Statistica Sinica **30** (2020), 111-134
doi:https://doi.org/10.5705/ss.202017.0472

ON THE ESTIMATION OF LOCALLY STATIONARY LONG-MEMORY PROCESSES

Ngai Hang Chan[1,2] and Wilfredo Palma[3]

[1]*Southwestern University of Finance and Economics,* [2]*The Chinese University of Hong Kong and* [3]*Pontificia Universidad Católica de Chile*

Abstract: This study establishes the statistical properties of a spectrum-based Whittle parameter estimation procedure for locally stationary long-range dependent processes. Both theoretical and empirical behaviors are investigated. In particular, a central limit theorem for the Whittle likelihood estimation method is derived under mild distributional conditions, extending its application to a wide range of non-Gaussian time series. The finite-sample properties of the estimators are examined using Monte Carlo experiments with gamma and gamma-normal noise distributions. These simulation studies demonstrate that the proposed method behaves properly, even for small to moderate sample sizes. Finally, the practical application of this methodology is illustrated using a well-known real-life data example.

Key words and phrases: Local stationarity, long-range dependence, nonstationarity, Whittle estimation.

1. Introduction

Locally stationary processes play an important role in modeling and analyzing nonstationary time series. This approach is based on the evolutionary spectra developed by Priestley (1965) and formally introduced in Dahlhaus (1996, 1997). In this context, the parameters of the spectral density vary smoothly over time, enabling nonstationary processes to be locally approximated by stationary models. Recent reviews of these processes are provided by Dahlhaus (2012) and Chapter 8 of Palma (2016). A large number of estimation and hypothesis-testing methods have been developed based on these seminal ideas; see, for example, Chandler and Polonik (2017), Paparoditis and Preuss (2015), Guinness and Fuentes (2015), Chen et al. (2018), Fiecas and Ombao (2016), Song, Banerjee and Kosorok (2016), Wu and Zhou (2011), Puchstein and Preuss (2016), Rosen, Wood and Stoffer (2012), Vogt and Dette (2015), Kreiss and Paparoditis (2015), Preuss, Puchstein and Dette (2015), Zhou (2014), Nason (2013), Preuss, Vetter and Dette (2013b) Guinness and Stein (2013), Giraitis, Kapetanios and Yates

(2014), Preuss, Vetter and Dette (2013a), Zhou (2013), Roueff and Von Sachs (2011), Dette, Preuss and Vetter (2011), Van Bellegem and Dahlhaus (2006) and Beran (2009), among others.

However, most of these methodologies are designed for handling short-memory Gaussian linear locally stationary processes such as time-varying ARMA models. An extension of these techniques Palma and Olea (2010) to the long-memory case is based on Gaussian input noise. Nevertheless, in many practical settings, time-series data may exhibit nonstationary behavior, along with long-range dependence and non-Gaussian distribution. Incorporating these three conditions involves important technical challenges, and the literature on this field is still far from complete. For instance, analyses of the linear functionals of these processes lead to highly nonstandard asymptotic results; see Leipus and Surgailis (2013), Wu and Zhou (2014), and Palma (2010).

This study addresses a novel parameter-estimation technique for non-Gaussian long-memory locally stationary processes. This method is based on a generalized version of the quasi-likelihood introduced by Whittle (1953). These estimates turn out to be asymptotically normally distributed. Note that relaxing the Gaussianity assumption increases the technical complexity of proving the large-sample theory. Most of these difficulties are related to handling higher-order cumulants of quadratic forms. In the Gaussian context, there is an explicit formula for these cumulants. However, there is no such formulation for the general case. Consequently, analyzing their asymptotic behavior becomes much more challenging. Furthermore, the Whittle estimates are computationally efficient because they can be calculated using the fast Fourier transform (FFT); see Palma and Olea (2010). Monte Carlo experiments have shown that the estimates have very good small-sample properties. Thus, this study provides a computationally efficient framework for modeling and conducting statistical inferences for non-Gaussian time-series data that exhibit nonstationarities.

The remainder of this paper is structured as follows. Section 2 discusses a class of non-Gaussian long-range dependent locally stationary processes and proposes a quasi-maximum likelihood estimator based on an extended version of the Whittle spectrum-based methodology. Section 3 investigates their large-sample properties, establishing a central limit theorem while Section 4 is devoted to proving these results. Section 5 reports the results from several Monte Carlo experiments to evaluate the finite-sample performance of the Whittle estimates. A real-life data example is presented in Section 6 to illustrate the application of the methodology and the difference with respect to assuming normality. Section 7 concludes the paper.

2. Methodology

Following Dahlhaus (1997) and Palma and Olea (2010), a class of locally stationary processes is given by the infinite moving average expansion

$$Y_{t,T} = \sigma\left(\frac{t}{T}\right)\sum_{j=0}^{\infty}\psi_j\left(\frac{t}{T}\right)\varepsilon_{t-j}, \tag{2.1}$$

where $\{\varepsilon_t\}$ is a zero-mean and unit-variance white-noise, and $\{\psi_j(u)\}$ are coefficients satisfying $\psi_0(u) = 1$, $\sum_{j=0}^{\infty}\psi_j(u)^2 < \infty$; for all $u \in [0,1]$. In this case, the transfer function of process given by (2.1) is $A(\lambda, t/T) = \sigma(t/T)\sum_{j=0}^{\infty}\psi_j(t/T)\, e^{-i\lambda j}$.

Given a sample $\{Y_{1,T}, \ldots, Y_{T,T}\}$ of the process in (2.1), we can estimate the vector of parameters of the model, denoted by θ, by minimizing the following Whittle log-likelihood function, as in Palma and Olea (2010):

$$\mathcal{L}_T(\theta) = \frac{1}{4\pi}\frac{1}{M}\int_{-\pi}^{\pi}\sum_{j=1}^{M}\left\{\log f_\theta(u_j,\lambda) + \frac{I_N(u_j,\lambda)}{f_\theta(u_j,\lambda)}\right\}d\lambda, \tag{2.2}$$

where $f_\theta(u,\lambda) = |A(u,\lambda)|^2$ is the time-varying spectral density of the process, $I_N(u,\lambda) = |d_N(u,\lambda)|^2/(2\pi N)$ is the periodogram with

$$d_N(u,\lambda) = \sum_{s=0}^{N-1} Y_{[uT]-N/2+s+1,T}\, e^{-i\lambda s},$$

$T = S(M-1)+N$, $u_j = t_j/T$, and $t_j = S(j-1)+N/2$, for $j = 1,\ldots,M$. In this extended version of the Whittle estimation procedure (2.2), the sample $\{Y_{1,T},\ldots,Y_{T,T}\}$ is subdivided into M blocks of length N, each shifting S places from block to block. Then, the spectrum is locally estimated using the periodogram on each of these M blocks and then averaged to form (2.2). Finally, the Whittle estimator of the parameter vector θ is given by

$$\widehat{\theta}_T = \arg\min \mathcal{L}_T(\theta), \tag{2.3}$$

minimized over a parameter space Θ.

3. Properties

This section examines the large-sample properties of the proposed estimators, establishing their asymptotic normality in Theorem 1. The first assumption concerns the time-varying spectral density of the process. The second is related to the higher-order cumulants of the process. The third assumption is concerned with the block sampling scheme. It is assumed that the parameter space Θ is compact. In what follows, K is always a positive constant that can vary from line to line.

A1. The time-varying spectral density of the process in (2.1) is strictly positive and satisfies

$$f_\theta(u,\lambda) \sim C_f(\theta,u)\ |\lambda|^{-2\,d_\theta(u)},$$

as $|\lambda| \to 0$, where $C_f(\theta,u)$ is a strictly positive function and $d_\theta(u) \in (0,1)$. There is an integrable function $g(\lambda)$, such that $|\nabla_\theta \log f_\theta(u,\lambda)| \le g(\lambda)$ for all $\theta \in \Theta$, $u \in [0,1]$, and $\lambda \in [-\pi,\pi]$. The function $A(u,\lambda)$ is twice differentiable with respect to u and satisfies

$$\int_{-\pi}^{\pi} A(u,\lambda)A(v,-\lambda)\exp(ik\lambda)\,d\lambda \sim C_1(\theta,u,v)\ k^{d_\theta(u)+d_\theta(v)-1},$$

as $k \to \infty$, where $|C_1(\theta,u,v)| \le K$ for $u,v \in [0,1]$ and $\theta \in \Theta$. The function $f_\theta(u,\lambda)^{-1}$ is twice differentiable with respect to θ, u and λ. Furthermore,

$$\psi_k(u) = \sigma(u)^{-1}\int_{-\pi}^{\pi} A(u,\lambda)\exp(ik\lambda)\,d\lambda \sim C_2(\theta,u)\ k^{d_\theta(u)-1},$$

as $k \to \infty$, where $\psi_0(u) = 1$ and $|C_2(\theta,u)| \le K$ for $u \in [0,1]$, and $\theta \in \Theta$.

A2. There is a constant c_k, such that $g_k(\lambda_1,\ldots,\lambda_{k-1}) = c_k$ for all $\lambda_1,\ldots,\lambda_k$.

A3. The sample size T and the subdivision integers N, S, and M all tend to infinity and satisfy $S/N \to 0$, $\sqrt{T}\log^2 N/N \to 0$, $\sqrt{T}/M \to 0$, and $N^3\log^2 N/T^2 \to 0$.

Next we establish several large-sample properties of the Whittle estimator described by (2.3). The proofs of these results are provided in Section 4.

