EFFICIENT ESTIMATION OF SEASONAL LONG-RANGE-DEPENDENT PROCESSES

WILFREDO PALMA AND NGAI HANG CHAN

Pontificia Universidad Católica de Chile and Chinese University of Hong Kong

First Version received November 2003

Abstract. This paper studies asymptotic properties of the exact maximum likelihood estimates (MLE) for a general class of Gaussian seasonal long-range-dependent processes. This class includes the commonly used Gegenbauer and seasonal autoregressive fractionally integrated moving average processes. By means of an approximation of the spectral density, the exact MLE of this class are shown to be consistent, asymptotically normal and efficient. Finite sample performance of these estimates is examined by Monte Carlo simulations and it is shown that the estimates behave very well even for moderate sample sizes. The estimation methodology is illustrated by a real-life Internet traffic example.

Keywords. Consistency; efficiency; long-range dependency; maximum likelihood estimates (MLE); cyclical.

1. INTRODUCTION

Several statistical methodologies have been developed to model time series exhibiting both cyclical and long-memory properties. Abrahams and Dempster (1979) and Jonas (1979) generalize Madelbrot's fractional Gaussian noise process to allow for an infinite spectrum at seasonal frequencies, Gray *et al.* (1989) propose the generalized fractional or Gegenbauer [generalized autoregressive moving average (GARMA)] processes, Porter-Hudak (1990) discusses seasonal fractionally integrated autoregressive moving average (SARFIMA) models, Hassler (1994) introduces the flexible seasonal fractionally integrated processes (flexible ARFISMA) and Woodward *et al.* (1998) introduce the k-GARMA processes.

Estimation methods and statistical properties of non-seasonal long-memory processes have been investigated by Yajima (1985) and Dahlhaus (1989), among others. On the contrary, seasonal long-range-dependent models have been studied by Giraitis and Leipus (1995), Chung (1996), Arteche and Robinson (2000), Velasco and Robinson (2000), Giraitis *et al.* (2001), and Ould Haye (2002) among others. Long-range-dependent data with seasonal behaviour have been reported in diverse fields. For example, inflation rates are studied by Hassler and Wolters (1995), revenue series are analysed by Ray (1993), monetary aggregates are considered by Porter-Hudak (1990), quarterly gross national product and

0143-9782/05/06 863–892 JOURNAL OF TIME SERIES ANALYSIS Vol. 26, No. 6

shipping data are discussed by Ooms (1995), and monthly flows of the Nile River are studied by Montanari *et al.* (2000).

This paper studies the asymptotic properties of the maximum likelihood estimate (MLE) of a general class of Gaussian seasonal long-memory processes with spectral densities specified by

$$f(\omega) = H(\omega)|\omega|^{-\alpha} \prod_{i=1}^{r} \prod_{j=1}^{m_i} |\omega - \omega_{ij}|^{-\alpha_i}, \tag{1}$$

where $\omega \in (-\pi, \pi]$, $0 \leq \alpha, \alpha_i < 1$, $i = 1, \dots, r$, $H(\omega)$ is a symmetric, strictly positive, continuous and bounded function and $\omega_{ij} \neq 0$ are known poles for $j = 1, \dots, m_i$, $i = 1, \dots, r$. To ensure the symmetry of f, we assume that for any $i = 1, \dots, r$, $j = 1, \dots, m_i$, there is one and only one $1 \leq j' \leq m_i$ such that $\omega_{ij} = -\omega_{ij'}$. The spectral density of many widely used models such as SARFIMA and k-factor GARMA satisfy specification (1). As a first example, consider a seasonal ARFIMA model with multiple periods $s_1, \dots, s_r$:

$$\phi(B)(1-B)^d \prod_{i=1}^{r} \Phi(B^{s_i})(1-B^{s_i})^{d_{s_i}} x_t = \theta(B) \prod_{i=1}^{r} \Theta(B^{s_i})\varepsilon_t, \tag{2}$$

where B is the backshift operator, and $\phi(B)$, $\Phi(B^{s_i})$, $\theta(B)$ and $\Theta(B^{s_i})$ are autoregressive and moving average polynomials. Observe that the spectral density of process (2) may be written as

$$f_{s_1,\dots,s_r}(\omega) = H(\omega)|\omega|^{-2d-2d_{s_1}-\cdots-2d_{s_r}} \prod_{i=1}^{r} \prod_{j=1}^{s_i} |\omega - \omega_{ij}|^{-2d_{s_i}}, \tag{3}$$

which is a special case of eqn (1). The corresponding parameters are

$$\omega_{ij} = 2\pi j/s_i, \quad \text{for } i = 1, \dots, r, j = 1, \dots, \left[\frac{s_i}{2}\right],$$
$$\omega_{ij} = \frac{2\pi([\frac{s_i}{2}] - j)}{s_i}, \quad \text{for } i = 1, \dots, r, j = \left[\frac{s_i}{2}\right] + 1, \dots, s_i,$$
$$\alpha = 2d + 2d_{s_1} + \cdots + 2d_{s_r} \text{ and } \alpha_i = 2d_{s_i}.$$

As a second example, note that the spectral density of a k-GARMA process is given by, Woodward *et al.* (1998),

$$f(\omega) = C|\theta(e^{i\omega})|^2 |\phi(e^{i\omega})|^{-2} \prod_{j=1}^{k} [\cos(\omega) - u_j]^{-2\lambda_j}, \tag{4}$$

where $C > 0$ is a constant, u_j are distinct values, $\lambda_j \in (0, \frac{1}{4})$ when $|u_j| = 1$ and $\lambda_j \in (0, \frac{1}{2})$ when $|u_j| \neq 1$. For $|u_j| \leq 1$, we may write $u_j = \cos(\omega_j)$ and this spectral density may be written in terms of eqn (1) as follows:

$$f(\omega) = H(\omega) \prod_{j=1}^{k} |\omega - \omega_j|^{-2\lambda_j} |\omega + \omega_j|^{-2\lambda_j},$$

where

$$H(\omega) = C|\theta(\mathrm{e}^{\mathrm{i}\omega})|^2 |\phi(\mathrm{e}^{\mathrm{i}\omega})|^{-2} \prod_{j=1}^{k} \left| \frac{\cos(\omega) - \cos(\omega_j)}{\omega^2 - \omega_j^2} \right|^{-2\lambda_j}$$

is a strictly positive, symmetric, continuous function with

$$\lim_{\omega \to \pm\omega_l} H(\omega) = C|\theta(\mathrm{e}^{\mathrm{i}\omega_l})|^2 |\phi(\mathrm{e}^{\mathrm{i}\omega_l})|^{-2} \prod_{j \neq l}^{k} \left| \frac{\cos(\omega_l) - \cos(\omega_j)}{\omega_l^2 - \omega_j^2} \right|^{-2\lambda_j} \left| \frac{\sin(\omega_l)}{2\omega_l} \right|^{-2\lambda},$$

for $\omega_l \neq 0$ and

$$\lim_{\omega \to \pm\omega_l} H(\omega) = 4^{\lambda_l} C|\theta(\mathrm{e}^{\mathrm{i}\omega_l})|^2 |\phi(\mathrm{e}^{\mathrm{i}\omega_l})|^{-2} \prod_{j \neq l}^{k} \left| \frac{\cos(\omega_l) - \cos(\omega_j)}{\omega_l^2 - \omega_j^2} \right|^{-2\lambda_j},$$

for $\omega_l = 0$. Observe that all these limits are finite and $H(\omega)$ is a bounded function.

