, d\lambda$$

$$+ 8\pi^2 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \Psi_n^{(2)}(\lambda_1, \lambda_2) f_2(\lambda_1) f_2(\lambda_2)\, d\lambda_1\, d\lambda_2,$$

where

$$\Phi_n(\lambda) = \frac{1}{2\pi n} \frac{\sin^2\left(\frac{\lambda n}{2}\right)}{\sin^2\left(\frac{\lambda}{2}\right)}$$

is the one-dimensional Féjer kernel and

$$\Psi_n^{(2)}(\lambda_1, \lambda_2) = \frac{1}{(2\pi)^2 n} \frac{\sin^2\left(\frac{(\lambda_1+\lambda_2)n}{2}\right)}{\sin^2\left(\frac{\lambda_1+\lambda_2}{2}\right)}, \qquad \Psi_n^{(3)}(\lambda_1, \lambda_2, \lambda_3) = \frac{1}{(2\pi)^3 n} \frac{\sin^2\left(\frac{(\lambda_1+\lambda_2)n}{2}\right)}{\sin^2\left(\frac{\lambda_1+\lambda_2}{2}\right)}$$

are two-dimensional and three-dimensional kernels (but not Féjer kernels), respectively.

Finally, according to [4, Proposition 1] it follows that as n tends to infinity

$$\text{Cov}\left(\sqrt{n}\overline{m}_1, \sqrt{n}\overline{m}_1\right) \to \frac{\alpha^2}{(\beta-1)^2(\beta-2)} f_2(0),$$

$$\text{Cov}\left(\sqrt{n}\overline{m}_1, \sqrt{n}\overline{m}_2\right) \to 4\pi^2 f_3(0,0) + \frac{2\alpha^3}{(\beta-1)^3(\beta-2)} f_2(0),$$

$$\text{Cov}\left(\sqrt{n}\overline{m}_2, \sqrt{n}\overline{m}_2\right) \to 8\pi^3 f_4(0,0,0) + \frac{16\alpha\pi^2}{\beta-1} f_3(0,0) + \frac{4\alpha^4}{(\beta-1)^4(\beta-2)} f_2(0)$$

$$+ \frac{2\alpha^4}{(\beta-1)^4(\beta-2)^2} f_2^2(0).$$

Therefore the elements of the asymptotic covariance matrix Σ are given by expressions

$$\sigma_{11}^2 = \frac{\alpha^2 \, \text{ctgh}\,(h\theta/2)}{(\beta-1)^2(\beta-2)}, \qquad \sigma_{12}^2 = 4\pi^2 f_3(0,0) + \frac{2\alpha^3 \, \text{ctgh}\,(h\theta/2)}{(\beta-1)^3(\beta-2)} = \sigma_{21}^2,$$

$$\sigma_{22}^2 = 8\pi^3 f_4(0,0,0) + \frac{16\alpha\pi^2}{\beta-1} f_3(0,0) + \frac{4\alpha^4 \, \text{ctgh}\,(h\theta/2)}{(\beta-1)^4(\beta-2)} + \frac{2\alpha^4 \, \text{ctgh}\,(h\theta/2)}{(\beta-1)^4(\beta-2)^2},$$

where $f_3(0,0)$ and $f_4(0,0,0)$ can be calculated from the expression (16).

Since $\alpha > 0$, $\beta > 4$ and since $\theta > 0$ is known, the matrix Σ is positive definite and therefore it is a well defined covariance matrix. This concludes the analysis of asymptotic properties of the estimator $(\overline{m}_1, \overline{m}_2)$. Asymptotic properties of the estimator $(\widehat{\alpha}, \widehat{\beta})$ are given in the following theorem.