Theorem 1 (Central Limit Theorem)**.** *Let θ_0 be the true value of the parameter θ. Under Assumptions* A1–A3*; the Whittle estimator $\widehat{\theta}_T$ satisfies the following central limit theorem:*

$$\sqrt{T}(\hat{\theta}_T - \theta) \to N\left(0,\Sigma\right),$$

where

$$\Sigma = \Gamma^{-1}(\Gamma + W)\Gamma^{-1} = \Gamma^{-1} + \Gamma^{-1}\,W\,\Gamma^{-1},$$
$$\Gamma = \frac{1}{4\pi}\int_0^1\int_{-\pi}^{\pi}\left[\nabla_\theta\ \log f_\theta(u,\lambda)\right]\left[\nabla_\theta\ \log f_\theta(u,\lambda)\right]'\,du\,d\lambda,$$

and

$$W = \frac{g_4}{8\pi}\int_0^1\left[\int_{-\pi}^{\pi}\nabla_\theta\ \log f_\theta(u,\lambda)\,d\lambda\right]\left[\int_{-\pi}^{\pi}\nabla_\theta\ \log f_\theta(u,\lambda)\,d\lambda\right]'du.$$

Corollary 1. *If the parameter vector $\theta = (\alpha, \beta)$ is separable, such that the scale parameter $\sigma_\theta = \sigma_{\alpha,\beta}$ depends only on the second component of the parameter space, so we can write σ_β. In addition Γ can be written as a block diagonal matrix*

$$\Gamma = \begin{pmatrix} \Gamma_\alpha & 0 \\ 0 & \Gamma_\beta \end{pmatrix}.$$

Then, we have that

$$\Gamma^{-1} = \begin{pmatrix} \Gamma_\alpha^{-1} & 0 \\ 0 & \Gamma_\beta^{-1} + \Gamma_\beta^{-1}\, W_\beta\, \Gamma_\beta^{-1} \end{pmatrix},$$

where

$$\Gamma_\alpha = \frac{1}{4\pi} \int_{-\pi}^{\pi} \int_0^1 [\nabla_\alpha \ \log f_\theta(u,\lambda)]\, [\nabla_\alpha \ \log f_\theta(u,\lambda)]'\, du\, d\lambda\,,$$

$$\Gamma_\beta = \frac{1}{4\pi} \int_{-\pi}^{\pi} \int_0^1 [\nabla_\beta \ \log f_\theta(u,\lambda)]\, [\nabla_\beta \ \log f_\theta(u,\lambda)]'\, du\, d\lambda\,,$$

and

$$W_\beta = \frac{g_4}{8\pi} \int_0^1 \left[\int_{-\pi}^{\pi} \nabla_\beta \ \log f_\theta(u,\lambda)\, d\lambda\right] \left[\int_{-\pi}^{\pi} \nabla_\beta \ \log f_\theta(u,\lambda)\, d\lambda\right]' du\,.$$

Proof. Note that, by the extension of Kolmogorov's formula provided by Theorem 3.2 of Dahlhaus (1996), $\int \log f_\theta(u,\lambda)\, d\lambda = 2\pi \log \sigma_\beta^2(u)$. As a result, $\int \nabla_\alpha \log f_\theta(u,\lambda)\, d\lambda = 0$ and, thus, the term W_α vanishes from the asymptotic variance.

Remark 1. Consider the LS-ARFIMA model given by

$$Y_{t,T} = \sigma\left(\frac{t}{T}\right)\, \Phi\left(\frac{t}{T}, B\right)^{-1}\, \Theta\left(\frac{t}{T}, B\right) (1-B)^{-d(t/T)}\, \varepsilon_t, \tag{3.1}$$

for $t = 1, \ldots, T$, where for $u \in [0,1]$, $\Phi(u,B) = 1 + \phi_1(u)B + \cdots + \phi_P(u)B^P$ is an autoregressive polynomial, $\Theta(u,B) = 1 + \theta_1(u)B + \cdots + \theta_Q(u)B^Q$ is a moving average polynomial, $d(u)$ is a long-memory parameter, $\sigma(u)$ is a noise scale factor, and $\{\varepsilon_t\}$ is a white-noise sequence with zero mean and unit variance. As a consequence of the preceding corollary, the asymptotic variance of the Whittle estimators associated with $d(u)$, $\Theta(u)$, and $\Phi(u)$ of the LS-ARFIMA model given in (3.1) are not affected by the distribution of the input noise. However, this is not the case for the estimators corresponding to the scale parameter $\sigma(u)$. This is illustrated in the following example.

Example 1. Consider an LS-ARFIMA model with time-varying spectral density $f_\theta(u,\lambda)$, such that $\sigma_\theta(u) = \beta$. In this case,

$$\Gamma_\beta = 2\int_0^1 \frac{du}{\beta^2} = \frac{2}{\beta^2}\,.$$

On the other hand, we have

$$\int_{-\pi}^{\pi} \log f_\theta(u,\lambda)\,d\lambda = 2\pi \log\left[\frac{\sigma^2(u)}{2\pi}\right],$$

for all $u \in [0,1]$. Therefore,

$$\int_{-\pi}^{\pi} \log f_\theta(u,\lambda)\,d\lambda = 2\pi \log\left[\frac{\sigma^2(u)}{2\pi}\right] = 4\pi\log\beta - 2\pi\log 2\pi,$$

such that,

$$\int_{-\pi}^{\pi} \nabla_\beta \log f_\theta(u,\lambda)\,d\lambda = \frac{4\pi}{\beta}.$$

Therefore,

$$W_\beta = \frac{g_4}{8\pi}\int_0^1 \left(\frac{4\pi}{\beta}\right)^2 du = \frac{2\pi g_4}{\beta^2}\,,$$

and

$$\Gamma_\beta + W_\beta = \frac{2}{\beta^2} + \frac{2}{\beta^2}\pi g_4 = \frac{2}{\beta^2}(1+\pi g_4).$$

Consequently,

$$\Gamma_\beta^{-1}(\Gamma_\beta + W_\beta)\Gamma_\beta^{-1} = \frac{\beta^2}{2}(1+\pi g_4).$$

Note that if the input noise corresponds to a centered $\Gamma(\alpha,\lambda)$-distribution, the excess kurtosis is $6/\alpha$ which means $\kappa_4 = (24\pi^2)/\alpha$ and $g_4 = 3/(\pi\alpha)$. Thus,

$$\Gamma_\beta^{-1}\left(\Gamma_\beta + W_\beta\right)\Gamma_\beta^{-1} = \frac{\beta^2}{2}\left(1+\frac{3}{\alpha}\right).$$

On the other hand, if the input noise corresponds to a centered log-normal distribution, the excess kurtosis is $e^4+2e^3+3e^2-6$. In this case,

$$\kappa_4 = 4\pi^2(e^4+2e^3+3e^2-6)\,,\quad g_4 = \frac{e^4+2e^3+3e^2-6}{2\pi}.$$

Therefore,

$$\Gamma_\beta^{-1}\left(\Gamma_\beta + W_\beta\right)\Gamma_\beta^{-1} = \frac{\beta^2}{4}\left(e^4+2e^3+3e^2-4\right).$$

4. Proofs

Consider the function $\phi : [0,1]\times[-\pi,\pi] \to \mathbb{R}$ and define the functional operator

$$J(\phi) = \int_0^1 \int_{-\pi}^{\pi} \phi(u,\lambda) f(u,\lambda)\, d\lambda du, \tag{4.1}$$

where $f(u,\lambda)$ is the time-varying spectral density of the limit process (1). Define the sample version of $J(\cdot)$ as

$$J_T(\phi) = \int_0^1 \int_{-\pi}^{\pi} \phi(u,\lambda) I_N(u,\lambda)\, d\lambda du = \frac{1}{M} \sum_{j=1}^{M} \int_{-\pi}^{\pi} \phi(u_j,\lambda) I_N(u_j,\lambda)\, d\lambda + o(1),$$

where M and u_j, for $j = 1, \ldots, M$, are as given in Section 2.

