Bernoulli **6**(4), 2000, 709–728

Parametric estimation for Gaussian long-range dependent processes based on the log-periodogram

CARENNE LUDEÑA

Departamento de Matemáticas, Instituto Venezolano de Investigaciones Científicas, Apartado 21827, Caracas 1020-A, Venezuela. E-mail: cludena@euclides.ivic.ve

We establish the consistency and asymptotic normality of a certain minimum contrast estimator, introduced by Taniguchi (1979), for Gaussian long-range dependent processes. The estimator is based on regression over the log-periodogram in a parametric setting.

Keywords: Gaussian processes; log-periodogram; long-range dependence; minimum contrast estimators; spectral estimates

1. Introduction

Let $\{X_n\}_{1 \leqslant n}$ be a centered, strongly dependent, stationary Gaussian process with spectral density $f(\lambda, \theta)$, $\lambda \in (-\pi, \pi]$. Here θ is assumed to belong to a compact set $\Theta \subset \mathbb{R}^q$. This paper is concerned with estimating the parameter vector θ based on a finite number of observations $X_1, \ldots, X_n$.

A number of estimators have been proposed for θ in this setting. Classical minimum contrast (parametric) methods, based on bilinear forms of the observations, such as the Whittle or maximum likelihood estimators, have been developed under quite general conditions (see Fox and Taqqu, 1986; Dahlhaus, 1989; Giraitis and Surgailis, 1990). On the other hand, using an approximation of the spectral density for $\lambda \to 0$, Geweke and Porter-Hudak (1990) proposed a semi-parametric least-squares estimator based on a regression over the log-periodogram for the exponent d of an ARIMA(p, d, q) when this exponent was negative. The idea was to approximate the low frequencies of the spectral density, under general conditions, without having to define a parametric model. This allowed the higher frequencies to be trimmed off, which reduced the problem of model misspecification, although of course convergence rates were slower. This semi-parametric estimator was justified later by Künsch (1986) for positive d, but Robinson (1995) was the first to give a thorough theoretical account for this estimation scheme.

We remark that although semi-parametric rates are slower, in the context of long-range dependence, regression over the log-periodogram is preferred as it provides simpler numerical minimization problems than the usual parametric methods.

Following an approach developed by Taniguchi (1979), in this paper we construct a

1350–7265

minimum contrast estimator which basically amounts to regression over the log-periodogram in a parametric setting – that is, considering the whole frequency band. In order to study its asymptotic behaviour, we have to show the asymptotic normality of an estimator of integral functionals of the log of the spectral density of the type $\int_{-\pi}^{\pi} \psi(\lambda)\log(f(\lambda))\,d\lambda$. This functional exists for all spectral densities of strongly dependent processes with mild restrictions on ψ.

When considering integrals with respect to the log of the spectral density, the numerical minimization problem is almost as simple as in the semi-parametric context and allows working with the optimal rates. As another attractive feature, this minimum contrast estimator satisfies certain robustness properties with respect to the parametric model (see Taniguchi 1979).

Including the whole frequency band introduces certain technical problems, namely, that evaluations of the periodogram at very low frequencies are not asymptotically uncorrelated. However, this bad behaviour is cancelled out by integration and we find the same convergence rates as in the weakly dependent case. Our methodology uses the expansion of the logarithm of the periodogram in Hermite polynomials. In fact, under additional conditions other functionals can be treated in essentially the same fashion. Although we rely heavily on the underlying Gaussianity assumption, this approach suggests methods of proof based on Appell polynomials for linear processes, provided the formal expansion of the functional exists (see Giraitis and Surgailis 1986).

2. Notation and hypotheses

Let $\mathscr{F} = \{f(\cdot, \theta)\}_{\theta\in\Theta}$ be a family of functions indexed over a compact parameter set, $\Theta \subset \mathbb{R}^q$, such that for each fixed $\theta \in \Theta$, $f(\cdot, \theta)\colon (-\pi, \pi] \to \mathbb{R}$ is a positive integrable even function. Assume $\mathscr{F}$ satisfies the following:

Assumption A1. *For each $\theta \in \Theta$, there exist $\alpha(\theta) \in (0, 1)$, $\rho(\theta) > 0$ and $C(\theta)$ (with $\sup_{\theta\in\Theta}|C(\theta)| < C$), such that as $\lambda \to 0^+$, for all $\delta < 0$,*

$$f(\lambda, \theta) = C(\theta)\lambda^{\alpha(\theta)-1-\delta} + O(\lambda^{\rho(\theta)+\alpha(\theta)-1}).$$

There exists a constant $D > 0$ such that

$$\inf_{\theta\in\Theta}\inf_{\lambda\in(-\pi,\pi]} f(\lambda, \theta) \geqslant D.$$

Assumption A2. *For each $\theta \in \Theta$, $f(\lambda, \theta)$ is continuously differentiable in λ for all $\lambda \in (-\pi, \pi)$, $\lambda \neq 0$, and there exists a positive constant C_0 (independent of θ) such that*

$$\left|\frac{\partial}{\partial\lambda}\log f(\lambda, \theta)\right| \leqslant C_0|\lambda|^{-1}.$$

Assumption A3. *For each θ in the interior of Θ, $\log f(\lambda, \theta)$ and $\log^2 f(\lambda, \theta)$ are twice continuously differentiable in θ.*

Assumption A4. *For each θ in the interior of Θ, $\partial/\partial\theta_j \log f(\lambda, \theta)$, $1 \leqslant j \leqslant q$, are continuously differentiable in λ for all $\lambda \in (-\pi, \pi)$, $\lambda \neq 0$, and there exists a $C_1 > 0$ (independent of θ) such that, for all $0 < \delta_1$,*

$$\left|\frac{\partial}{\partial\theta_j} \log f(\lambda, \theta)\right| \leqslant C_1(|\lambda|^{-\delta_1} + 1)$$

$$\left|\frac{\partial^2}{\partial\lambda\partial\theta_j} \log f(\lambda, \theta)\right| \leqslant C_1(|\lambda|^{-(1+\delta_1)} + 1).$$

Assumption A5. *For each θ in the interior of Θ, $\partial^2/\partial\theta_j\partial\theta_k \log f(\lambda, \theta)$, $1 \leqslant j, k \leqslant q$, are continuously differentiable in λ for all $\lambda \in (-\pi, \pi)$, $\lambda \neq 0$, and there exists a $C_2 > 0$ (independent of θ) such that, for all $0 < \delta_2$,*

$$\left|\frac{\partial^2}{\partial\theta_j\partial\theta_k} \log f(\lambda, \theta)\right| \leqslant C_2(|\lambda|^{-\delta_2} + 1)$$

$$\left|\frac{\partial^3}{\partial\lambda\partial\theta_j\partial\theta_k} \log f(\lambda, \theta)\right| \leqslant C_2(|\lambda|^{-(1+\delta_2)} + 1).$$

Assumption A6. *For $\theta, \mu \in \Theta$, $\theta \neq \mu$ implies $f(\lambda, \theta) \neq f(\lambda, \mu)$ on a set of positive Lebesgue measure.*

For notation purposes, we shall sometimes denote $f(\cdot, \theta)$ by f_θ.

Remark 1. Assume that the true spectral density of observations $X_1, \ldots, X_n$ is $f(\lambda, \theta_0)$, for some $\theta_0 \in \Theta$. Let $r_k = \mathrm{E}(X_0 X_k)$. Under Assumption A1, Fox and Taqqu (1987) showed that, for all $\delta > 0$, $r_k = O(|k|^{-\alpha(\theta_0)+\delta})$.

Example 1. Consider a fractional ARIMA(p, d, q) with $0 < \mathrm{d} < \frac{1}{2}$, and p, q positive integers. Here $\theta = (a_1, \ldots, a_p, b_1, \ldots, b_q, d, \sigma^2) \in \mathbb{R}^{p+q+2}$, where $(a_i)_{i=1,\ldots,p}$ and $(b_i)_{i=1,\ldots,q}$ are the coefficients of a causal invertible ARMA process. We assume $\theta \in \Theta$, where Θ is a compact subset. The resulting process has spectral density

$$f(\lambda, d) = \frac{\sigma^2}{2\pi}\left(2\sin\frac{\lambda}{2}\right)^{-2d} f_{p,q}(\lambda),$$

which satisfies Assumptions A1–A6. The function $f_{p,q} \in C^\infty((-\pi, \pi])$ is the spectral density of the causal invertible ARMA process.

