The Annals of Statistics
2010, Vol. 38, No. 5, 2958–2997
DOI: 10.1214/10-AOS812

AN EFFICIENT ESTIMATOR FOR LOCALLY STATIONARY GAUSSIAN LONG-MEMORY PROCESSES[1]

BY WILFREDO PALMA AND RICARDO OLEA

Pontificia Universidad Católica de Chile

This paper addresses the estimation of locally stationary long-range dependent processes, a methodology that allows the statistical analysis of time series data exhibiting both nonstationarity and strong dependency. A time-varying parametric formulation of these models is introduced and a Whittle likelihood technique is proposed for estimating the parameters involved. Large sample properties of these Whittle estimates such as consistency, normality and efficiency are established in this work. Furthermore, the finite sample behavior of the estimators is investigated through Monte Carlo experiments. As a result from these simulations, we show that the estimates behave well even for relatively small sample sizes.

1. Introduction. Even though stationarity is a very attractive theoretical assumption, in practice most time series data fail to meet this condition. As a consequence, several approaches to deal with nonstationarity have been proposed in the literature. Among these methodologies, differentiation and trend removal are popular choices. Other approaches include, for instance, the evolutionary spectral techniques first discussed by Priestley (1965). In a similar spirit, during the last decades a number of new time-varying dependence models have been proposed. One of these methodologies, the so-called locally stationary processes developed by Dahlhaus (1996, 1997), has been widely discussed in the recent time series literature, see, for example, Dahlhaus (2000), von Sachs and MacGibbon (2000), Jensen and Whitcher (2000), Guo et al. (2003), Genton and Perrin (2004), Orbe, Ferreira and Rodriguez-Poo (2005), Dahlhaus and Polonik (2006, 2009), Chandler and Polonik (2006), Fryzlewicz, Sapatinas and Subba Rao (2006) and Beran (2009), among others. This approach allows the stochastic process to be nonstationary, but assuming that the time variation of the model is sufficiently smooth so that it can be locally approximated by stationary processes.

On the other hand, during the last decades, long-range dependent data have arisen in disciplines as diverse as meteorology, hydrology, economics, etc., see, for example, the recent surveys by Doukhan, Oppenheim and Taqqu (2003) and Palma (2007). As a consequence, statistical methods for modeling that type of

Received April 2009; revised January 2010.

[1]Supported in part by Fondecyt Grant 1085239.

AMS 2000 subject classifications. Primary 62M10; secondary 60G15.

Key words and phrases. Nonstationarity, local stationarity, long-range dependence, Whittle estimation, consistency, asymptotic normality, efficiency.

data are of great interest to scientists and practitioners from many fields. At the same time, many of these long-memory data also display nonstationary behavior, see, for instance, Granger and Ding (1996), Jensen and Whitcher (2000) and Beran (2009). Nevertheless, most of the currently available methods for dealing with long-range dependence are incapable of modeling time series with these features. In particular, much of the theory of locally stationary processes applies only to time series with short memory, such as time-varying autoregressive moving average (ARMA) processes and not to time series exhibiting both nonstationarity and strong dependence. In order to treat that type of data, this paper addresses a class of strongly dependent locally stationary processes. In particular, these models include a Hurst parameter which evolves over time. Following Dahlhaus (1997), we propose a Whittle maximum likelihood estimation technique for fitting Gaussian long-memory locally stationary models. This is an extension of the spectrum-based likelihood estimator introduced by Whittle (1953). A great advantage of this estimation procedure is its computational efficiency, since it only requires the calculation of the periodogram by means of the fast Fourier transform. Additionally, we prove in this article that the proposed Whittle estimator is asymptotically consistent, normally distributed and efficient. Thus, this paper provides a framework for modeling and making statistical inferences about several types of nonstationarities that may be difficult to handle with other techniques. For instance, changes in the variance of a time series could be spotted by simple inspection of the data. However, variations on the dependence structure of the data are far more difficult to uncover and model.

The remainder of this paper is structured as follows. Section 2 discusses a class of long-memory locally stationary processes and proposes a quasi maximum likelihood estimator based on an extended version of the Whittle spectrum-based methodology. The consistency, asymptotic normality and efficiency of these quasi maximum likelihood estimators are established. Applications of the asymptotic results to some specific locally stationary processes are also presented in this section. Proofs of the theorems are provided in Section 3. Note that the techniques employed by Dahlhaus (1997) to show the asymptotic properties of the Whittle estimates are no longer valid for the class of long-memory locally stationary processes discussed in this paper. This difficulty is due to the fact that these processes have an unbounded time-varying spectral density at zero frequency. Consequently, several technical results must be introduced and proved. Section 4 reports the results from several Monte Carlo experiments which allow to gain some insight into the finite sample behavior of the Whittle estimates. Conclusions are presented in Section 5 while auxiliary lemmas are provided in a technical appendix. Additional examples and simulations along with a comparison of the Whittle estimator with a kernel maximum likelihood estimation approach and two real-life applications of the proposed methodology can be found in Palma and Olea (2010). The bandwidth selection problem for the locally stationary Whittle estimator is also discussed in that paper, from an empirical perspective.

2. Definitions and main results.

2.1. *Long-memory locally stationary processes.* A class of Gaussian locally stationary process with transfer function A^0 can be defined by the spectral representation

$$Y_{t,T} = \int_{-\pi}^{\pi} A^0_{t,T}(\lambda) e^{i\lambda t}\, dB(\lambda), \tag{1}$$

for $t = 1, \dots, T$, where $B(\lambda)$ is a Brownian motion on $[-\pi, \pi]$ and there is a positive constant K and a 2π-periodic function $A : (0, 1] \times \mathbb{R} \to \mathbb{C}$ with $A(u, -\lambda) = \overline{A(u, \lambda)}$ such that

$$\sup_{t,\lambda} \left| A^0_{t,T}(\lambda) - A\left(\frac{t}{T}, \lambda\right) \right| \le \frac{K}{T}, \tag{2}$$

for all T. The transfer function $A^0_{t,T}(\lambda)$ of this class of nontstationary processes changes smoothly over time so that they can be locally approximated by stationary processes. An example of this class of locally stationary processes is given by the infinite moving average expansion

$$Y_{t,T} = \sigma\left(\frac{t}{T}\right) \sum_{j=0}^{\infty} \psi_j\left(\frac{t}{T}\right) \varepsilon_{t-j}, \tag{3}$$

where $\{\varepsilon_t\}$ is a zero-mean and unit variance Gaussian white noise and $\{\psi_j(u)\}$ are coefficients satisfying $\sum_{j=0}^{\infty} \psi_j(u)^2 < \infty$ for all $u \in [0, 1]$. In this case, the transfer function of process (3) is given by $A^0_{t,T}(\lambda) = \sigma(\frac{t}{T}) \sum_{j=0}^{\infty} \psi_j(\frac{t}{T}) e^{-i\lambda j} = A(\frac{t}{T}, \lambda)$, so that condition (2) is satisfied. The model defined by (3) generalizes the Wold expansion for a linear stationary process allowing the coefficients of the infinite moving average expansion vary smoothly over time. A particular case of (3) is the generalized version of the fractional noise process described by the discrete-time equation

$$Y_{t,T} = \sigma\left(\frac{t}{T}\right)(1 - B)^{-d(t/T)} \varepsilon_t = \sigma\left(\frac{t}{T}\right) \sum_{j=0}^{\infty} \eta_j\left(\frac{t}{T}\right) \varepsilon_{t-j}, \tag{4}$$

for $t = 1, \dots, T$, where $\{\varepsilon_t\}$ is a Gaussian white noise sequence with zero mean and unit variance and the infinite moving average coefficients $\{\eta_j(u)\}$ are given by

$$\eta_j(u) = \frac{\Gamma[j + d(u)]}{\Gamma(j + 1)\Gamma[d(u)]}, \tag{5}$$

where $\Gamma(\cdot)$ is the Gamma function and $d(\cdot)$ is a smoothly time-varying long-memory coefficient. For simplicity, the locally stationary fractional noise process (4) will be denoted as LSFN.

A natural extension of the LSFN model is the locally stationary autoregressive fractionally integrated moving average (LSARFIMA) process defined by the equation

$$\Phi(t/T, B)Y_{t,T} = \sigma(t/T)\Theta(t/T, B)(1-B)^{-d(t/T)}\varepsilon_t, \tag{6}$$

for $t = 1, \dots, T$, 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 Gaussian white noise sequence with zero mean and unit variance. This class of models extends the well-known ARFIMA process, which is obtained when the components $\Phi(u, B)$, $\Theta(u, B)$, $d(u)$ and $\sigma(u)$ appearing in (6) do not depend on u. Note that by Theorem 4.3 of Dahlhaus (1996), under some regularity conditions on the polynomial $\Phi(u, B)$, the model defined by (6) satisfies (1) and (2), see Jensen and Whitcher (2000) for details.

2.2. *Estimation.* Let $\theta \in \Theta$ be a parameter vector specifying model (1) where the parameter space Θ is a subset of a finite-dimensional Euclidean space. Given a sample $\{Y_{1,T}, \dots, Y_{T,T}\}$ of the process (1) we can estimate θ by minimizing the Whittle log-likelihood function

$$\mathcal{L}_T(\theta) = \frac{1}{4\pi}\frac{1}{M}\int_{-\pi}^{\pi}\sum_{j=1}^{M}\left\{\log f_\theta(u_j, \lambda) + \frac{I_N(u_j, \lambda)}{f_\theta(u_j, \lambda)}\right\}d\lambda, \tag{7}$$

where $f_\theta(u, \lambda) = |A_\theta(u, \lambda)|^2$ is the time-varying spectral density of the limiting process specified by the parameter θ, $I_N(u, \lambda) = \frac{|D_N(u,\lambda)|^2}{2\pi H_{2,N}(0)}$ is a tapered periodogram with

$$D_N(u, \lambda) = \sum_{s=0}^{N-1} h\left(\frac{s}{N}\right) Y_{[uT]-N/2+s+1,T}\, e^{-i\lambda s}, \qquad H_{k,N} = \sum_{s=0}^{N-1} h\left(\frac{s}{N}\right)^k e^{-i\lambda s},$$

$T = S(M-1) + N$, $u_j = t_j/T$, $t_j = S(j-1) + N/2$, $j = 1, \dots, M$ and $h(\cdot)$ is a data taper. The intuition behind this extended version of the Whittle estimation procedure (7) is as follows: the sample $\{Y_{1,T}, \dots, Y_{T,T}\}$ is subdivided into M blocks of length N each shifting S places from block to block. For instance, if we split a time series of $T = 652$ observations into $M = 100$ blocks of length $N = 256$ each, shifting $S = 4$ positions forward each time we get the blocks $(Y_{1,652}, Y_{2,652}, \dots, Y_{256,652}), \dots, (Y_{397,652}, Y_{398,652}, \dots, Y_{652,652})$. Then, the spectrum is locally estimated by means of the data tapered periodogram on each one of these $M = 100$ blocks and then averaged to form (7). Finally, the Whittle estimator of the parameter vector θ is given by

$$\widehat{\theta}_T = \arg\min \mathcal{L}_T(\theta), \tag{8}$$

where the minimization is over a parameter space Θ. The analysis of the asymptotic properties of the Whittle locally stationary estimates (8) is discussed in detail next. Before stating these results, we introduce a set of the regularity conditions.

2.3. *Assumptions*. The first assumption below is concerned with the time-varying spectral density of the process. The second assumption is related to the data tapering function and 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 could be different from line to line.

A1. The time-varying spectral density of the limiting process (1) is strictly positive and satisfies

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

as $|\lambda| \to 0$, where $C_f(\theta,u) > 0$, $0 < \inf_{\theta,u} d_\theta(u)$, $\sup_{\theta,u} d_\theta(u) < \frac{1}{2}$ and $d_\theta(u)$ has bounded first derivative with respect to u. 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

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

as $k \to \infty$, where $|C(\theta,u,v)| \le K$ 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.

A2. The data taper $h(u)$ is a positive, bounded function for $u \in [0,1]$ and symmetric around $\frac{1}{2}$ with a bounded derivative.

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

EXAMPLE 2.1. As an illustration of the assumptions described above, consider the extension of the usual fractional noise process with time-varying Hurst parameter, described by (4) and (5). The spectral density of this LSFN process is given by

$$f_\theta(u,\lambda) = \frac{\sigma^2}{2\pi}\left(2\sin\frac{\lambda}{2}\right)^{-2d_\theta(u)}.$$

Note that this function is integrable over $\lambda \in [-\pi,\pi]$ for every $u \in [0,1]$ as long as $d_\theta(u) < \frac{1}{2}$ for all $u \in [0,1]$ and $\theta \in \Theta$. Furthermore, we have that $f_\theta(u,\lambda) \sim \frac{\sigma^2}{2\pi}|\lambda|^{-2d_\theta(u)}$, as $\lambda \to 0$. By assuming that $|\nabla_\theta d_\theta(u)| \le K$, we have that $|\nabla_\theta \log f_\theta(u,\lambda)| = |\nabla_\theta d_\theta(u)||\log(2\sin\frac{\lambda}{2})^2| \le K|\log|\lambda||$, which is an integrable function in $\lambda \in [-\pi,\pi]$. In addition, from (5) the function $A(u,\lambda)$ of this process satisfies

$$\int_{-\pi}^{\pi} A(u,\lambda)A(v,-\lambda)\exp(ik\lambda)\,d\lambda = \frac{\Gamma[1-d_\theta(u)-d_\theta(v)]\Gamma[k+d_\theta(u)]}{\Gamma[1-d_\theta(u)]\Gamma[d_\theta(u)]\Gamma[k+1-d_\theta(v)]},$$

for $k \ge 0$. Thus, by Stirling's approximation, we get

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

for $k \to \infty$. Besides, a simple calculation shows that $f_\theta(u,\lambda)^{-1}$ is twice differentiable with respect to u and λ as long as $d_\theta(u)$ is twice differentiable with respect to u. Thus, under these conditions the time-varying spectral density $f_\theta(u,\lambda)$ satisfies assumption A1. On the other hand, an example of data taper that satisfies assumption A2 is the cosine bell function

$$h(x) = \tfrac{1}{2}[1 - \cos(2\pi x)]. \tag{9}$$

Note that if $S = \mathcal{O}(N^a)$ and $M = \mathcal{O}(N^b)$ then $T = \mathcal{O}(N^{a+b})$ for $a + b \geq 1$. Thus, by choosing exponents a and b such that $(a,b) \in \mathcal{C} = \{a < 1, \frac{3}{2} < a + b < 2, a < b\}$, assumption A3 is fulfilled. Observe that the $\mathcal{C}$ is a nonempty set.

2.4. *Main results.* Some fundamental large sample properties of the Whittle quasi-likelihood estimators (8), including consistency, asymptotic normality and efficiency are established next. In addition, we establish an asymptotic result about the estimation of the time-varying long-memory parameter for a class of locally stationary processes. The proofs of these four results are provided in Section 3.