Proof. (Theorem 1) We first prove that, for all $\ell \geq 3$, we have

$$T^{\ell/2} \operatorname{cum}_\ell(J_T(\phi), \ldots, J_T(\phi)) \to 0,$$

as $T \to \infty$. For notational simplicity, we proceed with $\ell = 3$; the result for $\ell \geq 4$ follows analogously. Recall that

$$d_N(u,\lambda) = \sum_{t=1}^{N} Y_{uT+t,T}\, e^{-i\lambda t}$$

and

$$Y_{uT+t,T} = \sum_{s=-\infty}^{\infty} \psi_{t-s}\left(u + \frac{t}{T}\right) \varepsilon_s,$$

with the convention that $\psi_j(u) = 0$ for all $j < 0$, $0 \leq u \leq 1$, and $\psi_0 = 1$. Thus,

$$\begin{aligned} d_N(u,\lambda) &= \sum_s \left(\sum_{t=1}^{N} \psi_{t-s}\left(u + \frac{t}{T}\right) e^{-i\lambda t} \right) \varepsilon_s \\ &= \sum_s \left(\sum_{\ell=1-N}^{0} e^{-i\lambda(N+\ell)}\ \psi_{\ell+N-s}\left(u + \frac{\ell}{T} + \frac{N}{T}\right) \right) \varepsilon_s \\ &= \sum_s \varphi_{N-s}(u,N,T,\lambda)\varepsilon_s, \end{aligned}$$

where

$$\varphi_j(u,N,T,\lambda) = \sum_{\ell=1-N}^{0} \psi_{j+\ell}\left(u + \frac{\ell}{T} + \frac{N}{T}\right) e^{-i\lambda(\ell+N)}.$$

Hence,

$$d_N(u,\lambda) = \int_{-\pi}^{\pi} \varphi(u,N,T,\lambda,\omega) e^{i\omega(uT+N)} d\xi(\omega)$$

with

$$\varphi(u,N,T,\lambda,\omega)=\sum_j \varphi_j(u,N,T,\lambda)e^{-i\omega j}.$$

Consequently,

$$|d_N(u,\lambda)|^2=\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\varphi(u,\lambda,\omega)\overline{\varphi}(u,\lambda,\omega')e^{i(uT+N)(\omega-\omega')}\,d\xi(\omega)\,\overline{d\xi(\omega')},$$

where for, notation simplicity, we have dropped N and T from $\varphi(u,N,T,\lambda,\omega)$. Thus, the periodogram can be written as

$$I_N(u,\lambda)=\frac{1}{2\pi N}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\varphi(u,\lambda,\omega)\overline{\varphi}(u,\lambda,\omega')e^{i(uT+N)(\omega-\omega')}\,d\xi(\omega)\,\overline{d\xi(\omega')}$$

and

$$\begin{aligned}J_T(\phi)&=\int_{-\pi}^{\pi}\int_0^1\phi(u,\lambda)I_N(u,\lambda)\,du\,d\lambda\\&=\frac{1}{2\pi N}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_0^1\phi(u,\lambda)\varphi(u,\lambda,\omega)\overline{\varphi}(u,\lambda,\omega')e^{i(uT+N)(\omega-\omega')}\,du\,d\lambda\\&\quad d\xi(\omega)\overline{d\xi(\omega')}.\end{aligned}$$

By defining

$$h(\omega,\omega')=\int_{-\pi}^{\pi}\int_0^1\phi(u,\lambda)\varphi(u,\lambda,\omega')\overline{\varphi}(u,\lambda,\omega')\,du\,d\lambda,$$

we can write

$$J_T(\phi)=\frac{1}{2\pi N}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}h(\omega,\omega')\,e^{i(uT+N)(\omega-\omega')}\,d\xi(\omega)\,\overline{d\xi(\omega')}.$$

Now,

$$\begin{aligned}&\operatorname{cum}(J_T(\phi),J_T(\phi),J_T(\phi))\\&=\left(\frac{1}{2\pi N}\right)^3\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}h(\omega_1,\omega_2)h(\omega_3,\omega_4)h(\omega_5,\omega_6)\\&\quad\times\operatorname{cum}(d\xi(\omega_1)\overline{d\xi(\omega_2)},d\xi(\omega_3)\ d\xi(\omega_4),d\xi(\omega_5)\ d\xi(\omega_6))\\&=\left(\frac{1}{2\pi N}\right)^3\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}h(\omega_1,\omega_2)h(\omega_3,\omega_4)h(\omega_5,\omega_6)\\&\quad\times\left\{g_6\delta\left(\sum_1^6\omega_j\right)+g_4g_2\sum_{\nu_4,\nu_2}c_{\nu_4,\nu_2}\delta\left(\sum_{\nu_4}\omega_j\right)\delta\left(\sum_{\nu_2}\omega_j\right)\right.\qquad(4.2)\\&\quad\left.+g_3^2\sum_{\nu_3,\nu_3}c_{\nu_3,\nu_3}\delta\left(\sum_{\nu_3}\omega_j\right)\delta\left(\sum_{\nu_3}\omega_j\right)\right\}d\omega_1,\ldots,\,d\omega_6\\&=\left(\frac{1}{2\pi N}\right)^3(A_N+B_N+C_N),\end{aligned}$$

say, where ν_j denotes a partition with j components, along with its respective constant. Consequently, from Lemma 2, we have that $\mathrm{cum}(J_T(\phi), J_T(\phi), J_T(\phi)) = (1/2\pi N)^3 \times O(1)$. Therefore,

$$T^{3/2}\mathrm{cum}(J_T(\phi), J_T(\phi), J_T(\phi)) = O\left(\frac{\sqrt{T}}{N}\right)^3.$$

Note that, by assumption, $\sqrt{T}/N \to 0$, Hence, $T^{3/2}\mathrm{cum}(J_T(\phi), J_T(\phi), J_T(\phi)) \to 0$ as $N, T \to \infty$. On the other hand, for $\ell = 2$, we have that

$$T\,\mathrm{cum}[J_T(\phi), J_T(\phi)] = \frac{T}{M^2}\sum_{j,k=1}^{M}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\phi(u_j,\lambda)\phi(u_k,\mu) \times \mathrm{cum}[I_N(u_j,\lambda), I_N(u_k,\mu)]\,d\lambda\,d\mu \;\equiv A_T + B_T,$$

where A_T corresponds to the term involving the covariance, and B_T corresponds to the fourth cumulant term:

$$B_T = \frac{T}{(2\pi MN)^2}\sum_{j,k=1}^{M}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\phi(u_j,\lambda)\phi(u_k,\mu) \times \mathrm{cum}[d_N(u_j,\lambda), d_N(u_j,-\lambda), d_N(u_k,\mu), d_N(u_k,-\mu)]\,d\lambda\,d\mu.$$

By Proposition 2 of Palma and Olea (2010), A_T converges to

$$\lim_{T\to\infty} TA_T = 4\pi\int_0^1\int_{-\pi}^{\pi}\phi(u,\lambda)\phi(u,\lambda)\,f(u,\lambda)^2\;d\lambda\,du.$$

On the other hand, by defining

$$b_N(u,\omega,\lambda) = \sum_{t=0}^{N-1} A\left(u + \frac{t}{T}, \omega\right)e^{it(\omega-\lambda)},$$

we may write

$$d_n(u,\lambda) = \int_{-\pi}^{\pi} b_N(u,\omega,\lambda)e^{iuT\omega}\,d\xi(\omega).$$

Thus,

$$B_T = \frac{g_4 T}{(2\pi MN)^2}\sum_{j,k=1}^{M}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\phi(u_j,\lambda)\phi(u_k,\mu)\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} b_N(u_j,\omega_1,\lambda)e^{i(u_j-u_k)T\omega_1} \times b_N(u_j,\omega_2,-\lambda)e^{i(u_j-u_k)T\omega_2}b_N(u_k,\omega_3,\mu)b_N(u_k,-\omega_1-\omega_2-\omega_3,-\mu) \times\; d\lambda\,d\mu\,d\omega_1\,d\omega_2\,d\omega_3.$$

Now, by integrating successively with respect to ω_3, ω_2, and ω_1, we obtain

$$B_T = \frac{2\pi g_4 T}{(MN)^2} \sum_{t,s=0}^{N-1} \sum_{\substack{j,k=1 \\ S(j-k)=s-t}}^{M} \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \phi(u_j,\lambda)\phi(u_k,\mu) \sum_{\ell,m,n,p=0}^{\infty} \psi_\ell\left(u_k + \frac{t}{T}\right)$$
$$\times\, \psi_m\left(u_k + \frac{s}{T}\right)\psi_n\left(u_k + \frac{t}{T}\right)\psi_p\left(u_k + \frac{s}{T}\right) e^{i\lambda(\ell-m)+i\mu(n-p)}\, d\lambda\, d\mu.$$

Given that $0 \le t/T \le N/T \to 0$ as $N, T \to \infty$, we have that

$$\lim_{T\to\infty} B_T = \lim_{T\to\infty} \frac{2\pi g_4 T}{(MN)^2} \sum_{t,s=0}^{N-1} \sum_{\substack{j,k=1 \\ S(j-k)=s-t}}^{M} \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \phi(u_j,\lambda)\phi(u_k,\mu) \sum_{\ell,m,n,p=0}^{\infty} \psi_\ell(u_k)$$
$$\times\, \psi_m(u_k)\psi_n(u_k)\psi_p(u_k) e^{i\lambda(\ell-m)+i\mu(n-p)}\, d\lambda\, d\mu$$
$$\lim_{T\to\infty} \frac{2\pi g_4 T}{(MN)^2} \sum_{t,s=0}^{N-1} \sum_{\substack{j,k=1 \\ S(j-k)=s-t}}^{M} \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \phi_1\left(u_k + \frac{S(j-k)}{T}, \lambda\right)\phi_2(u_k,\mu)$$
$$\times\, A(u_k,\lambda)A(u_k,-\lambda)A(u_k,\mu)A(u_k,-\mu)\, d\lambda\, d\mu$$
$$\lim_{T\to\infty} \frac{2\pi g_4 T}{(MN)^2} \sum_{t=0}^{N-1} \sum_{p=\frac{t}{S}}^{\frac{N-t}{S}} \sum_{k=1}^{M-|p|} \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \phi_1(u_{k+p},\lambda)\phi_2(u_k,\mu) f(u_k,\lambda) f(u_k,\mu)\, d\lambda\, d\mu.$$

Now, by an argument analogous to that in the proof of Proposition 2 of Palma and Olea (2010), we conclude that

$$\begin{aligned}\lim_{T\to\infty} B_T &= 2\pi g_4 \int_0^1 \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \phi(u,\lambda)\phi(u,\mu) f(u,\lambda) f(u,\mu)\, d\lambda\, d\mu\, du \\ &= 2\pi g_4 \int_0^1 \left[\int_{-\pi}^{\pi} \phi(u,\lambda) f(u,\lambda)\, d\lambda\right]\left[\int_{-\pi}^{\pi} \phi(u,\mu) f(u,\mu)\, d\mu\right] du.\end{aligned}$$