Consider the set $\mathscr{G} = \{h \in L^1(-\pi, \pi] : \int_{-\pi}^{\pi} \log^2(h(\lambda))\,\mathrm{d}\lambda < \infty\}$. Over this set we can define (Taniguchi 1979) the functional $D_2 : \mathscr{G} \times \mathscr{G} \to \mathbb{R}$ given by

$$D_2(f, g) = \int_{-\pi}^{\pi} \log^2\left(\frac{f}{g}(\lambda)\right) \mathrm{d}\lambda.$$

Now we can define the functional $T_2(\cdot) : \mathscr{G} \to \Theta$, based on D_2, given by

$$T_2(g) = \arg\min_{\theta\in\Theta} D_2(f_\theta,\, g)$$

$$= \arg\min_{\theta\in\Theta} \int_{-\pi}^{\pi} (\log^2 f_\theta(\lambda) - 2\log f_\theta(\lambda)\log g(\lambda))\,\mathrm{d}\lambda. \qquad (1)$$

Remark 2. $T_2(g)$ may have multiple values so we shall assume that it stands for any one of those values.

Let $\mathscr{H}$ be a given class of functions. If

$$\int_{-\pi}^{\pi} \phi(\lambda)\log(g_n(\lambda))\,\mathrm{d}\lambda \to \int_{-\pi}^{\pi} \phi(\lambda)\log(g(\lambda))\,\mathrm{d}\lambda$$

for every function $\phi(\cdot) \in \mathscr{H}$, we will say that g_n converges log–$\mathscr{H}$ to g ($g_n \overset{l.\mathscr{H}}{\to} g$).

Throughout this paper, $a \vee b = \max(a, b)$, C stands for a generic constant which may change value from line to line, and by 'θ_0–probability' we mean that generated by a centred Gaussian process with spectral density $f(\lambda, \theta_0)$.

3. Main result

Given observations $X_1, \ldots, X_n$ with true spectral density $f_0(\lambda) = f(\lambda, \theta_0)$, $\theta_0 \in \Theta$, we want to construct a minimum contrast estimator of the parameter θ_0 based on the best approximation over the class $\mathscr{F}$ for f_0. In order to do this we shall choose the value of θ that minimizes the functional $D_2(f_\theta, \hat{f}_0)$, where $\hat{f}_0$ is an estimator of the true spectral density f_0. The main result in this section discusses the asymptotic behaviour of this estimator.

The periodogram I_n is defined by $I_n(\lambda) = |w(\lambda)|^2$, with

$$w(\lambda) = \frac{1}{(2\pi n)^{1/2}} \sum_{j=1}^{n} X_j \mathrm{e}^{\mathrm{i}\lambda j}. \qquad (2)$$

Although it is a bad pointwise estimator of f_0, sufficiently smooth functionals of the periodogram yield consistent estimators, as they average out this bad behaviour.

Consider, for a given $n \in \mathbb{N}$, the set of frequencies $\lambda_j = (2\pi j)/n$. Our estimator of f_0 will be given by the step approximation of the periodogram

$$g_n(\lambda) = I_n(\lambda_j) \qquad \text{if } \frac{2\pi(j-1)}{n} < \lambda \leqslant \frac{2\pi j}{n}.$$

Notice that $g_n \in \mathscr{G}$ almost everywhere in θ_0-probability. In all that follows we shall assume g_n is positive.

Let $\kappa = \mathrm{e}^{\gamma}$, with γ denoting Euler's constant. We have the following theorem which is analogous to Theorems 4, 5 and 6 in Taniguchi (1979) in the weakly dependent case:

Theorem 1. *Assume that the family $\mathscr{F}$ satisfies Assumptions A1–A6. Then $T_2(f_0)$ exists, is unique and $T_2(f_0) = \theta_0$. If additionally θ_0 lies in the interior of Θ, then $T_2(\kappa g_n) \xrightarrow{P} \theta_0$. Finally, if*

$$\sigma_f = \left\{ \int_{-\pi}^{\pi} \left(\frac{\partial^2 \log f(\lambda, \theta)}{\partial\theta_j \partial\theta_k} \log f_0(\lambda) - \frac{1}{2} \frac{\partial^2 \log^2 f(\lambda, \theta)}{\partial\theta_j \partial\theta_k} \right) \bigg|_{\theta=\theta_0} \mathrm{d}\lambda \right\} \tag{3}$$

is a non-singular matrix, then

$$\sqrt{n}(T_2(\kappa g_n) - \theta_0) \xrightarrow{\mathscr{D}} \mathscr{N}\left(0, \frac{2\pi^3}{3} \int_{-\pi}^{\pi} \sigma_f(\lambda)\sigma'_f(\lambda)\,\mathrm{d}\lambda\right), \tag{4}$$

where $\xrightarrow{\mathscr{D}}$ stands for convergence in distribution and $\sigma_f(\lambda)$ is the q-vector defined by

$$\sigma_f(\lambda)_j = (\sigma_f)^{-1} \frac{\partial \log f(\lambda, \theta)}{\partial\theta_j} \bigg|_{\theta=\theta_0}.$$

Remark 3. This minimum contrast estimator is not efficient. Indeed, the lower bound for the variance of any estimator of θ is given by (see, for example, Dzhaparidze 1985)

$$\sigma^2 = 4\pi \int_{-\pi}^{\pi} \sigma_f(\lambda)\sigma'_f(\lambda)\,\mathrm{d}\lambda. \tag{5}$$

This lower bound is, for example, achieved by the maximum likelihood estimator or the Whittle pseudo-likelihood estimator. Thus, the relative efficiency of the estimator we propose with respect to these estimators is $\pi^2/6$.

4. Integrals with respect to the log-periodogram

In order to prove Theorem 1 we require asymptotics for integrals with respect to the log-periodogram. Actually, we will consider discrete versions of the integral. In this section, we will drop the parameter θ from the notation.

Assume $\psi \in L^2((-\pi, \pi])$ is an even function that satisfies the following:

Assumption B1. *There exists a positive constant K_1 and $\beta < \frac{1}{3}$ such that*

$$|\psi(\lambda)| \leqslant K_1 |\lambda|^{-\beta}.$$

Assumption B2. *ψ is continuously differentiable in λ, for all $\lambda \neq 0$, and there exists a positive constant K_2 such that*

$$\left| \frac{\mathrm{d}}{\mathrm{d}\lambda} \psi(\lambda) \right| \leqslant K_2 |\lambda|^{-\beta-1}.$$

We are interested in the asymptotics of the function

$$l_n = \frac{4\pi}{n} \sum_{j=1}^{[n/2]-1} \psi(\lambda_j) \log(\kappa I_n(\lambda_j)). \tag{6}$$

for $\lambda_j = 2\pi j/n$ *and* $\kappa = \mathrm{e}^{\gamma}$. *We have the following result:*

Theorem 2. *Assume that the true spectral density of the observations* $f_0(\lambda)$ *satisfies Assumptions A1 and A2. Then*

$$\sqrt{n}\left(\frac{4\pi}{n}\sum_{j=1}^{[n/2]-1}\psi(\lambda_j)\log(\kappa I_n(\lambda_j)) - \int_{-\pi}^{\pi}\psi(\lambda)\log(f_0(\lambda))\,\mathrm{d}\lambda\right) \xrightarrow{\mathscr{D}} \mathscr{N}\left(0, \frac{2\pi^3}{3}\int_{-\pi}^{\pi}\psi^2(\lambda)\,\mathrm{d}\lambda\right)$$

with $\kappa = \mathrm{e}^{\gamma}$.

Remark 4. That the variance in Theorem 2 does not depend on the spectral density $f_0(\lambda)$ is not surprising. Recall that the Fréchet derivative of the nonlinear functional of the spectral density $L_\psi(f) = \int_{-\pi}^{\pi} \psi(\lambda) \log f(\lambda)\,\mathrm{d}\lambda$ at the point f is given by $DL_\psi(f)(\lambda) = \psi(\lambda) 1/f(\lambda)$. Thus, this result only states that the variance of the estimator is a constant times $\|DL_\psi(f)f\|_2^2$.

Before proving Theorem 2 we shall require some more notation and some preliminary technical results, whose proofs are given in the Appendix.