THEOREM 2.1 (Consistency). *Let θ_0 be the true value of the parameter θ. Under assumptions A1–A3, the estimator $\widehat{\theta}_T$ satisfies $\widehat{\theta}_T \to \theta_0$, in probability, as $T \to \infty$.*

THEOREM 2.2 (Normality). *Let θ_0 be the true value of the parameter θ. If assumptions A1–A3 hold, then the Whittle estimator $\widehat{\theta}_T$ satisfies a central limit theorem*

$$\sqrt{T}(\widehat{\theta}_T - \theta_0) \to N[0, \Gamma(\theta_0)^{-1}],$$

in distribution, as $T \to \infty$, where

$$\Gamma(\theta) = \frac{1}{4\pi} \int_0^1 \int_{-\pi}^{\pi} [\nabla \log f_\theta(u,\lambda)][\nabla \log f_\theta(u,\lambda)]' \, d\lambda \, du. \tag{10}$$

THEOREM 2.3 (Efficiency). *Assuming that conditions A1–A3 hold, the Whittle estimator $\widehat{\theta}_T$ is asymptotically Fisher efficient.*

REMARK 2.1. Recall that for a stationary fractional noise process FN(d), the asymptotic variance of the maximum likelihood estimate of the long-memory parameter, $\widehat{d}$, satisfies

$$\lim_{T\to\infty} T \operatorname{Var}(\widehat{d}) = \frac{6}{\pi^2}.$$

On the other hand, suppose that we consider a LSFN process where the long-memory parameter varies according to, for example, $d(u) = \alpha_0 + \alpha_1 u$. Thus, in order to estimate $d(u)$, the parameters α_0 and α_1 must be estimated. Let $\widehat{\alpha}_0$ and

$\widehat{\alpha}_1$ be their Whittle estimators, respectively, so that $\widehat{d}(u) = \widehat{\alpha}_0 + \widehat{\alpha}_1 u$. According to Theorem 2.2, the asymptotic variance of this estimate of $d(u)$ satisfies

$$\lim_{T\to\infty} T \operatorname{Var}[\widehat{d}(u)] = \frac{24}{\pi^2}(1 - 3u + 3u^2),$$

and then integrating over u we get

$$\lim_{T\to\infty} T \int_0^1 \operatorname{Var}[\widehat{d}(u)]\, du = \frac{12}{\pi^2}.$$

Since two parameters are being estimated, on the average, the asymptotic variance of the estimate $\widehat{d}(u)$ is twice the asymptotic variance of $\widehat{d}$ from a stationary FN process. This result can be generalized to the case where three or more coefficients are estimated and to more complex trends, as established on the following theorem.

THEOREM 2.4. *Consider a LSFN process* (4) *with time-varying long-memory parameter* $d_\beta(u) = \sum_{j=1}^{p} \beta_j g_j(u)$, *where* $\{g_j(u)\}$ *are basis functions as defined in* (12) *below. Let* $\widehat{d}(u) = \sum_{j=1}^{p} \widehat{\beta}_j g_j(u)$ *be the estimator of* $d_\beta(u)$ *for* $u \in [0, 1]$. *Then under assumptions* A1–A3 *we have that*

$$\lim_{T\to\infty} T \int_0^1 \operatorname{Var}[\widehat{d}(u)]\, du = \frac{6p}{\pi^2}. \tag{11}$$

REMARK 2.2. Note that according to Theorem 2.4 the limiting average of the variances of $d(u)$ given by (11) does not depend on the basis functions $g_j(\cdot)$ for $j = 1, \dots, p$.

2.5. *Illustrations.* As an illustration of the asymptotic results discussed above, consider the class of LSARFIMA models defined by (6). The evolution of these models can be specified in terms of a general class of functions. For example, let $\{g_j(u)\}$, $j = 1, 2, \dots,$ be a basis for a space of smoothly varying functions and let $d_\theta(u)$ be the time-varying long-memory parameter in model (6). Then we could write $d_\theta(u)$ in terms of the basis $\{g_j(u)\}$ as follows:

$$\ell[d_\theta(u)] = \sum_{j=0}^{k} \alpha_j g_j(u), \tag{12}$$

for unknown values of k and $\theta = (\alpha_0, \alpha_1, \dots, \alpha_k)'$, where $\ell(\cdot)$ is a known link function. In this situation, estimating θ involves determining k and estimating the coefficients $\alpha_0, \dots, \alpha_k$. Important examples of this approach are the classes of polynomials generated by the basis $\{g_j(u) = u^j\}$, Fourier expansions generated by the basis $\{g_j(u) = e^{iuj}\}$ and wavelets generated by, for instance, the Haar or Daubechies systems. Extensions of these cases can also be considered. For example, the basis functions could also include parameters as in the case $\{g_j(u) = e^{iu\beta_j}\}$, where $\{\beta_j\}$ are unknown values.

In order to illustrate the application of the theoretical results established in Section 2.4, we discuss next a number of combinations of polynomial and harmonic evolutions of the long-memory parameter, the noise variance, the autoregressive and moving average components of the LSARFIMA process (6). Additional examples are provided in Section 2 of Palma and Olea (2010).

EXAMPLE 2.2. Consider first the case $P = Q = 0$ in model (6) where $d(u)$ and $\sigma(u)$ are specified by

$$\ell_1[d(u)] = \sum_{j=0}^{p} \alpha_j g_j(u), \qquad \ell_2[\sigma(u)] = \sum_{j=0}^{q} \beta_j h_j(u),$$

for $u \in [0, 1]$, where $\ell_1(\cdot)$ and $\ell_2(\cdot)$ are differentiable link functions, $g_j(\cdot)$ and $h_j(\cdot)$ are basis functions. The parameter vector in this case is $\theta = (\alpha_0, \dots, \alpha_p, \beta_0, \dots, \beta_q)'$ and the matrix Γ can be written as

$$\Gamma = \begin{pmatrix} \Gamma_\alpha & 0 \\ 0 & \Gamma_\beta \end{pmatrix}, \tag{13}$$

where

$$\Gamma_\alpha = \frac{\pi^2}{6}\left[\int_0^1 \frac{g_i(u) g_j(u)\,du}{[\ell_1'(d(u))]^2}\right]_{i,j=0,\dots,p},$$

$$\Gamma_\beta = 2\left[\int_0^1 \frac{h_i(u) h_j(u)\,du}{[\sigma(u)\ell_2'(\sigma(u))]^2}\right]_{i,j=0,\dots,q}.$$

EXAMPLE 2.3. As a particular case of the parameter specification of the previous example, consider the case $P = Q = 0$ in model (6) where $d(u)$ and $\sigma(u)$ are both specified by polynomials,

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

for $u \in [0, 1]$. Similar to Example 2.2, in this case the parameter vector is $\theta = (\alpha_0, \dots, \alpha_p, \beta_0, \dots, \beta_q)'$, $\ell_1(u) = \ell_2(u) = u$ and the matrix Γ given by (10) can be written as in (13) with

$$\Gamma_\alpha = \left[\frac{\pi^2}{6(i + j + 1)}\right]_{i,j=0,\dots,p},$$

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

The above integrals can be evaluated by standard calculus procedures; see, for example, Gradshteyn and Ryzhik [(2000), page 64] or by numerical integration.

EXAMPLE 2.4. Considering now a similar setup as Example 2.3 with $p = q = 1$, but with link function $\ell(\cdot) = \log(\cdot)$ such that

$$\log[d(u)] = \alpha_0 + \alpha_1 u, \qquad \log[\sigma(u)] = \beta_0 + \beta_1 u,$$

for $u \in [0, 1]$. Then Γ can be written as (13) with

$$\Gamma_\alpha = \frac{\pi^2}{6}\frac{e^{2\alpha_0}}{4\alpha_1^3}\begin{bmatrix} 2\alpha_1^2(e^{2\alpha_1}-1) & \alpha_1\big((2\alpha_1-1)e^{2\alpha_1}+1\big) \\ \alpha_1\big((2\alpha_1-1)e^{2\alpha_1}+1\big) & (2\alpha_1^2-2\alpha_1+1)e^{2\alpha_1}+1 \end{bmatrix},$$

$$\Gamma_\beta = \begin{bmatrix} 2 & 1 \\ 1 & 2/3 \end{bmatrix}.$$

EXAMPLE 2.5. Following with the assumption $P = Q = 0$ in model (6), consider that $d(u)$ and $\sigma(u)$ are defined by the harmonic expansions

$$d(u) = \alpha_0 + \alpha_1\cos(\lambda_1 u) + \cdots + \alpha_p\cos(\lambda_p u),$$
$$\sigma(u) = \beta_0 + \beta_1\cos(\omega_1 u) + \cdots + \beta_q\cos(\omega_q u),$$

for $u \in [0, 1]$, where $\lambda_0 = 0$, $\lambda_i^2 \neq \lambda_j^2$ for all $i, j = 0, \dots, p$, $i \neq j$, $\omega_0 = 0$ and $\omega_i^2 \neq \omega_j^2$ for all $i, j = 0, \dots, q$, $i \neq j$. For simplicity, the values of the frequencies $\{\lambda_j\}$ and $\{\omega_j\}$ are assumed to be known. As in Example 2.3, in this case the parameter vector is $\theta = (\alpha_0, \dots, \alpha_p, \beta_0, \dots, \beta_q)'$ and the matrix Γ appearing in (10) can be written as in (13) with

$$\Gamma_\alpha = \frac{\pi^2}{12}\left[\frac{\sin(\lambda_i - \lambda_j)}{\lambda_i - \lambda_j} + \frac{\sin(\lambda_i + \lambda_j)}{\lambda_i + \lambda_j}\right]_{i,j=0,\dots,p}$$

and

$$\Gamma_\beta = \frac{\pi^2}{12}\left[\frac{\sin(\omega_i - \omega_j)}{\omega_i - \omega_j} + \frac{\sin(\omega_i + \omega_j)}{\omega_i + \omega_j}\right]_{i,j=0,\dots,q}.$$

EXAMPLE 2.6. Consider now the case $P = Q = 1$ in model (6) where $\sigma(u) = 1$ and $d(u)$, $\Phi(u, B)$, $\Theta(u, B)$ are specified by

$$d(u) = \alpha_1 u,$$
$$\Phi(u, B) = 1 + \phi(u)B, \qquad \phi(u) = \alpha_2 u,$$
$$\Theta(u, B) = 1 + \theta(u)B, \qquad \theta(u) = \alpha_3 u,$$

for $u \in [0, 1]$. In this case, the parameter vector is $\theta = (\alpha_1, \alpha_2, \alpha_3)'$, with $0 < \alpha_1 < \frac{1}{2}$, $|\alpha_j| < 1$, $j = 1, 2$ and the matrix Γ from (10) can be written as

$$\Gamma = \begin{pmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} \\ \gamma_{21} & \gamma_{22} & \gamma_{23} \\ \gamma_{31} & \gamma_{32} & \gamma_{33} \end{pmatrix},$$

where

$$\gamma_{11} = \frac{1}{2\alpha_1^3}\log\frac{1+\alpha_1}{1-\alpha_1} - \frac{1}{\alpha_1^2}, \qquad \gamma_{12} = \frac{g(\alpha_1\alpha_2)}{(\alpha_1\alpha_2)^{3/2}} - \frac{1}{\alpha_1\alpha_2},$$

$$\gamma_{13} = \frac{1}{2\alpha_1}\left\{\left[\frac{1}{2} - \frac{1}{\alpha_1}\right] - \left[1 - \frac{1}{\alpha_1^2}\right]\log(1+\alpha_1)\right\},$$

$$\gamma_{22} = \frac{1}{2\alpha_2^3}\log\frac{1+\alpha_2}{1-\alpha_2} - \frac{1}{\alpha_2^2},$$

$$\gamma_{23} = \frac{1}{2\alpha_2}\left\{\left[1 - \frac{1}{\alpha_2^2}\right]\log(1+\alpha_2) - \left[\frac{1}{2} - \frac{1}{\alpha_2}\right]\right\}, \qquad \gamma_{33} = \frac{\pi^2}{18},$$

with $g(x) = \operatorname{arctanh}(\sqrt{x})$ for $x \in (0,1)$ and $g(x) = \arctan(\sqrt{-x})$ for $x \in (-1,0)$.

3. Proofs. This section is devoted to the proof of Theorems 2.1–2.4. Before presenting the proofs of these results, we introduce and prove three useful propositions which are of independent interest. These propositions involve the large sample properties of the functional operator defined next. Consider the function $\phi : [0,1] \times [-\pi,\pi] \to \mathbb{R}$ and define the functional operator

$$J(\phi) = \int_0^1 \int_{-\pi}^{\pi} \phi(u,\lambda) f(u,\lambda)\, d\lambda\, du, \tag{14}$$

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

$$J_T(\phi) = \frac{1}{M}\sum_{j=1}^{M} \int_{-\pi}^{\pi} \phi(u_j,\lambda) I_N(u_j,\lambda)\, d\lambda, \tag{15}$$

where M and u_j, $j = 1,\dots,M$ are given in Section 2. Furthermore, define the matrix

$$Q(u) = \left(\int_{-\pi}^{\pi} \phi(u,\lambda) e^{i\lambda(s-t)\, d\lambda}\right)_{s,t=1,\dots,N}, \tag{16}$$

and the block-diagonal matrix $Q(\phi) = \operatorname{diag}[Q(u_1),\dots,Q(u_M)]$. For notational simplicity, sometimes in what follows we have dropped θ from $d_\theta(u)$ so that it becomes $d(u)$.

REMARK 3.1. Since the function $A(u,\lambda)$ and the spectral density $f(u,\lambda)$ of a locally stationary long-memory process are unbounded at zero frequency, the techniques used next to prove the large sample properties of $J(\phi)$ and the quasi-likelihood estimators are different from those used in the short-memory context. For instance, the function $A(u,\lambda)$ does not satisfy the key assumption A.1 of Dahlhaus (1997) or the coefficients $\psi_j(\frac{t}{T})$ of (3) fail to meet conditions (2) and (3) of Dahlhaus and Polonik (2009). Due to the unboundeness of $f(u,\lambda)$ at the origin, our proofs exploit the properties of the Fourier transforms

$$\widehat{f}(u,\cdot) = \int_{-\pi}^{\pi} f(u,\lambda)e^{i\lambda\cdot}\,d\lambda,$$

$$\widehat{f}(u,v,\cdot) := \int_{-\pi}^{\pi} A(u,\lambda)A(v,-\lambda)e^{i\lambda\cdot}\,d\lambda.$$

3.1. *Propositions.*

PROPOSITION 1. *Let $f(u,\lambda)$ be a time-varying spectral density satisfying assumption A1 and assume that the function $\phi(u,\lambda)$ appearing in (14) is symmetric in λ and twice differentiable with respect to u. Let $\widehat{f}(u,k)$ and $\widehat{\phi}(u,k)$ be their Fourier coefficients, respectively. If there is a positive constant K such that*

$$|\widehat{f}(u,k)\widehat{\phi}(u,k)| \le K\left(\frac{\log k}{k^2}\right),$$

for all $u \in [0,1]$ and $k > 1$, then, under assumptions A2 and A3 we have that

$$E[J_T(\phi)] = J(\phi) + \mathcal{O}\left(\frac{\log^2 N}{N}\right) + \mathcal{O}\left(\frac{1}{M}\right).$$