In this paper, we prove that the exact MLE for a Gaussian process with spectral density satisfying eqn (1) is consistent, asymptotically normal and efficient in the sense of Fisher, i.e. its variance attains the Cramér–Rao lower bound. The method for obtaining these results is based on an approximation of the spectral density proposed by Hannan (1973). Details of the proof are discussed in the beginning of Section 2. Note that expression (1) is fairly general and encompasses a wide variety of seasonal long-memory models (see, e.g. Leipus and Viano, 2000). Because of the poles other than zero in eqn (1), results of Dahlhaus (1989) for the exact MLE of non-seasonal long-range-dependent processes are not directly applicable. In particular, Dahlhaus' conditions (A2), and (A4) to (A9) do not hold in the presence of multiple singularities. As a result, one may have to reprove all the technical lemmas used by Fox and Taqqu (1987) and Dahlhaus (1989) to deal with the seasonal case (1). On the contrary, using the idea of Hannan (1973), the spectral density (1) can be approximated by a well-behaved function around the seasonal poles. Moreover, this idea may prove to be useful for establishing limiting properties of estimates for other general classes of time-series models.

This paper proceeds as follows. In Section 2, the consistency, a central limit theorem and the asymptotic efficiency of MLE of a model satisfying (1) are derived. Finite sample performance of MLE is investigated in Section 3, where the results from several Monte Carlo studies are exhibited. An Internet traffic data example that illustrates the proposed methodology is discussed in Section 4 and concluding remarks are presented in Section 5. Proofs of the technical lemmas needed to prove Theorems 1–3 are given in the Appendix.

2. MAXIMUM LIKELIHOOD ESTIMATION

Large-sample properties of MLE of Gaussian seasonal long-range-dependent processes are investigated in this section. Based on an argument used by Hannan (1973), the spectral density (1) with multiple singularities is approximated by the following density having a pole only at zero frequency:

$$f_\eta(\omega) = H(\omega)|\omega|^{-\alpha} \prod_{i=1}^{r} [(1-\eta)^2 + \eta \prod_{j=1}^{m_i} |\omega - \omega_{ij}|]^{-\alpha_i}. \quad (5)$$

Observe that this spectral density corresponds to eqn (1) when $\eta = 1$. For $0 < \eta < 1$, eqn (5) does not have poles at frequencies other than zero, and as a result, the MLE for such a process is consistent, asymptotically normal and efficient. Using this observation together with the fact that f_η converges to f as η tends to 1 in the sense specified by Lemma 1, we show that the MLE based on the spectral density (1) is asymptotically normal and efficient. Specifically, Lemmas 1 and 2 deal with the convergence of the approximated density and related results for the covariance matrices. Lemma 3 shows an inequality that is used to establish Lemma 6. Lemmas 4 and 5 deal with the convergence of matrix product traces, whereas Lemmas 6 and 7 are used to establish the limiting behaviour of variance–covariance matrices as the sample size increases. The consistency of the MLE is proved in Theorem 1, asymptotic normality is derived in Theorem 2, and the efficiency of the MLE is established in Theorem 3.

Herein, we assume that the spectral density depends on the parameter $\boldsymbol{\theta}$, denoted by $f_{\boldsymbol{\theta}}$, and that the parameter space $\Theta \subseteq \mathbb{R}^J$ is compact. The symbol K denotes a generic positive constant, the actual value of which may vary from step to step. Furthermore, to shorten the proofs in this paper, we made use of some of the technical arguments given by Dahlhaus (1989) that are unaffected by the failure of Assumptions (A2) and (A4)–(A9).

2.1. *Consistency*

This section focuses on proving the convergence of the MLE, $\hat{\boldsymbol{\theta}}_n$, to the true parameter $\boldsymbol{\theta}_0$ (see Theorem 1). In what follows, $\|\cdot\|$ and $|\cdot|$ denote the spectral and the Euclidean norm, respectively, i.e.

$$\|A\| = \sup_{\|x\|=1} \sqrt{x'A'Ax} \quad \text{and} \quad |A| = \sqrt{\mathrm{tr}(AA')}.$$

In addition, if A is a vector of matrices $A = [A_1, \ldots, A_p]'$ then $|||A|||$ denotes $\sqrt{\sum_{j=1}^{p} \|A_j\|^2}$.

Let

$$[T(f_{\boldsymbol\theta})]_{r,s=1,\ldots,n} = \int_{-\pi}^{\pi} f_{\boldsymbol\theta}(\omega)e^{i\omega(r-s)}d\omega$$

be the covariance matrix of the observation series $\mathbf{X} = (x_1, \ldots, x_n)'$. Unless otherwise stated, we write $T(f_{\boldsymbol\theta})$ as T and $T(f_{\eta,\boldsymbol\theta})$ as T_η herein. Let $\mathcal{L}_n(\boldsymbol\theta)$ be $-1/n$ times the loglikelihood function based on the spectral density f (omitting a constant):

$$\mathcal{L}_n(\boldsymbol\theta) = \frac{1}{2n}\log\det T(f_{\boldsymbol\theta}) + \frac{1}{2n}\mathbf{X}'T(f_{\boldsymbol\theta})^{-1}\mathbf{X}.$$

With the previous notations and Lemmas 1 and 2 in the Appendix we are now ready to derive the consistency of the MLE.

THEOREM 1. (*Consistency*) $\hat{\boldsymbol\theta}_n \to_p \boldsymbol\theta_0$ *as* $n \to \infty$.