Theorem 4.1. *Let $\{X_t, t \geq 0\}$ be the RGOU process with the unknown parameter (α, β), where $\alpha > 0$ and $\beta > 4$.*

(i) *$(\widehat{\alpha}, \widehat{\beta}) \xrightarrow{P} (\alpha, \beta)$, i.e. $(\widehat{\alpha}, \widehat{\beta})$ is the consistent estimator of the unknown parameter (α, β).*

(ii) *$\sqrt{n}\,[\Sigma(\alpha, \beta, \theta)]^{-1/2}\,(\widehat{\alpha} - \alpha, \widehat{\beta} - \beta) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{I})$, where $\Sigma(\alpha, \beta, \theta)$ is the asymptotic covariance matrix.*

(iii) *$\sqrt{n}\,\left[\Sigma(\widehat{\alpha}, \widehat{\beta}, \theta)\right]^{-1/2}\,(\widehat{\alpha} - \alpha, \widehat{\beta} - \beta) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{I})$, where θ is the known value of the autocorrelation parameter.*

(iv) *$\sqrt{n}\,\left[\Sigma(\widehat{\alpha}, \widehat{\beta}, \widehat{\theta})\right]^{-1/2}\,(\widehat{\alpha} - \alpha, \widehat{\beta} - \beta) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{I})$, where $\widehat{\theta}$ is the consistent estimator (28) of the autocorrelation parameter θ.*

Proof. Part(i). Estimators $\overline{m}_1$ and $\overline{m}_2$ given by (29) are consistent estimators of the first and the second moment, respectively. Since estimators $\widehat{\alpha}$ and $\widehat{\beta}$ are their continuous transformations, i.e. $(\widehat{\alpha}, \widehat{\beta}) = (g_1(\overline{m}_1, \overline{m}_2), g_2(\overline{m}_1, \overline{m}_2))$, according to the continuity mapping theorem it follows that $\widehat{\alpha} = g_1(\overline{m}_1, \overline{m}_2) \xrightarrow{P} g_1(\mathsf{E}[\overline{m}_1, \overline{m}_2]) = \alpha$, and $\widehat{\beta} = g_2(\overline{m}_1, \overline{m}_2) \xrightarrow{P} g_2(\mathsf{E}[\overline{m}_1, \overline{m}_2]) = \beta$, as $n \to \infty$. From here it follows that $(\widehat{\alpha}, \widehat{\beta}) \xrightarrow{P} (\alpha, \beta)$, $n \to \infty$, i.e. $(\widehat{\alpha}, \widehat{\beta})$ is the consistent estimator of the unknown parameter (α, β).

Part (ii). Since $\widehat{\alpha}$ and $\widehat{\beta}$ are the continuous transformations of $\overline{m}_1$ and $\overline{m}_2$, according to the multivariate delta method (see Serfling [27], Theorem 3.3.A.), the estimator $(\widehat{\alpha}, \widehat{\beta})$ is asymptotically normal, i.e.

$$\sqrt{n}(\widehat{\alpha} - \alpha, \widehat{\beta} - \beta) \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, D\Sigma D^\tau\right), \tag{33}$$

where $\alpha = g_1\left(\mathsf{E}[\overline{m}_1, \overline{m}_2]\right)$, $\beta = g_2\left(\mathsf{E}[\overline{m}_1, \overline{m}_2]\right)$ and (2×2) matrix

$$D = \begin{bmatrix} (\beta - 1)(2\beta - 3) & -\frac{(\beta-1)(\beta-2)^2}{\alpha} \\ \frac{2(\beta-1)^2(\beta-2)}{\alpha} & -\frac{(\beta-1)^2(\beta-2)^2}{\alpha^2} \end{bmatrix}.$$

Explicit forms of the elements $\sigma_{ij}^2(\alpha, \beta, \theta)$ of the covariance matrix $D\Sigma D^\tau = \Sigma(\alpha, \beta, \theta)$ are omitted in this proof due to their complicated form.

Part(iii). Part (ii) of this theorem and the positive definiteness of the covariance matrix $\Sigma(\alpha, \beta, \theta)$ imply that

$$[\Sigma(\alpha, \beta, \theta)]^{-1/2}\sqrt{n}\big(\widehat{\alpha} - \alpha, \widehat{\beta} - \beta\big) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{I}), \tag{34}$$

where $\mathbf{I}$ is the (2×2) identity matrix. Since $(\widehat{\alpha}, \widehat{\beta})$ is the consistent estimator of the parameter (α, β), it follows that

$$\big[\Sigma\big(\widehat{\alpha}, \widehat{\beta}, \theta\big)\big]^{-1/2}[\Sigma(\alpha, \beta, \theta)]^{1/2} \xrightarrow{\mathsf{P}} \mathbf{I}. \tag{35}$$