PROOF. From definition (15), we can write

$$\begin{aligned} E[J_T(\phi)] &= \frac{1}{M}\sum_{j=1}^{M}\int_{-\pi}^{\pi}\phi(u_j,\lambda)E[I_N(u_j,\lambda)]\,d\lambda \\ &= \frac{1}{2\pi M H_{2,N}(0)}\sum_{j=1}^{M}\int_{-\pi}^{\pi}\phi(u_j,\lambda)E|D_N(u_j,\lambda)|^2\,d\lambda \\ &= \frac{1}{2\pi M H_{2,N}(0)}\sum_{j=1}^{M}\int_{-\pi}^{\pi}\phi(u_j,\lambda)\sum_{t,s=0}^{N-1}h\left(\frac{t}{N}\right)h\left(\frac{s}{N}\right) \\ &\qquad\times c(u_j,t,s)e^{i\lambda(s-t)}\,d\lambda, \end{aligned}$$

where

$$c(u,t,s) = E\big(Y_{[uT]-N/2+t+1,T}Y_{[uT]-N/2+s+1,T}\big).$$

Thus,

$$\begin{aligned} E[J_T(\phi)] &= \frac{1}{2\pi M H_{2,N}(0)} \\ &\quad\times\sum_{j=1}^{M}\sum_{t,s=0}^{N-1}h\left(\frac{t}{N}\right)h\left(\frac{s}{N}\right)c(u_j,t,s)\int_{-\pi}^{\pi}\phi(u_j,\lambda)e^{i\lambda(s-t)}\,d\lambda \end{aligned}$$

$$= \frac{1}{2\pi M H_{2,N}(0)} \sum_{j=1}^{M} \sum_{t,s=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) c(u_j, t, s) \widehat{\phi}(u_j, s-t)$$

$$= \frac{1}{2\pi M H_{2,N}(0)} \sum_{j=1}^{M} \sum_{t=0}^{N-1} \sum_{k=0}^{N-t} h\left(\frac{t}{N}\right) h\left(\frac{t}{N} + \frac{k}{N}\right) c(u_j, t, t+k) \times \widehat{\phi}(u_j, k)(2 - \delta_k),$$

where $\delta_k = 1$ for $k = 0$ and $\delta_k = 0$ for $k \neq 0$. By assumption A2 and Taylor's theorem,

$$h\left(\frac{t}{N} + \frac{k}{N}\right) = h\left(\frac{t}{N}\right) + h'(\xi_{t,k,N}) \frac{k}{N},$$

for some $\xi_{t,k,N} \in (\frac{t}{N}, \frac{t+k}{N})$, for $k \geq 0$. Thus,

$$\begin{aligned} E[J_T(\phi)] = {}& \frac{1}{2\pi M H_{2,N}(0)} \sum_{j=1}^{M} \sum_{t=0}^{N-1} \sum_{k=0}^{N-t} h^2\left(\frac{t}{N}\right) c(u_j, t, t+k) \times \widehat{\phi}(u_j, k)(2 - \delta_k) \\ & + \frac{1}{2\pi M H_{2,N}(0)} \sum_{j=1}^{M} \sum_{t=0}^{N-1} \sum_{k=0}^{N-t} h\left(\frac{t}{N}\right) h'(\xi_{t,k,N}) c(u_j, t, t+k) \times \widehat{\phi}(u_j, k)(2 - \delta_k). \end{aligned} \tag{17}$$

Under assumption A1, we can expand $c(u, t, t+k)$ by Taylor's theorem as

$$\begin{aligned} c(u, t, t+k) &= \widehat{f}(u, k) + \widehat{f}(u, k)\varphi_1(u, k)\left(\frac{t + 1 - N/2}{T}\right) \\ &\quad + \widehat{f}(u, k)\varphi_2(u, k)\left(\frac{k}{T}\right) + R(u, t, k, N, T), \end{aligned}$$

where

$$\varphi_1(u, k) = \frac{C_1(\theta, u, u)}{C(\theta, u, u)} + 2d'(u) \log k,$$

$$C_1(\theta, u, u) = \frac{\partial C(\theta, u, u+v)}{\partial u}\bigg|_{v=0},$$

$$\varphi_2(u, k) = \frac{C_2(\theta, u, u)}{C(\theta, u, u)} + d'(u) \log k,$$

$$C_2(\theta, u, u) = \frac{\partial C(\theta, u, u+v)}{\partial v}\bigg|_{v=0},$$

$d'(u) = \frac{\partial d_\theta(u)}{\partial u}$, $C(\theta, u, v)$ is defined in assumption A1 and the remainder term is given by

$$R(u,t,k,N,T) = \mathcal{O}\left\{\widehat{f}(u,k)\left[\left(\frac{k}{T}\right)^2 + \left(\frac{t}{T}\right)^2\right]\log^2 k\right\}.$$

Thus, since by assumption A1 $|d'(u)| \le K$ for all $u \in [0,1]$, we have $|\varphi_j(u,k)| \le K \log k$ for $j = 1, 2$ and $k > 1$. Now we can write

$$\begin{aligned}
&\sum_{k=0}^{N-t} c(u_j, t, t+k)\widehat{\phi}(u_j,k)(2-\delta_k) \\
&\quad= \sum_{k=0}^{N-t} \widehat{f}(u_j,k)\widehat{\phi}(u_j,k)(2-\delta_k) \\
&\qquad+ \sum_{k=0}^{N-t} \widehat{f}(u_j,k)\widehat{\phi}(u_j,k)\varphi_1(u_j,k)(2-\delta_k)\left(\frac{t+1-N/2}{T}\right) \\
&\qquad+ \sum_{k=0}^{N-t} \widehat{f}(u_j,k)\widehat{\phi}(u_j,k)\varphi_2(_j,k)(2-\delta_k)\frac{k}{T} \\
&\qquad+ \sum_{k=0}^{N-t} R(u_j,t,k,N,T)\widehat{\phi}(u_j,k)(2-\delta_k).
\end{aligned} \tag{18}$$

Since by assumption $|\widehat{f}(u,k)\widehat{\phi}(u,k)| \le K \log k/k^2$, for $k > 1$, uniformly in $u \in [0,1]$, we conclude that there is a finite limit $A(u) < \infty$ such that

$$A(u) = \lim_{N\to\infty} \sum_{k=0}^{N} \widehat{f}(u,k)\widehat{\phi}(u,k)(2-\delta_k).$$

Consequently,

$$\begin{aligned}
&\sum_{t=0}^{N-1} h^2\left(\frac{t}{N}\right) \sum_{k=0}^{N-t} \widehat{f}(u_j,k)\widehat{\phi}(u_j,k)(2-\delta_k) \\
&\quad= \sum_{t=0}^{N-1} h^2\left(\frac{t}{N}\right) \sum_{k=0}^{N-1} \widehat{f}(u_j,k)\widehat{\phi}(u_j,k)(2-\delta_k) \\
&\qquad- \sum_{t=0}^{N-1} h^2\left(\frac{t}{N}\right) \sum_{k=N-t+1}^{N-1} \widehat{f}(u_j,k)\widehat{\phi}(u_j,k)(2-\delta_k) \\
&\quad= \sum_{t=0}^{N-1} h^2\left(\frac{t}{N}\right)\left[A(u_j) - \sum_{k=N}^{\infty} \widehat{f}(u_j,k)\widehat{\phi}(u_j,k)(2-\delta_k)\right] \\
&\qquad+ \mathcal{O}(\log^2 N),
\end{aligned}$$

by Lemma 7. Hence,

$$\sum_{t=0}^{N-1} h^2\left(\frac{t}{N}\right) \sum_{k=0}^{N-t} \widehat{f}(u_j,k)\widehat{\phi}(u_j,k)(2-\delta_k)$$
$$= A(u_j) \sum_{t=0}^{N-1} h^2\left(\frac{t}{N}\right) - \sum_{t=0}^{N-1} h^2\left(\frac{t}{N}\right) \sum_{k=N}^{\infty} \widehat{f}(u_j,k)\widehat{\phi}(u_j,k)(2-\delta_k)$$
$$+ \mathcal{O}(\log^2 N).$$

But,

$$\left|\sum_{k=N}^{\infty} \widehat{f}(u_j,k)\widehat{\phi}(u_j,k)(2-\delta_k)\right| < K \sum_{k=N}^{\infty} \frac{\log k}{k^2} = \mathcal{O}\left(\frac{\log N}{N}\right),$$

and consequently,

$$\left|\sum_{t=0}^{N-1} h^2\left(\frac{t}{N}\right) \sum_{k=N}^{\infty} \widehat{f}(u_j,k)\widehat{\phi}(u_j,k)(2-\delta_k)\right| = \mathcal{O}(\log N).$$

Therefore,

$$\sum_{t=0}^{N-1} h^2\left(\frac{t}{N}\right) \sum_{k=0}^{N-t} \widehat{f}(u_j,k)\widehat{\phi}(u_j,k)(2-\delta_k) = A(u_j) \sum_{t=0}^{N-1} h^2\left(\frac{t}{N}\right) + \mathcal{O}(\log^2 N).$$

On the other hand, by analyzing the term involving the second summand of (18) we get

$$\sum_{t=0}^{N-1} h^2\left(\frac{t}{N}\right) \sum_{k=0}^{N-t} \varphi_1(u_j,k)\widehat{f}(u_j,k)\widehat{\phi}(u_j,k)(2-\delta_k)\left(\frac{t+1-N/2}{T}\right)$$
$$= \sum_{t=0}^{N-1} h^2\left(\frac{t}{N}\right)\left[\sum_{k=0}^{N-1} \varphi_1(u_j,k)\widehat{f}(u_j,k)\widehat{\phi}(u_j,k)(2-\delta_k)\right]\left(\frac{t+1-N/2}{T}\right)$$
$$+ \sum_{t=0}^{N-1} h^2\left(\frac{t}{N}\right) \sum_{k=N-t+1}^{N-1} \varphi_1(u_j,k)\widehat{\phi}(u_j,k)\widehat{f}(u_j,k)\left(\frac{t+1-N/2}{T}\right)$$
$$= \left[\sum_{t=0}^{N-1} h^2\left(\frac{t}{N}\right)\left(\frac{t+1-N/2}{T}\right)\right]\left[\sum_{k=0}^{N-1} \varphi_1(u_j,k)\widehat{f}(u_j,k)\widehat{\phi}(u_j,k)(2-\delta_k)\right]$$
$$+ \mathcal{O}\left(\frac{N\log^2 N}{T}\right),$$

by Lemma 8. Now, since $h(\cdot)$ is symmetric around $1/2$, we have

$$\sum_{t=0}^{N-1} h^2\left(\frac{t}{N}\right)\left(\frac{t+1-N/2}{T}\right) = \mathcal{O}\left(\frac{1}{T}\right).$$

Besides, $|\sum_{k=0}^{N-1}\varphi_1(u_j,k)\widehat{f}(u_j,k)\widehat{\phi}(u_j,k)(2-\delta_k)| \le K\sum_{k=1}^{N}\frac{(\log k)^2}{k^2} < \infty$. Consequently,

$$\sum_{t=0}^{N-1} h^2\Big(\frac{t}{N}\Big)\sum_{k=0}^{N-t}\varphi_1(u_j,k)\widehat{f}(u_j,k)\widehat{\phi}(u_j,k)(2-\delta_k)\Big(\frac{t+1-N/2}{T}\Big)$$
$$= \mathcal{O}\Big(\frac{N}{T}\log^2 N\Big).$$

The third term of (18) can be bounded as follows:

$$\Bigg|\sum_{k=0}^{N-t}\varphi_2(u_j,k)\widehat{f}(u_j,k)\widehat{\phi}(u_j,k)k\Bigg| \le K\sum_{k=1}^{N}\frac{\log k}{k} \le K\log^2 N,$$

and then

$$\Bigg|\sum_{t=0}^{N-1} h^2\Big(\frac{t}{N}\Big)\sum_{k=0}^{N-t}\varphi_2(u_j,k)\widehat{f}(u_j,k)\widehat{\phi}(u_j,k)\frac{k}{T}\Bigg| \le K\frac{N}{T}\log^2 N.$$

The last term of (18) can be bounded as follows:

$$\Bigg|\sum_{k=0}^{N-t} R(u_j,t,k,N,T)\widehat{\phi}(u_j,k)(2-\delta_k)\Bigg| \le K\log^2 N\Big(\frac{N}{T}\Big)^2,$$

and then

$$\Bigg|\sum_{t=0}^{N-1}\sum_{k=0}^{N-t} h^2\Big(\frac{t}{N}\Big)R(u_j,t,k,N,T)\widehat{\phi}(u_j,k)(2-\delta_k)\Bigg| \le K\frac{N^3}{T^2}\log^2 N.$$

Note that by assumption A3, the term above converges to zero as $N,T\to\infty$. Therefore, the first term in (17) can be written as

$$\frac{1}{2\pi M H_{2,N}(0)}\sum_{j=1}^{M}\sum_{t,k=0}^{N-1} h^2\Big(\frac{t}{N}\Big)c(u_j,t,t+h)\widehat{\phi}(u_j,h)(2-\delta_k)$$
$$= \frac{1}{2\pi M}\sum_{j=1}^{M} A(u_j) + \mathcal{O}\Big(\frac{\log^2 N}{N}\Big).$$

Now, by Lemma 1 we can write $A(u) = 2\pi\int_{-\pi}^{\pi}\phi(u,\omega)f(u,\omega)\,d\omega$, and then

$$\frac{1}{2\pi M}\sum_{j=1}^{M} A(u_j) = J(\phi) + \mathcal{O}\Big(\frac{1}{M}\Big). \tag{19}$$

On the other hand, the second term in (17) can be bounded as follows:

$$
\begin{aligned}
&|c(u_j, t, t+k)\widehat{\phi}(u_j, k)(2-\delta_k)| \\
&\quad \le K\biggl\{\widehat{f}(u_j, k)\widehat{\phi}(u_j, k) + \frac{N}{T}|\widehat{f}(u_j, k)\widehat{\phi}(u_j, k)\varphi_1(u_j, k)| \\
&\qquad + \frac{N}{T}|\widehat{f}(u_j, k)\widehat{\phi}(u_j, k)\varphi_2(u_j, k)| + \frac{N^2}{T^2}\log^2 N|\widehat{f}(u_j, k)\widehat{\phi}(u_j, k)|\biggr\}.
\end{aligned}
$$

Since $|\varphi_i(u_j, k)| \le K \log k$ for $i = 1, 2$, $j = 1, \dots, M$ and $k > 1$, we conclude that

$$
|c(u_j, t, t+k)\widehat{\phi}(u_j, k)(2-\delta_k)| \le K\frac{N}{T}\frac{\log k}{k^2}.
$$

Therefore, since $|h'(u)| \le K$ for $u \in [0, 1]$ by assumption A2, we have

$$
\left|\sum_{k=0}^{N-t} c(u_j, t, t+k)\widehat{\phi}(u_j, k)\frac{k}{N}h'(\xi_{t,k,N})\right| \le \frac{K}{T}\sum_{k=1}^{N}\frac{\log k}{k} \le K\frac{\log^2 N}{T}.
$$

Consequently,

$$
\left|\sum_{t=0}^{N-1} h\left(\frac{t}{N}\right)\sum_{k=0}^{N-t} c(u_j, t, t+k)\widehat{\phi}(u_j, k)(2-\delta_k)\frac{k}{N}h'(\xi_{t,k,N})\right| \le KN\frac{\log^2 N}{T}.
$$

Hence, the second term of (17) is bounded by $K(\log^2 N)/T$. From this and (19), the required result is obtained. □

PROPOSITION 2. *Let $f(u, \lambda)$ be a time-varying spectral density satisfying assumption A1. Let $\phi_1, \phi_2 : [0, 1] \times [-\pi, \pi] \to \mathbb{R}$ be two functions such that $\phi_1(u, \lambda)$ and $\phi_2(u, \lambda)$ are symmetric in λ, twice differentiable with respect to u and their Fourier coefficients satisfy $|\widehat{\phi}_1(u, k)|, |\widehat{\phi}_2(u, k)| \le K|k|^{-2d(u)-1}$ for $u \in [0, 1]$ and $|k| > 1$. If assumptions A2 and A3 hold, then*

$$
\lim_{T\to\infty} T \operatorname{cov}[J_T(\phi_1), J_T(\phi_2)] = 4\pi\int_0^1\int_{-\pi}^{\pi}\phi_1(u, \lambda)\phi_2(u, \lambda)f(u, \lambda)^2\, d\lambda\, du.
$$

PROOF. We can write

$$
\begin{aligned}
&T \operatorname{cov}[J_T(\phi_1), J_T(\phi_2)] \\
&\quad = \frac{T}{M^2}\sum_{j,k=1}^{M}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\phi_1(u_j, \lambda)\phi_2(u_k, \mu) \\
&\qquad\qquad \times \operatorname{cov}[I_N(u_j, \lambda), I_N(u_k, \mu)]\, d\lambda\, d\mu.
\end{aligned}
$$

But,

$$
\begin{aligned}
&\operatorname{cov}[I_N(u_j,\lambda), I_N(u_k,\mu)] \\
&\quad = \frac{1}{[2\pi H_{2,N}(0)]^2} \operatorname{cov}(|D_N(u_j,\lambda)|^2, D_N(u_k,\mu)|^2) \\
&\quad = \frac{1}{[2\pi H_{2,N}(0)]^2} \sum_{t,s,p,m=0}^{N-1} h\left(\frac{t}{N}\right)h\left(\frac{s}{N}\right)h\left(\frac{p}{N}\right)h\left(\frac{m}{N}\right) \\
&\qquad\qquad \times e^{i\lambda(s-t)+i\mu(m-p)} \\
&\qquad\qquad \times \operatorname{cov}\big(Y_{[u_jT]-N/2+s+1,T} Y_{[u_jT]-N/2+t+1,T}, \\
&\qquad\qquad\qquad Y_{[u_kT]-N/2+p+1,T} Y_{[u_kT]-N/2+m+1,T}\big).
\end{aligned}
$$