4.1. Using Hermite polynomials

First of all write, for each λ, $I_n(\lambda) = X_n^2(\lambda) + Y_n^2(\lambda)$, where $X_n(\lambda)$ stands for the real part of (2) and $Y_n(\lambda)$ for its imaginary part. Thus $\log(I_n(\lambda))$ is actually a function of $X_n(\lambda)$ and $Y_n(\lambda)$.

Using this simple fact and taking advantage of the Gaussian framework, we shall expand the functions $(\log(x^2 + y^2) - (\log 2 - \gamma))^p \in L^2(\mathrm{e}^{-(x^2+y^2)/2})$, with $p \in \mathbb{N}$, on the basis of the two-dimensional Hermite polynomials $H_{m,l}(x, y) = H_m(x) H_l(y)$. Let $c_{m,l}^{(p)}$ be the corresponding coefficients:

$$\begin{aligned} c_{m,l}^{(p)} &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(\log(x^2+y^2) - (\log 2 - \gamma))^p H_m(x) H_l(y)\,\mathrm{e}^{-(x^2+y^2)/2}\,\mathrm{d}x\,\mathrm{d}y \\ &= \frac{1}{2\pi}\int_0^{2\pi}\int_0^{\infty}(\log(r^2) - (\log 2 - \gamma))^p r H_m(r\cos\theta) H_l(r\sin\theta)\,\mathrm{e}^{-r^2/2}\,\mathrm{d}r\,\mathrm{d}\theta. \end{aligned} \tag{7}$$

Notice that $c_{m,l}^{(p)} \equiv 0$ if either m or l is odd. Define

$$h^{(p)} = \left(\frac{1}{2}\right)\int_0^{\infty}(\log(u) - (\log 2 - \gamma))^p\,\mathrm{e}^{-u/2}\,\mathrm{d}u.$$

It is straightforward to show that $c_{0,0}^{(p)} = h^{(p)}$. We have, in particular, $c_{0,0}^{(1)} = 0$ and $c_{0,0}^{(2)} = \pi^2/6$. Observe also that by the Cauchy–Schwarz inequality,

$$|c_{m,l}^{(p)}| \leqslant (c_{0,0}^{(p)})^{1/2}\sqrt{m!}\sqrt{l!}. \tag{8}$$

Consider the set of frequencies $\lambda_{j_1}, \ldots, \lambda_{j_\ell} \in \{\lambda_1, \ldots, \lambda_{[n/2]-1}\}$, with $j_s \neq j_t$ for $s \neq t$, $1 \leqslant s, t \leqslant \ell$. The following two lemmas give bounds for the covariance of the vector $(X_n(\lambda_{j_1}), Y_n(\lambda_{j_1}), \ldots, X_n(\lambda_{j_\ell})Y_n(\lambda_{j_\ell}))$.

Lemma 1. *Assume that the true spectral density of the observations $f_0(\lambda)$ satisfies Assumptions A1 and A2. Then* $\mathrm{var}(X_n(\lambda_j)) = a_n(\lambda_j) + b_n(\lambda_j)$, $\mathrm{var}(Y_n(\lambda_j)) = a_n(\lambda_j) - b_n(\lambda_j)$ and $\mathrm{cov}(X_n(\lambda_j), Y_n(\lambda_j)) = c_n(\lambda_j)$, *where*

$$a_n(\lambda_j) = \frac{1}{4\pi}\sum_{|k|\leqslant n}\left(1 - \frac{|k|}{n}\right) r_k \cos\lambda_j k$$

$$b_n(\lambda_j) = \frac{1}{4\pi n}\frac{\cos(\lambda_j)}{\sin(\lambda_j)}\sum_{k=0}^{n-1} r_k \sin\lambda_j k$$

$$c_n(\lambda_j) = \frac{1}{4\pi n}\sum_{k=0}^{n-1} r_k \sin\lambda_j k.$$

Then there exist constants $C, D > 0$ such that

$$\left|\frac{2a_n(\lambda_j) - f_0(\lambda_j)}{f_0(\lambda_j)}\right| \leqslant \frac{C}{j} \qquad \text{for all } 1 \leqslant j \leqslant [n/2]-1, \tag{9}$$

$$a_n(\lambda_j) \geqslant D \qquad \text{for all } 1 \leqslant j \leqslant [n/2]-1. \tag{10}$$

Lemma 2. *Let $n^\nu \leqslant j < k \leqslant [n/2]-1$, for $0 < \nu < 1$. Assume that the true spectral density of the observations $f_0(\lambda)$ satisfies Assumptions A1 and A2. Then there exists a constant $C > 0$ (independent of j,k) such that*

$$|\mathrm{cov}(X_n(\lambda_j), X_n(\lambda_k))| \leqslant \frac{C\log k + 1}{j} f_0^{1/2}(\lambda_j) f_0^{1/2}(\lambda_k),$$

$$|\mathrm{cov}(Y_n(\lambda_j), Y_n(\lambda_k))| \leqslant \frac{C\log k + 1}{j} f_0^{1/2}(\lambda_j) f_0^{1/2}(\lambda_k),$$

$$|\mathrm{cov}(X_n(\lambda_j), Y_n(\lambda_k))| \leqslant \frac{C\log k + 1}{j} f_0^{1/2}(\lambda_j) f_0^{1/2}(\lambda_k),$$

$$|\mathrm{cov}(Y_n(\lambda_j), X_n(\lambda_k))| \leqslant \frac{C\log k + 1}{j} f_0^{1/2}(\lambda_j) f_0^{1/2}(\lambda_k).$$

The above lemma is due to Theorem 2 of Robinson (1995), if $n^{\nu_1} < j$, $k < n^{\nu_2}$, for $0 < \nu_1 < \nu_2 < 1$. If $j = O(n)$ or $k = O(n)$ the proof follows analogously, so it is omitted.

Set $\varepsilon_{j,n} = b_n(\lambda_j)/a_n(\lambda_1)$. The following lemma deals with the asymptotic behaviour of this sequence.

Lemma 3. *There exist $c > 0$ and n_0 such that, for all $n > n_0$ and for all $1 \leqslant j \leqslant [n/2] - 1$,*

$$|\varepsilon_{n,j}| \leqslant 1 - c. \tag{11}$$

Also, there exists a constant $C > 0$ such that, for all $1 \leqslant j \leqslant [n/2] - 1$,

$$|\varepsilon_{n,j}| \leqslant C \frac{\log j + 1}{j}. \tag{12}$$

Remark 5. As a consequence of equations (10), (11) and Remark 1 we have that, for all $1 \leqslant j \leqslant [n/2] - 1$, $|c_n(\lambda_j)/(a_n(\lambda_j)(1 - \varepsilon_{n,j}^2)^{1/2})| \to 0$ uniformly as $n \to \infty$.

Introduce the normalized random variables

$$\tilde{X}_n(\lambda_j) = \frac{X_n(\lambda_j)}{\sqrt{a_n(\lambda_j) + b_n(\lambda_j)}}, \qquad \tilde{Y}_n(\lambda_j) = \frac{Y_n(\lambda_j)}{\sqrt{a_n(\lambda_j) - b_n(\lambda_j)}}.$$

Note that $\tilde{X}_n(\lambda_j)$ and $\tilde{Y}_n(\lambda_j)$ are standard Gaussian variables with covariance $c_n(\lambda_j)/(a_n(\lambda_j)(1 - \varepsilon_{n,j}^2)^{1/2}$. Put $Z_{n,j} = Y_n^2(\lambda_j) - X_n^2(\lambda_j)/(X_n^2(\lambda_j) + Y_n^2(\lambda_j))$. It follows that $Z_{n,j}$ is almost surely bounded by 1. Now write

$$\begin{aligned}&\log(\tilde{X}_n^2(\lambda_j) + \tilde{Y}_n^2(\lambda_j))\\ &\quad = \log(X_n^2(\lambda_j) + Y_n^2(\lambda_j)) - \log(a_n(\lambda_j)) - \log(1 - \varepsilon_{n,j}^2) + \log(1 + \varepsilon_{n,j} Z_{n,j}).\end{aligned} \tag{13}$$

As a result of equations (9)–(13), it turns out that the asymptotics of the logarithm of the periodogram can be obtained from those of the normalized periodogram. Based on Lemma 1, the next lemma shows how to calculate the moments of $\log(\tilde{X}_n^2(\lambda_j) + \tilde{Y}_n^2(\lambda_j)) - (\log 2 - \gamma)$, for all $1 \leqslant j \leqslant [n/2] - 1$.