If we multiply $\big[\Sigma\big(\alpha, \beta, \theta\big)\big]^{-1/2}\sqrt{n}(\widehat{\alpha} - \alpha, \widehat{\beta} - \beta)$ from the left by the matrix

$$\big[\Sigma\big(\widehat{\alpha}, \widehat{\beta}, \theta\big)\big]^{-1/2}[\Sigma(\alpha, \beta, \theta)]^{1/2},$$

then (34), (35) and the Slutsky Lemma (see Serfling [27], Theorem 1.5.4.1.) imply that

$$\big[\Sigma\big(\widehat{\alpha}, \widehat{\beta}, \theta\big)\big]^{-1/2}[\Sigma(\alpha, \beta, \theta)]^{1/2}[\Sigma(\alpha, \beta, \theta)]^{-1/2}\sqrt{n}\big(\widehat{\alpha} - \alpha, \widehat{\beta} - \beta\big) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{I}).$$

Since $[\Sigma(\alpha, \beta, \theta)]^{1/2}[\Sigma(\alpha, \beta, \theta)]^{-1/2} = \mathbf{I}$, it follows that

$$\big[\Sigma\big(\widehat{\alpha}, \widehat{\beta}, \theta\big)\big]^{-1/2}\sqrt{n}\big(\widehat{\alpha} - \alpha, \widehat{\beta} - \beta\big) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{I}).$$

Part (iv). Since $\widehat{\theta}$ given by (28) is the consistent estimator of the unknown autocorrelation parameter θ, the proof follows by the same argument as the proof of (iii).   □

APPENDIX A. MINIMUM CONTRAST ESTIMATION OF LINEAR PROCESSES

We present here some parameter estimation theory for stationary linear processes

$$X_t = \int_{u \in \mathbb{R}} \widehat{a}(t - u)\,\xi(du), \qquad t \in \mathbb{R}^d, \tag{36}$$

where $\xi(A)$ is a set indexed Lévy process with finite second moments and stationary intensity proportional to the Lebesgue measure. For conditions which ensure that (36) is well-defined, see, e.g. [2, 6] and references therein. With the choice of an appropriate "Green function" $\widehat{a}(t)$, this very general class includes the solutions of many interesting differential equations with random noise $\xi(du)$, e.g., generalized OU processes in $\mathbb{R}$ [2].

Under the assumption that $\mathsf{E}\,|X_t|^p < \infty$, the spectral density of order $2 \le k \le p$ of the process (36) can be written down in the closed form as

$$f_k(\lambda_1, \ldots, \lambda_{k-1}) = d_k a\left(-\sum_{i=1}^{k-1}\lambda_i\right)\prod_{i=1}^{k-1}a(\lambda_i) = d_k\prod_{i=1}^{k}a(\lambda_i)\delta\left(\sum_{j=1}^{k}\lambda_j\right) \tag{37}$$

(if this complex-valued function belongs to $L_1(\mathbb{R}^{k-1})$), where $\lambda_i \in \mathbb{R}$, $1 \le i \le k - 1$, d_k is the k'th cumulant of $\xi([0, 1])$. We will denote $f(\lambda) = f_2(\lambda) = d_2 a(\lambda)a(-\lambda)$.

A.1. **Minimum contrast estimation based on the Whittle contrast function.**
The class of Whittle estimators is the most popular in applications (see, e.g. [16], [23] and references therein). We begin with the following assumption.
A.I. Let X_t, $t \in D_T = \left[-\frac{T}{2}, \frac{T}{2}\right]$, be observations of a real-valued measurable zero mean stationary linear process X_t, $t \in \mathbb{R}^1$, with the spectral density $f(\lambda, \theta)$, $\lambda \in \mathbb{R}^1$, $\theta \in \Theta \subset \mathbb{R}^m$, Θ is a compact, the true value of the parameter $\theta_0 \in \text{int}\,\Theta$, the interior of Θ. Suppose that $f(\lambda; \theta_1) \not\equiv f(\lambda; \theta_2)$ for $\theta_1 \neq \theta_2$, a.e. in $\mathbb{R}^1$ w.r.t. the Lebesgue measure.