Now, an application of Theorem 2.3.2 of Brillinger (1981) yields

$$
\begin{aligned}
&\operatorname{cov}[I_N(u_j,\lambda), I_N(u_k,\mu)] \\
&\quad = \frac{1}{[2\pi H_{2,N}(0)]^2} \\
&\qquad \times \sum_{t,s,p,m=0}^{N-1} h\left(\frac{t}{N}\right)h\left(\frac{s}{N}\right)h\left(\frac{p}{N}\right)h\left(\frac{m}{N}\right) e^{i\lambda(s-t)+i\mu(m-p)} \\
&\qquad\qquad \times \big\{\operatorname{cov}\big(Y_{[u_jT]-N/2+t+1,T}, Y_{[u_kT]-N/2+m+1,T}\big) \\
&\qquad\qquad\quad \times \operatorname{cov}\big(Y_{[u_jT]-N/2+s+1,T}, Y_{[u_kT]-N/2+p+1,T}\big) \\
&\qquad\qquad\quad + \operatorname{cov}\big(Y_{[u_jT]-N/2+t+1,T}, Y_{[u_kT]-N/2+p+1,T}\big) \\
&\qquad\qquad\qquad \times \operatorname{cov}\big(Y_{[u_jT]-N/2+s+1,T}, Y_{[u_kT]-N/2+p+1,T}\big)\big\} \\
&\quad = \frac{1}{[2\pi H_{2,N}(0)]^2} \\
&\qquad \times \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} H_N\left(A^0_{t_j-N/2+1+\cdot,T}(x)h\left(\frac{\cdot}{N}\right), \lambda - x\right) \\
&\qquad\qquad \times H_N\left(A^0_{t_k-N/2+1+\cdot,T}(x)h\left(\frac{\cdot}{N}\right), x - \mu\right) \\
&\qquad\qquad \times H_N\left(A^0_{t_j-N/2+1+\cdot,T}(y)h\left(\frac{\cdot}{N}\right), -y - \lambda\right) \\
&\qquad\qquad \times H_N\left(A^0_{t_k-N/2+1+\cdot,T}(y)h\left(\frac{\cdot}{N}\right), y + \mu\right) \\
&\qquad\qquad \times e^{i\lambda(s-t)+i\mu(m-p)}\, dx\, dy
\end{aligned}
$$

$$+\frac{1}{[2\pi H_{2,N}(0)]^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} H_N\biggl(A^0_{t_j-N/2+1+\cdot,T}(x)h\biggl(\frac{\cdot}{N}\biggr),\lambda-x\biggr)$$
$$\times H_N\biggl(A^0_{t_k-N/2+1+\cdot,T}(x)h\biggl(\frac{\cdot}{N}\biggr),x+\mu\biggr)$$
$$\times H_N\biggl(A^0_{t_j-N/2+1+\cdot,T}(y)h\biggl(\frac{\cdot}{N}\biggr),-y-\lambda\biggr)$$
$$\times H_N\biggl(A^0_{t_k-N/2+1+\cdot,T}(y)h\biggl(\frac{\cdot}{N}\biggr),y-\mu\biggr)$$
$$\times e^{i\lambda(s-t)+i\mu(m-p)}\,dx\,dy.$$

Thus,

$$T\operatorname{cov}(J_T(\phi_1),J_T(\phi_2))=\frac{T}{[2\pi M H_{2,N}(0)]^2}\bigl[B_N^{(1)}+B_N^{(2)}\bigr], \tag{20}$$

where

$$B_N^{(1)}=\sum_{j,k=1}^{M}\int_{\Pi}\phi_1(u_j,\lambda)\phi_2(u_k,\mu)H_N\biggl(A^0_{t_j-N/2+1+\cdot,T}(x)h\biggl(\frac{\cdot}{N}\biggr),\lambda-x\biggr)$$
$$\times H_N\biggl(\overline{A^0_{t_k-N/2+1+\cdot,T}(x)}h\biggl(\frac{\cdot}{N}\biggr),x-\mu\biggr)H_N$$
$$\times\biggl(A^0_{t_j-N/2+1+\cdot,T}(y)h\biggl(\frac{\cdot}{N}\biggr),-y-\lambda\biggr)$$
$$\times H_N\biggl(\overline{A^0_{t_k-N/2+1+\cdot,T}(y)}h\biggl(\frac{\cdot}{N}\biggr),y+\mu\biggr)$$
$$\times e^{i(x+y)(t_j-t_k)}\,dx\,dy\,d\mu\,d\lambda,$$

with $\Pi=[-\pi,\pi]^4$, and

$$B_N^{(2)}=\sum_{j,k=1}^{M}\int_{\Pi}\phi_1(u_j,\lambda)\phi_2(u_k,\mu)H_N\biggl(A^0_{t_j-N/2+1+\cdot,T}(x)h\biggl(\frac{\cdot}{N}\biggr),\lambda-x\biggr)$$
$$\times H_N\biggl(\overline{A^0_{t_k-N/2+1+\cdot,T}(x)}h\biggl(\frac{\cdot}{N}\biggr),x+\mu\biggr)$$
$$\times H_N\biggl(A^0_{t_j-N/2+1+\cdot,T}(y)h\biggl(\frac{\cdot}{N}\biggr),-y-\lambda\biggr)$$
$$\times H_N\biggl(\overline{A^0_{t_k-N/2+1+\cdot,T}(y)}h\biggl(\frac{\cdot}{N}\biggr),y-\mu\biggr)$$
$$\times e^{i(x+y)(t_j-t_k)}\,dx\,dy\,d\mu\,d\lambda.$$

The term $B_N^{(1)}$ can be written as follows:

$$\begin{aligned}
B_N^{(1)} &= \sum_{j,k=1}^{M} \int_{\Pi} \phi_1(u_j,\lambda)\phi_2(u_k,\mu) \\
&\qquad \times A(u_j,x)A(u_k,-x)A(u_j,y)A(u_k,-y) \\
&\qquad \times H_N(\lambda-x)H_N(x-\mu)H_N(\mu+y)H_N(-y-\lambda) \\
&\qquad \times e^{i(x+y)(t_j-t_k)}\,dx\,dy\,d\lambda\,d\mu + R_N \\
&= \sum_{j,k=1}^{M} \int_{\Pi} \phi_1(u_j,x)A(u_j,x)A(u_k,-x)\phi_2(u_k,y)A(u_j,y) \\
&\qquad \times A(u_k,-y)H_N(\lambda-x)H_N(x-\mu)H_N(\mu+y) \\
&\qquad \times H_N(-y-\lambda)e^{i(x+y)(t_j-t_k)}\,dx\,dy\,d\lambda\,d\mu \\
&\quad + \Delta_N^{(1)} + \Delta_N^{(2)} + R_N,
\end{aligned} \tag{21}$$

with

$$\begin{aligned}
\Delta_N^{(1)} &= \sum_{j,k=1}^{M} \int_{\Pi} [\phi_1(u_j,\lambda)-\phi_1(u_j,x)]\phi_2(u_k,\mu)A(u_j,x)A(u_k,-x)A(u_j,y) \\
&\qquad \times A(u_k,-y)H_N(\lambda-x)H_N(x-\mu)H_N(\mu+y) \\
&\qquad \times H_N(-y-\lambda)e^{i(x+y)(t_j-t_k)}\,dx\,dy\,d\lambda\,d\mu, \\
\Delta_N^{(2)} &= \sum_{j,k=1}^{M} \int_{\Pi} \phi_1(u_j,x)[\phi_2(u_k,\mu)-\phi_2(u_k,y)]A(u_j,x)A(u_k,-x)A(u_j,y) \\
&\qquad \times A(u_k,-y)H_N(\lambda-x)H_N(x-\mu)H_N(\mu+y) \\
&\qquad \times H_N(-y-\lambda)e^{i(x+y)(t_j-t_k)}\,dx\,dy\,d\lambda\,d\mu,
\end{aligned}$$

and by Lemma 2 the remainder term R_N can be bounded as follows:

$$\begin{aligned}
|R_N| \le \frac{N}{T}\Bigg| \sum_{j,k=1}^{M} \int_{\Pi} &\phi_1(u_j,\lambda)\phi_2(u_k,\mu)A(u_j,x)A(u_k,-x) \\
&\times A(u_j,y)y^{-d(u_k)}L_N(y+\mu) \\
&\times H_N(\lambda-x)H_N(x-\mu) \\
&\times H_N(-y-\lambda)e^{i(x+y)(t_j-t_k)}\,dx\,dy\,d\lambda\,d\mu\Bigg|.
\end{aligned} \tag{22}$$

By integrating with respect to μ the term $\Delta_N^{(1)}$ can be written as

$$\Delta_N^{(1)} = \sum_{j,k=1}^{M}\sum_{t,s=0}^{N-1} h\left(\frac{t}{N}\right)h\left(\frac{s}{N}\right)\widehat{\phi}_2(u_k, t-s) \times \int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}[\phi_1(u_j,\lambda)-\phi_1(u_j,x)]A(u_j,x)A(u_k,-x) \times A(u_j,y)A(u_k,-y)H_N(\lambda-x)H_N(-y-\lambda) \times e^{i(x+y)(t_j-t_k)-ixt-iys}\,dx\,dy\,d\lambda,$$

and by integrating with respect to y we get

$$\begin{aligned}\Delta_N^{(1)} &= \sum_{j,k=1}^{M}\sum_{t,s,p=0}^{N-1} h\left(\frac{t}{N}\right)h\left(\frac{s}{N}\right)h\left(\frac{p}{N}\right) \\ &\quad\times \widehat{\phi}_2(u_k,t-s)\widehat{f}(u_j,u_k,t_j-t_k-s+p) \\ &\quad\times \int_{-\pi}^{\pi}\int_{-\pi}^{\pi}[\phi_1(u_j,\lambda)-\phi_1(u_j,x)] \\ &\qquad\times A(u_j,x)A(u_k,-x)H_N(\lambda-x) \\ &\qquad\times e^{ix(t_j-t_k-t)+i\lambda p}\,dx\,d\lambda \\ &= \sum_{j,k=1}^{M}\sum_{t,s,p=0}^{N-1} h\left(\frac{t}{N}\right)h\left(\frac{s}{N}\right)h\left(\frac{p}{N}\right)\widehat{\phi}_2(u_k,t-s) \\ &\quad\times \widehat{f}(u_j,u_k,t_j-t_k-s+p) \\ &\quad\times \varepsilon_N(u_j,u_k,p,t_j-t_k-t),\end{aligned}$$

where $\widehat{f}(u,v,k)$ and $\varepsilon_N(r)$ are given by $\widehat{f}(u,v,k)=\int_{-\pi}^{\pi}A(u,\lambda)A(v,-\lambda)e^{i\lambda k}\,d\lambda$, and

$$\begin{aligned}&\varepsilon_N(u_j,u_k,p,r) \\ &\quad= \sum_{m=0}^{N-1} h\left(\frac{m}{N}\right)\widehat{\phi}_1(u_j,p-m)\int_{-\pi}^{\pi}A(u_j,x)A(u_k,-x)e^{ix(r+m)}\,dx \\ &\qquad - 2\pi h\left(\frac{p}{m}\right)\int_{-\pi}^{\pi}\phi_1(u_j,x)A(u_j,x)A(u_k,x)e^{ix(r+p)}\,dx \\ &\quad= \int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\phi_1(u_j,\lambda)A(u_j,x)A(u_k,-x)e^{i(p\lambda+rx)}\sum_{m=0}^{N-1}h\left(\frac{m}{N}\right)e^{im(x-\lambda)}\,d\lambda\,dx \\ &\qquad - 2\pi\int_{-\pi}^{\pi}h\left(\frac{m}{N}\right)\phi_1(u_j,x)A(u_j,x)A(u_k,-x)e^{i(r+p)x}\,dx.\end{aligned}$$

But $h(\frac{m}{M}) = h(\frac{p}{M}) + h'(\xi_{p,m})\frac{m-p}{N}$ for some $\xi_{p,m} \in [0, 1]$. Thus,

$$
\begin{aligned}
&\varepsilon_N(u_j, u_k, p, r) \\
&\quad = h\left(\frac{p}{N}\right)\left\{\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \phi_1(u_j, \lambda)A(u_j, x)A(u_k, -x)e^{i(p\lambda+rx)} \right.\\
&\qquad\qquad \times \sum_{m=0}^{N-1} e^{im(x-\lambda)}\, d\lambda\, dx \\
&\qquad\qquad \left. - 2\pi \int_{-\pi}^{\pi} \phi_1(u_j, x)A(u_j, x)A(u_k, -x)e^{i(r+p)x}\, dx \right\} \\
&\qquad + \sum_{m=0}^{N-1} h'(\xi_{p,m})\frac{m-p}{N}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \phi_1(u_j, \lambda)A(u_j, x)A(u_k, -x) \\
&\qquad\qquad \times e^{im(x-\lambda)+i(p\lambda+rx)}\, d\lambda\, dx \\
&\quad = h\left(\frac{p}{N}\right)\varepsilon_N^{(1)}(u_j, u_k, p, r) + \varepsilon_N^{(2)}(u_j, u_k, p, r),
\end{aligned}
$$

where the term $\varepsilon_N^{(1)}(u_j, u_k, p, r)$ is given by

$$
\varepsilon_N^{(1)}(u_j, u_k, p, r) = \int_{-\pi}^{\pi} g(\omega)e^{ir\omega}\sum_{m=0}^{N-1} e^{im\omega}\, d\omega - 2\pi g(0),
$$

with $g(\omega) = \int_{-\pi}^{\pi} \phi_1(u_j, \lambda)A(u_j, \omega+\lambda)A(u_k, -\omega-\lambda)e^{i(p+r)\lambda}\, d\lambda$. Observe that by Lemma 1, for every u_j, u_k, p, r, $\varepsilon_N^{(1)}(u_j, u_k, p, r) \to 0$ as $N \to \infty$, consequently we can write

$$
\begin{aligned}
\varepsilon_N^{(1)}(u_j, u_k, p, r) &= \int_{-\pi}^{\pi} g(\omega)\sum_{m=N}^{\infty} e^{i(m+r)\omega}\, d\omega \\
&= \sum_{m=N}^{\infty} \widehat{\phi}_1(u_j, p-m)\widehat{f}(u_j, u_k, r+m).
\end{aligned}
$$

On the other hand, by assumption A1, $|\widehat{f}(u_j, u_k, r+m)| \le K|r+m|^{d(u_j)+d(u_k)-1}$. Thus, the term $\varepsilon_N^{(2)}(u_j, u_k, p, r)$ is bounded by

$$
\begin{aligned}
&\left|\varepsilon_N^{(2)}(u_j, u_k, p, r)\right| \\
&\quad \le \frac{K}{N}\sum_{m=0}^{N-1} |m-p|\widehat{\phi}_1(u_j, m-p)\widehat{f}(u_j, u_k, r+m) \\
&\quad \le \frac{K}{N}\sum_{m=0}^{N-1} |m-p|^{-2d(u_j)}|r+m|^{d(u_j)+d(u_k)-1}
\end{aligned}
$$

$$
\begin{aligned}
&\le K\left\{\sum_{m=0}^{N-1}\left|\frac{m}{N}-\frac{p}{N}\right|\left|\frac{r}{N}-\frac{m}{N}\right|^{d(u_j)+d(u_k)-1}\frac{1}{N}\right\}N^{d(u_k)-d(u_j)-1}\\
&\le K\int_0^1\left|x-\frac{p}{m}\right|^{-2d(u_j)}\left|\frac{r}{N}+x\right|^{d(u_j)+d(u_k)-1}dx\,N^{d(u_k)-d(u_j)-1}\\
&\le KN^{d(u_k)-d(u_j)-1}\left(\frac{r}{N}\right)^{d(u_j)+d(u_k)-1}\int_0^1\left|x-\frac{p}{m}\right|^{-2d(u_j)}dx\\
&\le KN^{-2d(u_j)}r^{d(u_j)+d(u_k)-1},
\end{aligned}
$$

where for notational simplicity we have dropped θ from $d_\theta(\cdot)$. Thus,

$$
\begin{aligned}
\varepsilon_N(u_j,u_k,p,r)=h\left(\frac{p}{N}\right)\sum_{m=N}^{\infty}\widehat{\phi}_1(u_j,p-m)\widehat{f}(u_j,u_k,r+m)&\\
+\mathcal{O}\big(N^{-2d(u_j)}r^{d(u_j)+d(u_k)-1}\big)&.
\end{aligned}
$$