Lemma 4. *Assume that the true spectral density of the observations $f_0(\lambda)$ satisfies Assumptions A1 and A2. Then*

$$\mathrm{E}(\log(\tilde{X}_n^2(\lambda_j) + \tilde{Y}_n^2(\lambda_j)) - (\log 2 - \gamma))^p = c_{0,0}^{(p)} + O\left(\frac{c_n(\lambda_j)}{a_n(\lambda_j)}\right)^2.$$

Remark 6. Remark 5 yields $c_n(\lambda_j)/(a_n(\lambda_j)(1 - \varepsilon_{n,j}^2)^{1/2} = o(1)$ uniformly for all $1 \leqslant j \leqslant [n/2] - 1$, so that the bounds in Lemma 4 are uniform over the whole frequency range.

For a given $p \geqslant 2$, consider any collection of positive a_i, $1 \leqslant i \leqslant \ell$, such that $\sum_i a_i = p$. Let s be the number of $a_i = 1$. Consider the vector $(\lambda_{j_1}, \ldots, \lambda_{j_\ell})$, $n^\nu \leqslant j_i \leqslant [n/2]$ with $\lambda_{j_k} \neq \lambda_{j_i}$ if $k \neq i$. We have the following lemma.

Lemma 5. *Assume that the true spectral density of the observations $f_0(\lambda)$ satisfies Assumptions A1 and A2. Assume $n^\nu \leqslant j_i \leqslant [n/2]-1$, $i=1,\ldots,\ell$, for $0<\nu<1$. Let $j=\min(j_1,\ldots,j_\ell)$. Then, there exists an n such that*

$$\mathrm{E}\prod_{i=1}^{\ell}(\log(\tilde{X}_n^2(\lambda_{j_i})+\tilde{Y}_n^2(\lambda_{j_i}))-(\log 2-\gamma))^{a_i}=\prod_{i=1}^{\ell}c_{0,0}^{(a_i)}+O\left(\frac{C\log n}{j}\right)^{s\vee 2}.$$

We are now ready to prove Theorem 2.

Proof of Theorem 2. First define $\tilde{I}_n(\lambda_j)=I_n(\lambda_j)/a_n(\lambda_j)$. We define the normalized version of l_n (see (6)) as

$$\tilde{l}_n=\frac{4\pi}{n}\sum_{j=1}^{[n/2]-1}\psi(\lambda_j)\log(\kappa\tilde{I}_n(\lambda_j)).$$

As is clear from (13), convergence in distribution of $\sqrt{n}(l_n-\mathrm{E}l_n)$ will be implied by that of $\sqrt{n}(\tilde{l}_n-\mathrm{E}\tilde{l}_n)$, if we show that $R_n=1/\sqrt{n}\sum_{j=1}^{[n/2]-1}\psi(\lambda_j)\log(1+\varepsilon_{n,j}Z_{n,j})$ tends to zero in probability. The proof of the theorem will thus be divided into three parts. First we will show convergence in distribution for the centred $\tilde{l}_n$, then control the bias term $\sqrt{n}(\mathrm{E}l_n-\int_{-\pi}^{\pi}\psi(\lambda)\log f_0(\lambda)\,\mathrm{d}\lambda)$, and finally show the aforementioned convergence to zero in probability.

Asymptotic distribution of $\tilde{l}_n-\mathrm{E}\tilde{l}_n$. Choose $\beta/(1-\beta)<\nu<(1-2\beta)/2(1-\beta)$. As a consequence of Lemma 4 we have

$$\begin{aligned}\mathrm{E}\left[n^{-1/2}4\pi\sum_{j=1}^{n^\nu}\psi(\lambda_j)(\log(\tilde{I}_n(\lambda_j))-(\log 2-\gamma))\right]^2 &\leqslant Cn^{\nu-1}\sum_{j=1}^{n^\nu}\psi^2(\lambda_j)\left(\frac{\pi^2}{6}+o(1)\right)\\ &\leqslant Cn^{\nu-1+2\beta}\sum_{j=1}^{n^\nu}\frac{1}{j^{2\beta}}\\ &=O(n^{2\nu(1-\beta)+2\beta-1})=o(1).\end{aligned}$$

Thus $n^{-1/2}4\pi\sum_{j=1}^{n^\nu}\psi(\lambda_j)(\log(\tilde{I}_n(\lambda_j))-(\log 2-\gamma))\xrightarrow{P}0$.

On the other hand, consider

$$U_{n,\nu}=n^{-1/2}4\pi\sum_{j=n^\nu}^{[n/2]}\psi(\lambda_j)(\log(\tilde{I}_n(\lambda_j))-(\log 2-\gamma)).$$

Clearly $\mathrm{E}U_{n,\nu}=0$. Also, as a consequence of Lemmas 4 and 5, we have

$$\mathrm{E}U_{n,\nu}^2=n^{-1}(4\pi)^2\sum_{j=n^\nu}^{[n/2]-1}\psi^2(\lambda_j)\left(\frac{\pi^2}{6}+o(1)\right)+2n^{-1}(4\pi)^2\sum_{j_1=n^\nu}^{[n/2]}\sum_{j_2=j_1+1}^{[n/2]}\psi(\lambda_{j_1})\psi(\lambda_{j_2})\left(\frac{C\log j_2}{j_1}\right)^2.$$

The first term in the above sum converges to the stated variance. The second term is bounded by

$$Cn^{\beta}(\log n)^2 \left(\sum_{j_1=n^{\nu}}^{[n/2]-1} \frac{1}{j_1^{\beta+1}} \right) \leqslant Cn^{(\beta(1-\nu)\vee 0)-\nu}(\log n)^2.$$

Now assume that $p \geqslant 2$ and even. We have

$$\begin{aligned}
\mathrm{E}U^p_{n,1/2} &= \sum_{\ell=1}^{p} \frac{n^{-p/2}(4\pi)^p}{\ell!} \sum_{a_1,\ldots,a_\ell}^{1} \sum_{\substack{n^{1/2}\leqslant j_1,\ldots,j_\ell\leqslant[n/2]\\ j_m\neq j_k}} \prod_{i=1}^{\ell} \psi^{a_i}(\lambda_{j_i}) \mathrm{E} \prod_{i=1}^{\ell} (\log(\tilde{I}_n(\lambda_{j_i})) - (\log 2 - \gamma))^{a_i} \\
&= n^{-p/2}(4\pi)^p \frac{p!}{(p/2)!2^p} \sum_{\substack{n^{1/2}\leqslant j_1,\ldots,j_{p/2}\leqslant[n/2]\\ j_m\neq j_k}} \prod_{i=1}^{p/2} \psi^2(\lambda_{j_i}) \mathrm{E} \prod_{i=1}^{p/2} (\log(\tilde{I}_n(\lambda_{j_i})) - (\log 2 - \gamma))^2 \\
&\quad + \sum_{\substack{\ell=1,\ldots,p\\ \ell\neq p/2}} \frac{n^{-p/2}(4\pi)^p}{\ell!} \sum_{a_1,\ldots,a_\ell}^{1*} \sum_{\substack{n^{1/2}\leqslant j_1,\ldots,j_\ell\leqslant[n/2]\\ j_m\neq j_k}} \prod_{i=1}^{\ell} \psi^{a_i}(\lambda_{j_i}) \qquad (14) \\
&\quad \times \mathrm{E} \prod_{i=1}^{\ell} (\log(\tilde{I}_n(\lambda_{j_i})) - (\log 2 - \gamma))^{a_i}.
\end{aligned}$$

Here $\sum^1$ is the sum over all possible $a_1, \ldots, a_\ell$ with $a_i \neq 0$ and such that $\sum_i a_i = p$; $\sum^{1*}$ excludes the case $a_i = 2$ for $i = 1, \ldots, p/2$.