Consider the Whittle contrast process (or objective function)

$$U_T(\theta) = \frac{1}{4\pi} \int_{\mathbb{R}^1} \left(\log f(\lambda; \theta) + \frac{I_T(\lambda)}{f(\lambda; \theta)} \right) w(\lambda) \, d\lambda, \tag{38}$$

where $I_T(\lambda)$ is the periodogram of the second order

$$I_T(\lambda) = \frac{1}{2\pi T} \left| \int_{D_T} X_t e^{-it\lambda} \, dt \right|^2, \qquad \lambda \in \mathbb{R}^1, \tag{39}$$

$w(\lambda)$ is an even function such that all considered integrals are well defined and which will satisfy some conditions given below; in some cases we can choose $w(\lambda) = (1 + \lambda^2)^{-1}$.

To state the result on consistency of the minimum contrast estimator based on the contrast process (38) we will need the following conditions.

A.II. $f(\lambda; \theta_0) w(\lambda) \frac{1}{f(\lambda; \theta)} \in L_1(\mathbb{R}^1) \cap L_2(\mathbb{R}^1)$, for all $\theta \in \Theta$.

A.III. There exists a function $v(\lambda)$, $\lambda \in \mathbb{R}^1$, such that the function $h(\lambda; \theta) = v(\lambda) \frac{1}{f(\lambda; \theta)}$ is uniformly continuous in $\mathbb{R}^1 \times \Theta$ and $f(\lambda; \theta_0) \frac{w(\lambda)}{v(\lambda)} \in L_1(\mathbb{R}^1) \cap L_2(\mathbb{R}^1)$.

Theorem A.1. *Let the assumptions A.I to A.III be satisfied. Then the minimum contrast estimator $\widehat{\theta}_T$ defined as*

$$\widehat{\theta}_T = \arg\min_{\theta \in \Theta} U_T(\theta) \tag{40}$$

is a consistent estimator of the parameter θ, that is, $\widehat{\theta}_T \to \theta_0$ in P_0-probability as $T \to \infty$.

For asymptotic normality of the estimator (40) we need, in addition:

A.IV. The function $\frac{1}{f(\lambda; \theta)}$ is twice differentiable in a neighborhood of the point θ_0 and

(i) $f(\lambda; \theta_0) w(\lambda) \frac{\partial^2}{\partial \theta_i \partial \theta_j} \frac{1}{f(\lambda; \theta)} \in L_1(\mathbb{R}^1) \cap L_2(\mathbb{R}^1)$, $i, j = 1, \ldots, m$, $\theta \in \Theta$;

(ii) the limiting normal law holds for $T^{1/2}(J_T(\phi_i) - \mathsf{E} \, J_T(\phi_i))$, where

$$J_T(\phi) = \int_{\mathbb{R}^1} I_T(\lambda) \phi(\lambda) \, d\lambda$$

and $\phi_i(\lambda) = w(\lambda) \frac{\partial}{\partial \theta_i} \frac{1}{f(\lambda; \theta)}$;

(iii) $T^{1/2} \int_{\mathbb{R}^1} (\mathsf{E} \, I_T(\lambda) - f(\lambda; \theta_0)) w(\lambda) \frac{\partial}{\partial \theta_i} \frac{1}{f(\lambda; \theta)} \, d\lambda \to 0$ as $T \to \infty$, $i = 1, \ldots, m$, $\theta \in \Theta$;

(iv) the second order derivatives $\frac{\partial^2}{\partial \theta_i \partial \theta_j} \frac{1}{f(\lambda; \theta)}$, $i = 1, \ldots, m$, are continuous in θ.