Hence, $\Delta_N^{(1)}$ can be written as

$$
\begin{aligned}
\Delta_N^{(1)} = \sum_{j,k=1}^{M}\sum_{t,s,p=0}^{N-1} & h\left(\frac{t}{N}\right)h\left(\frac{s}{N}\right)h\left(\frac{p}{N}\right)\widehat{\phi}_2(u_k,t-s)\widehat{f}(u_j,u_k,t_j-t_k)\\
&\times\left\{h\left(\frac{p}{N}\right)\sum_{m=N+1}^{\infty}\widehat{\phi}_1(u_j,p-m)\widehat{f}(u_j,u_k,t_j-t_k+m)\right.\\
&\left.\qquad+O\big(N^{-2d(u_j)}|t_j-t_k-t|^{d(u_j)+d(u_k)-1}\big)\right\}\\
:= \Delta_N^{(1.1)}+\Delta_N^{(1.2)}&,
\end{aligned}
$$

say. Therefore, $|\Delta_N^{(1)}|\le|\Delta_N^{(1.1)}|+|\Delta_N^{(1.2)}|$. Observe that since $\phi_2(u,\lambda)\sim C|\lambda|^{2d(u)}$ as $\lambda\to 0$ and $d(u)>0$ for all $u\in[0,1]$, we conclude that $\phi_2(u,0)=\sum_{k=-\infty}^{\infty}\widehat{\phi}_2(u,k)=0$. Thus,

$$
\begin{aligned}
&\sum_{t,s=0}^{N-1}h\left(\frac{t}{N}\right)h\left(\frac{s}{N}\right)\widehat{\phi}_2(u,t-s)\\
&\qquad=\sum_{t=0}^{N-1}\sum_{k=1-N}^{N-1}h\left(\frac{t}{N}\right)h\left(\frac{t}{N}+\frac{k}{N}\right)\widehat{\phi}_2(u,k),
\end{aligned}
$$

where for simplicity we assume that $h(x)=0$ for x outside $[0,1]$. Now, by an application of Taylor's theorem we can write $h(\frac{t}{N}+\frac{k}{N})=h(\frac{t}{N})+h'(\xi(t,k))\frac{k}{N}$

for some $\xi(t,k) \in (\frac{t}{N} - \frac{|k|}{N}, \frac{t}{N} + \frac{|k|}{N})$. Hence,

$$\sum_{t=0}^{N-1} \sum_{k=1-N}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{t}{N} + \frac{k}{N}\right) \widehat{\phi}_2(u,k)$$
$$= \sum_{t=0}^{N-1} h\left(\frac{t}{N}\right)^2 \sum_{k=1-N}^{N-1} \widehat{\phi}_2(u,k)$$
$$+ \sum_{t=0}^{N-1} \sum_{k=1-N}^{N-1} h\left(\frac{t}{N}\right) h'(\xi(t,k)) \widehat{\phi}_2(u,k) \frac{k}{N}.$$

Note that $\sum_{k=1-N}^{N-1} \widehat{\phi}_2(u,k) = 2\sum_{k=N}^{\infty} \widehat{\phi}_2(u,k)$. Therefore, $|\sum_{k=1-N}^{N-1} \widehat{\phi}_2(u,k)| \le K \sum_{k=N}^{\infty} k^{-2d(u)-1} \le K N^{-2d(u)}$. Consequently, $|\sum_{t=0}^{N-1} h(\frac{t}{N})^2 \sum_{k=1-N}^{N-1} \widehat{\phi}_2(u, k)| \le K N^{1-2d(u)}$. On the other hand, $| \sum_{t=0}^{N-1} \sum_{k=1-N}^{N-1} h(\frac{t}{N}) h'(\xi(t,k)) \widehat{\phi}_2(u,k) \times \frac{k}{N}| \le K \sum_{k=1}^{N} k^{-2d(u)} \le K N^{1-2d(u)}$. Hence,

$$\left| \sum_{t=0}^{N-1} \sum_{k=1-N}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{t}{N} + \frac{k}{N}\right) \widehat{\phi}_2(u,k) \right| \le K N^{1-2d(u)}.$$

Thus, we conclude that

$$\begin{aligned}
\left|\Delta_N^{(1.1)}\right| &\le K \sum_{j,k=1, j\neq k}^{M} \sum_{p=0}^{N-1} N^{1-2d(u_k)} |S(j-k)+p|^{d(u_j)+d(u_k)-1} \\
&\qquad \times \sum_{m=N+1}^{\infty} |p-m|^{-2d(u_j)-1} |S(j-k)+m|^{d(u_j)+d(u_k)-1} \\
&\le K \sum_{j,k=1, j\neq k}^{M} \sum_{p=0}^{N-1} \left| \frac{S}{N}(j-k) + \frac{p}{N} \right|^{d(u_j)+d(u_k)-1} \\
&\qquad \times \frac{1}{N} \sum_{m=N+1}^{\infty} \left| \frac{p}{N} - \frac{m}{N} \right|^{-2d(u_j)-1} \\
&\qquad \times \left| \frac{S}{N}(j-k) + \frac{m}{N} \right|^{d(u_j)+d(u_k)-1} \frac{1}{N} \\
&\le K \sum_{j,k=1, j\neq k}^{M} \int_0^1 \int_1^{\infty} \left| \frac{S}{N}(j-k) + x \right|^{d(u_j)+d(u_k)-1} |x-y|^{-2d(u_j)-1} \\
&\qquad \times \left| \frac{S}{N}(j-k) + y \right|^{d(u_j)+d(u_k)-1} dy\, dx.
\end{aligned}$$

Let $\delta > 0$ and define $I_1(\delta) = \{j, k = 1, M : k < j \vee k - j > \frac{N}{S}(1+\delta)\}$ and $I_2(\delta) = \{j, k = 1, M : 0 < k - j \le \frac{N}{S}(1+\delta)\}$. Therefore, the sum above can be written as $\sum_{j,k=1, j\neq k}^{M} \cdot = \sum_{I_1(\delta)} \cdot + \sum_{I_2(\delta)} \cdot := |\Delta_N^{(1.1.1)}| + |\Delta_N^{(1.1.2)}|$, say. Observe that over $I_1(\delta)$ we have that $|\frac{S}{N}(j-k) + x|^{-\alpha} \le K|\frac{S}{N}(j-k)|^{-\alpha}$ for $\alpha > 0$. Hence,

$$\begin{aligned}
|\Delta_N^{(1.1.1)}| &\le K \sum_{I_1(\delta)} \left|\frac{S}{N}(j-k)\right|^{d(u_j)+d(u_k)-1} \\
&\qquad \times \int_0^1 \int_1^\infty |x-y|^{-2d(u_j)-1} \\
&\qquad\qquad \times \left|\frac{S}{N}(j-k) + y\right|^{d(u_j)+d(u_k)-1} dy\, dx \\
&\le K \sum_{j,k=1, j\neq k}^{M} \left|\frac{S}{N}(j-k)\right|^{d(u_j)+d(u_k)-1} \\
&\qquad \times \int_0^1 \int_1^\infty |x-y|^{-2d(u_j)-1} \\
&\qquad\qquad \times \left|\frac{S}{N}(j-k) + y\right|^{d(u_j)+d(u_k)-1} dy\, dx.
\end{aligned}$$

Since the integrands in the above expression are all positive, an application of Tonelli's theorem yields

$$\begin{aligned}
|\Delta_N^{(1.1.1)}| &\le K \sum_{j,k=1, j\neq k}^{M} \left|\frac{S}{N}(j-k)\right|^{d(u_j)+d(u_k)-1} \\
&\qquad \times \int_1^\infty \int_0^1 |x-y|^{-2d(u_j)-1} \\
&\qquad\qquad \times \left|\frac{S}{N}(j-k) + y\right|^{d(u_j)+d(u_k)-1} dx\, dy \\
&\le K \sum_{j,k=1, j\neq k}^{M} \left|\frac{S}{N}(j-k)\right|^{d(u_j)+d(u_k)-1} \\
&\qquad \times \int_1^\infty \left[(y-1)^{-2d(u_j)} - y^{-2d(u_j)}\right] \\
&\qquad\qquad \times \left|\frac{S}{N}(j-k) + y\right|^{d(u_j)+d(u_k)-1} dy.
\end{aligned}$$

Then, by Lemma 3 we conclude that

$$
\begin{aligned}
|\Delta_N^{(1.1.1)}| &\le K \sum_{j,k=1, j\neq k}^{M} \left|\frac{S}{N}(j-k)\right|^{2d(u_j)+2d(u_k)-2} \\
&\le K\Bigg[\sum_{\substack{j,k=1, j\neq k \\ d(u_j)+d(u_k)\le 1/2}}^{M} \left|\frac{S}{N}(j-k)\right|^{2d(u_j)+2d(u_k)-2} \\
&\qquad + \sum_{\substack{j,k=1, j\neq k \\ d(u_j)+d(u_k)> 1/2}}^{M} \left|\frac{S}{N}(j-k)\right|^{2d(u_j)+2d(u_k)-2}\Bigg].
\end{aligned}
$$

For the first summand above, we have the upper bound

$$
\sum_{j,k=1, j\neq k}^{M} |j-k|^{-1}\left(\frac{N}{S}\right)^2 \le K\left(\frac{N}{S}\right)^2 M \log M,
$$

while the second summand can be bounded as follows:

$$
\begin{aligned}
&\sum_{\substack{j,k=1, j\neq k \\ d(u_j)+d(u_k)> 1/2}}^{M} \left|\frac{SM}{N}\left(\frac{j}{M}-\frac{k}{M}\right)\right|^{2d(u_j)+2d(u_k)-2} \\
&\qquad \le K\left(\frac{T}{N}\right)^{-\varepsilon} \sum_{j,k=1}^{M}\left|\frac{j}{M}-\frac{k}{M}\right|^{\varepsilon-1} \\
&\qquad \le K\left(\frac{T}{N}\right)^{-\varepsilon} M^2 \int_0^1\int_0^1 |x-y|^{\varepsilon-1}\,dx\,dy \\
&\qquad \le K\left(\frac{T}{N}\right)^{-\varepsilon} M^2 \le K M^2.
\end{aligned}
$$

Thus,

$$
|\Delta_N^{(1.1.1)}| \le K\left(\frac{N}{S}\right)^2 M\log M + M^2. \tag{23}
$$

On the other hand, if $z = \frac{S}{N}(k-j)$ then $0 < z \le 1+\delta$ for $j,k \in I_2(\delta)$. Thus, an application of Lemma 9 yields for $2 > \delta > 0$

$$
|\Delta_N^{(1.1.2)}| \le K \sum_{I_2(\delta)} \left|1 - \frac{S}{N}(k-j)\right|^{2d-1},
$$

where $d := \inf_{0 \le u \le 1} d(u) > 0$. Hence, by defining $p = k - j$ and $P = N/S$ we can write

$$|\Delta_N^{(1.1.2)}| \le KM \sum_{p=1}^{P(1+\delta)} \left|1 - \frac{p}{P}\right|^{2d-1}$$
$$\le KM \frac{N}{S} \int_0^{1+\delta} |1-x|^{2d-1}\, dx \le KM \frac{N}{S}.$$

Note that from assumption A3, $N/S \to \infty$. Thus, by combining the above bound and (23) we conclude that

$$|\Delta_N^{(1.1)}| \le K \left(\frac{N}{S}\right)^2 M \log M + M^2. \tag{24}$$

A similar bound can be found for $|\Delta_N^{(1.2)}|$ and consequently for $|\Delta_N^{(1)}|$. Furthermore, an analogous argument yields a similar bound for the term $|\Delta_N^{(2)}|$ appearing in (21). Now, we focus on obtaining an upper bound for the remaining term R_N from (22). By integrating that expression with respect to λ we get

$$|R_N| \le \frac{N}{T} \Bigg| \sum_{j,k=1}^{M} \sum_{t,s=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) \widehat{\phi}_1(u_j, s-t)$$
$$\times \int_\Pi \phi_2(u_k, \mu) A(u_j, x) A(u_k, -x)$$
$$\times A(u_j, y) y^{-d(u_k)} L_N(y+\mu) H(x-\mu)$$
$$\times e^{ix(t_j - t_k + t) + iy(t_j - t_k + s)}\, dx\, dy\, d\mu \Bigg|,$$

where the function $L_N(\cdot)$ is defined as

$$L_N(x) = \begin{cases} N, & |x| \le 1/N, \\ 1/|x|, & 1/N < |x| \le \pi. \end{cases}$$

Hence,

$$|R_N| \le \frac{N}{T} \Bigg| \sum_{j,k=1}^{M} \sum_{t,s,p=0}^{N-1} h\left(\frac{t}{N}\right) h\left(\frac{s}{N}\right) h\left(\frac{p}{N}\right)$$
$$\times \widehat{\phi}_1(u_j, s-t) \widehat{f}(u_j, u_k, t_j - t_k + t - p)$$
$$\times \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_2(u_k, \mu) A(u_j, y) y^{-d(u_k)}$$
$$\times L_N(\mu + y) e^{iy(t_j - t_k + s) + ip\mu}\, dy\, d\mu \Bigg|$$

$$\leq K\frac{N}{T}\sum_{j,k=1}^{M}\sum_{t,s,p=0}^{N-1}|\widehat{\phi}_1(u_j,s-t)||\widehat{f}(u_j,u_k,t_j-t_k+t-p)|$$
$$\times\left|\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\phi_2(u_k,\mu)A(u_j,y)y^{-d(u_k)}\right.$$
$$\left.\times L_N(\mu+y)e^{iy(t_j-t_k+s)+ip\mu}\,dy\,d\mu\right|$$
$$\leq K\frac{N}{T}\sum_{j,k=1}^{M}\sum_{t,s,p=0}^{N-1}|\widehat{\phi}_1(u_j,s-t)||\widehat{f}(u_j,u_k,t_j-t_k+t-p)|$$
$$\times\left|\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}L_N(\mu+y)y^{-d(u_j)-d(u_k)}\,dy\,d\mu\right|$$
$$\leq K\frac{N\log N}{T}\sum_{j,k=1}^{M}\sum_{t,s,p=0,t\neq s}^{N-1}|s-t|^{-2d(u_j)-1}$$
$$\times|S(j-k)+t-p|^{d(u_j)+d(u_k)-1}$$
$$\leq K\frac{N^2\log N}{T}\sum_{j,k=1}^{M}\sum_{t,s,p=0,t\neq s}^{N-1}|s-t|^{-2d(u_j)-1}S^{d(u_j)+d(u_k)-1}$$
$$\times|j-k|^{d(u_j)+d(u_k)-1}$$
$$\leq K\frac{N^3\log N}{T}M^2\sum_{j,k=1}^{M}(SM)^{d(u_j)+d(u_k)-1}\left|\frac{j}{M}-\frac{k}{M}\right|^{d(u_j)+d(u_k)-1}\frac{1}{M^2}.$$