Lemma 1. *Let $(d, d_1, \ldots, d_\ell) \in (0, 1/2)^{\ell+1}$ for $\ell \ge 3$, where $d = d(u)$, $d_i = d(u_i)$, with $u, u_i \in [0,1]$ for $i = 1, \ldots, \ell$, define the function*

$$C(h_1, \ldots, h_\ell) = \sum_{j=0}^{\infty} \psi_j(u) \prod_{i=1}^{\ell} \psi_{j+h_i}(u_i),$$

and let $d_0 = \max\{d, d_\ell\}$. Then, for $h_1, \ldots, h_\ell \to \infty$,

$$|C(h_1, \ldots, h_\ell)| \le C_0\, h_1^{d_1-1} \cdots h_{d_{\ell-1}}^{d_{\ell-1}-1} h_\ell^{2d_0-1}.$$

Proof. We first establish that the function C is well defined. Observe that, by Assumption A1, for large $h_1, \ldots, h_\ell$, we have $|\psi_{j+h_i}(u_i)| \le K j^{d_i-1}$ and

$$|C(h_1, h_2, \ldots, d_\ell)| \le K \sum_{j=1}^{\infty} j^{d-1+\sum_{i=1}^{\ell}(d_i-1)} \le K \sum_{j=1}^{\infty} j^{d+d_1+d_2+\cdots+d_\ell-\ell-1}.$$

Furthermore,

$$\sum_{j=n}^{\infty} j^{d+d_1+d_2+\cdots+d_\ell-\ell-1} \le K n^{d+d_1+d_2+\cdots+d_\ell-\ell} \le K n^{(1-\ell)/2}.$$

Consequently, given that this sum is convergent for all $\ell \ge 2$, we conclude that

$$|C(h_1, \ldots, h_\ell)| < \infty.$$

Finally, because for $i = 1, \ldots, \ell-1$, $|\psi_{j+h_i}(u_i)| \le K h_i^{d_i-1}$, we have

$$\begin{aligned}|C(h_1, \ldots, h_\ell)| &\le K h_1^{d_1-1} \cdots h_{\ell-1}^{d_\ell-1} \sum_{j=1}^{\infty} j^{d-1}(j+h_\ell)^{d_\ell-1} \\ &\le K h_1^{d_1-1} \cdots h_{\ell-1}^{d_\ell-1} \sum_{j=1}^{\infty} j^{d_0-1}(j+h_\ell)^{d_0-1} \\ &\le K h_1^{d_1-1} \cdots h_{\ell-1}^{d_\ell-1} h_\ell^{2d_0-1},\end{aligned}$$

as required.

Lemma 2. *Let $Z_N = A_N + B_N + C_N$, where these terms are defined as in* (4.2)*. Then, under the assumptions of Theorem* 1*, we have that $Z_N = O(1)$.*

Proof. We proceed by first proving the result for A_N; that is, we consider the case where the frequencies satisfy the condition $\omega_1 + \cdots + \omega_6 = 0$ such that that $\omega_6 = -\omega_1 - \omega_2 - \cdots - \omega_5$. Define the integral

$$\begin{aligned}I = &\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} h(\omega_1, \omega_2) h(\omega_3, \omega_4) \\ &h(\omega_5, -\omega_1 - \omega_2 - \cdots - \omega_5) d\omega_1, \ldots, d\omega_5.\end{aligned}$$

Let $\omega_0 = \omega_2 + \cdots + \omega_5$, and write

$$I_1 = \int_{-\pi}^{\pi} h(\omega_1, \omega_2) h(\omega_5, -\omega_1 - \omega_0)\, d\omega_1,$$

such that

$$I = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} h(\omega_3, \omega_4) I_1(\omega_2, \ldots, \omega_5)\, d\omega_1, \ldots,\, d\omega_5.$$

Hence, by writing $I_1 = I_1(\omega_2, \ldots, \omega_5)$, for simplicity, we have

$$\begin{aligned}I_1 &= \int_{-\pi}^{\pi} h(\omega_1, \omega_2) h(\omega_5, -\omega_1 - \omega_0)\, d\omega_1 \\ &= \int_{-\pi}^{\pi}\int_0^1\int_{-\pi}^{\pi}\int_0^1\int_{-\pi}^{\pi} \phi(u_1, \lambda_1)\varphi(u_1, \lambda_1, \omega_1)\varphi(u_1, \lambda_1, \omega_2)\phi(u_3, \lambda_3)\end{aligned}$$

$$\times\varphi(u_3,\lambda_3,\omega_5)\varphi(u_3,\lambda_3,-\omega_1-\omega_0)\,d\omega_1\,du_1\,d\lambda_1\,du_3\,d\lambda_3\,.$$

Observe that

$$\begin{aligned}&\int_{-\pi}^{\pi}\varphi(u_1,\lambda_1,\omega_1)\varphi(u_3,\lambda_3,-\omega_1-\omega_0)\,d\omega_1\\&=\int_{-\pi}^{\pi}\sum_j\varphi_j(u_1,\lambda_1)e^{-i\omega j}\\&\quad\times\sum_k\varphi_k(u_3,\lambda_3)e^{i\omega_1 k+i\omega_0 k}\,d\omega_1\\&=\sum_j\sum_k e^{i\omega_0 k}\varphi_j(u_1,\lambda_1)\varphi_k(u_3,\lambda_3)\int_{-\pi}^{\pi}e^{i\omega_1(k-j)}\,d\omega_1\\&=2\pi\sum_j e^{i\omega_0 j}\varphi_j(u_1,\lambda_1)\varphi_j(u_3,\lambda_3).\end{aligned}$$

Therefore,

$$\begin{aligned}I_1=&\int_0^1\int_0^1\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\phi(u_1,\lambda_1)\varphi(u_1,\lambda_1,\omega_2)\phi(u_3,\lambda_3)\varphi(u_3,\lambda_3,\omega_5)\\&\times 2\pi\sum_j e^{i\omega_0 j}\varphi_j(u_1,\lambda_1)\varphi_j(u_3,\lambda_3)\,du_1\,du_3\,d\lambda_1\,d\lambda_3.\end{aligned}$$

Now, integrating with respect to λ_1 yields

$$\sum_j\left[\int_{-\pi}^{\pi}\phi(u_1,\lambda_1)\varphi(u_1,\lambda_1,\omega_2)\varphi_j(u_1,\lambda_1)d\lambda_1\right]e^{i\omega_0 j}\varphi_j(u_3,\lambda_3).$$

Note that

$$\begin{aligned}&\int_{-\pi}^{\pi}\phi(u_1,\lambda_1)\varphi(u_1,\lambda_1,\omega_2)\varphi_j(u_1,\lambda_1)d\lambda_1\\&=\sum_k\int_{-\pi}^{\pi}\varphi_k(u_1,\lambda_1)e^{-i\omega_2 k}\phi(u_1,\lambda_1)\varphi_j(u_1,\lambda_1)d\lambda_1.\end{aligned}$$

In addition,

$$\varphi_k(u_1,\lambda_1)=\sum_{\ell=1-N}^{0}\psi_{k+\ell}^{(1)}(u_1),e^{-i\lambda_1\ell}\,,$$

$$\varphi_j(u_1,\lambda_1)=\sum_{p=1-N}^{0}\psi_{j+p}^{(1)}(u_1)e^{-i\lambda_1 p}.$$

Thus, by dropping u_i from $\psi_{k+\ell}^{(i)}(u_i)$ and writing $\gamma_\phi(h)=\int_{-\pi}^{\pi}\phi(u,\lambda)e^{i\lambda h}\,d\lambda$, for simplicity, we have

$$\begin{aligned}&\int_{-\pi}^{\pi}\phi(u_1,\lambda_1)\varphi(u_1,\lambda_1,\omega_2)\varphi_j(u_1,\lambda_1)\,d\lambda_1\\&=\sum_k\sum_{\ell,p}\psi^{(1)}_{k+\ell}\psi^{(1)}_{j+p}e^{-i\omega_2 k}\int_{-\pi}^{\pi}\phi(u_1,\lambda_1)e^{i\lambda_1(\ell-p)}\,d\lambda_1\\&=\sum_k\sum_{\ell,p}\psi^{(1)}_{k+\ell}\psi^{(1)}_{j+p}e^{-i\omega_2 k}\gamma_\phi(\ell-p)\,.\end{aligned}$$

Consequently,

$$\begin{aligned}&\sum_j\left[\int_{-\pi}^{\pi}\phi(u_1,\lambda_1)\varphi(u_1,\lambda_1,\omega_2)\varphi_j(u_1,\lambda_1)\,d\lambda_1\right]e^{i\omega_0 j}\varphi_j(u_3,\lambda_3)\\&=\sum_j\sum_{k,\ell,p}\psi^{(1)}_{k+\ell}\,\psi^{(1)}_{j+p}\varphi(u_3,\lambda_3)\gamma_\phi(\ell-p)e^{i\omega_0 j-i\omega_2 k}.\end{aligned}$$