By Lemma 5,

$$\mathrm{E} \prod_{i=1}^{p/2} (\log(\tilde{I}_n(\lambda_{j_i})) - (\log 2 - \gamma))^2 = (\pi^2/6)^{p/2} + (C \log n / n^{1/2})^2.$$

Thus, the first term in the above sum converges to

$$\frac{p!}{(p/2)!2^p} \left(n^{-1}(4\pi)^2 \sum_{j=n^{1/2}}^{[n/2]} \psi^2(\lambda_j) \frac{\pi^2}{6} \right)^{p/2} + o(n^{-p/2}).$$

In order to show the covergence to zero of the second term of the right-hand side of (14), we shall group the terms in $\sum^{1*}$ based on the number of a_i which are equal to one. Assume, for some fixed term, that the number of $a_i = 1$ is s, $0 \leqslant s \leqslant \ell$, with ℓ the number of positive a_i. For notational purposes assume, for $s > 0$, that $a_1 = a_2 = \cdots = a_s = 1$. By Lemma 5 there exists a constant $C = C(a_1, \ldots, a_\ell, p)$ such that

$$\sum_{\substack{n^{1/2}\leqslant j_1,\ldots,j_\ell\leqslant[n/2]\\ j_m\neq j_k}} \sum_{i=1}^{\ell} \psi^{a_i}(\lambda_{j_i})\mathrm{E}\prod_{i=1}^{\ell}(\log(\tilde{I}_n(\lambda_{j_i})) - (\log 2 - \gamma))^{a_i}$$

$$\leqslant Cn^{\beta p}(\log n)^{\ell}\left(\sum_{n^{1/2}\leqslant j\leqslant[n/2]}\frac{1}{j^{\beta+2}}\right)^{s}\prod_{i=s}^{\ell}\sum_{n^{1/2}\leqslant j\leqslant[n/2]}\frac{1}{j^{\beta a_i}}. \tag{15}$$

We shall study this last expression according to the value of β. If $\beta < 0$ then each of the sums on the right-hand side of (15) diverges, so that it is bounded by

$$Cn^{\beta p}(\log n)^{\ell} n^{-\beta} n^{(\ell - s/2)-\beta(p-1)} \leqslant C(\log n)^{\ell} n^{(l-s/2)}. \tag{16}$$

If $\beta = 0$ then the right-hand side of (15) is bounded by

$$Cn^{\beta p}(\log n)^{\ell+1} n^{(\ell-s/2)} \leqslant C(\log n)^{\ell+1} n^{(l-s/2)}. \tag{17}$$

The case where $0<\beta<\frac{1}{3}$ is the most delicate as the sums on the right-hand side of (15) may or may not converge according to the value of a_i. Assume for notational purposes that $\beta a_i \leqslant 1$ for $s+1 \leqslant i \leqslant t$, for some $t \geqslant 0$ (if there are no a_i in this set then $t=0$). If $t=0$ we assume $\prod_{i=s+1}^{t} = 1$. On the other hand, we have $\beta a_i > 1$ for all $t+1 \leqslant i \leqslant \ell$, so that the respective sums converge. Observe also that $2 \leqslant a_i$, for all $s+1 \leqslant i \leqslant \ell$. With this notation,

$$\begin{aligned}
\prod_{i=s+1}^{\ell}\sum_{n^{v}\leqslant j\leqslant[n/2]}\frac{1}{j^{\beta a_i}} &= \prod_{i=s+1}^{t}\sum_{n^{v}\leqslant j\leqslant[n/2]}\frac{1}{j^{\beta a_i}} \times \prod_{i=t+1}^{\ell}\sum_{n^{v}\leqslant j\leqslant[n/2]}\frac{1}{j^{\beta a_i}}\\
&\leqslant Cn^{(t-s)-\beta}\sum_{i=s+1}^{t} a_i\\
&\leqslant Cn^{(t-s)-2\beta(t-s)}\\
&\leqslant Cn^{(\ell-s)(1-2\beta)}.
\end{aligned}$$

Thus, as the sum corresponding to $i=1$ is convergent in this case, the right-hand side of (15) is bounded by

$$Cn^{\beta p}(\log n)^{\ell} n^{(\ell-s/2)(1-2\beta)}. \tag{18}$$

As we are considering the case $\ell = p/2$ separately, we have $\ell - s < p/2 - s/2$ in the second sum in (14). This concludes the proof as we must normalize by $n^{-p/2}$ in equations (16), (17) and (18).

If $p \geqslant 3$ and odd, bounding as above, $\mathrm{E}U^{p}_{n,1/2} = o(1)$ as we do not have the term given in (14) with all $a_i = 2$.

This together with the fact that $\mathrm{E}(U_{n,1/2} - U_{n,v})^2 \to 0$ as $n \to \infty$, shows the required convergence in distribution.

Controlling the bias term. Because of the definition of l_n and $\tilde{l}_n$, in order to control the bias term we have to verify that

$$\sqrt{n}\left(\frac{4\pi}{n}\sum_{j=1}^{[n/2]-1}\psi(\lambda_j)\log f_0(\lambda_j)-\int_{-\pi}^{\pi}\psi(\lambda)\log f_0(\lambda)\,\mathrm{d}\lambda\right)\to 0, \tag{19}$$

$$\sqrt{n}\left(\frac{4\pi}{n}\sum_{j=1}^{[n/2]-1}\psi(\lambda_j)(\log f_0(\lambda_j)-\log(2a_n(\lambda_j)))\right)\to 0, \tag{20}$$

$$\sqrt{n}\left(\frac{4\pi}{n}\sum_{j=1}^{[n/2]-1}\psi(\lambda_j)\log(1-\varepsilon_{n,j}^2)\right)\to 0. \tag{21}$$

Choose $0<\nu<(1-2\beta)/2(1-\beta)$. Under Assumptions A1, A2, B1 and B2 we have

$$n^{-1/2}4\pi\sum_{j=1}^{n^\nu}\psi(\lambda_j)\log f_0(\lambda_j)\leqslant C\log(d)n^{\nu(1-\beta)+\beta-1/2}=o(1),$$

$$n^{-1/2}2\pi\sum_{j=1}^{n^\nu}\psi(\lambda_j)\log a_n(\lambda_j)\leqslant C\log(D)n^{\nu(1-\beta)+\beta-1/2}=o(1),$$

$$n^{1/2}\int_0^{n^{\nu-1}}\psi(\lambda)\log f_0(\lambda)\,\mathrm{d}\lambda\leqslant C\log(d)n^{1/2+(\nu-1)(1-\beta)}=o(1),$$

$$n^{-1/2}4\pi\sum_{j=n^\nu}^{[n/2]-1}\psi(\lambda_j)(\log f_0(\lambda_j)-\log 2a_n(\lambda_j))\leqslant Cn^{-1/2}4\pi\sum_{j=n^\nu}^{[n/2]-1}\frac{|\psi(\lambda_j)|}{j}$$

$$=o(n^{-1/2+(\beta\vee 0)})=o(1),$$

$$\sqrt{n}\left(\frac{2\pi}{n}\sum_{j=n^\nu}^{[n/2]-1}\psi(\lambda_j)\log f_0(\lambda_j)-\int_{n^{\nu-1}}^{\pi}\psi(\lambda)\log f_0(\lambda)\,\mathrm{d}\lambda\right)$$

$$\leqslant Cn^{-1/2}4\pi\sum_{j=n^\nu}^{[n/2]-1}\frac{\log n}{n}(\lambda_j)^{-1-\beta}\leqslant Cn^{-3/2}\log n n^{1+(\beta\vee 0)}=o(1),$$

which shows convergence to zero of the left-hand sides of (19) and (20).

Convergence of the left-hand side of (21) follows readily from (11) and (12), as

$$n^{-1/2}4\pi\sum_{j=1}^{n^\nu}\psi(\lambda_j)\leqslant n^{\nu-1/2}.$$

Convergence to zero in probability of R_n. As $|Z_{n,j}|\leqslant 1$ a.s., the stated result follows as, by Lemma 4, there exists a $c>0$ such that $\log(1+\varepsilon_{n,j}Z_{n,j})\leqslant c$ a.s., and if $n^{\nu_1}\leqslant j\leqslant[n/2]-1$, with $0<\nu_1<1$, then there exists a $C>0$ such that $\log(1+\varepsilon_{n,j}Z_{n,j})\leqslant C\log(j+1)/j$ almost everywhere. With these bounds we can then proceed as for the bias terms. □

We can now prove Theorem 1.

Proof of Theorem 1. Following the proof of Theorem 4 in Taniguchi (1979), define $h(\theta) = D_2(f(\cdot, \theta), f(\cdot, \theta_0))$ and $h_n(\theta) = D_2(f(\cdot, \theta), \kappa g_n(\cdot))$. Let $h^{(1)}(\theta)$ and $h_n^{(1)}$ be the respective gradient vectors, and let $h^{(2)}(\theta)$ and $h_n^{(2)}$ be the respective matrices of second derivatives. Let $\theta_m \to \theta$; then under Assumptions A1–A3 by the dominated convergence theorem we have

$$|h(\theta_m) - h(\theta)| = \left| \int_{-\pi}^{\pi} \log^2(f_{\theta_m}) - \log^2(f_\theta) + 2(\log(f_{\theta_m}) - \log(f_\theta))\log f_0 \, d\lambda \right| \to 0.$$

Thus, h is continuous and reaches a (not necessarily unique) minimum over the compact set Θ. If $f_0 = f(\cdot, \theta_0)$, then $h(\theta_0) = 0$. As $h(\theta) \geqslant 0$ this gives $T_2(f_0) = \theta_0$ in this case. Notice that under Assumption A6 this minimum is unique.