Since by assumption A3, $T/N^2 \to 0$, we conclude that

$$|R_N|\leq K\frac{N^3M^2}{T^{2-d}}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}|x-y|^{2d-1}\,dx\,dy\leq KN^3M^2T^{d-2}. \tag{25}$$

Thus, from (24) and (25), we conclude

$$\frac{TB_N^{(1)}}{[2\pi MH_{2,N}(0)]^2}$$
$$=\frac{T}{[2\pi MH_{2,N}(0)]^2}$$
$$\times\sum_{j,k=1}^{M}\int_{\Pi}\phi_1(u_j,x)A(u_j,x)A(u_k,-x)\phi_2(u_k,y)$$
$$\times A(u_j,y)A(u_k,-y)H_N(\lambda-x)H_N(x-\mu)H_N(\mu+y)$$
$$\times H_N(-y-\lambda)e^{i(x+y)(t_j-t_k)}\,dx\,dy\,d\lambda\,d\mu+C_N,$$

where

$$C_N = \mathcal{O}\left(\frac{\log M}{S} + \frac{T}{N^2} + NT^{d-1}\right). \tag{26}$$

Therefore, by assumption A3 we conclude that $C_N = o(1)$. By following successive decompositions as in (21), we replace $\phi_2(u_k, y)$ by $\phi_2(u_k, x)$, $A(u_k, -y)$ by $A(u_k, -x)$ and $A(u_j, y)$ by $A(u_j, x)$, respectively. Thus,

$$\begin{aligned}
\frac{TB_N^{(1)}}{[2\pi MH_{2,N}(0)]^2} = {} & \frac{T}{[2\pi MH_{2,N}(0)]^2} \\
& \times \sum_{j,k=1}^{M} \int_{\Pi} \phi_1(u_j, x)A(u_j, x)A(u_k, -x)\phi_2(u_k, x)A(u_j, x) \\
& \qquad \times A(u_k, -x)H_N(\lambda - x)H_N(x - \mu)H_N(\mu + y) \\
& \qquad \times H_N(-y - \lambda)e^{i(x+y)(t_j - t_k)}\,dx\,dy\,d\lambda\,d\mu + o(1).
\end{aligned}$$

By integrating with respect to μ and λ, we get

$$\begin{aligned}
&\frac{TB_N^{(1)}}{[2\pi MH_{2,N}(0)]^2} \\
&\quad = \frac{T}{[MH_{2,N}(0)]^2} \sum_{j,k=1}^{M} \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \phi_1(u_j, x)A(u_j, x) \\
&\qquad\qquad \times A(u_k, -x)\phi_2(u_k, x)A(u_j, x)A(u_k, -x) \\
&\qquad\qquad \times |H_{2,N}(x + y)|^2 e^{i(x+y)(t_j - t_k)}\,dx\,dy + o(1) \\
&\quad = \frac{T}{[MH_{2,N}(0)]^2} \sum_{j,k=1}^{M} \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} [\phi_1(u_j, x)f(u_j, x)][\phi_2(u_k, x)f(u_k, x)] \\
&\qquad\qquad \times |H_{2N}(x + y)|^2 e^{i(x+y)[s(j-k)]}\,dx\,dy + o(1) \\
&\quad = \frac{T}{[MH_{2,N}(0)]^2} \sum_{j,k=1}^{M} \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} [\phi_1(u_j, x)f(u_j, x)][\phi_2(u_k, x)f(u_k, x)] \\
&\qquad\qquad \times |H_{2N}(z)|^2 e^{iz[s(j-k)]}\,dx\,dz + o(1) \\
&\quad = \frac{T}{[MH_{2,N}(0)]^2} \\
&\qquad \times \sum_{j,k=1}^{M} \int_{-\pi}^{\pi} [\phi_1(u_j, x)f(u_j, x)][\phi_2(u_k, x)f(u_k, x)] \\
&\qquad\qquad \times \sum_{s,t=0}^{N-1} h^2\left(\frac{t}{N}\right)h^2\left(\frac{s}{N}\right)\int_{-\pi}^{\pi} e^{iz[s(j-k)+t-s]}\,dx\,dz + o(1)
\end{aligned}$$

$$= \frac{2\pi T}{[MH_{2,N}(0)]^2} \sum_{s,t=0}^{N-1} \sum_{\substack{j,k=1\\ S(j-k)=s-t}}^{M} h^2\left(\frac{t}{N}\right) h^2\left(\frac{s}{N}\right) \times \int_{-\pi}^{\pi} [\phi_1(u_j,x) f(u_j,x)] \times [\phi_2(u_k,x) f(u_k,x)]\,dx + o(1).$$

By assumption A3, for $S < N$ we can write

$$\frac{T}{[2\pi MH_{2,N}(0)]^2} B_N^{(1)} = \frac{2\pi T}{[MH_{2,N}(0)]^2} \times \sum_{t=0}^{N-1} \sum_{p=-t/S}^{(N-t)/S} h^2\left(\frac{t}{N}\right) h^2\left(\frac{t}{N} + \frac{pS}{N}\right) \times \sum_{j=1}^{M-|p|} \int_{-\pi}^{\pi} [\phi_1(u_j,x) f(u_j,x)] \times [\phi_1(u_{j+p},x) f(u_{j+p},x)]\,dx + o(1).$$

Observe that by the assumptions of this proposition the products $\phi_1(u,x) f(u,x)$ and $\phi_2(u,x) f(u,x)$ are differentiable with respect to u. Furthermore, note that by assumption A3, $\lim_{T,S\to\infty} \frac{S|p|}{T} = 0$ for any $|p| \le \frac{N}{S}$. Consequently,

$$\frac{S}{T} \sum_{j=1}^{M-|p|} \int_{-\pi}^{\pi} [\phi_1(u_j,x) f(u_j,x)][\phi_2(u_{j+p},x) f(u_{j+p},x)]\,dx \to \int_0^1 \int_{-\pi}^{\pi} \phi_1(u,x)\phi_2(u,x) f(u,x)^2\,dx\,du,$$

for any $|p| < \frac{N}{S}$ as $M, N, S, T \to \infty$. On the other hand,

$$\frac{2\pi T^2 N^2}{S^2[MH_{2,N}(0)]^2} \sum_{t=0}^{N-1} \sum_{p=-t/S}^{(N-t)/S} h^2\left(\frac{t}{N}\right) h^2\left(\frac{t}{N} + \frac{pS}{N}\right) \frac{S}{N^2} \to 2\pi \int_0^1 \int_{-x}^{1-x} h^2(x) h^2(x+y)\,dx\,dy \left(\int_0^1 h^2(x)\,dx\right)^{-2} = 2\pi,$$

as $M, N, S, T \to \infty$. Therefore, in this case

$$\frac{T}{[2\pi MH_{2,N}(0)]^2} B_N^{(1)} \to 2\pi \int_0^1 \int_{-\pi}^{\pi} \phi_1(u,x)\phi_2(u,x) f(u,x)^2\,dx\,du,$$

as $M, N, S, T \to \infty$. Similarly, we have that

$$\frac{T}{[2\pi M H_{2,N}(0)]^2} B_N^{(2)} \to 2\pi \int_0^1 \int_{-\pi}^{\pi} \phi_1(u,x)\phi_2(u,x) f(u,x)^2 \, dx \, du,$$

as $M, N, S, T \to \infty$. Therefore, by virtue of (20) this proposition is proved. □

PROPOSITION 3. *Let* $\operatorname{cum}_p(\cdot)$ *be the pth order cumulant with* $p \geq 3$. *Then,* $T^{p/2} \operatorname{cum}_p(J_T(\phi)) \to 0$, *as* $T \to \infty$.

PROOF. Observe that $J_T(\phi)$ can be written as

$$J_T(\phi) = \frac{1}{2\pi M H_{2,N}(0)} Y' Q(\phi) Y,$$

where the block-diagonal matrix $Q(\phi)$ is defined in (16) and $Y \in \mathbb{R}^{NM}$ is a Gaussian random vector defined by $Y = (Y(u_1)', \ldots, Y(u_M)')'$, $Y(u) = (Y_1(u), \ldots, Y_N(u))$, $Y_t(u) = h(\frac{t}{N}) Y_{[uT]-N/2+t+1,T}$ with $Y_{[uT]-N/2+t+1,T}$ satisfying (1). For simplicity, denote the matrix $Q(\phi)$ as Q. Since Y is Gaussian,

$$\operatorname{cum}_p[J_T(\phi)] = \frac{2^{p-1}(p-1)!}{(2\pi M H_{2,N}(0))^p} \operatorname{tr}(RQ)^p,$$

where $R = \operatorname{Var}(Y)$. Let $|A| = [\operatorname{tr}(AA')]^{1/2}$ be the Euclidean norm of matrix A and let $\|A\| = \sup_{\|x\|=1} (Ax)'Ax$ be the spectral norm of A. Now, since $|\operatorname{tr}(QB)| \leq |Q||B|$ and $|QB| \leq \|Q\||B|$ we get $|\operatorname{tr}(RQ)^p| \leq \|RQ\|^{p-2}|RQ|^2$.

On the other hand, for fixed λ, decompose the function $\phi(\cdot,\lambda)$ as $\phi(\cdot,\lambda) = \phi_+(\cdot,\lambda) - \phi_-(\cdot,\lambda)$ where $\phi_+(\cdot,\lambda), \phi_-(\cdot,\lambda) \geq 0$. Thus, we can write $Q = Q(\phi) = Q(\phi_+ - \phi_-) = Q(\phi_+) - Q(\phi_-) := Q_+ - Q_-$, say. Now, by Lemma 6 we conclude that

$$\|RQ\| = \|RQ_+ - RQ_-\| \leq \|RQ_+\| + \|RQ_-\| \leq K(MN^{1-2d}T^{2d-1}),$$

and by Proposition 2 we have that $|RQ|^2 \leq K\frac{M^2N^2}{T}$. Thus,

$$|\operatorname{tr}(RQ)^p| \leq K(MN^{1-2d}T^{2d-1})\frac{M^2N^2}{T}.$$

Consequently,

$$|T^{p/2}\operatorname{cum}_p[J_T(\phi)]| \leq KM^{1-p/2}\left(\frac{N}{T}\right)^{(1-2d)(p-2)}\left(\frac{\sqrt{T}}{N}\right)^{p-2}.$$

Since $p \geq 3$ and by assumption A2, $N/T \to 0$ and $\sqrt{T}/N \to 0$ as $T, N \to \infty$, the required result is obtained. □

3.2. *Proof of theorems.*

PROOF OF THEOREM 2.1. To prove the consistency of the Whittle estimator, it suffices to show that

$$\sup_{\theta}|\mathcal{L}_T(\theta) - \mathcal{L}(\theta)| \to 0,$$

in probability, as $T \to \infty$, where $\mathcal{L}(\theta) := \frac{1}{4\pi}\int_0^1\int_{-\pi}^{\pi}[\log f_\theta(u,\lambda) + \frac{f_{\theta_0}(u,\lambda)}{f_\theta(u,\lambda)}]\,d\lambda\,du$. Define $g_\theta(u,\lambda) = f_\theta(u,\lambda)^{-1}$. By assumption A1, $g_\theta(u,\lambda)$ is continuous in θ, λ and u. Thus, g_θ can be approximated by the Cesaro sum of its Fourier series

$$g_\theta^{(L)}(u,\lambda) = \frac{1}{4\pi^2}\sum_{\ell=-L}^{L}\sum_{m=-L}^{L}\left(1-\frac{|\ell|}{L}\right)\left(1-\frac{|m|}{L}\right) \times \widehat{g}_\theta(\ell,m)\exp(-i2\pi u_j\ell - i\lambda m),$$

such that $\sup_\theta|g_\theta(u,\lambda) - g_\theta^{(L)}(u,\lambda)| < \varepsilon$; see, for example, Theorem 1.5(ii) of Körner (1988). Following Theorem 3.2 of Dahlhaus (1997), we can write

$$\begin{aligned}
&\sup_{\theta}|\mathcal{L}_T(\theta) - \mathcal{L}(\theta)| \\
&\quad\le \mathcal{O}\left(\frac{1}{M}\right) + \frac{\varepsilon}{4\pi}\frac{1}{M}\sum_{j=1}^{M}\int_{-\pi}^{\pi}[I_N(u_j,\lambda) + f(u_j,\lambda)]\,d\lambda \\
&\qquad + \frac{1}{16\pi^3}\sum_{\ell=-L}^{L}\sum_{m=-L}^{L}\left(1-\frac{|\ell|}{L}\right)\left(1-\frac{|m|}{L}\right)\sup_\theta|\widehat{g}_\theta(\ell,m)| \\
&\qquad\qquad \times \left|\frac{1}{M}\sum_{j=1}^{M}\int_{-\pi}^{\pi}\exp(-i2\pi u_j\ell - i\lambda m) \times \{I_N(u_j,\lambda) - f(u_j,\lambda)\}\,d\lambda\right|,
\end{aligned}$$

where

$$\widehat{g}_\theta(\ell,m) = \int_0^1\int_{-\pi}^{\pi} g_\theta(u,\lambda)\exp(i2\pi u\ell + i\lambda m)\,du\,d\lambda.$$

Consequently, $|\widehat{g}_\theta(\ell,m)| \le 2\pi\sup_{(\theta,u,\lambda)}|g_\theta(u,\lambda)|$. However, by assumption A1, $|g_\theta(u,\lambda)|$ is continuous in θ, u and λ. Thus, since the parameter space is compact we have that $|\widehat{g}_\theta(\ell,m)| \le K$, for some positive constant K. Now, by defining for fixed $\ell, m = 1,\dots,L$, $\phi(u,\lambda) = \cos(2\pi u\ell)\cos(\lambda m)$ or $\phi(u,\lambda) = \sin(2\pi u\ell)\cos(\lambda m)$ in Proposition 1 and $\phi_1(u,\lambda) = \phi_2(u,\lambda) = \cos(2\pi u\ell)\cos(\lambda \times$

$m)$ or $\phi_1(u,\lambda)=\phi_2(u,\lambda)=\sin(2\pi u\ell)\cos(\lambda m)$ in Proposition 2, we deduce that

$$\begin{aligned}
&\frac{1}{16\pi^3}\sum_{\ell=-L}^{L}\sum_{m=-L}^{L}\left(1-\frac{|\ell|}{L}\right)\left(1-\frac{|m|}{L}\right)\sup_\theta|\widehat{g}_\theta(\ell,m)| \\
&\qquad\times\left|\frac{1}{M}\sum_{j=1}^{M}\int_{-\pi}^{\pi}\exp(-i2\pi u_j\ell-i\lambda m)\right. \\
&\qquad\qquad\left.\times[I_N(u_j,\lambda)-f(u_j,\lambda)]\,d\lambda\right| \\
&\quad\le\frac{1}{16\pi^3}\sum_{\ell=-L}^{L}\sum_{m=-L}^{L}\left(1-\frac{|\ell|}{L}\right)\left(1-\frac{|m|}{L}\right)\sup_\theta|\widehat{g}_\theta(\ell,m)| \\
&\qquad\times\left\{\left|\frac{1}{M}\sum_{j=1}^{M}\int_{-\pi}^{\pi}\cos(2\pi u_j\ell)\cos(\lambda m)\right.\right. \\
&\qquad\qquad\left.\times[I_N(u_j,\lambda)-f(u_j,\lambda)]\,d\lambda\right| \\
&\qquad+\left|\frac{1}{M}\sum_{j=1}^{M}\int_{-\pi}^{\pi}\sin(2\pi u_j\ell)\cos(\lambda m)\right. \\
&\qquad\qquad\left.\left.\times[I_N(u_j,\lambda)-f(u_j,\lambda)]\,d\lambda\right|\right\}\to 0
\end{aligned}\tag{27}$$

and

$$\frac{1}{M}\sum_{j=1}^{M}\int_{-\pi}^{\pi}\{I_N(u_j,\lambda)+f(u_j,\lambda)\}\,d\lambda\to 2\int_0^1\int_{-\pi}^{\pi}f(u,\lambda)\,d\lambda\,du,\tag{28}$$

in probability, as $M\to\infty$. Now, from the limits (27) and (28), this theorem follows. □