A.V. The matrices $W_1(\theta) = \left(w_{ij}^{(1)}(\theta)\right)_{i,j=1,\ldots,m}$, $W_2(\theta) = \left(w_{ij}^{(2)}(\theta)\right)_{i,j=1,\ldots,m}$, $V(\theta) = (v_{ij}(\theta))_{i,j=1,\ldots,m}$ are positive definite, where

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

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

$$v_{ij}(\theta) = \frac{1}{8\pi} \frac{d_4}{d_2^2} \int_{\mathbb{R}^1} w(\lambda) \frac{\partial}{\partial \theta_i} \log f(\lambda; \theta) d\lambda \int_{\mathbb{R}^1} w(\lambda) \frac{\partial}{\partial \theta_j} \log f(\lambda; \theta) \, d\lambda.$$

Theorem A.2. *Let the assumptions A.I to A.V be satisfied. Then as $T \to \infty$*

$$T^{1/2}\left(\widehat{\theta}_T - \theta_0\right) \xrightarrow{\mathcal{D}} \mathcal{N}_m\left(0, W_1^{-1}(\theta_0)(W_2(\theta_0) + V(\theta_0))W_1^{-1}(\theta_0)\right),$$

where $\mathcal{N}_m(\cdot, \cdot)$ denotes the m-dimensional Gaussian law.

A.2. **Minimum contrast estimation based on the Ibragimov contrast function.**
Consider the minimum contrast functional motivated by the paper [18], see also [3, 5].

We assume condition **A.I** and introduce the following condition

B. I. There exists a nonnegative even function $w(\lambda)$, $\lambda \in \mathbb{R}$, such that $w(\lambda)f(\lambda; \theta)$ is in $L_1(\mathbb{R})$ for $\forall \theta \in \Theta$.

Under the condition B.I, we set $\sigma^2(\theta) = \int_{\mathbb{R}} f(\lambda; \theta)w(\lambda)\, d\lambda$ and consider the factorization of the spectral density $f(\lambda; \theta) = \sigma^2(\theta)\psi(\lambda; \theta)$, $\lambda \in \mathbb{R}$, $\theta \in \Theta$. For the function $\psi(\lambda, \theta)$, $\lambda \in \mathbb{R}$, $\theta \in \Theta$, we have $\int_{\mathbb{R}} \psi(\lambda; \theta)w(\lambda)\, d\lambda = 1$ and we additionally suppose

B. II. The derivatives $\nabla_\theta \psi(\lambda; \theta)$ exist and

$$\nabla_\theta \int_{\mathbb{R}} \psi(\lambda; \theta)w(\lambda)\, d\lambda = \int_{\mathbb{R}} \nabla_\theta \psi(\lambda; \theta)w(\lambda)\, d\lambda = 0,$$

that is we can differentiate under the integral sign in the above integral.

Consider the following contrast process (or objective function):

$$U_T(\theta) = -\int_{\mathbb{R}} I_T(\lambda)w(\lambda)\log\psi(\lambda; \theta)\, d\lambda, \qquad \theta \in \Theta. \tag{41}$$

B. III. $f(\lambda; \theta_0)\, w(\lambda) \log \psi(\lambda; \theta) \in L_1(\mathbb{R}^1) \cap L_2(\mathbb{R}^1)$, $\forall \theta \in \Theta$.

B. IV. There exists a function $v(\lambda)$, $\lambda \in \mathbb{R}^1$, such that $h(\lambda; \theta) := v(\lambda)\log\psi(\lambda; \theta)$ is uniformly continuous in $\mathbb{R}^1 \times \Theta$; $f(\lambda; \theta_0)\frac{w(\lambda)}{v(\lambda)} \in L_1(\mathbb{R}^1) \cap L_2(\mathbb{R}^1)$.

Theorem A.3. *Let conditions AI, B.I - B.IV be satisfied. Then the minimum contrast estimator $\widehat{\theta}_T$ defined as*

$$\widehat{\theta}_T = \arg\min_{\theta \in \Theta} U_T(\theta), \tag{42}$$

is a consistent estimator of the parameter θ, that is, $\widehat{\theta}_T \to \theta_0$ in P_0-probability as $T \to \infty$, and the estimator $\widehat{\sigma}_T^2 = \int_{\mathbb{R}^n} I_T(\lambda)w(\lambda)d\lambda$ is a consistent estimator of the parameter $\sigma^2(\theta)$, that is, $\widehat{\sigma}_T^2 \to \sigma^2(\theta_0)$ in P_0-probability as $T \to \infty$.