Let us now study the asymptotic behaviour of the estimator of $T_2(f_0)$. Let $\psi(\lambda) = \log f(\lambda, \theta)$. Then under Assumptions A1, A2, B1 and B2,

$$\sqrt{n}\left(\frac{4\pi}{n} \sum_{j=1}^{[n/2]} \psi(\lambda_j) \log f(\lambda_j, \theta_0) - \int_{-\pi}^{\pi} \psi(\lambda) \log f(\lambda, \theta_0) \, d\lambda \right) \to 0. \tag{22}$$

Also,

$$\sqrt{n}\left(\int_{-\pi}^{\pi} \psi(\lambda) g_n(\lambda) \, d\lambda - \frac{4\pi}{n} \sum_{j=1}^{[n/2]} \psi(\lambda_j) \log I_n(\lambda_j) \right) \to 0 \tag{23}$$

in θ_0-probability.

Indeed, choose ν as in Theorem 2. In order to verify (23), notice that

$$\sqrt{n}\left(\int_{-\pi}^{\pi} \psi(\lambda) g_n(\lambda) - \frac{4\pi}{n} \sum_{j=1}^{[n/2]} \psi(\lambda_j) \log I_n(\lambda_j) \right)$$
$$= \frac{4\pi}{n^{1/2}} \sum_{j > n^\nu} \psi'(\tilde{\lambda}_j)(\lambda_j - \lambda_j^1) \log I_n(\lambda_j) + \frac{4\pi}{n^{1/2}} \sum_{j \leqslant n^\nu} [\psi(\lambda_{j_1}) - \psi(\lambda_j)] \log I_n(\lambda_j),$$

where $\lambda_j \leqslant \lambda_j^1 < \lambda_j^1 < \lambda_{j+1}$ and $\tilde{\lambda}_j \in [\lambda_j, \lambda_j^1]$. To see that this last sum tends to zero in θ_0-probability, we calculate its expectation and variance as in the proof of Theorem 2 and use the fact that $\mathrm{E}[\log(I_n(\lambda_j))]$ and $\mathrm{var}[\log(I_n(\lambda_j))]$ are bounded for each λ_j by a slowly varying function at most. Then we use the fact that $\psi(\lambda_j)$ satisfies Assumptions B1 and B2 to complete the proof.

Let $\mathscr{H}$ be a class of functions that satisfy Assumptions B1 and B2. Then, from Theorem 2 and equations (22), (23), we have that $\kappa g_n(\cdot) \xrightarrow{l.\mathscr{H}} f(\cdot, \theta_0)$ in θ_0-probability. We now require the following lemma, which controls the fluctuations of $h_n(\theta)$ and its derivatives in probability. As $h^{(1)}$, $h_n^{(1)}$ are q-vectors and $h^{(2)}$, $h_n^{(2)}$ $q \times q$ matrices, the next lemma is understood to apply componentwise.

Lemma 6. *Suppose Assumptions A1–A6 are satisfied. Then:*

1. *for all fixed* $\theta \in \Theta$, $h_n^{(k)}(\theta) - H^{(K)}(\theta) \to 0$ *in probability componentwise for* $k = 0, 1, 2$;
2. *for all* $B > 0$, *componentwise*

$$\lim_{\eta \to 0} \sup P\left(\sup_{|\theta_1 - \theta_2| > \eta} |h_n^{(k)}(\theta_1) - h_n^{(k)}(\theta_2)| > B \right) \to 0$$

for $k = 0, 1, 2$.

The proof of Lemma 6 is given in the Appendix.

To show consistency of $T_2(\kappa g_n)$, assume $T_2(f_0)$ is unique and lies in the interior of Θ. We want to prove that $P(|T_2(\kappa g_n) - T_2(f_0)| > \varepsilon) \to 0$. As for each n, $T_2(\kappa g_n)$ may not be unique, consider instead $\tilde{T}_2(\kappa g_n) = \inf_\Theta T_2(\kappa g_n)$. In what follows, we shall assume $f_0(\lambda) = f(\lambda, \theta_0)$ and write $T_2(f_0) = \theta_0$. Notice that as θ_0 is the unique value that minimizes $h(\cdot)$ and because $h(\theta_0) = 0$, we have $h(\theta) = (\theta - \theta_0)^{\mathrm{t}} h^{(2)}(\tilde{\theta})(\theta - \theta_0)$. We have also that $h^{(2)}(\theta_0)$ is positive definite so that, under Assumption A5, $h^{(2)}(\tilde{\theta})$ is positive definite over a certain subset of Θ, which does not depend on θ. Thus, there exists a certain $\delta > 0$ such that we have $\inf_{|\theta - \theta_0| > \varepsilon} h(\theta) \geqslant \inf(\delta, \varepsilon^2 \| h^{(2)}(\theta_0)\|/2)$, with $\|\cdot\|$ the L^2 matrix norm. Now, assume $\varepsilon > 0$ is small enough. There exists a constant $B > 0$ such that

$$\begin{aligned}
&P(|\tilde{T}_2(\kappa g_n) - \theta_0| > \varepsilon) \\
&\quad = P\left(\inf_{|\theta - \theta_0| > \varepsilon} h_n(\theta) - h_n(\theta_0) < 0 \right) \\
&\quad \leqslant P\left(\sup_{|\theta - \theta_0| > \varepsilon} (h_n(\theta) - h(\theta) - h_n(\theta_0)) > \inf_{|\theta - \theta_0| > \varepsilon} h(\theta) \right) \\
&\quad \leqslant \sum_{j=1}^{q} P\left(|(h_n^{(1)}(\theta_0))_j| > \frac{B\varepsilon}{2q} \right) + \sum_{j,k=1}^{q} P\left(\sup_\theta |(h_n^{(2)}(\theta) - h^{(2)}(\theta))_{j,k}| > \frac{B}{2q^2} \right).
\end{aligned}$$

Consistency of $\tilde{T}_2(\kappa g_n)$ now follows from Lemma 6 and the continuity (thus uniform) of each component of $h^{(2)}(\cdot)$. In order to show consistency of the estimator, consider $\overline{T}_2(\kappa g_n) = \sup_\Theta T_2(\kappa g_n)$, and repeat the above arguments.

Once we have shown consistency, we can prove convergence in distribution. As in the proof of Theorem 5 of Taniguchi (1979), we can write, in matrix notation,

$$(T_2(\kappa g_n) - \theta_0)\left\{\int_{-\pi}^{\pi} -2\frac{\partial^2 \log f(\lambda, \theta_j)}{\partial\theta\partial\theta_k}\bigg|_{\theta=\theta_0} \log f(\lambda, \theta_0) + \frac{\partial^2 \log^2 f(\lambda, \theta)}{\partial\theta_j\partial\theta_k}\bigg|_{\theta=\theta_0} \mathrm{d}\lambda + A + B\right\}$$

$$= 2\left(\int_{-\pi}^{\pi} \frac{\partial f(\lambda, \theta)}{\partial\theta_j}\bigg|_{\theta=\theta_0} (\log \kappa g_n(\lambda) - \log f(\lambda, \theta_0))\,\mathrm{d}\lambda\right),$$

with

$$A = \left\{\int_{-\pi}^{\pi} -2\frac{\partial^2 \log f(\lambda, \theta)}{\partial\theta_j\partial\theta_k}\bigg|_{\theta=\mu_1} \log \kappa g_n(\lambda)\,\mathrm{d}\lambda + \int_{-\pi}^{\pi} 2\frac{\partial^2 \log f(\lambda, \theta)}{\partial\theta_j\partial\theta_k}\bigg|_{\theta=\theta_0} \log f(\lambda, \theta_0)\,\mathrm{d}\lambda\right\}$$

and

$$B = \left\{\int_{-\pi}^{\pi} \frac{\partial^2 \log^2 f(\lambda, \theta)}{\partial\theta_j\partial\theta_k}\bigg|_{\theta=\mu_2} \mathrm{d}\lambda - \int_{-\pi}^{\pi} \frac{\partial^2 \log^2 f(\lambda, \theta)}{\partial\theta_j\partial\theta_k}\bigg|_{\theta=\theta_0} \mathrm{d}\lambda\right\}.$$

Because $T_2^{\phi}(\kappa g_n) - \theta_0 \xrightarrow{P} 0$ in θ_0-probability, under Assumption A4, we have, by the dominated convergence theorem, $B \xrightarrow{P} 0$ in θ_0-probability (that is, all entries of the matrix B).