PROOF. Following Basawa *et al.* (1976) and Ling and Li (1997), it suffices to prove that: (i) $\nabla\mathcal{L}_n(\boldsymbol\theta_0) \to_p 0$, as $n \to \infty$; (ii) there exists a positive definite matrix $M(\boldsymbol\theta_0)$ such that for all $\epsilon > 0$,

$$P(\nabla^2\mathcal{L}_n(\boldsymbol\theta_0) > M(\boldsymbol\theta_0)) > 1 - \epsilon, \quad \text{for all } n > n(\epsilon);$$

and (iii) there exists a constant $0 < M < \infty$ such that

$$E|\nabla^3\mathcal{L}_n(\boldsymbol\theta)| < M \quad \text{for all } \boldsymbol\theta \in \boldsymbol\Theta.$$

We prove that conditions (i)–(iii) hold in this case. (i)

$$E[\nabla\mathcal{L}_n(\boldsymbol\theta_0)] = \frac{1}{2n}\mathrm{tr}\{T(f_{\boldsymbol\theta_0})^{-1}T(\nabla f_{\boldsymbol\theta_0})\} - \frac{1}{2n}\mathrm{tr}\left\{T(f_{\boldsymbol\theta_0})A^{(3)}_{\boldsymbol\theta_0}\right\},$$

where $A^{(3)}_{\boldsymbol\theta_0}$ is defined as in Lemma 6. But,

$$T(f_{\boldsymbol\theta_0})A^{(3)}_{\boldsymbol\theta_0} = T(\nabla f_{\boldsymbol\theta_0})T(f_{\boldsymbol\theta_0)})^{-1}.$$

Hence $E\,[\nabla\mathcal{L}_n(\boldsymbol\theta_0)] = 0$. On the other hand,

$$\mathrm{var}[\nabla\mathcal{L}_n(\boldsymbol\theta_0)] = \frac{1}{2n^2}\mathrm{tr}\{T(f_{\boldsymbol\theta_0})^{-1}T(\nabla f_{\boldsymbol\theta_0})T(f_{\boldsymbol\theta_0})^{-1}T(\nabla f_{\boldsymbol\theta_0})\}.$$

But, by Lemma 5 we have

$$\frac{1}{n}\mathrm{tr}\{T(f_{\boldsymbol\theta_0)})^{-1}T(\nabla f_{\boldsymbol\theta_0})T(f_{\boldsymbol\theta_0)})^{-1}T(\nabla f_{\boldsymbol\theta_0})\} \to \frac{1}{2\pi}\int_{-\pi}^{\pi}[\nabla\log f_{\boldsymbol\theta_0}(\omega)][\nabla\log f_{\boldsymbol\theta_0}(\omega)]'d\omega,$$

as $n \to \infty$. Therefore, $\mathrm{var}[\nabla\mathcal{L}_n(\boldsymbol\theta_0)] \to 0$ and the result holds by virtue of the Chebyshev's inequality. (ii) Observe that

$$E[\nabla^2 \mathcal{L}_n(\boldsymbol{\theta}_0)] = \frac{1}{2n} \operatorname{tr}\{T(f_{\boldsymbol{\theta}_0})^{-1} T(\nabla f_{\boldsymbol{\theta}_0}) T(f_{\boldsymbol{\theta}_0})^{-1} T(\nabla f_{\boldsymbol{\theta}_0})\}.$$

Hence, by Lemma 5 we have

$$E[\nabla^2 \mathcal{L}_n(\boldsymbol{\theta}_0)] \to \frac{1}{4\pi} \int_{-\pi}^{\pi} [\nabla \log f_{\boldsymbol{\theta}_0}(\omega)[\nabla \log f_{\boldsymbol{\theta}_0}(\omega)]' \mathrm{d}\omega \equiv \Gamma(\boldsymbol{\theta}_0).$$

On the other hand,

$$\operatorname{var}[\nabla^2 \mathcal{L}_n(\boldsymbol{\theta}_0)] = \frac{1}{2n^2} \operatorname{tr}\{[T(f_{\boldsymbol{\theta}_0)})(2A^{(1)}_{\boldsymbol{\theta}_0} - A^{(2)}_{\boldsymbol{\theta}_0})]^2\}$$

where the matrices $A^{(1)}_{\boldsymbol{\theta}_0}$ and $A^{(2)}_{\boldsymbol{\theta}_0})$ are defined as in Lemma 6. Then, an application of Lemma 5 establishes that $\operatorname{var}[\nabla^2 \mathcal{L}_n(\boldsymbol{\theta}_0)] \to 0$. Thus, $\nabla^2 \mathcal{L}_n(\boldsymbol{\theta}_0) \to_p \Gamma(\boldsymbol{\theta}_0)$. Besides, since $\Gamma(\boldsymbol{\theta}_0)$ is positive-definite, we can choose the positive-definite matrix $M(\boldsymbol{\theta}_0) \equiv \Gamma(\boldsymbol{\theta}_0) - \kappa I$ with $0 < \kappa < \lambda_{\min} (\Gamma(\boldsymbol{\theta}_0))$ and the result holds.

(iii) Note that

$$\begin{aligned}
\nabla^3 \mathcal{L}_n(\boldsymbol{\theta}) = & \frac{1}{n} \operatorname{tr}\{T(f_{\boldsymbol{\theta}})^{-1} T(\nabla f_{\boldsymbol{\theta}}) T(f_{\boldsymbol{\theta}})^{-1} T(\nabla f_{\boldsymbol{\theta}}) T(f_{\boldsymbol{\theta}})^{-1} T(\nabla f_{\boldsymbol{\theta}})\} \\
& + \frac{1}{2n} \operatorname{tr}\{T(f_{\boldsymbol{\theta}})^{-1} T(\nabla^3 f_{\boldsymbol{\theta}})\} \\
& - \frac{3}{2n} \operatorname{tr}\{T(f_{\boldsymbol{\theta}})^{-1} T(\nabla f_{\boldsymbol{\theta}}) T(f_{\boldsymbol{\theta}})^{-1} T(\nabla^2 f_{\boldsymbol{\theta}})\} \\
& + \frac{1}{2n} \mathbf{X}'[3T(f_{\boldsymbol{\theta}})^{-1} T(\nabla^2 f_{\boldsymbol{\theta}}) T(f_{\boldsymbol{\theta}})^{-1} T(\nabla f_{\boldsymbol{\theta}}) T(f_{\boldsymbol{\theta}})^{-1}. \\
& + 3T(f_{\boldsymbol{\theta}})^{-1} T(\nabla f_{\boldsymbol{\theta}}) T(f_{\boldsymbol{\theta}})^{-1} T(\nabla^2 f_{\boldsymbol{\theta}}) T(f_{\boldsymbol{\theta}})^{-1} \\
& - 6T(f_{\boldsymbol{\theta}})^{-1} T(\nabla f_{\boldsymbol{\theta}}) T(f_{\boldsymbol{\theta}})^{-1} T(\nabla f_{\boldsymbol{\theta}}) T(f_{\boldsymbol{\theta}})^{-1} T(\nabla f_{\boldsymbol{\theta}}) T(f_{\boldsymbol{\theta}})^{-1} \\
& - T(f_{\boldsymbol{\theta}})^{-1} T(\nabla^3 f_{\boldsymbol{\theta}}) T(f_{\boldsymbol{\theta}})^{-1}]\mathbf{X}.
\end{aligned}$$