For asymptotic normality of the estimator (42) further conditions are needed.

B. V. The function $\psi(\lambda; \theta)$ is twice differentiable in a neighborhood of the point θ_0 and

(i) $f(\lambda; \theta)w(\lambda)\frac{\partial^2}{\partial\theta_i\partial\theta_j}\log\psi(\lambda, \theta) \in L_1(\mathbb{R}) \cap L_2(\mathbb{R})$, $i, j = 1, \ldots, m$, $\theta \in \Theta$;

(ii) the limiting normal law holds for $T^{1/2}(J_T(\phi_i) - \mathsf{E}\, J_T(\phi_i))$,

where $J_T(\phi) = \int_{\mathbb{R}^1} I_T(\lambda)\phi(\lambda)\, d\lambda$ and $\phi_i(\lambda) = w(\lambda)\frac{\partial}{\partial\theta_i}\log\psi(\lambda, \theta)$;

(iii) $T^{1/2}\int_{\mathbb{R}^1}(EI_T(\lambda) - f(\lambda; \theta_0))w(\lambda)\frac{\partial}{\partial\theta_i}\log\psi(\lambda; \theta)\, d\lambda \to 0$ as $T \to \infty$, for all $i = 1, \ldots, m$, $\theta \in \Theta$;

(iv) the second order derivatives $\frac{\partial^2}{\partial\theta_i\partial\theta_j}\log\psi(\lambda; \theta)$, $i = 1, \ldots, m$, are continuous in θ.

B. VI. The matrices $S(\theta) = (s_{ij}(\theta))_{i,j=1,\ldots,m}$ and $A(\theta) = (a_{ij}(\theta))_{i,j=1,\ldots,m}$ are positive definite, where

$$s_{ij}(\theta) = \sigma^2(\theta)\int_{\mathbb{R}} w(\lambda)\left[\frac{\partial^2}{\partial\theta_i\partial\theta_j}\psi(\lambda, \theta) - \frac{1}{\psi(\lambda, \theta)}\frac{\partial}{\partial\theta_i}\psi(\lambda, \theta)\frac{\partial}{\partial\theta_j}\psi(\lambda, \theta)\right] d\lambda,$$

$$a_{ij}(\theta) = 4\pi\left(\sigma^2(\theta)\right)^2\int_{\mathbb{R}} w^2(\lambda)\frac{\partial}{\partial\theta_i}\psi(\lambda; \theta)\frac{\partial}{\partial\theta_j}\psi(\lambda; \theta)\, d\lambda$$

$$+ 2\pi\frac{d_4}{d_2^2}\left(\sigma^2(\theta)\right)^2\int_{\mathbb{R}} w(\lambda)\frac{\partial}{\partial\theta_i}\psi(\lambda; \theta)\, d\lambda \int_{\mathbb{R}} w(\lambda)\frac{\partial}{\partial\theta_j}\psi(\lambda; \theta)\, d\lambda.$$

Theorem A.4. *Let the conditions AI, B.I–B.VI be satisfied. Then as $T \to \infty$*

$$T^{1/2}\left(\widehat{\theta}_T - \theta_0\right) \xrightarrow{\mathcal{D}} N_m\left(0, S^{-1}(\theta_0)A(\theta_0)S^{-1}(\theta_0)\right),$$

where $N_m(\cdot, \cdot)$ denotes the m-dimensional Gaussian law.

A.3. **Conditions on asymptotic normality of the estimates: more details.** The results on consistency of estimators (Theorems A.1 and A.3) are consequences of corresponding theorems stated for the general case in [23] for the Whittle functional and in [3] for the case of Ibragimov functional.

The set of conditions for asymptotic normality formulated in the previous two sections are also based on the results of the mentioned above papers (with corresponding simplifications for the case of linear processes). However, one of the conditions (see, A.IV(ii) in the case of Whittle functional and B.V(ii) for the case of Ibragimov functional) we presented here in a more general form, exactly what is needed in the corresponding chain of (rather standard) arguments in the proof.