Under Assumption A5, we have that $\log \kappa g_n(\cdot)$ is bounded in probability and, again using the dominated converge theorem, we have that A tends to zero in θ_0-probability (that is, all entries of the matrix A).

Thus, we have

$$\sqrt{n}(T_2^{\phi}(\kappa g_n) - \theta_0) = \sqrt{n}\int_{-\pi}^{\pi} \sigma_f(\lambda)(\log \kappa g_n(\lambda) - \log f(\lambda, \theta_0))\,\mathrm{d}\lambda$$
$$+\ b_n\int_{-\pi}^{\pi} -\frac{\partial \log f(\lambda, \theta)}{\partial\theta_j}\bigg|_{\theta_0} \sqrt{n}(\log \kappa g_n(\lambda) - \log f(\lambda, \theta_0))\,\mathrm{d}\lambda,$$

where $b_n \to 0$ in θ_0-probability. Finally, Theorem 2 concludes the proof. □

Appendix

Proof of Lemma 1. Put $\lambda_j = 2\pi j/n$, $1 \leqslant j \leqslant [n/2]$. Let $F_n(\lambda) = (1/n)|\sum_{j=1}^{n} \mathrm{e}^{\mathrm{i}j\lambda}|^2$ be the Féjer kernel of degree n.

Define, for $\lambda \in (-\pi, \pi]$,

$$a_n(\lambda) = \frac{1}{4\pi}\sum_{|k|\leqslant n}\left(1 - \frac{|k|}{n}\right) r_k \cos\lambda k$$

$$= \frac{1}{4\pi}\int_{-\pi}^{\pi} F_n(\mu) f(\lambda - \mu)\,\mathrm{d}\mu,$$

$$b_n(\lambda) = \frac{1}{8\pi n}\left(\frac{\cos\lambda}{\sin\lambda}\sin 2n\lambda + \cos 2n\lambda - 1\right)\sum_{k=0}^{n-1} r_k \cos\lambda k$$

$$+ \left(\frac{\cos\lambda}{\sin\lambda}(\cos 2n\lambda + 1) + \sin 2n\lambda\right)\sum_{k=0}^{n-1} r_k \sin\lambda k,$$

$$c_n(\lambda) = -\left[\frac{1}{8\pi \mathrm{n}}\left(\frac{\cos\lambda}{\sin\lambda}\sin 2n\lambda + \cos 2\eta\lambda + 1\right)\sum_{k=0}^{n-1} r_k \sin\lambda k\right.$$

$$\left. - \left(\frac{\cos\lambda}{\sin\lambda}(\cos 2n\lambda - 1) + \sin 2n\lambda\right)\sum_{k=0}^{n-1} r_k \cos\lambda k\right].$$

Notice that

$$c_n(\lambda_j) = \frac{1}{4\pi n}\sum_{k=0}^{n-1} r_k \sin\lambda_j k,$$

and $b_n(\lambda_j) = \cot(\lambda_j)c_n(\lambda_j)$. On the other hand, it can be shown after some calculation, recalling the definition of $w(\lambda)$ given in (2), that

$$\mathrm{E}w(\lambda_j)\overline{w}(\lambda_j) = \mathrm{E}(X_n^2(\lambda_j) + Y_n^2(\lambda_j)) = 2a_n(\lambda_j), \tag{24}$$

$$\begin{aligned}\mathrm{E}w(\lambda_j)w(\lambda_j) &= \mathrm{E}(X_n^2(\lambda_j) - Y_n^2(\lambda_j)) + 2\mathrm{i}\mathrm{E}(X_n(\lambda_j)Y_n(\lambda_j)) \\ &= 2b_n(\lambda_j) + 2\mathrm{i}c_n(\lambda_j).\end{aligned} \tag{25}$$

Equations (24) and (25) yield

$$\mathrm{E}(X_n(\lambda_j))^2 = a_n(\lambda_j) + b_n(\lambda_j),$$

$$\mathrm{E}(X_n(\lambda_j))^2 = a_n(\lambda_j) - b_n(\lambda_j),$$

$$\mathrm{cov}(X_n(\lambda_j),\, Y_n(\lambda_j)) = c_n(\lambda_j).$$

Assumption A1 and the fact that F_n is positive and integrates to 2π yield $a_n(\lambda_j) \geqslant D/2$. Finally, for completeness we include the proof of equation (9) (cf. Robinson 1995), throughout which $f(\lambda)$ stands for the spectral density:

$$\mathrm{E}w(\lambda_j)\overline{w}(\lambda_j) - f(\lambda_j) = \frac{1}{2\pi}\int_{-\pi}^{\pi} F_n(\lambda - \lambda_j)(f(\lambda) - f(\lambda_j))\,\mathrm{d}\lambda$$

$$= \frac{1}{2\pi}\int_{-\pi}^{-\lambda_j/2} + \int_{-\lambda_j/2}^{\lambda_j/2} + \int_{\lambda_j/2}^{3\lambda_j/2} + \int_{3\lambda_j/2}^{\pi} F_n(\lambda - \lambda_j)(f(\lambda) - f(\lambda_j))\,\mathrm{d}\lambda$$

$$\leqslant \max_{\lambda > \lambda_j/2} f(\lambda_j)\frac{1}{2\pi}\int_{-\pi}^{-\lambda_j/2} + \int_{\lambda_j/2}^{\pi} F_n(\lambda - \lambda_j)\,\mathrm{d}\lambda + \max_{\lambda > \lambda_j/2} F_n(\lambda - \lambda_j)\frac{1}{2\pi}\int_{-\lambda_j/2}^{\lambda_j/2}(f(\lambda) + f(\lambda_j))\,\mathrm{d}\lambda$$

$$+ \max_{\lambda > \lambda_j/2} |f'(\lambda)|\frac{1}{2\pi}\int_{\lambda_j/2}^{3\lambda_j/2} F_n(\lambda - \lambda_j)|\lambda - \lambda_j|\,\mathrm{d}\lambda.$$

The properties of the Féjer kernel yield the desired result under Assumptions A1 and A2. □

Proof of lemma 3. First, we will prove the inequality in (12). Robinson (1995) shows in the proof of Robinson's (1995) Theorem 2, that there exists a constant $C > 0$ such that, for $2 \leqslant j \leqslant [n/2] - 1$, $\mathrm{E}w(\lambda_j)w(\lambda_j) \leqslant Cf_0(\lambda_j)\log j/j$. This, together with equation (9), yields the desired result.

On the other hand, Hurvich and Beltrao (1993) have shown that, for fixed j,

$$\lim_{n\to\infty} b_n(\lambda_j)f_0(\lambda_j) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\sin^2(\lambda/2)}{(2\pi j - \lambda)(2\pi j + \lambda)}\left|\frac{\lambda}{2\pi j}\right|^{\alpha(\theta_0)-1}, \tag{26}$$

$$\lim_{n\to\infty} a_n(\lambda_j)f_0(\lambda_j) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\sin^2(\lambda/2)}{(2\pi j - \lambda)^2}\left|\frac{\lambda}{2\pi j}\right|^{\alpha(\theta_0)-1}. \tag{27}$$

By (12), there exists a j_0 such that, for all $j \geqslant j_0$, $\varepsilon_n(\lambda_j) \leqslant \frac{1}{4}$. Also, from (26) and (27) it follows that there exists a $c > 0$ such that, for each fixed j, there exists an n_j such that $b_n(\lambda_j)/a_n(\lambda_j) < 1 - c$. Now, choose $n_0 = \max_{1\leqslant j\leqslant j_0} n_j$. This yields (11). □