PROOF OF THEOREM 2.2. Let $\widehat{\theta}_T$ be the parameter value that minimizes the Whittle log-likelihood function $\mathcal{L}_T(\theta)$ given by (7) and let θ_0 be the true value of the parameter. By the mean value theorem, there exists a vector $\bar{\theta}_T$ satisfying $\|\bar{\theta}_T-\theta_0\|\le\|\widehat{\theta}_T-\theta_0\|$, such that

$$\nabla\mathcal{L}_T(\widehat{\theta}_T)-\nabla\mathcal{L}_T(\theta_0)=[\nabla^2\mathcal{L}_T(\bar{\theta}_T)](\widehat{\theta}_T-\theta_0).\tag{29}$$

Therefore, it suffices to show that (a) $\nabla^2\mathcal{L}_T(\theta_0)\to\Gamma(\theta_0)$, as $T\to\infty$; (b) $\nabla^2\mathcal{L}_T(\bar{\theta}_T)-\nabla^2\mathcal{L}_T(\theta_0)\to 0$ in probability, as $T\to\infty$; and (c) $\sqrt{T}\nabla\mathcal{L}_T(\theta_0)\to$

$N[0, \Gamma(\theta_0)]$, in distribution, as $T \to \infty$. To this end, observe that

$$\begin{aligned}
\nabla^2 \mathcal{L}_T(\theta) &= \frac{1}{4\pi}\frac{1}{M}\sum_{j=1}^{M}\int_{-\pi}^{\pi}[I_N(u_j,\lambda) - f_\theta(u_j,\lambda)]\nabla^2 f_\theta(u_j,\lambda)^{-1} \\
&\qquad - \nabla f_\theta(u_j,\lambda)[\nabla f_\theta(u_j,\lambda)^{-1}]'\,d\lambda \\
&= \frac{1}{4\pi}\frac{1}{M}\left\{\sum_{j=1}^{M}\int_{-\pi}^{\pi}\phi(u_j,\lambda)[I_N(u_j,\lambda) - f_\theta(u_j,\lambda)]\right. \\
&\qquad \left. + \sum_{j=1}^{M}\int_{-\pi}^{\pi}\nabla \log f_\theta(u_j,\lambda)[\nabla \log f_\theta(u_j,\lambda)]'\,d\lambda\right\} \\
&= \frac{1}{4\pi}[J_T(\phi) - J(\phi)] + \Gamma(\theta) + \mathcal{O}\left(\frac{1}{M}\right),
\end{aligned}$$

where $\phi(u,\lambda) = \nabla^2 f_\theta(u,\lambda)^{-1}$. Hence, an application of Proposition 1 and Proposition 2 yields parts (a) and (b). On the other hand, part (c) can be proved by means of the cumulant method. That is, by showing that all the cumulants of $\sqrt{T}\nabla\mathcal{L}_T(\theta_0)$ converge to zero, excepting the second order cumulant. To this end, note that

$$\begin{aligned}
\nabla \mathcal{L}_T(\theta_0) &= \frac{1}{4\pi}\frac{1}{M}\sum_{j=1}^{M}\int_{-\pi}^{\pi}[I_N(u_j,\lambda) - f_{\theta_0}(u_j,\lambda)]\nabla f_{\theta_0}(u_j,\lambda)^{-1}\,d\lambda \\
&= \frac{1}{4\pi}J_T(\phi) - \frac{1}{4\pi}\sum_{j=1}^{M}\int_{-\pi}^{\pi} f_{\theta_0}(u_j,\lambda)\nabla f_{\theta_0}(u_j,\lambda)^{-1}\,d\lambda \qquad (30)\\
&= \frac{1}{4\pi}[J_T(\phi) - J(\phi)] + \mathcal{O}\left(\frac{1}{M}\right),
\end{aligned}$$

where $\phi(u,\lambda) = \nabla f_{\theta_0}(u,\lambda)^{-1}$. Hence, by Proposition 1 and assumption A3, the first-order cumulant of $\sqrt{T}\nabla\mathcal{L}_T(\theta_0)$ satisfies

$$\begin{aligned}
\sqrt{T}E[\nabla\mathcal{L}_T(\theta_0)] &= \mathcal{O}\left(\frac{\sqrt{T}\log^2 N}{N}\right) + \mathcal{O}\left(\frac{\sqrt{T}}{M}\right) \qquad (31)\\
&\to 0,
\end{aligned}$$

as $T \to \infty$. Furthermore, by (30) we have that the second-order cumulant of $\sqrt{T}\nabla\mathcal{L}_T(\theta_0)$ can be written as

$$T\,\mathrm{cov}[\nabla\mathcal{L}_T(\theta_0), \nabla\mathcal{L}_T(\theta_0)] = \frac{1}{16\pi^2}T\,\mathrm{cov}[J_T(\phi), J_T(\phi)].$$

Therefore, by Proposition 2 we have that

$$\lim_{T\to\infty} T\operatorname{cov}[\nabla\mathcal{L}_T(\theta_0), \nabla\mathcal{L}_T(\theta_0)]$$

$$= \frac{1}{4\pi}\int_0^1\int_{-\pi}^{\pi} \nabla f_{\theta_0}(u,\lambda)^{-1}[\nabla f_{\theta_0}(u,\lambda)^{-1}]' f_{\theta_0}(u,\lambda)^2\,d\lambda\,du$$

$$= \frac{1}{4\pi}\int_0^1\int_{-\pi}^{\pi} \nabla\log f_{\theta_0}(u,\lambda)[\nabla\log f_{\theta_0}(u,\lambda)]'\,d\lambda\,du = \Gamma(\theta_0).$$

Finally, for $p > 2$, Proposition 3 gives $T^{p/2}\operatorname{cum}_p[\nabla\mathcal{L}_T(\theta_0)] \to 0$, as $T \to \infty$, proving part (c). □

PROOF OF THEOREM 2.3. By observing that the Fisher information matrix evaluated at the true parameter, $\Gamma_T(\theta_0)$, is given by

$$\Gamma_T(\theta_0) = T\operatorname{cov}[\nabla\mathcal{L}_T(\theta_0), \nabla\mathcal{L}_T(\theta_0)],$$

the result is an immediate consequence of Proposition 2. □

PROOF OF THEOREM 2.4. Let $V^{(T)} = [V_{ij}^{(T)}]_{i,j=1,\dots,p} = \operatorname{Var}(\widehat{\beta})$, then

$$\int_0^1 \operatorname{Var}[\widehat{d}(u)]\,du = \int_0^1 \sum_{i=1}^p\sum_{j=1}^p g_i(u)V_{ij}^{(T)}g_j(u)\,du$$

$$= \sum_{i=1}^p\sum_{j=1}^p V_{ij}^{(T)}\int_0^1 g_i(u)g_j(u)\,du$$

$$= \sum_{i=1}^p\sum_{j=1}^p V_{ij}^{(T)}b_{ij},$$

where $b_{ij} = \int_0^1 g_i(u)g_j(u)\,du = b_{ji}$. Therefore, by Theorem 2.2

$$\lim_{T\to\infty} T\int_0^1 \operatorname{Var}[\widehat{d}(u)]\,du = \sum_{i=1}^p\sum_{j=1}^p \lim_{T\to\infty}[TV_{ij}^{(T)}]b_{ij} = \sum_{i=1}^p\sum_{j=1}^p a_{ij}b_{ij},$$

where $A = (a_{ij})_{i,j=1,\dots,p} = \Gamma^{-1}$ and

$$\Gamma_{ij} = \frac{1}{4\pi}\int_0^1\int_{-\pi}^{\pi} \frac{\partial}{\partial\beta_i}\log f(u,\lambda)\,\frac{\partial}{\partial\beta_j}\log f(u,\lambda)\,d\lambda\,du.$$

But, $\log f(u,\lambda) = \log(\sigma^2) - \log(2\pi) - d_\beta(u)\log|1-e^{i\lambda}|^2$. Thus,

$$\frac{\partial}{\partial\beta_i}\log f(u,\lambda) = -g_i(u)\log|1-e^{i\lambda}|^2.$$

Hence, $\Gamma_{ij} = \int_0^1 g_i(u) g_j(u)\,du \times \frac{1}{4\pi}\int_{-\pi}^{\pi} (\log|1-e^{i\lambda}|^2)^2\,d\lambda = \frac{\pi^2}{6} b_{ij}$. Therefore, $\Gamma = \frac{\pi^2}{6} B$ and $A = \frac{6}{\pi^2} B^{-1}$. Consequently, since A and B are symmetric matrices $\lim_{T\to\infty} T \int_0^1 \operatorname{Var}[\widehat{d}(u)]\,du = \operatorname{tr}(AB) = \frac{6}{\pi^2}\operatorname{tr}(I_p) = \frac{6p}{\pi^2}$. □

4. Simulations. In order to gain some insight into the finite sample performance of the Whittle estimator discussed in Section 2, we report next a number of Monte Carlo experiments for the LSARFIMA model

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

for $t = 1, \dots, T$ with $d(u) = \alpha_0 + \alpha_1 u$, $\sigma(u) = \beta_0 + \beta_1 u$ and Gaussian white noise $\{\varepsilon_t\}$ with unit variance. The samples of this LSARFIMA process are generated by means of the innovation algorithm; see, for example, Brockwell and Davis (1991), page 172. In this implementation, the covariances of the process $\{Y_{t,T}\}$ is given by

$$E[Y_{s,T}Y_{t,T}] = \sigma\left(\frac{s}{T}\right)\sigma\left(\frac{t}{T}\right)\frac{\Gamma[1-d(s/T)-d(t/T)]\Gamma[s-t+d(s/T)]}{\Gamma[1-d(s/T)]\Gamma[d(s/T)]\Gamma[s-t+1-d(t/T)]} \times \left[1+\vartheta^2 - \vartheta\frac{s-t-d(t/T)}{s-t-1+d(s/T)} - \vartheta\frac{s-t+d(s/T)}{s-t+1-d(t/T)}\right],$$

for $s, t = 1, \dots, T$, $s \geq t$. Let $\theta = (\alpha_0, \alpha_1, \beta_0, \beta_1, \vartheta)'$ be the parameter vector. The Whittle estimates in these Monte Carlo simulations have been computed by using the cosine bell data taper (9). Figure 1 displays the contour curves for the empirical mean squared error (MSE) for the Whittle estimator $\widehat{\theta}$ defined in this case as the average of $\|\widehat{\theta} - \theta\|^2$ over 100 replications of $\widehat{\theta}$, where θ is the true value of the parameter. These contour curves correspond to $\theta = (0.20, 0.25, 0.5, 0.3, 0.5)$, for sample sizes $T = 512$ and $T = 1024$, respectively. In these graphs, the darkest regions represent the minimal empirical MSE while clear regions indicate greater MSE values. Note that for the case $T = 512$, shown in the left panel, the minimal empirical MSE region is located around $N \approx 105$ and $S \approx 35$. For the sample size $T = 1024$, displayed on the right panel, the minimal empirical MSE is reached close to $N \approx 200$ y $S \approx 45$. As noted in these graphs, there is a degree of flexibility for selecting N and S as long they belong to the areas with minimal empirical MSE. Contour curves for other parameters θ such as those presented in Tables 1 and 2 are similar to Figure 1 and produce similar empirical optimal regions for N and S. Tables 1 and 2 report the results from the Monte Carlo simulations for several parameter values, based on 1000 replications. These tables show the average of the estimates as well as their theoretical and empirical standard deviations (SD). The theoretical SD are based on Theorem 2.2 with matrix Γ_θ given by

$$\Gamma_\theta = \begin{pmatrix} \Gamma_\alpha & 0 & \gamma_{\alpha\vartheta} \\ 0 & \Gamma_\beta & 0 \\ \gamma'_{\alpha\vartheta} & 0 & \gamma_\vartheta \end{pmatrix},$$

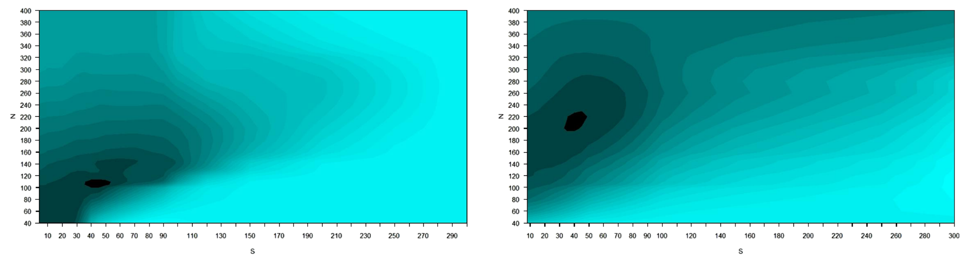

FIG. 1. *Contour curves of the empirical MSE of Whittle estimator. Left*: *sample size* $T = 512$. *Right*: *sample size* $T = 1024$.

TABLE 1
Whittle estimation: *sample size* $T = 512$, *block size* $N = 105$ *and shift* $S = 35$

Case	**Parameters** α_0	α_1	β_0	β_1	ϑ	**Estimates** $\widehat{\alpha}_0$	$\widehat{\alpha}_1$	$\widehat{\beta}_0$	$\widehat{\beta}_1$	$\widehat{\vartheta}$
1	0.15	0.20	0.5	0.3	0.5	0.130	0.177	0.497	0.299	0.473
2	0.15	0.20	0.8	−0.2	0.5	0.124	0.167	0.795	−0.201	0.463
3	0.20	0.25	0.5	0.3	0.5	0.161	0.219	0.497	0.301	0.455
4	0.20	0.25	0.8	−0.2	0.5	0.163	0.218	0.797	−0.201	0.453
5	0.30	−0.20	0.5	0.3	0.5	0.291	−0.183	0.498	0.299	0.506
6	0.30	−0.20	0.8	−0.2	0.5	0.287	−0.183	0.797	−0.203	0.501
7	0.15	0.20	0.5	0.3	−0.4	0.138	0.189	0.496	0.301	−0.407
8	0.15	0.20	0.8	−0.2	−0.4	0.138	0.190	0.799	−0.206	−0.410
9	0.20	0.25	0.5	0.3	−0.4	0.195	0.228	0.498	0.299	−0.409
10	0.20	0.25	0.8	−0.2	−0.4	0.193	0.229	0.795	−0.197	−0.412
11	0.30	−0.20	0.5	0.3	−0.4	0.286	−0.197	0.498	0.298	−0.404
12	0.30	−0.20	0.8	−0.2	−0.4	0.279	−0.180	0.796	−0.203	−0.404

Case	**Theoretical SD** $\sigma(\widehat{\alpha}_0)$	$\sigma(\widehat{\alpha}_1)$	$\sigma(\widehat{\beta}_0)$	$\sigma(\widehat{\beta}_1)$	$\sigma(\widehat{\vartheta})$	**Estimated SD** $\widehat{\sigma}(\widehat{\alpha}_0)$	$\widehat{\sigma}(\widehat{\alpha}_1)$	$\widehat{\sigma}(\widehat{\beta}_0)$	$\widehat{\sigma}(\widehat{\beta}_1)$	$\widehat{\sigma}(\widehat{\vartheta})$
1	0.115	0.119	0.035	0.069	0.109	0.117	0.146	0.045	0.089	0.106
2	0.115	0.119	0.047	0.075	0.109	0.115	0.146	0.057	0.100	0.103
3	0.115	0.119	0.035	0.069	0.109	0.107	0.132	0.043	0.091	0.096
4	0.115	0.119	0.047	0.075	0.109	0.110	0.131	0.056	0.098	0.102
5	0.115	0.119	0.035	0.069	0.109	0.131	0.140	0.043	0.090	0.108
6	0.115	0.119	0.047	0.075	0.109	0.125	0.140	0.057	0.099	0.107
7	0.074	0.119	0.035	0.069	0.051	0.089	0.155	0.045	0.091	0.058
8	0.074	0.119	0.047	0.075	0.051	0.088	0.150	0.054	0.096	0.053
9	0.074	0.119	0.035	0.069	0.051	0.090	0.142	0.044	0.091	0.053
10	0.074	0.119	0.047	0.075	0.051	0.088	0.142	0.057	0.099	0.054
11	0.074	0.119	0.035	0.069	0.051	0.089	0.140	0.046	0.093	0.055
12	0.074	0.119	0.047	0.075	0.051	0.093	0.146	0.057	0.099	0.056