But,

$$\begin{aligned}
|\mathbf{X}' T(\nabla f_{\boldsymbol{\theta}}) \mathbf{X}| &= \left| \int_{-\pi}^{\pi} \nabla f_{\boldsymbol{\theta}} \Big| \sum_{j=1}^{n} \mathrm{e}^{\mathrm{i}\omega j} X_j \Big|^2 \mathrm{d}\omega \right| \\
&\leq \int_{-\pi}^{\pi} |\nabla f_{\boldsymbol{\theta}}| \left| \sum_{j=1}^{n} \mathrm{e}^{\mathrm{i}\omega j} X_j \right|^2 \mathrm{d}\omega = \mathbf{X}' T(|\nabla f_{\boldsymbol{\theta}}|) \mathbf{X},
\end{aligned}$$

where

$$|\nabla f_{\boldsymbol{\theta}}| = (|\partial f_{\boldsymbol{\theta}} / \partial \theta_1|, \ldots, |\partial f_{\boldsymbol{\theta}} / \partial \theta_J|).$$

Analogously,

$$|\mathbf{X}' T(\nabla^2 f_{\boldsymbol{\theta}}) \mathbf{X}| \leq \mathbf{X}' T(|\nabla^2 f_{\boldsymbol{\theta}}|) \mathbf{X} \quad \text{and} \quad |\mathbf{X}' T(\nabla^3 f_{\boldsymbol{\theta}}) \mathbf{X}| \leq \mathbf{X}' T(|\nabla^3 f_{\boldsymbol{\theta}}|) \mathbf{X}.$$

Therefore,

$$
\begin{aligned}
|\nabla^3 \mathcal{L}_n(\boldsymbol{\theta}| \leq & \frac{1}{n} \operatorname{tr}\{T(f_{\boldsymbol{\theta}})^{-1} T(|\nabla f_{\boldsymbol{\theta}}|) T(f_{\boldsymbol{\theta}})^{-1} T(|\nabla f_{\boldsymbol{\theta}}|) T(f_{\boldsymbol{\theta}})^{-1} T(|\nabla f_{\boldsymbol{\theta}}|)\} \\
& + \frac{1}{2n} \operatorname{tr}\{T(f_{\boldsymbol{\theta}})^{-1} T(|\nabla^3 f_{\boldsymbol{\theta}}|)\} \\
& + \frac{3}{2n} \operatorname{tr}\{T(f_{\boldsymbol{\theta}})^{-1} T(|\nabla f_{\boldsymbol{\theta}}|) T(f_{\boldsymbol{\theta}})^{-1} T(|\nabla^2 f_{\boldsymbol{\theta}}|) \\
& + \frac{1}{2n} \mathbf{X}' \Big[3T(f_{\boldsymbol{\theta}})^{-1} T(|\nabla^2 f_{\boldsymbol{\theta}}|) T(f_{\boldsymbol{\theta}})^{-1} T(|\nabla f_{\boldsymbol{\theta}}|) T(f_{\boldsymbol{\theta}})^{-1} \\
& + 3T(f_{\boldsymbol{\theta}})^{-1} T(|\nabla f_{\boldsymbol{\theta}}|) T(f_{\boldsymbol{\theta}})^{-1} T(|\nabla^2 f_{\boldsymbol{\theta}}|)) T(f_{\boldsymbol{\theta}})^{-1} \\
& + 6T(f_{\boldsymbol{\theta}})^{-1} T(|\nabla f_{\boldsymbol{\theta}}|) T(f_{\boldsymbol{\theta}})^{-1} T(|\nabla f_{\boldsymbol{\theta}}|) T(f_{\boldsymbol{\theta}})^{-1} T(|\nabla f_{\boldsymbol{\theta}}|) T(f_{\boldsymbol{\theta}})^{-1} \\
& + T(f_{\boldsymbol{\theta}})^{-1} T(|\nabla^3 f_{\boldsymbol{\theta}}|) T(f_{\boldsymbol{\theta}})^{-1} \Big] \mathbf{X},
\end{aligned}
$$

and

$$
\begin{aligned}
E|\nabla^3 \mathcal{L}_n(\boldsymbol{\theta}| \leq & \frac{4}{n} \operatorname{tr}\{T(f_{\boldsymbol{\theta}})^{-1} T(|\nabla f_{\boldsymbol{\theta}}|) T(f_{\boldsymbol{\theta}})^{-1} T(|\nabla f_{\boldsymbol{\theta}}|) T(f_{\boldsymbol{\theta}})^{-1} T(|\nabla f_{\boldsymbol{\theta}}|)\} \\
& + \frac{1}{n} \operatorname{tr}\{T(f_{\boldsymbol{\theta}})^{-1} T(|\nabla^3 f_{\boldsymbol{\theta}}|)\} \\
& + \frac{9}{2n} \operatorname{tr}\{T(f_{\boldsymbol{\theta}})^{-1} T(|\nabla f_{\boldsymbol{\theta}}|) T(f_{\boldsymbol{\theta}})^{-1} T(|\nabla^2 f_{\boldsymbol{\theta}}|)\}\}.
\end{aligned}
$$

Now, by an argument similar to Dahlhaus (1989, p. 1758), we conclude that the first summand is bounded by

$$
\begin{aligned}
\frac{1}{n} \operatorname{tr} \Bigg[\Bigg\{ T\Bigg(K|\omega|^{-\alpha+\epsilon} \prod_{i=1}^{r} \prod_{j=1}^{m_i} |\omega - \omega_{ij}|^{-\alpha_i+\epsilon_i} \Bigg)^{-1} \\
\times T\Bigg(K|\omega|^{-\alpha-\epsilon} \prod_{i=1}^{r} \prod_{j=1}^{m_i} |\omega - \omega_{ij}|^{-\alpha_i-\epsilon_i} \Bigg) \Bigg\}^3 \Bigg].
\end{aligned}
$$

where $\epsilon, \epsilon_1, \ldots, \epsilon_r > 0$ can be chosen arbitrarily small. By Lemma 5, this term converges to

$$
K \int_{-\pi}^{\pi} |\omega|^{-3\epsilon} \prod_{i=1}^{r} \prod_{j=1}^{m_i} |\omega - \omega_{ij}|^{-3\epsilon_i} \leq K_1 < \infty.
$$

The other two summands can be bounded similarly by constants K_2 and K_3, respectively. The result follows by choosing $M = \max\{K_1, K_2, K_3\}$. □

2.2. *Asymptotic normality*

A central limit theorem of the MLE is established in this section. The main result is stated in Theorem 2. The proof of this result requires Lemmas 3–7 shown in the Appendix.