Proof of Lemma 4. Given $p \geqslant 1$ we choose n such that $c_n(\lambda_j)/(a_n(\lambda_j)(1 - \varepsilon_{n,j}^2)^{1/2}) < 1/(2p - 1)$. Then, by Proposition 3.1(ii) in Taqqu (1977), term-by-term integration is allowed and, using the formula for the expectation of products of Hermite polynomials (see, for example, Taqqu 1977)

$$\mathrm{E}(\log(\tilde{I}_n(\lambda_j)) - (\log 2 - \gamma))^p = c_{0,0}^{(p)} + \sum_{q=2}^{\infty} \sum_{\substack{k_1+k_2=2q,\\ 0\leqslant k_i\leqslant q,\\ k_i \text{ even}}} \frac{1}{k_1!k_2!} c_{k_1,k_2}^{(p)} \mathrm{E} H_{k_1}(\tilde{X}_n(\lambda_j)) H_{k_2}(\tilde{Y}_n(\lambda_j))$$

$$= c_{0,0}^{(p)} + \sum_{q=2}^{\infty} \frac{c_{q,q}^{(p)}}{q!} \left(\frac{c_n(\lambda_j)}{a_n(\lambda_j)\sqrt{1-\varepsilon_{n,j}^2}} \right)^q .$$

As a consequence of inequality (8) the above series converges as $c_n(\lambda_j)/(a_n(\lambda_j)(1-\varepsilon_{n,j}^2)^{1/2}) < 1/(2p-1) \leqslant 1$ and it is bounded by $C(c_n(\lambda_j)/a_n(\lambda_j))^2$, for a certain positive constant C. □

Proof of Lemma 5. Given ℓ, choose n such that $\log n/n^\nu < 1/(2\ell-1)$. Then by Proposition 3.1(ii) in Taqqu (1977), term-by-term integration is allowed and

$$\mathrm{E}\prod_{i=1}^{\ell}(\log(\tilde{I}_n(\lambda_{j_i})) - (\log 2 - \gamma))^{a_i} = \sum_{q=0}^{\infty} \sum_{\substack{k_1+\cdots+k_{2\ell}=2q\\ 0\leqslant k_i\leqslant q,\\ k_i \text{ even}}} \frac{\prod_{i=1}^{\ell} c_{k_{2i-1},k_{2i}}^{(a_i)}}{k_1! \ldots k_{2\ell}!} \mathrm{E}\prod_{i=1}^{\ell} H_{k_{2i-1}}(\tilde{X}_n(\lambda_{j_i})) H_{k_{2i}}(\tilde{Y}_n(\lambda_{j_i}))$$

$$= \prod_{i=1}^{\ell} c_{0,0}^{(a_i)} + \sum_{q\geqslant 2} \sum_{\substack{k_1+\cdots+k_{2\ell}=2q\\ 0\leqslant k_i\leqslant q,\\ k_i \text{ even}}} \prod_{i=1}^{\ell} c_{k_{2i-1},k_{2i}}^{(a_i)} A_{k_1,\ldots,k_{2\ell}}(\lambda_{j_1}, \ldots, \lambda_{j_\ell}),$$

where, if $j = \min(j_1, \ldots, j_\ell)$,

$$\left| \prod_{i=1}^{\ell} c_{k_{2i-1},k_{2i}}^{(a_i)} A_{k_1,\ldots,k_{2\ell}}(\lambda_{j_1}, \ldots, \lambda_{j_\ell}) \right| \leqslant (2\ell-1)^q \left(\frac{\log n}{j}\right)^q \left(\prod_{i=1}^{\ell} c_{0,0}^{(2a_i)} \right)^{1/2}$$

because of properties of the expectation of products of Hermite polynomials (see, for example, Taqqu 1977), Lemma 2 and inequality (8). On the other hand, assume for simplicity that $a_1 = \ldots = a_s = 1$. As $\sum_i k_i = 2q$ and all the k_i are even, if $q < s$ it is not possible for each pair (k_{2i-1}, k_{2i}), $i = 1, \ldots, s$, to have at least one non-zero term. As $c_{0,0}^{(1)} = 0$ we have that if the number of a_i which are equal to one is s, then

$$\mathrm{E}\prod_{i=1}^{\ell}(\log(\tilde{I}_n(\lambda_{j_i})) - (\log 2 - \gamma))^{a_i} = \sum_{q\geqslant 2\vee s} \sum_{\substack{k_1+\cdots+k_{2\geqslant}=2q\\ 0\leqslant k_i\leqslant q, k_i \text{ even}}} \prod_{i=1}^{\ell} c_{k_{2i-1},k_{2i}}^{(a_i)} A_{k_1,\ldots,k_{2\ell}}(\lambda_{j}1, \ldots, \lambda_{j_\ell})$$

$$\leqslant \left(\prod_{i=1}^{\ell} c_{0,0}^{2a_i)} \right)^{1/2} \sum_{q\geqslant 2\vee s} \binom{2q+2\ell-1}{2q} (2\ell-1)^q \left(\frac{\log n}{j}\right)^q .$$

This last series converges as $\log n/n^\nu < 1/(2\ell-1)$ and thus the lemma is proved. □

Proof of Lemma 6. The first part of this lemma follows directly from Theorem 2. In order to verify the convergence to zero in the second formula, observe, for $k = 0$,

$$\begin{aligned}|h_n(\theta_1) - h_n(\theta_2)| &\leqslant \left| 2\sum_{j=1}^{[n/2]} \log(\kappa I_n(\lambda_j)) \int_{\lambda_{j-1}}^{\lambda_j} (\log f(\lambda, \theta_1) - \log f(\lambda, \theta_2))\,\mathrm{d}\lambda \right| \\ &\quad + \left| \int_{-\pi}^{\pi} (\log^2 f(\lambda, \theta_1) - \log^2 f(\lambda, \theta_2))\,\mathrm{d}\lambda \right| \\ &\leqslant C\eta(\theta_1, \theta_2)\left(\left| \frac{4\pi}{n}\sum_{j=1}^{[n/2]} \log(\kappa I_n(\lambda_j)) \right| + 1 \right).\end{aligned}$$

Here $\eta(\theta_1, \theta_2) \to 0$ if $|\theta_1 - \theta_2| \to 0$. The result follows because the second moment of $\log(\kappa I_n(\lambda_j))$ is uniformly bounded. For $k = 1, 2$ the proof follows analogously. □

Acknowledgements

The author is indebted to the referees, whose helpful comments greatly improved the presentation of this paper. The author especially appreciates the suggestions of one of the referees who helped correct a previous flaw in the proof of Theorem 2.

References

Dahlhaus, R. (1989) Efficient parameter estimation for self-similar processes. *Ann. Statist.*, **17**, 1749–1766.

Dzhaparidze, K. (1985) *Parameter Estimation and Hypothesis Testing in Spectral Analysis of Stationary Time Series*. Berlin: Springer-Verlag.

Geweke, J. and Porter-Hudak, S. (1990) The estimation and application of long memory time series models. *J. Time Series Anal.*, **4**, 221–238.

Giraitis, L. and Surgailis, D. (1986) Multivariate Appell polynomials and the central limit theorem. In E. Eberlein and M.S. Taqqu (eds), *Dependence in Probability and Statistics*. Boston: Birkhäuser.

Giraitis, L. and Surgailis, D. (1990) A central limit theorem for quadratic forms in strongly dependent linear variables and its application to asymptotic normality of Whittle's estimate. *Probab. Theory Related Fields*, **86**, 87–104.

Fox, R. and Taqqu, M.S. (1986) Large sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *Ann. Statist.*, **14**, 517–532.

Fox, R. and Taqqu, M.S. (1987) Central limit theorems for quadratic forms in random variables having long range dependence. *Probab. Theory Related Fields*, **74**, 213–240.

Künsch, H.R. (1986) Discrimination between monotonic trends and long range dependence *J. Appl. Probab.*, **23**, 1025–1030.

Hurvich, C.M. and Beltrao, K.I. (1993) Asymptotics for the low-frequency ordinates of the periodogram of a long-memory time series. *J. Time Series Anal.*, **14**, 455–472.

Robinson, P.M. (1995) Log-periodogram regression of time series with long range dependence. *Ann. Statist.*, **23**, 1048–1072.

Taqqu, M.S. (1977) Law of the iterated logarithm for sums of non-linear functions of Gaussian variables that exhibit a long range dependence. *Z. Wahrscheinlichkeitstheorie Verw. Geb.*, **40**, 203–238.

Taniguchi, M. (1979) On estimation of parameters of Gaussian stationary processes. *J. Appl. Probab.*, **16**, 575–591.

Received March 1996 and revised October 1999