Now, by integrating I_1 with respect to ω_2, we have

$$\begin{aligned}&\int_{-\pi}^{\pi}I_1(\omega_2,\ldots,\omega_5)\,d\omega_2\\&=2\pi\int_0^1\int_0^1\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\sum_{jk}\sum_{\ell,p}\psi^{(1)}_{k+\ell}\psi^{(1)}_{j+p}\,\varphi_j(u_3,\lambda_3)\\&\quad\times\gamma_\phi(\ell-p)e^{i\omega_2(j-k)+i\omega_3 j+i\omega_4 j+i\omega_5 j}\phi(u_3,\lambda_3)\varphi(u_3,\lambda_3,\omega_5)\,du_1\,du_3\,d\lambda_3\,d\omega_2\\&=(2\pi)^2\int_0^1\int_0^1\int_{-\pi}^{\pi}\sum_j\sum_{\ell,p}\psi^{(1)}_{j+\ell}\psi^{(1)}_{j+p}\,\varphi_j(u_3,\lambda_3)\gamma_\phi(\ell-p)e^{ij(\omega_3+\omega_4+\omega_5)}\\&\quad\times\phi(u_3,\lambda_3)\varphi(u_3,\lambda_3,\omega_5)\,du_1\,du_3\,d\lambda_3.\end{aligned}$$

Integrating with respect to λ_3 in the above expression yields

$$\begin{aligned}&\int_{-\pi}^{\pi}\varphi_j(u_3,\lambda_3)\phi(u_3,\lambda_3)\varphi(u_3,\lambda_3,\omega_5)\,d\lambda_3\\&=\sum_k\int_{-\pi}^{\pi}\varphi_j(u_3,\lambda_3)\phi(u_3,\lambda_3)\varphi_k(u_3,\lambda_3)e^{-i\omega_5 k}\,d\lambda_3,\end{aligned}$$

where

$$\varphi_j(u_3,\lambda_3)=\sum_{q=1-N}^{0}\psi^{(3)}_{j+q}e^{-i\lambda_3 q},\quad \varphi_k(u_3,\lambda_3)=\sum_{n=1-N}^{0}\psi^{(3)}_{k+n}e^{-i\lambda_3 n}.$$

Hence, we obtain

$$\begin{aligned}&\int_{-\pi}^{\pi}\varphi_j(u_3,\lambda_3)\phi(u_3,\lambda_3)\varphi(u_3,\lambda_3,\omega_5)\,d\lambda_3=\sum_k\sum_{q,n}\psi^{(3)}_{j+q}\psi^{(3)}_{k+n}e^{-i\omega_5 k}\times\\&\int_{-\pi}^{\pi}\phi(u_3,\lambda_3)e^{i\lambda_3(n-q)}\,d\lambda_3=\sum_k\sum_{q,n}\psi^{(3)}_{j+q}\psi^{(3)}_{k+n}e^{-i\omega_5 k}\gamma_{\phi_3}(n-q)\,.\end{aligned}$$

Define I_2 as

$$I_2(\omega_3,\omega_4,\omega_5)=\int_{-\pi}^{\pi} I_1(\omega_2,\ldots,\omega_5)\,d\omega_2.$$

Consequently,

$$\begin{aligned}I_2(\omega_3,\omega_4,\omega_5)=(2\pi)^2\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\sum_{jk}\sum_{\ell pqn}&\psi_{j+\ell}^{(1)}\psi_{j+p}^{(1)}\gamma_{\phi_1}(\ell-p)\\ &\times\psi_{j+q}^{(3)}\psi_{k+n}^{(3)}\gamma_{\phi_j}(n-p)e^{i\omega_3 j+ij\omega_4+i\omega_5(j-k)}\,du_1\,du_3.\end{aligned}$$

Furthermore, by defining I_3 as

$$I_3(\omega_3,\omega_4)=\int_{-\pi}^{\pi} I_2(\omega_3,\omega_4,\omega_5)\,d\omega_5,$$

we obtain

$$\begin{aligned}I_3(\omega_3,\omega_4)=(2\pi)^3\int_0^1\int_0^1\sum_{j}\sum_{\ell pqn}&\psi_{j+\ell}^{(1)}\psi_{j+p}^{(1)}\gamma_{\phi_1}(\ell-p)\\ &\times\psi_{j+q}^{(3)}\psi_{j+n}^{(3)}\gamma_{\phi_3}(n-p)e^{i\omega_3 j+ij\omega_4}\,du_1\,du_3.\end{aligned}$$

On the other hand,

$$h(\omega_3,\omega_4)=\int_{-\pi}^{\pi}\int_0^1 \phi(u_2,\lambda_2)\varphi(u_2,\lambda_2,\omega_3)\varphi(u_2,\lambda_2,\omega_4)\,du_2\,d\lambda_2.$$

Thus, integrating with respect to ω_3, yields

$$\begin{aligned}\int_{-\pi}^{\pi}\varphi(u_2,\lambda_2,\omega_3)e^{i\omega_3 j}\,d\omega_3&=\sum_k\varphi_k(u_2,\lambda_2)\int_{-\pi}^{\pi}e^{-i\omega_3 k+i\omega_3 j}\,d\omega_3\\ &=2\pi\varphi_j(u_2,\lambda_2).\end{aligned}$$

Integrating with respect to ω_4, we get

$$\int_{-\pi}^{\pi}\varphi(u_2,\lambda_2,\omega_4)d\omega_4=2\pi\varphi_j(u_2,\lambda_2)\,.$$

Thus, by defining I_4 as

$$\begin{aligned}I_4&=\int_{-\pi}^{\pi} I_4(\omega_3,\omega_4)\,d\omega_3\,d\omega_4\\ &=(2\pi)^5\int_{-\pi}^{\pi}\sum_j\sum_{\ell pqn}\psi_{j+\ell}^{(1)}\psi_{j+p}^{(1)}\gamma_{p_1}(\ell-p)\psi_{j+q}^{(3)}\psi_{j+n}^{(3)}\gamma_{\phi_3}(h-p)\varphi_j(u_2,\lambda_2)\\ &\quad\times du_1\,du_2\,du_3\,d\lambda_2\,\hat{\phi}(u_2,\lambda_2)\int_{-\pi}^{\pi}\phi(u_2,\lambda_2)\varphi_j(u_2,\lambda_2)\,d\lambda_2\\ &=\sum_s\sum_t\psi_{j+s}^{(2)}\psi_{j+t}^{(2)}\int_{-\pi}^{\pi}e^{i\lambda_2(5-t)}\phi(u_2,\lambda_2)\,d\lambda_2\sum_{s,t}\psi_{j+s}^{(2)}\psi_{j+t}^{(2)}\gamma_{\phi_2}(s-t),\end{aligned}$$

we obtain,

$$I = (2\pi)^5 \sum_j \sum_{\ell,p,q,n,s,t} \int_0^1 \int_0^1 \int_0^1 \psi^{(1)}_{j+\ell}\psi^{(1)}_{j+p}\gamma_{\phi_1}(\ell - p)\psi^{(2)}_{j+s}\psi^{(2)}_{j+t}\gamma_{\phi_2}(s-t)$$
$$\times \psi^{(3)}_{j+q}\psi^{(3)}_{j+n}\gamma_{\phi_3}(n-q)\, du_1\, du_2\, du_3.$$

Furthermore, we can write

$$I = (2\pi)^5 \sum_{j=\infty}^{\infty} \sum_{\ell,p,q,n,s,t=1}^{N} \int_0^1 \int_0^1 \int_0^1 \psi^{(1)}_{j+s}\psi^{(1)}_{j+p}\gamma_{\phi_1}(\ell - p)\psi^{(2)}_{j+s}\psi^{(2)}_{j+t}$$
$$\times \gamma_{\phi_2}(s-t)\psi^{(3)}_{j+q}\psi^{(3)}_{j+n}\gamma_{\phi_3}(n-q)\, du_1\, du_2\, du_3$$
$$= (2\pi)^5 \sum_{\ell \ldots t=0}^{N} \int_0^1 \int_0^1 \int_0^1 \left(\sum_j \psi^{(1)}_{j+\ell}\psi^{(1)}_{j+p}\psi^{(2)}_{j+n}\psi^{(2)}_{j+q}\psi^{(3)}_{j+t}\psi^{(3)}_{j+s} \right)$$
$$\times \gamma_{\phi_1}(\ell - p)\gamma_{\phi_2}(t-s)\gamma_{\phi_3}(n-q)\, du_1\, du_2\, du_3.$$

Now, by writing $h_1 = \ell - p$, $h_2 = q - n$, and $h_3 = t - s$ and applying Lemma 1, we conclude that $I = O(1)$. The proofs of the remaining cases are analogous.

5. Monte Carlo Studies

The following simulation results are based on 1,000 repetitions from the LS-FN model defined by the discrete-time equation

$$Y_{t,T} = \sigma\left(\frac{t}{T}\right)(1-B)^{-d(t/T)}\varepsilon_t, \tag{5.1}$$

for $t = 1, \ldots, T$, where $\{\varepsilon_t\}$ is a white-noise sequence with zero mean and unit variance. In these Monte Carlo experiments, the evolution of the long-memory parameter is specified by $d(u) = \alpha_0 + \alpha_1 u$ and the standard deviation is assumed to be constant, $\sigma(u) = \beta$. In addition, the white-noise sequence $\{\varepsilon_t\}$ follows either a gamma distribution or a log-normal distribution. For comparison purposes, the tables also include the case of Gaussian white-noise. In all these cases, the input noises have a zero mean and unit variance.