Note that conditions for the asymptotic normal law for spectral functionals of the second order $\int_R I_T(\lambda)\phi(\lambda)\,d\lambda$, where I_T is the periodogram (39), have been investigated by many authors, in particular, in the context of of quasi-likelihood estimation.

For the processes all of whose moments exist one can apply the methods of moments or cumulants to prove the convergence to the normal law.

Another approach is to reduce the problem of convergence of the spectral functionals $\int_{\mathbb{R}} I_T(\lambda)\phi(\lambda)\,d\lambda$ to the convergence of the empirical covariance function (and its integrals), which, in its own turn can be reduced to the problem of validity of a central limit theorem for the process $Y_u = X_{t+u}X_t$, and at this point one can apply the results on the central limit theorem for stationary processes. Thus, assuming the process to be weakly dependent (spectral density to be square integrable) we can demand instead of the assumptions A.IV(ii) and B.V(ii) the asymptotic normality of empirical covariance function and its weighted integrals, and we note that sufficient conditions for this can be given, for example, for strongly mixing processes via conditions on mixing coefficients (demanding, e.g., the exponential decay for the Rosenblatt mixing coefficient α).

As a result, conditions guaranteeing A.IV(ii) and B.V(ii) can be formulated as follows.

(1) In the case when all moments of the process X_t exist we demand (see, e.g., [6]) $f(\lambda;\theta_0) \in \mathbf{L}_p(\mathbb{R}^1)$, $\phi \in \mathbf{L}_q(\mathbb{R}^1)$, for some p, q such that $\frac{1}{p}+\frac{1}{q}=\frac{1}{2}$, $i=1,\ldots,m$, $\theta \in \Theta$, where $\phi = w(\lambda)\frac{\partial}{\partial\theta_i}\frac{1}{f(\lambda;\theta)}$ for the case of Whittle functional (A.IV(ii) condition) and $\phi = w(\lambda)\frac{\partial}{\partial\theta_i}\log\psi(\lambda,\theta)$ for the case of Ibragimov functional (B.V(ii) condition).

(2) In the case when q moments of the process X_t exist, for some $q \geq 4$, and the process is strongly mixing, we can demand $\int \alpha_X^{1/2}(t)\,dt < \infty$.

Next important conditions for the asymptotic normality of our estimates are A.IV(iii) and B.V(iii), more details on these can be found, e.g., in [3, 5].

REFERENCES

1. M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, The National Bureau of standards, 1964.
2. V. V. Anh, C. C. Heyde, and N. N. Leonenko, *Dynamic models of long-memory processes driven by Lévy noise*, J. Appl. Probab. **39** (2002), no. 4, 730–747.
3. V. V. Anh, N. N. Leonenko, and L. M. Sakhno, *Quasi-likelihood-based higher-order spectral estimation of random fields with possible long-range dependence*, J. Appl. Probab. **41A** (2004), 35–53.
4. V. V. Anh, N. N. Leonenko, and L. M. Sakhno, *On a class of minimum contrast estimators for fractional stochastic processes and fields*, J. Statist. Plann. Inference **123** (2004), 161–185.
5. V. V. Anh, N. N. Leonenko, and L. M. Sakhno, *Minimum contrast estimation of random processes based on information of second and third orders*, J. Statist. Plann. Inference **137** (2007), 1302–1331.
6. F. Avram, N. N. Leonenko, and L. M. Sakhno, *On a Szegö type limit theorem, the Hölder–Young–Brascamp–Lieb inequality, and the asymptotic theory of integrals and quadratic forms of stationary fields*, ESAIM Probab. Stat. **14** (2010), 210–255.