THEOREM 2 (Central limit theorem). *The maximum likelihood estimate*, $\hat{\boldsymbol{\theta}}_n$, *satisfies the following limiting distribution as* $n \to \infty$:

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) \to_D N(0, \Gamma(\boldsymbol{\theta}_0)^{-1}),$$

$$\Gamma(\boldsymbol{\theta}_0) = \frac{1}{4\pi} \int_{-\pi}^{\pi} [\nabla \log f_{\boldsymbol{\theta}}(\omega)][\nabla \log f_{\boldsymbol{\theta}}(\omega)]' \mathrm{d}\omega. \tag{6}$$

PROOF. It suffices to show that:

(i) $$\sqrt{n}\nabla \mathcal{L}_n(\boldsymbol{\theta}_0) \to_D N(0, \Gamma(\boldsymbol{\theta}_0));$$

(ii) $$|\nabla^2 \mathcal{L}_n(\hat{\boldsymbol{\theta}}_n) - \nabla^2 \mathcal{L}_n(\boldsymbol{\theta}_0)| \to_p 0 \quad \text{as } \hat{\boldsymbol{\theta}}_n \to_p \boldsymbol{\theta}_0;$$

and

(iii) $$\nabla^2 \mathcal{L}_n(\boldsymbol{\theta}_0) \to_p \Gamma(\boldsymbol{\theta}_0), \quad \text{with } \eta = \eta(n) = 1 - n^{-\delta},$$

for some δ such that $\delta p > 1$ with p as in Lemma 1. Part (i) follows from the product of cumulants (see Taniguchi and Kakizawa, 2000 p. 54 and p. 168ff):

$$n \operatorname{cov}(\nabla \mathcal{L}_n(\boldsymbol{\theta}_0), \nabla \mathcal{L}_n(\boldsymbol{\theta}_0)) = \frac{1}{2n} \operatorname{tr}\{T(f_{\boldsymbol{\theta}_0})^{-1} T(\nabla f_{\boldsymbol{\theta}_0}) T(f_{\boldsymbol{\theta}_0})^{-1} T(\nabla f_{\boldsymbol{\theta}_0})\}.$$

By Lemma 5, this term converges to $\Gamma(\boldsymbol{\theta}_0)$. By the same lemma,

$$n^{p/2} \operatorname{cum}\{\mathcal{L}_n(\boldsymbol{\theta}_0)_{j_1}, \ldots, \mathcal{L}_n(\boldsymbol{\theta}_0)_{j_p}\},$$

converges to zero, as $n \to \infty$. Result (ii) is a consequence of the equicontinuity of the quadratic form $Z_n^{(i)}$, (see Lemma 7) and from Lemma 5. Finally, part (iii) follows from the proof of Theorem 1, since $E[\nabla^2 \mathcal{L}_n(\boldsymbol{\theta}_0)] \to \Gamma(\boldsymbol{\theta}_0)$ and $\operatorname{var} \nabla^2 \mathcal{L}_n(\boldsymbol{\theta}_0)$ tends to zero as $n \to \infty$. □

THEOREM 3 (Efficiency). *The maximum likelihood estimate*, $\hat{\boldsymbol{\theta}}_n$, *is asymptotically an efficient estimate of* $\boldsymbol{\theta}_0$.

PROOF. It suffices to prove that the Fisher information matrix $n^{-1}\Gamma_n(\boldsymbol{\theta}_0)$ converges to the variance of the estimates, $\Gamma(\boldsymbol{\theta}_0)$, as $n \to \infty$. But, this follows directly from part (ii) of the proof of Theorem 1, since

$$\begin{aligned} n^{-1}\Gamma_n(\boldsymbol{\theta}_0) &= n\, E\{[\nabla\mathcal{L}_n(\boldsymbol{\theta}_0)][\nabla\mathcal{L}_n(\boldsymbol{\theta}_0)]'\} \\ &= \frac{1}{2n}\mathrm{tr}\{T(f_{\boldsymbol{\theta}_0})^{-1}T(\nabla f_{\boldsymbol{\theta}_0})T(f_{\boldsymbol{\theta}_0})^{-1}T(\nabla f_{\boldsymbol{\theta}_0})\}, \end{aligned}$$

which tends to $\Gamma(\boldsymbol{\theta}_0)$ as $n \to \infty$. □

3. MONTE CARLO EXPERIMENTS

In order to assess the finite sample performance of the MLEs, a number of Monte Carlo simulations are conducted for the class of SARFIMA models described by eqn (2).

Tables I–VI show the results from simulations for SARFIMA$(0, d, 0) \times (0, d_s, 0)_s$ processes for different values of $\mathbf{d}$, $\mathbf{d_s}$, sample size n, seasonal period s and $\sigma_\varepsilon^2 = 1$. We choose this class of models to examine the finite sample behaviour of the estimates. Furthermore, the finite sample performance of the MLE is compared with the Whittle estimate and the Kalman filter approach with truncations $m = 60$ and $m = 120$. This simulation setup is considered for finite sample comparison purposes because the asymptotic theory developed in this paper applies only to the exact MLE case. Finite sample performances of a number of other estimation techniques for fractional seasonal models are studied in two recent papers by Reisen *et al.* (2005a, b).

The results are based on 1000 repetitions, with seasonal series generated using the Durbin–Levinson algorithm with zero-mean Gaussian noise. The autocovariance function was computed by the convolution method of Bertelli and Caporin (2002). In order to explore the effect of the estimation of the mean we considered two situations: *known mean* where the process is assumed to be have zero mean and *unknown mean* where the expected value of the process is estimated by the sample mean and then centred before the computations.

The exact MLE method has been implemented computationally by means of the Durbin–Levinson algorithm (see Brockwell and Davis, 1991, Sect. 5.2) with autocovariance calculated by the Bertelli and Caporin (2002) approach.