Figure 1 displays a simulated time series of 1,000 observations with gamma input noise. The histogram and the estimated density of this series are shown in Figures 3 and 4, respectively. Note the skewness in the empirical distribution of $\{Y_{t,T}\}$. Similar behavior is observed for the log-normal case; see Figures 2, 5, and 6.

Tables 3 and 2 report the results from several Monte Carlo simulations of the Whittle likelihood estimates for the LS-FN models. The sample sizes are $T = 500$

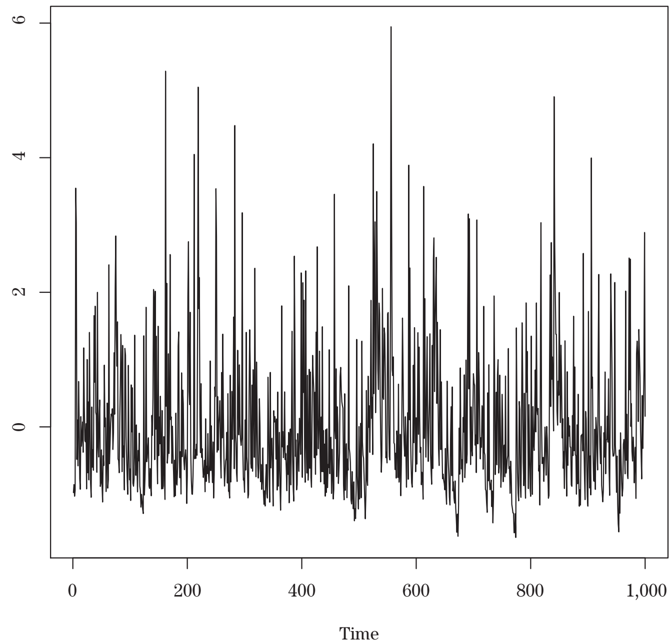

Figure 1. Time series with a gamma input noise.

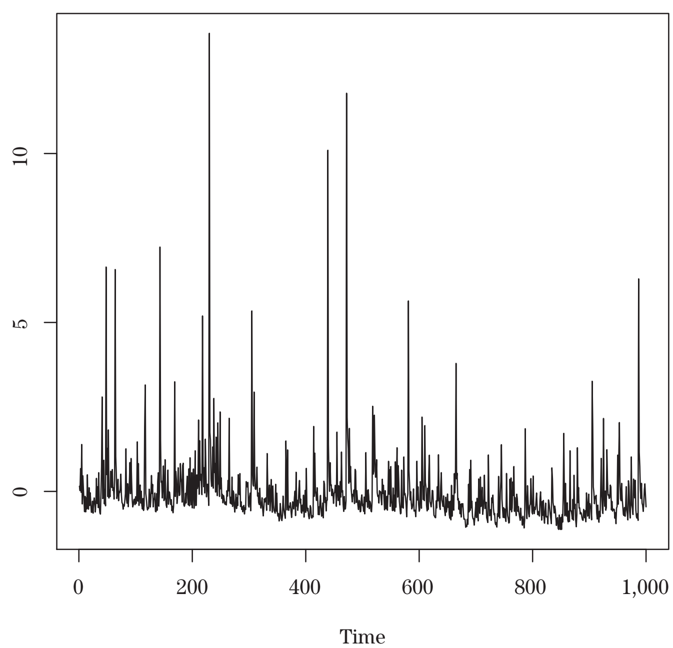

Figure 2. Time series with a log-normal input noise.

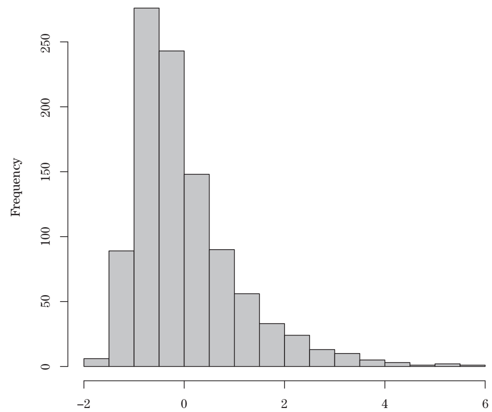

Figure 3. Histogram of time series with a gamma input noise.

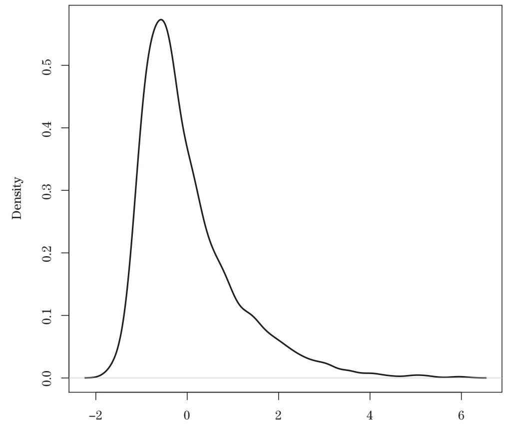

Figure 4. Estimated density of time series with a gamma input noise.

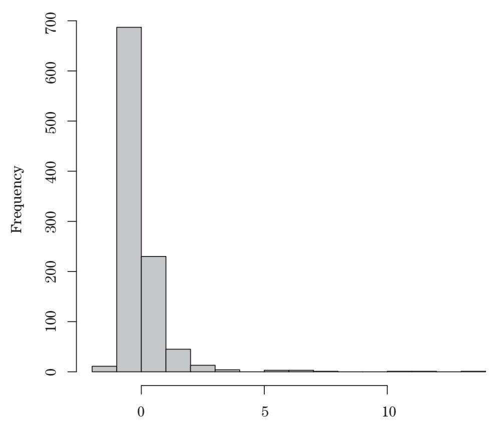

Figure 5. Histogram of time series with a log-normal input noise.

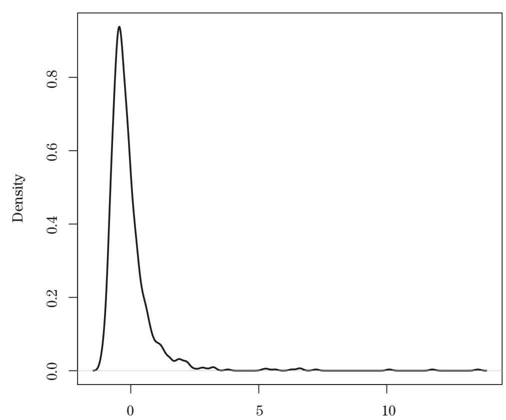

Figure 6. Estimated density of time series with a log-normal input noise.

and $T = 1,000$, respectively. The values of N and S are similar to those in Palma and Olea (2010). These tables show the averages and empirical standard deviations for the Whittle estimates, as well as the corresponding theoretical standard deviations. Given that $d(u) = \alpha_0 + \alpha_1 u$, in this case, the value of the time-varying parameter $d(u)$ moves from 0.10 to 0.30.

The theoretical standard deviations shown in these two tables are obtained using the following general procedure. Assume the following specification:

$$d(u) = \alpha_0 + \alpha_1 u + \cdots + \alpha_p u^p, \qquad \sigma(u) = \beta_0 + \beta_1 u + \cdots + \beta_q u^q,$$

for $u \in [0,1]$. In this case, the parameter vector is $\theta = (\alpha_0, \ldots, \alpha_p, \beta_0, \ldots, \beta_q)'$,

Table 1. Whittle estimation: Sample size $T = 500$, block size $N = 125$, and shift $S = 30$.

	Parameters			Estimates		
Case	α_0	α_1	β_0	$\widehat{\alpha}_0$	$\widehat{\alpha}_1$	$\widehat{\beta}_0$
Gamma	0.10	0.20	1.0	0.0983	0.1868	0.9946
Log-normal	0.10	0.20	1.0	0.1026	0.1817	1.0204
Normal	0.10	0.20	1.0	0.1002	0.1804	0.9987
	Theoretical SD			Estimated SD		
Case	$\sigma(\widehat{\alpha}_0)$	$\sigma(\widehat{\alpha}_1)$	$\sigma(\widehat{\beta}_0)$	$\widehat{\sigma}(\widehat{\alpha}_0)$	$\widehat{\sigma}(\widehat{\alpha}_1)$	$\widehat{\sigma}(\widehat{\beta}_0)$
Gamma	0.0697	0.1207	0.0632	0.0830	0.1552	0.0636
Log-normal	0.0697	0.1207	0.2376	0.0888	0.1633	0.2175
Normal	0.0697	0.1207	0.0316	0.0847	0.1607	0.0321

Table 2. Whittle estimation: Sample size $T = 1{,}000$, block size $N = 160$, and shift $S = 50$.