TABLE 2
Whittle estimation: sample size $T = 1024$, block size $N = 200$ and shift $S = 45$

	Parameters					**Estimates**				
Case	α_0	α_1	β_0	β_1	ϑ	$\widehat{\alpha}_0$	$\widehat{\alpha}_1$	$\widehat{\beta}_0$	$\widehat{\beta}_1$	$\widehat{\vartheta}$
1	0.15	0.20	0.5	0.3	0.5	0.127	0.193	0.498	0.301	0.475
2	0.15	0.20	0.8	−0.2	0.5	0.131	0.195	0.796	−0.198	0.479
3	0.20	0.25	0.5	0.3	0.5	0.179	0.239	0.497	0.304	0.473
4	0.20	0.25	0.8	−0.2	0.5	0.176	0.241	0.798	−0.199	0.475
5	0.30	−0.20	0.5	0.3	0.5	0.286	−0.189	0.500	0.298	0.493
6	0.30	−0.20	0.8	−0.2	0.5	0.286	−0.197	0.799	−0.202	0.488
7	0.15	0.20	0.5	0.3	−0.4	0.143	0.198	0.498	0.302	−0.404
8	0.15	0.20	0.8	−0.2	−0.4	0.144	0.197	0.797	−0.197	−0.404
9	0.20	0.25	0.5	0.3	−0.4	0.195	0.245	0.500	0.300	−0.405
10	0.20	0.25	0.8	−0.2	−0.4	0.197	0.243	0.797	−0.199	−0.405
11	0.30	−0.20	0.5	0.3	−0.4	0.293	−0.197	0.500	0.299	−0.402
12	0.30	−0.20	0.8	−0.2	−0.4	0.293	−0.200	0.797	−0.199	−0.403

	Theoretical SD					**Estimated SD**				
Case	$\sigma(\widehat{\alpha}_0)$	$\sigma(\widehat{\alpha}_1)$	$\sigma(\widehat{\beta}_0)$	$\sigma(\widehat{\beta}_1)$	$\sigma(\widehat{\vartheta})$	$\widehat{\sigma}(\widehat{\alpha}_0)$	$\widehat{\sigma}(\widehat{\alpha}_1)$	$\widehat{\sigma}(\widehat{\beta}_0)$	$\widehat{\sigma}(\widehat{\beta}_1)$	$\widehat{\sigma}(\widehat{\vartheta})$
1	0.081	0.084	0.025	0.049	0.077	0.089	0.106	0.032	0.064	0.081
2	0.081	0.084	0.033	0.053	0.077	0.097	0.106	0.040	0.069	0.085
3	0.081	0.084	0.025	0.049	0.077	0.093	0.097	0.031	0.062	0.078
4	0.081	0.084	0.033	0.053	0.077	0.090	0.095	0.038	0.067	0.073
5	0.081	0.084	0.025	0.049	0.077	0.107	0.103	0.030	0.061	0.091
6	0.081	0.084	0.033	0.053	0.077	0.101	0.104	0.040	0.068	0.079
7	0.052	0.084	0.025	0.049	0.036	0.066	0.110	0.031	0.061	0.039
8	0.052	0.084	0.033	0.053	0.036	0.065	0.113	0.039	0.066	0.040
9	0.052	0.084	0.025	0.049	0.036	0.066	0.100	0.030	0.060	0.040
10	0.052	0.084	0.033	0.053	0.036	0.058	0.087	0.040	0.068	0.037
11	0.052	0.084	0.025	0.049	0.036	0.066	0.103	0.029	0.060	0.039
12	0.052	0.084	0.033	0.053	0.036	0.064	0.101	0.039	0.068	0.039

where $\gamma_{\alpha\vartheta} = [\frac{\log(1-\vartheta)}{\vartheta}, \frac{\log(1-\vartheta)}{2\vartheta}]'$, $\gamma_\vartheta = \frac{1}{1-\vartheta^2}$, and the matrices Γ_α and Γ_β are given in Example 2.3. The bandwidth parameters N and S for each table are based on values found in Figure 1 for $\theta = (0.20, 0.25, 0.5, 0.3, 0.5)$. As mentioned above, these values are very similar for the other parameters reported in Tables 1 and 2. Observe from these tables that the estimated parameters are close to their true values. Besides, the empirical standard deviations are close to their theoretical counterparts. These simulations suggest that the finite sample performance of the proposed estimators seems to be very good in terms of bias and standard deviations. This, despite the fact that in many of these simulations we have tested the method with large values of the long-memory parameter, that is, close to $\frac{1}{2}$. In Table 1, for example, for the combination $\alpha_0 = 0.20$, $\alpha_1 = 0.25$, the maximum value of $d(u)$ is 0.45. Additional Monte Carlo experiments with other model specifications are reported in Palma and Olea (2010). Those simulations explore the empirical optimal selection of N and S and the finite sample performance of the Whittle estimators. Note, however, that further research is needed to establish optimal selection of N and S from a theoretical perspective. A comparison of the performances of the Whittle method with a kernel maximum likelihood estimation approach proposed by Beran (2009) and two data illustrations are also discussed in that paper.

5. Final remarks. A class of locally stationary long-memory processes has been addressed in this paper, which is capable of modeling nonstationary time series data exhibiting time-varying long-range dependence. A computationally efficient Whittle estimation method has been proposed and it has been shown that these estimators possess very desirable asymptotic properties such as consistency, normality and efficiency. Moreover, several Monte Carlo simulations indicate that the estimates perform well even for relatively small sample sizes.

APPENDIX

This appendix contains nine auxiliary lemmas used to prove the theorems stated in Section 2 and the propositions stated in Section 3. Proof of these results are provided in Palma and Olea (2010).

LEMMA 1. *Let $f(u,\lambda)$ be a time-varying spectral density satisfying assumption A1 and let $\phi:[0,1]\times[-\pi,\pi]\to\mathbb{R}$ be a function such that $\phi(u,\lambda)$ is continuously differentiable in λ. Consider the function defined by*

$$g(u,\lambda)=\int_{-\pi}^{\pi}\phi(u,\lambda+\omega)f(u,\omega)\,d\omega,$$

and its Fourier coefficients $\widehat{g}(u,k)=\int_{-\pi}^{\pi}g(u,\lambda)e^{-ik\lambda}\,d\lambda$. Under assumption A1, for every $u\in[0,1]$ we have that $\lim_{n\to\infty}\sum_{k=-n}^{n}\widehat{g}(u,k)=2\pi g(u,0)$.

LEMMA 2. *Consider the function $\phi:[0,1]\times[-\pi,\pi]\to\mathbb{C}$, such that $\partial\phi(u,\gamma)/\partial u$ exists and $|\partial\phi(u,\gamma)/\partial u|\le K|\gamma|^{-2d(u)}$, where $0\le d(u)\le d$ for all $u\in[0,1]$. Then, for any $0\le t\le N$ we have that*

$$H_N\left[\phi\left(\frac{\cdot}{T},\gamma\right)h\left(\frac{\cdot}{N}\right),\lambda\right]=\phi\left(\frac{t}{T},\gamma\right)H_N(\lambda)+\mathcal{O}\left[\frac{N}{T}|\gamma|^{-2d}L_N(\lambda)\right].$$

LEMMA 3. *Consider $d_1,d_2\in[0,1/2)$ and for any $\ell\in\mathbb{Z}$ define the integral $I(\ell)=\int_1^{\infty}[(x-1)^{-2d_1}-x^{-2d_1}]|\ell+x|^{d_1+d_2-1}\,dx$. Then $I(\ell)=\mathcal{O}(|\ell|^{d_1+d_2-1})$.*

LEMMA 4. *Let $\phi(u,\lambda)$ be a positive function, symmetric in λ, such that $\phi(u,\lambda)\ge C|\lambda|^{2d(u)}$, for $\lambda\in[-\pi,\pi]$, where $d(u)$ is a positive bounded function for $u\in[0,1]$ and $C>0$. Let $Q(u)$ for $u\in[0,1]$ be the matrix defined in (16). Then there exists $K>0$ such that $X'Q(u)^{-1}X\le KX'XN^{2d(u)}$, for all vector $X\in\mathbb{R}^N$.*

LEMMA 5. *Let $\phi(u,\lambda)$ be a positive function, symmetric in λ, such that $\phi(u,\lambda) \geq C|\lambda|^{2d(u)}$, for $\lambda \in [-\pi,\pi]$, where $d(u)$ is a positive bounded function for $u \in [0,1]$ and $C > 0$. Let $Q(u)$ for $u \in [0,1]$ and $Q(\phi)$ be the matrices defined in (16). Then there exists $K > 0$ such that*

$$|X'[Q(\phi)^{-1} - Q(\varphi)]X| \leq K X'X N^{2d+1/2},$$

where $\varphi(u,\cdot) = \phi(u,\cdot)^{-1}/4\pi^2$, $d = \sup d(u) < \infty$ and $X \in \mathbb{R}^{NM}$.

LEMMA 6. *Let $\phi(u,\lambda)$ be a positive function, symmetric in λ, such that $\phi(u,\lambda) \geq C|\lambda|^{2d(u)}$, for $\lambda \in [-\pi,\pi]$, where $d(u)$ is a positive bounded function for $u \in [0,1]$ and $C > 0$. Let $Q(\phi)$ be the block-diagonal matrix defined in (16). Then there exists $K > 0$ such that*

$$\sup_X \left| \frac{X'RX}{X'Q(\phi)^{-1}X} \right| \leq KMN^{1-2d}T^{2d-1},$$

where $d = \sup d(u) < \frac{1}{2}$ and $X \in \mathbb{R}^{NM}$.

LEMMA 7. *Let $f(\lambda)$ and $\phi(\lambda)$ be two real-valued functions defined over $\lambda \in [-\pi,\pi]$ with Fourier coefficients $\widehat{f}(k)$ and $\widehat{\phi}(k)$, respectively, satisfying $|\widehat{f}(k)\widehat{\phi}(k)| \leq K/k^2$, for some positive constant K and $|k| > 0$. Let $C(N)$ be given by $C(N) = \sum_{t=0}^{N-1} h^2(\frac{t}{N}) \sum_{k=N-t}^{N-1} \widehat{f}(k)\widehat{\phi}(k)$ with bounded data taper, $|h(u)| < K$, for all $u \in [0,1]$. Then there exits a positive constant K such that $|C(N)| \leq K \log^2 N$.*

LEMMA 8. *Define $D(N,T) = \frac{1}{N} \sum_{t=0}^{N-1} \sum_{k=N-t+1}^{N-1} \frac{\varphi(k)}{k^2-d^2} (\frac{t-N/2}{T})$ with function $|\varphi(k)| < C \log N$ for all $0 \leq k \leq N$, $N > 1$, where C is a positive constant. Then there exists a constant $K > 0$ such that $|D(N,T)| \leq K \frac{\log^2 N}{T}$.*

LEMMA 9. *Let $z \in [0, 1+\delta]$ with $2 > \delta > 0$ and $2\beta > 2\alpha > 0$. Then, the positive double integral $I(z) = \int_0^1 |z-x|^{\alpha-1} \int_1^\infty (y-x)^{-\beta}(y-z)^{\alpha-1}\,dy\,dx$, satisfies $I(z) \leq K|1-z|^{2\alpha-\beta}$.*

Acknowledgments. We are deeply thankful to the Associate Editor and two anonymous referees for their careful reading of the manuscript and for their constructive comments which led to substantial improvements.

REFERENCES

BERAN, J. (2009). On parameter estimation for locally stationary long-memory processes. *J. Statist. Plann. Inference* **139** 900–915. MR2479836

BRILLINGER, D. R. (1981). *Time Series: Data Analysis and Theory*, 2nd ed. Holden-Day, Oakland, CA. MR0595684

BROCKWELL, P. J. and DAVIS, R. A. (1991). *Time Series: Theory and Methods*, 2nd ed. Springer, New York. MR1093459

CHANDLER, G. and POLONIK, W. (2006). Discrimination of locally stationary time series based on the excess mass functional. *J. Amer. Statist. Assoc.* **101** 240–253. MR2268042

DAHLHAUS, R. (1996). On the Kullback–Leibler information divergence of locally stationary processes. *Stochastic Process. Appl.* **62** 139–168. MR1388767

DAHLHAUS, R. (1997). Fitting time series models to nonstationary processes. *Ann. Statist.* **25** 1–37. MR1429916

DAHLHAUS, R. (2000). A likelihood approximation for locally stationary processes. *Ann. Statist.* **28** 1762–1794. MR1835040

DAHLHAUS, R. and POLONIK, W. (2006). Nonparametric quasi-maximum likelihood estimation for Gaussian locally stationary processes. *Ann. Statist.* **34** 2790–2824. MR2329468

DAHLHAUS, R. and POLONIK, W. (2009). Empirical spectral processes for locally stationary time series. *Bernoulli* **15** 1–39. MR2546797

DOUKHAN, P., OPPENHEIM, G. and TAQQU, M. S., EDS. (2003). *Theory and Applications of Long-Range Dependence*. Birkhäuser, Boston, MA. MR1956041

FRYZLEWICZ, P., SAPATINAS, T. and SUBBA RAO, S. (2006). A Haar–Fisz technique for locally stationary volatility estimation. *Biometrika* **93** 687–704. MR2261451

GENTON, M. and PERRIN, O. (2004). On a time deformation reducing nonstationary stochastic processes to local stationarity. *J. Appl. Probab.* **41** 236–249. MR2036285

GRADSHTEYN, I. S. and RYZHIK, I. M. (2000). *Table of Integrals, Series, and Products*, 6th ed. Academic Press, San Diego, CA. MR1773820

GRANGER, C. W. J. and DING, Z. (1996). Varieties of long memory models. *J. Econometrics* **73** 61–77. MR1410001

GUO, W., DAI, M., OMBAO, H. C. and VON SACHS, R. (2003). Smoothing spline ANOVA for time-dependent spectral analysis. *J. Amer. Statist. Assoc.* **98** 643–652. MR2011677

JENSEN, M. J. and WHITCHER, B. (2000). Time-varying long memory in volatility: Detection and estimation with wavelets. Technical report, EURANDOM.

KÖRNER, T. W. (1988). *Fourier Analysis*. Cambridge Univ. Press, Cambridge. MR0924154

ORBE, S., FERREIRA, E. and RODRIGUEZ-POO, J. (2005). Nonparametric estimation of time varying parameters under shape restrictions. *J. Econometrics* **126** 53–77. MR2118278

PALMA, W. (2007). *Long-Memory Time Series: Theory and Methods*. Wiley, Hoboken, NJ. MR2297359

PALMA, W. and OLEA, R. (2010). Supplement to "An efficient estimator for locally stationary Gaussian long-memory processes." DOI: 10.1214/10-AOS812.

PRIESTLEY, M. B. (1965). Evolutionary spectra and non-stationary processes. *J. Roy. Statist. Soc. Ser. B* **27** 204–237. MR0199886

VON SACHS, R. and MACGIBBON, B. (2000). Non-parametric curve estimation by wavelet thresholding with locally stationary errors. *Scand. J. Statist.* **27** 475–499. MR1795776

WHITTLE, P. (1953). Estimation and information in stationary time series. *Ark. Mat.* **2** 423–434. MR0060797

DEPARTMENT OF STATISTICS
PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
VICUÑA MACKENNA 4860
MACUL, SANTIAGO
CHILE
E-MAIL: wilfredo@mat.puc.cl