The Whittle method has been implemented by minimizing the expression, cf. Giraitis *et al.* (2001):

$$S(\boldsymbol{\theta}) = \frac{2}{n}\sum_{k=1}^{[n/2]}\frac{I(\omega_k)}{f_{\boldsymbol{\theta}}(\omega_k)},$$

where $\boldsymbol{\theta} = (\mathbf{d},\mathbf{d_s})$ and $\omega_k = 2\pi k/n$, with periodogram given by

$$I(\omega_k) = \frac{1}{2\pi n}\left|\sum_{t=1}^{n} x_t \mathrm{e}^{\mathrm{i}t\omega_k}\right|^2,$$

TABLE I

ESTIMATES OF d AND d_s FOR $d = 0.3$, $d_s = 0.1$ AND $s = 6$

		Estimators			
n	Stat.	Exact	Whittle	Kalman (60)	Kalman (120)
Known mean					
256	Mean($\hat{\mathbf{d}}$)	0.2946	0.2617	0.3026	0.3005
	Std($\hat{\mathbf{d}}$)	0.0476	0.0596	0.0512	0.0497
	Mean($\hat{\mathbf{d}}_s$)	0.0952	0.0611	0.0983	0.0990
	Std($\hat{\mathbf{d}}_s$)	0.0469	0.0493	0.0500	0.0494
512	Mean($\hat{\mathbf{d}}$)	0.2968	0.2815	0.3048	0.3030
	Std($\hat{\mathbf{d}}$)	0.0339	0.0398	0.0366	0.0370
	Mean($\hat{\mathbf{d}}_s$)	0.0979	0.0779	0.0983	0.1007
	Std($\hat{\mathbf{d}}_s$)	0.0341	0.0370	0.0365	0.0365
Unknown mean					
256	Mean($\hat{\mathbf{d}}$)	0.2788	0.2626	0.2861	0.2891
	Std($\hat{\mathbf{d}}$)	0.0506	0.0591	0.0499	0.0512
	Mean($\hat{\mathbf{d}}_s$)	0.0853	0.0610	0.0833	0.0834
	Std($\hat{\mathbf{d}}_s$)	0.0461	0.0474	0.0465	0.0476
512	Mean($\hat{\mathbf{d}}$)	0.2890	0.2800	0.2841	0.2857
	Std($\hat{\mathbf{d}}$)	0.0350	0.0393	0.0351	0.0344
	Mean($\hat{\mathbf{d}}_s$)	0.0909	0.0787	0.0909	0.0921
	Std($\hat{\mathbf{d}}_s$)	0.0351	0.0378	0.0377	0.0370

TABLE II

ESTIMATES OF d AND d_s FOR $d = 0.3$, $d_s = 0.1$ AND $s = 10$

		Estimators			
n	Stat.	Exact	Whittle	Kalman (60)	Kalman (120)
Known mean					
256	Mean($\hat{\mathbf{d}}$)	0.2933	0.2589	0.3007	0.2989
	Std($\hat{\mathbf{d}}$)	0.0461	0.0594	0.0505	0.0479
	Mean($\hat{\mathbf{d}}_s$)	0.0966	0.0630	0.0999	0.1008
	Std($\hat{\mathbf{d}}_s$)	0.0477	0.0497	0.0509	0.0510
512	Mean($\hat{\mathbf{d}}$)	0.2962	0.2806	0.3058	0.2996
	Std($\hat{\mathbf{d}}$)	0.0332	0.0383	0.0381	0.0369
	Mean($\hat{\mathbf{d}}_s$)	0.0957	0.0798	0.0985	0.0975
	Std($\hat{\mathbf{d}}_s$)	0.0364	0.0379	0.0381	0.0360
Unknown mean					
256	Mean($\hat{\mathbf{d}}$)	0.2753	0.2581	0.2734	0.2764
	Std($\hat{\mathbf{d}}$)	0.0525	0.0595	0.0462	0.0462
	Mean($\hat{\mathbf{d}}_s$)	0.0864	0.0618	0.0905	0.0921
	Std($\hat{\mathbf{d}}_s$)	0.0450	0.0477	0.0490	0.0484
512	Mean($\hat{\mathbf{d}}$)	0.2877	0.2799	0.2819	0.2844
	Std($\hat{\mathbf{d}}$)	0.0355	0.0383	0.0367	0.0357
	Mean($\hat{\mathbf{d}}_s$)	0.0904	0.0762	0.0941	0.0955
	Std($\hat{\mathbf{d}}_s$)	0.0349	0.0368	0.0361	0.0363

and spectral density

$$f_{\boldsymbol{\theta}}(\omega) = \frac{1}{2\pi}|1 - \mathrm{e}^{\mathrm{i}t\omega}|^{-2d}|1 - \mathrm{e}^{\mathrm{i}t\omega s}|^{-2d_s}.$$

TABLE III

ESTIMATES OF d AND d_s FOR $d = 0.2$, $d_s = 0.2$ AND $s = 6$

n	Stat.	Estimators: Exact	Whittle	Kalman (60)	Kalman (120)
Known mean					
256	Mean($\hat{\mathbf{d}}$)	0.1962	0.1604	0.2004	0.2001
	Std($\hat{\mathbf{d}}$)	0.0492	0.0594	0.0516	0.0508
	Mean($\hat{\mathbf{d}}_s$)	0.1931	0.1635	0.1975	0.1998
	Std($\hat{\mathbf{d}}_s$)	0.0457	0.0581	0.0496	0.0486
512	Mean($\hat{\mathbf{d}}$)	0.1975	0.1786	0.1997	0.2056
	Std($\hat{\mathbf{d}}$)	0.0350	0.0414	0.0379	0.0370
	Mean($\hat{\mathbf{d}}_s$)	0.1952	0.1828	0.2003	0.2027
	Std($\hat{\mathbf{d}}_s$)	0.0341	0.0388	0.0367	0.0393
Unknown mean					
256	Mean($\hat{\mathbf{d}}$)	0.1789	0.1588	0.1725	0.1755
	Std($\hat{\mathbf{d}}$)	0.0519	0.0587	0.0549	0.0556
	Mean($\hat{\mathbf{d}}_s$)	0.1851	0.1659	0.1913	0.1910
	Std($\hat{\mathbf{d}}_s$)	0.0480	0.0591	0.0493	0.0499
512	Mean($\hat{\mathbf{d}}$)	0.1889	0.1810	0.1860	0.1877
	Std($\hat{\mathbf{d}}$)	0.0356	0.0389	0.0375	0.0383
	Mean($\hat{\mathbf{d}}_s$)	0.1928	0.1812	0.1923	0.1912
	Std($\hat{\mathbf{d}}_s$)	0.0324	0.0397	0.0315	0.0317