	Parameters			Estimates		
Case	α_0	α_1	β_0	$\widehat{\alpha}_0$	$\widehat{\alpha}_1$	$\widehat{\beta}_0$
Gamma	0.10	0.20	1.0	0.0932	0.2001	0.9991
Log-normal	0.10	0.20	1.0	0.0923	0.2040	1.0204
Normal	0.10	0.20	1.0	0.0955	0.1942	0.9986
	Theoretical SD			Estimated SD		
Case	$\sigma(\widehat{\alpha}_0)$	$\sigma(\widehat{\alpha}_1)$	$\sigma(\widehat{\beta}_0)$	$\widehat{\sigma}(\widehat{\alpha}_0)$	$\widehat{\sigma}(\widehat{\alpha}_1)$	$\widehat{\sigma}(\widehat{\beta}_0)$
Gamma	0.0493	0.0854	0.0447	0.0598	0.1109	0.0445
Log-normal	0.0493	0.0854	0.1680	0.0612	0.1140	0.1467
Normal	0.0493	0.0854	0.0223	0.0590	0.1117	0.0214

and the asymptotic variance–covariance matrix of the Whittle estimates of θ corresponds to Σ^{-1}/T, is given by Theorem 1,

$$\Sigma = \Gamma^{-1}(\Gamma + W)\Gamma^{-1} = \Gamma^{-1} + \Gamma^{-1}\, W\, \Gamma^{-1},$$

with

$$\Gamma = \begin{pmatrix} \Gamma_\alpha & 0 \\ 0 & \Gamma_\beta \end{pmatrix},$$

$$\Gamma_\alpha = \left[\frac{\pi^2}{6\,(i+j+1)}\right]_{i,j=0,\ldots,p}, \Gamma_\beta = 2\left[\int_0^1 \frac{u^{i+j}\, du}{(\beta_0 + \beta_1 u + \cdots + \beta_q u^q)^2}\right]_{i,j=0,\ldots,q},$$

and

$$W = \begin{pmatrix} I & 0 \\ 0 & W_\beta \end{pmatrix},$$

where W_β is given in Example 1.

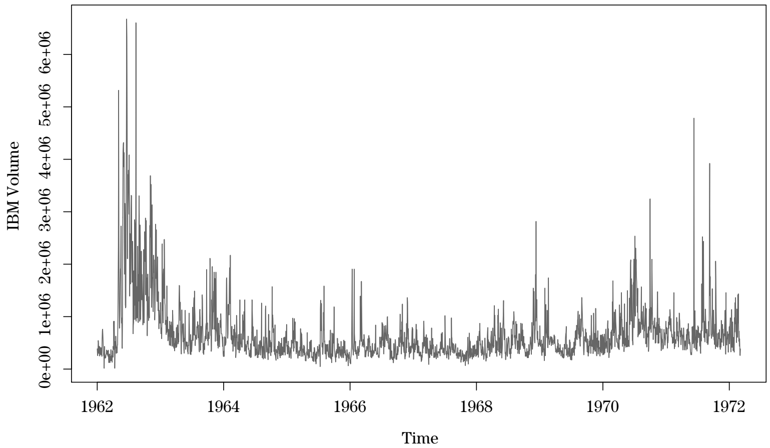

Figure 7. IBM daily transaction volume from 1962 to 1972.

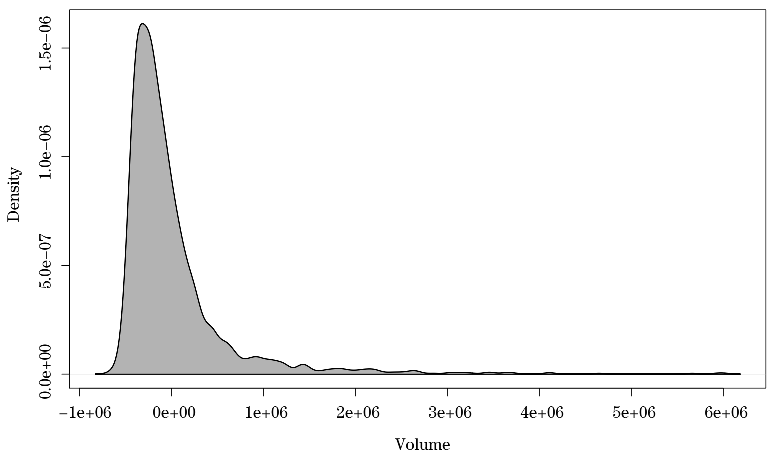

Figure 8. Estimated density of centered IBM daily transactions volume data.

Observe from Tables 1 and 2 that for both sample sizes and the three distributions considered (gamma, log-normal and Normal), the estimates are close to the true parameter values. Furthermore, the empirical standard deviations are close to their theoretical counterparts provided by Theorem 1.

6. Data Illustration

To illustrate the proposed methodology, consider the IBM daily transaction volume data for the period January 2, 1962, to December 31, 1972, depicted in Figure 7. This decade-long period has been selected to avoid market fluctuations

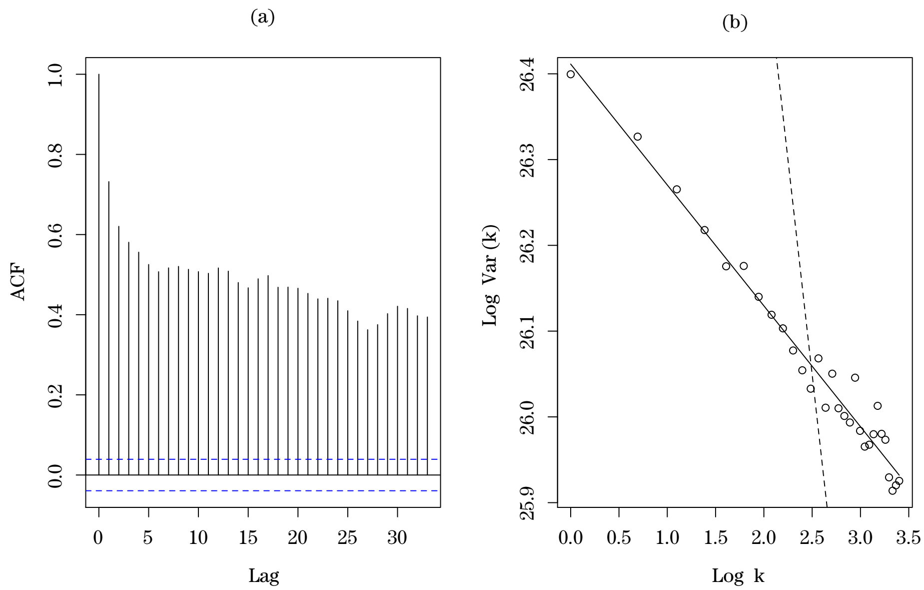

Figure 9. ACF of IBM daily transaction volume data. Panels: (a) sample ACF, (b) variance plot.

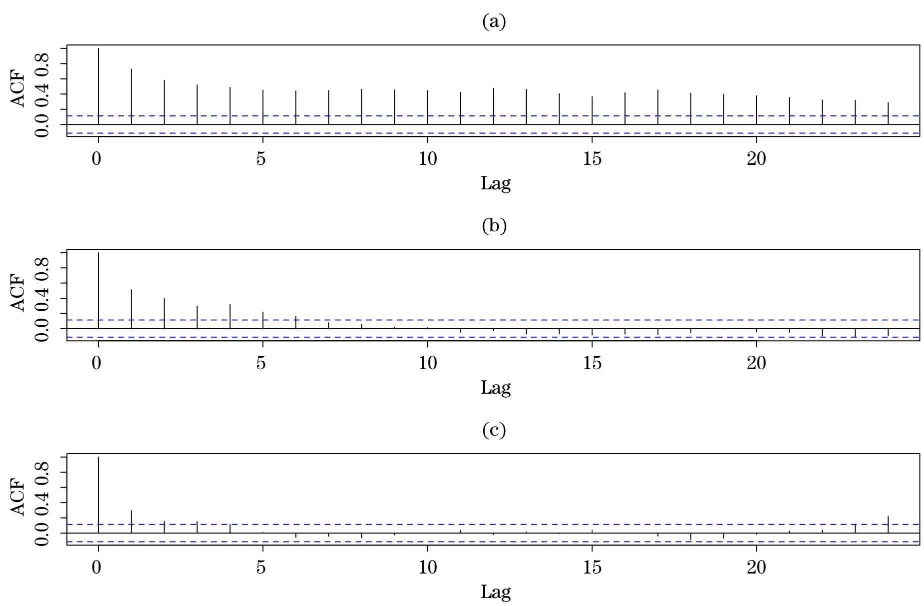

Figure 10. ACF of IBM daily transaction volume data. Panels: (a) observations 1 to 300, (b) observations 1,300 to 1,300 and (c) observations 2,000 to 2,000.

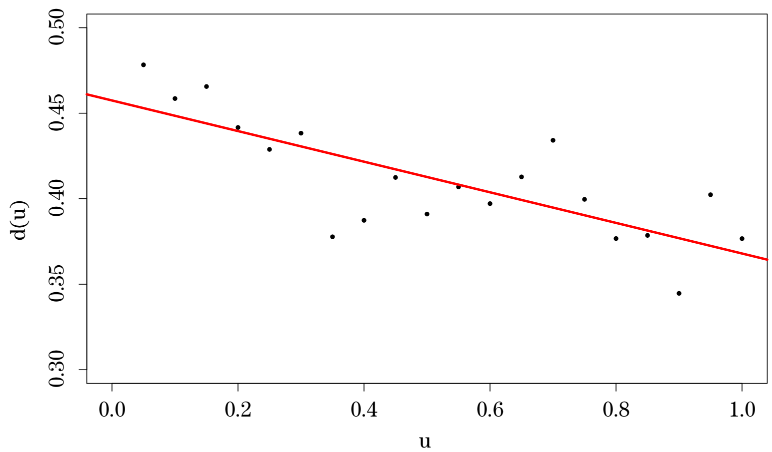

Figure 11. Rolling-window estimation of the long-memory parameter $d(u)$ as a function of u.

due to the oil crisis in 1973. Figure 8 shows the histogram of the centered series. Note that the data do not seem to be normally distributed. On the other hand, Figure 9 displays the sample autocorrelation function (ACF) of this series, along with a variance plot. Both graphs suggest the presence of long-range dependence. Moreover, Figure 10 shows three segments of data: a) observations 1 to 300; (b) observations 1,000 to 2,300 and (c) observations 2,000 to 2,300. Note that the sample ACF seems to decrease over time. Based on this plot, Figure 11 reports a rolling window estimation of the long-memory parameter $d(u)$ as a function of the standardized time $u = t/T$. These rolling-windows correspond to estimates of d on successive blocks of 350 observations and shifts of 100 days. Note that the long-memory parameter seems to decrease over time. Consequently, a locally stationary process was selected using the AIC. The fitted model is a LS-ARFIMA$(1, d, 0)$, defined as follows

$$Y_{t,T} = \beta\,(1 - \phi B)^{-1}\,(1 - B)^{-d(t/T)}\varepsilon_t,$$

with $d(u) = \alpha_0 + \alpha_1\,u$, and a centered log-normal input noise ε_t. The results of the model fitting are reported in Table 3. Observe that the standard errors vary significantly when assuming normality or log-normality.