7. O. E. Barndorff-Nielsen, J. L. Jensen, and M. Sørensen, *Some stationary processes in discrete and continuous time*, Adv. Appl. Prob. **30** (1998), 989–1007.
8. M. B. Bibby, M. Skovgaard, and M. Sørensen, *Diffusion-type models with given marginal distribution and autocorrelation function*, Bernoulli **11** (2005), 191–220.
9. O. E. Barndorff-Nielsen and C. Halgreen, *Infinite divisibility of the hyperolic and generalized inverse Gaussian distributions*, Wahrsch Verw. Gebite **38** (1997), 309–312.
10. O. E. Barndorff-Nielsen and N. Shephard, *Non-Gaussian Ornstein–Uhlenbeck-based models and some of thier uses in financial econometrics (with discussion)*, J. R. Stat. Soc. **63** (2001), 167–241.
11. J. Bertoin, *Lévy Processes*, Cambridge University Press, Cambridge, 1996.
12. P. J. Brockwell, Davis R.A., and Y. Yang, *Estimation for nonnegative Lévy driven Ornstein–Uhlenbeck processes*, J. Appl. Prob. **44** (2007), 977–989.
13. V. Genon-Catalot, T. Jeantheau, and C. Laredo, *Stochastic volatility models as hidden Markov models and statistical applications*, Bernoulli **6** (2000), 1051–1079.
14. P. Hall and C. C. Heyde, *Martingale Limit Theory and its Application*, Academic Press, New York, 1980.
15. H. Hassani, *Sum of the sample autocorrelation function*, Random Oper. Stoch. Equ. **17** (2009), 125–130.
16. C. C. Heyde, *Quasi-Likelihood And Its Applications: A General Approach to Optimal Parameter Estimation*, Springer, New York, 1997.
17. C. C. Heyde and N. N. Leonenko, *Student processes*, Adv. in Appl. Probab. **37** (2005), 342–365.
18. I. A. Ibragimov, *On maximum likelihood estimation of parameters of the spectral density of stationary time series*, Theory Probab. Appl. **12** (1967), 115–119.
19. G. Jongbloed, F. H. Van der Meulen, and A. W. Van der Vaart, *Nonparametric inference for Lévy-driven Ornstein–Uhlenbeck processes*, Bernoulli **11** (2005), 759–791.
20. Z. J. Jurek, *Remarks on the self-decomposability and new examples*, Demonstr. Mathematica **34** (2001), no. 2, 241–250.
21. Z. J. Jurek and J. D. Mason, *Operator-Limit Distributions in Probability Theory*, John Wiley and Sons, New York, 1993.
22. S. Karlin and H. M. Taylor, *A First Course in Stochastic Processes*, Academic Press, New York, 1981.
23. N. N. Leonenko and L. M. Sakhno, *On the Whittle estimators for some classes of continuous parameter random processes and fields*, Stat. Probability Letters **76** (2006), 781–795.
24. H. Masuda, *On multidimensional Ornstein–Uhlenbeck process driven by a general Lévy process*, Bernoulli **10** (2004), 97–120.
25. K. Pearson, *Tables for Statisticians and Biometricians*, Cambridge University Press, Cambridge, 1914.
26. K. Sato, *Lévy Processes and Infinitely Divisible Distributions*, Cambridge University Press, Cambridge, 1999.
27. R. J. Serfling, *Approximation Theorems of Mathematical Statistics*, John Wiley & Sons, New York, 1980.
28. K. Spiliopoulos, *Method of moments estimation of Ornstein–Uhlenbeck processes driven by general Lévy process*, Ann. I.S.U.P. **53** (2009), 3–17.
29. S. Sun and X. Zhang, *Empirical likelihood estimation of discretely sampled processes of OU type*, Sci. China Ser. A **52** (2009), 908–931.
30. L. Valdivieso, W. Schoutens, and F. Tuerlinckx, *Maximum likelihood estimation in processes of Ornstein–Uhlenbeck type*, Stat. Inference Stoch. Process. **12** (2009), 1–19.
31. V. Witkovsky, *Exact distribution of positive linear combinations of inverted chi-square random variables with odd degrees of freedom*, Statist. Probab. Letters **56** (2002), 45–50.

School of Mathematics, Cardiff University, Senghennydd Road, Cardiff CF244AG, UK
E-mail address: LeonenkoN@Cardiff.ac.uk

Department of Probability, Statistics and Actuarial Mathematics, Mechanics and Mathematics Faculty, Taras Shevchenko National University of Kyiv, 64, Volodymyrs'ka St., 01601 Kyiv, Ukraine
E-mail address: lms@univ.kiev.ua

Department of Mathematics, University of Osijek, Gajev Trg 6, HR-31 000 Osijek, Croatia
E-mail address: nsuvak@mathos.hr