TABLE IV

ESTIMATES OF d AND d_s FOR $d = 0.2$, $d_s = 0.2$ AND $s = 10$

n	Stat.	Estimators: Exact	Whittle	Kalman (60)	Kalman (120)
Known mean					
256	Mean($\hat{\mathbf{d}}$)	0.1955	0.1579	0.2019	0.1997
	Std($\hat{\mathbf{d}}$)	0.0484	0.0594	0.0527	0.0509
	Mean($\hat{\mathbf{d}}_s$)	0.1937	0.1666	0.2013	0.2014
	Std($\hat{\mathbf{d}}_s$)	0.0460	0.0588	0.0500	0.0502
512	Mean($\hat{\mathbf{d}}$)	0.1979	0.1775	0.1999	0.1978
	Std($\hat{\mathbf{d}}$)	0.0319	0.0411	0.0368	0.0359
	Mean($\hat{\mathbf{d}}_s$)	0.1955	0.1834	0.2007	0.2000
	Std($\hat{\mathbf{d}}_s$)	0.0324	0.0383	0.0358	0.0357
Unknown mean					
256	Mean($\hat{\mathbf{d}}$)	0.1798	0.1581	0.1702	0.1733
	Std($\hat{\mathbf{d}}$)	0.0503	0.0572	0.0481	0.0476
	Mean($\hat{\mathbf{d}}_s$)	0.1867	0.1679	0.1947	0.1948
	Std($\hat{\mathbf{d}}_s$)	0.0456	0.0560	0.0528	0.0525
512	Mean($\hat{\mathbf{d}}$)	0.1896	0.1789	0.1933	0.1947
	Std($\hat{\mathbf{d}}$)	0.0363	0.0394	0.0308	0.0314
	Mean($\hat{\mathbf{d}}_s$)	0.1924	0.1852	0.1957	0.1948
	Std($\hat{\mathbf{d}}_s$)	0.0325	0.0378	0.0365	0.0371

The approximate Kalman filter MLEs are based on a finite state–space representation of the truncated MA(∞) expansion

TABLE V

Estimates of $\boldsymbol{d}$ and $\boldsymbol{d_s}$ for $\boldsymbol{d} = 0.1$, $\boldsymbol{d_s} = 0.3$ and $s = 6$

		Estimators			
n	Stat.	Exact	Whittle	Kalman (60)	Kalman (120)
Known mean					
256	Mean($\hat{\mathbf{d}}$)	0.0976	0.0625	0.1003	0.1061
	Std($\hat{\mathbf{d}}$)	0.0466	0.0497	0.0516	0.0479
	Mean($\hat{\mathbf{d}}_s$)	0.2921	0.2815	0.3106	0.3064
	Std($\hat{\mathbf{d}}_s$)	0.0424	0.0607	0.0500	0.0472
512	Mean($\hat{\mathbf{d}}$)	0.0975	0.0755	0.1007	0.0969
	Std($\hat{\mathbf{d}}$)	0.0331	0.0403	0.0383	0.0373
	Mean($\hat{\mathbf{d}}_s$)	0.2983	0.2965	0.3111	0.3088
	Std($\hat{\mathbf{d}}_s$)	0.0317	0.0400	0.0365	0.0360
Unknown mean					
256	Mean($\hat{\mathbf{d}}$)	0.0823	0.0603	0.0771	0.0808
	Std($\hat{\mathbf{d}}$)	0.0470	0.0500	0.0399	0.0408
	Mean($\hat{\mathbf{d}}_s$)	0.2833	0.2804	0.2993	0.2972
	Std($\hat{\mathbf{d}}_s$)	0.0426	0.0609	0.0483	0.0489
512	Mean($\hat{\mathbf{d}}$)	0.0924	0.0776	0.0774	0.0804
	Std($\hat{\mathbf{d}}$)	0.0354	0.0379	0.0338	0.0336
	Mean($\hat{\mathbf{d}}_s$)	0.2897	0.2970	0.3046	0.3033
	Std($\hat{\mathbf{d}}_s$)	0.0323	0.0399	0.0384	0.0367

TABLE VI

Estimates of $\boldsymbol{d}$ and $\boldsymbol{d_s}$ for $\boldsymbol{d} = 0.1$, $\boldsymbol{d_s} = 0.3$ and $s = 10$

		Estimators			
n	Stat.	Exact	Whittle	Kalman (60)	Kalman (120)
Known mean					
256	Mean($\hat{\mathbf{d}}$)	0.0969	0.0617	0.0997	0.0987
	Std($\hat{\mathbf{d}}$)	0.0456	0.0495	0.0512	0.0485
	Mean($\hat{\mathbf{d}}_s$)	0.2934	0.2903	0.3171	0.3155
	Std($\hat{\mathbf{d}}_s$)	0.0408	0.0608	0.0504	0.0491
512	Mean($\hat{\mathbf{d}}$)	0.0969	0.0748	0.0974	0.0967
	Std($\hat{\mathbf{d}}$)	0.0325	0.0399	0.0374	0.0364
	Mean($\hat{\mathbf{d}}_s$)	0.2951	0.3013	0.3149	0.3114
	Std($\hat{\mathbf{d}}_s$)	0.0289	0.0398	0.0359	0.0355
Unknown mean					
256	Mean($\hat{\mathbf{d}}$)	0.0803	0.0604	0.0669	0.0710
	Std($\hat{\mathbf{d}}$)	0.0475	0.0496	0.0436	0.0421
	Mean($\hat{\mathbf{d}}_s$)	0.2873	0.2927	0.3153	0.3107
	Std($\hat{\mathbf{d}}_s$)	0.0395	0.0580	0.0441	0.0438
512	Mean($\hat{\mathbf{d}}$)	0.0886	0.0765	0.0796	0.0819
	Std($\hat{\mathbf{d}}$)	0.0354	0.0378	0.0414	0.0412
	Mean($\hat{\mathbf{d}}_s$)	0.2896	0.2991	0.3036	0.2997
	Std($\hat{\mathbf{d}}_s$)	0.0322	0.0414	0.0418	0.0383

$$x_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}, \tag{7}$$

where ψ_j are the coefficients of

$$\psi(z) = \sum_{j=0}^{\infty} \psi_j z^j = (1-z)^{-d}(1-z^s)^{-d_s}$$

to the approximate MA(m) given by

$$x_t = \sum_{j=0}^{m} \psi_j \varepsilon_{t-j}, \tag{8}$$

Extending the canonical state–space representation of the MA(m) model (8) given by Chan and Palma (1998) in the context of ARFIMA processes we have:

$$W_{t+1} = \begin{bmatrix} 0 & & I_{m-1} \\ 0 & \cdots & 0 \end{bmatrix} W_t + \begin{bmatrix} \psi_1 \\ \vdots \\ \psi_m \end{bmatrix} \varepsilon_t, \tag{9}$$

$$x_t = [1 \quad 0 \quad 0 \quad \cdots \quad 0] W_t + \varepsilon_t, \tag{10}$$

with

$$F = \begin{bmatrix} 0 & & I_{m-1} \\ 0 & \cdots & 0 \end{bmatrix}, \quad G = [1 \quad 0 \quad 0 \quad \cdots \quad 0], \tag{11}$$