A residual analysis indicates that the Box–Ljung test statistic is 15.165 with 10 degrees of freedom, producing a P-value $= 0.1262$. Finally, Figure 12 shows a histogram of the residuals.

Table 3. Whittle estimation, IBM transaction data.

Parameter	Estimate	Std. Error	z-value	P-value
ϕ	0.1315	0.0361	3.6402	0.0003
α_0	0.4158	0.0302	13.7810	0.0000
α_1	-0.2140	0.0339	-6.3054	0.0000
β (Normal)	370,748.0189	5,253.0320	70.5779	0.0000
β (Lognormal)	370,748.0189	39,447.3300	9.3985	0.0000

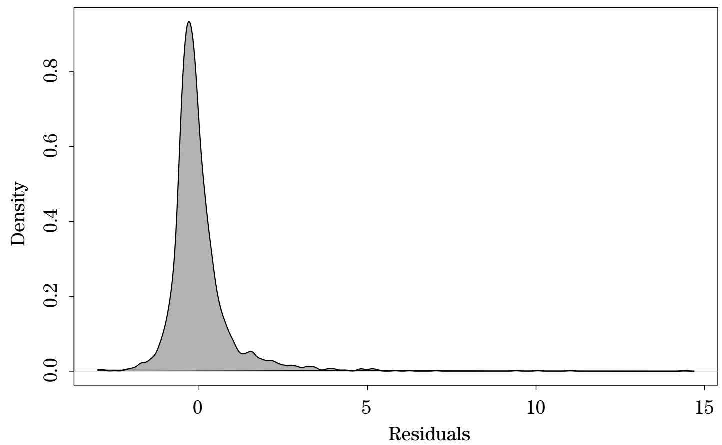

Figure 12. Estimated probability density of the residuals.

7. Conclusion

This study examines asymptotic and finite-sample properties of a spectrum-based Whittle parameter estimation procedure for locally stationary long-range dependent models. In particular, it establishes a central limit theorem for the Whittle estimator. This result extends previous works on Gaussian time series to the non-Gaussian cases. Finite-sample performance of the estimates is investigated using several Monte Carlo experiments, including gamma and log-normal distributions. In these simulation studies, the proposed estimator exhibits a very good performance and the empirical precisions obtained are close to their theoretical counterparts, as specified by Theorem 1. In addition, a real-life data application is presented. Thus, we have shown that the Whittle methodology can be applied to a much broader class of situations.

Acknowledgements

The authors would like to thank the two anonymous referees and the Editor for their constructive feedback and suggestions. This research is supported in part by the General Research Fund of HKSAR-RGC-GRF Nos. 14300514 and 14325216, HKSAR-RGCCRF: CityU8/CRG/12G, the Theme-based Research Scheme of HKSAR-RGC-TBS T32-101/15-R (Chan), and Fondecyt Grant 11608 61 (Palma).

References

Beran, J. (2009). On parameter estimation for locally stationary long-memory processes. *J. Statist. Plann. Inference* **139**, 900–915.

Chandler, G. and Polonik, W. (2017). Residual empirical processes and weighted sums for time-varying processes with applications to testing for homoscedasticity. *J. Time Series Anal.* **38**, 72–98.

Chen, X. B., Gao, J., Li, D. and Silvapulle, P. (2018). Nonparametric estimation and forecasting for time-varying coefficient realized volatility models. *J. Bus. Econom. Statist.* **36**, 88–100.

Dahlhaus, R. (1996). On the Kullback-Leibler information divergence of locally stationary processes. *Stochastic Process. Appl.* **62**, 139–168.

Dahlhaus, R. (1997). Fitting time series models to nonstationary processes. *Ann. Statist.* **25**, 1–37.

Dahlhaus, R. (2012). Locally stationary processes. *Handbook of Statist.* **30**, Elsevier, 351–412.

Dette, H., Preuss, P. and Vetter, M. (2011). A measure of stationarity in locally stationary processes with applications to testing. *Journal of the American Statistical Association* **106**, 1113–1124.

Fiecas, M. and Ombao, H. (2016). Modeling the evolution of dynamic brain processes during an associative learning experiment. *J. Amer. Statist. Assoc.* in press.

Giraitis, L., Kapetanios, G. and Yates, T. (2014). Inference on stochastic time-varying coefficient models. *Journal of Econometrics* **179**, 46–65.

Guinness, J. and Fuentes, M. (2015). Likelihood approximations for big nonstationary spatial temporal lattice data. *Statistica Sinica* **25**, 329–349.

Guinness, J. and Stein, M. L. (2013). Transformation to approximate independence for locally stationary Gaussian processes. *Journal of Time Series Analysis* **34**, 574–590.

Kreiss, J. P. and Paparoditis, E. (2015). Bootstrapping locally stationary processes. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **77**, 267–290.

Leipus, R. and Surgailis, D. (2013). Asymptotics of partial sums of linear processes with changing memory parameter. *Lith. Math. J.* **53**, 196–219.

Nason, G. (2013). A test for second-order stationarity and approximate confidence intervals for localized autocovariances for locally stationary time series. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **75**, 879–904.

Palma, W. (2010). On the sample mean of locally stationary long-memory processes. *J. Statist. Plann. Inference* **140**, 3764–3774.

Palma, W. (2016). *Time Series Analysis*. Wiley Series in Probability and Statistics. Hoboken, New Jersey: John Wiley & Sons, Hoboken, New Jersey.

Palma, W. and Olea, R. (2010). An efficient estimator for locally stationary Gaussian long-memory processes. *Ann. Statist.* **38**, 2958–2997.

Paparoditis, E. and Preuss, P. (2015). On local power properties of frequency domain-based tests for stationarity in press.

Preuss, P., Puchstein, R. and Dette, H. (2015). Detection of multiple structural breaks in multivariate time series. *Journal of the American Statistical Association* **110**, 654–668.

Preuss, P., Vetter, M. and Dette, H. (2013a). A test for stationarity based on empirical processes. *Bernoulli* **19**, 2715–2749.

Preuss, P., Vetter, M. and Dette, H. (2013b). Testing semiparametric hypotheses in locally stationary processes. *Scandinavian Journal of Statistics* **40**, 417–437.

Priestley, M. B. (1965). Evolutionary spectra and non-stationary processes. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **27**, 204–237.

Puchstein, R. and Preuss, P. (2016). Testing for stationarity in multivariate locally stationary processes. *Journal of Time Series Analysis* **37**, 3–29.

Rosen, O., Wood, S. and Stoffer, D. (2012). Adaptspec: Adaptive spectral estimation for non-stationary time series. *J. Amer. Statist. Assoc.* **107**, 1575–1589.

Roueff, F. and Von Sachs, R. (2011). Locally stationary long memory estimation. *Stochastic Processes and their Applications* **121**, 813–844.

Song, R., Banerjee, M. and Kosorok, M. (2016). Asymptotics for change-point models under varying degrees of mis-specification. *The Annals of Statistics* **44**, 153–182.

Van Bellegem, S. and Dahlhaus, R. (2006). Semiparametric estimation by model selection for locally stationary processes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **68**, 721–746.

Vogt, M. and Dette, H. (2015). Detecting gradual changes in locally stationary processes. *Ann. Statist.* **43**, 713–740.

Whittle, P. (1953). Estimation and information in stationary time series. *Ark. Mat.* **2**, 423–434.

Wu, W. and Zhou, Z. (2014). Simultaneous quantile inference for non-stationary long memory time series. Preprint, University of Toronto.

Wu, W. B. and Zhou, Z. (2011). Gaussian approximations for non-stationary multiple time series. *Statist. Sinica* **21**, 1397–1413.

Zhou, Z. (2013). Heteroscedasticity and autocorrelation robust structural change detection. *Journal of the American Statistical Association* **108**, 726–740.

Zhou, Z. (2014). Inference of weighted v-statistics for nonstationary time series and its applications. *The Annals of Statistics* **42**, 87–114.

Department of Statistics, The Chinese University of Hong Kong, Shatin, NT, Hong Kong.

E-mail: nhchan@sta.cuhk.edu.hk

Department of Statistics, Pontificia Universidad Catolica de Chile, Edificio Rolando Chuaqui, Campus San Joaqun. Avda., Vic. Mackenna 4860, Macul, Chile.

E-mail: wilfredo@mat.puc.cl

(Received December 2016; accepted March 2018)