$$W_t = [x(t|t-1), x(t+1|t-1), \ldots, x(t+m-1|t-1)]', \tag{12}$$

$$x(t+j|t-1) = E[x_{t+j}|x_{t-1}, x_{t-2}, \ldots]. \tag{13}$$

The Kalman approach estimates are obtained by maximizing the Gaussian log-likelihood function (14). The log-likelihood function can be evaluated by directly applying the Kalman recursive equations in Proposition 12.2.2 of Brockwell and Davis (1991) to the state–space system (9) to (11). The log-likelihood function (excepting a constant) is given by

$$\mathcal{L}(\boldsymbol{\theta}) = -\frac{1}{2}\left\{\sum_{t=1}^{n} \log \Delta_t + \sum_{t=1}^{n} \frac{(x_t - \hat{x}_t)^2}{\Delta_t}\right\}, \tag{14}$$

where $\hat{x}_t$ is the one-step predictor of x_t and Δ_t its variance given by eqn (12.2.6) of Brockwell and Davis (1991).

The Monte Carlo experiments were conducted with a Pentium IV 2.8GHz. HT machine using a Fortran program in a Windows XP platform. Table VII shows the CPU average times (in seconds) for one evaluation of the respective method. The average times are computed over 2000 calls from the optimization routine. For both sample sizes studied, the fastest method is the Whittle. The Kalman filter approach with $m = 60$ is faster than the exact MLE method for $n = 512$ while the Kalman method with $m = 120$ is the slowest algorithm for both sample sizes.

TABLE VII

CPU Average Times (in Seconds) for Computing $\hat{\mathbf{d}}$ and $\hat{\mathbf{d}}_s$ with $\mathbf{d} = 0.3$, $\mathbf{d_s} = 0.1$ and $s = 6$

Sample size	Exact	Whittle	Kalman (60)	Kalman (120)
256	0.00302	0.00012	0.00439	0.02786
512	0.00693	0.00023	0.00441	0.03022

From Tables I–VI, it seems that for the *known mean* case the exact MLE and the Kalman methods display little bias for both sample sizes. On the contrary, the Whittle method presents a noticeable downward bias for both estimators $\hat{\mathbf{d}}$ and $\hat{\mathbf{d}}_s$. The sample standard deviations of the estimates are close to their theoretical counterparts (see Table VIII), for the four methods considered. However, the exact MLE seems to have slightly lower sample standard deviations than the other methods, for both long-memory parameters and both sample sizes.

In the *unknown mean* case, all the estimates seems to display a downward bias, which is stronger for the Whittle method. Nevertheless, the bias displayed by this estimate is similar to the *known mean* case, since the Whittle method is not affected by the estimation of the mean. Similar to the previous case, the estimated standard deviations are comparable to the theoretical values and the exact MLE displays slightly lower sample standard deviations than the other methods for most long-memory parameters and sample size combinations.

The theoretical values of the standard deviations of the estimated parameters given in Table VIII are based on formula (6). In general, analytic expressions for the integral in eqn (6) are difficult to obtain for an arbitrary period s. For a SARFIMA(0, d, 0) $\times$ (0, d_s, 0)$_s$ model, the matrix $\Gamma(\boldsymbol{\theta})$ can be written as

$$\Gamma(\boldsymbol{\theta}) = \begin{pmatrix} \frac{\pi^2}{6} & \Gamma_{12} \\ \Gamma_{12} & \frac{\pi^2}{6} \end{pmatrix}, \tag{15}$$

with

$$\Gamma_{12} = \frac{1}{\pi}\int_{-\pi}^{\pi}\left\{\log\left|2\sin\left(\frac{\omega}{2}\right)\right|\right\}\left\{\log\left|2\sin\left(s\frac{\omega}{2}\right)\right|\right\}\mathrm{d}\omega.$$

An interesting feature of the asymptotic variance of the parameters is that for a SARFIMA(0, d, 0) $\times$ (0, d_s, 0)$_s$ process, the variance of $\hat{\mathbf{d}}$ is the same as the variance of $\hat{\mathbf{d}}_s$. An explicit expression for this integral can be given for $s = 2$. In this case,

TABLE VIII

Asymptotic Standard Deviation of $\hat{\mathbf{d}}$ and $\hat{\mathbf{d}}_s$

n	$s = 6$	$s = 10$
256	0.050330	0.050332
512	0.035588	0.035590

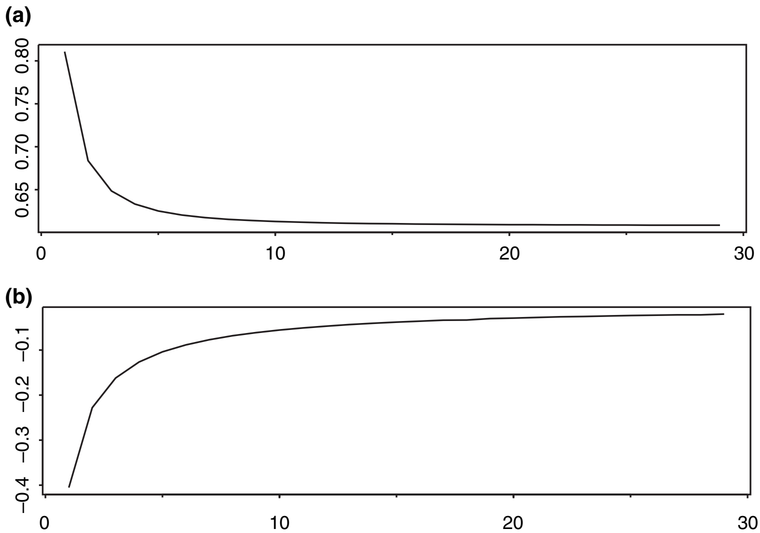

FIGURE 1. (a) Values of $\sigma^2(\hat{\mathbf{d}}_s)$ as a function of the period s; (b) values of $\text{cov}(\hat{\mathbf{d}}, \hat{\mathbf{d}}_s)$ as a function of s.

$$\Gamma(\boldsymbol{\theta}) = \frac{\pi^2}{12}\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

For other values of s, the integral may be evaluated numerically. Figure 1 shows the evolution of $\sigma^2(\hat{\mathbf{d}})$ as a function of the period s, see panel (a), and the evolution of $\text{cov}(\hat{\mathbf{d}}, \hat{\mathbf{d}}_s)$ as s increases, see panel (b). Both curves are based on the numerical evaluation of eqn (15) and then inverting this matrix to obtain the asymptotic variance–covariance matrix of the parameters.

Observe that $\sigma^2(\hat{\mathbf{d}}_s)$ [equivalently $\sigma^2(\hat{\mathbf{d}})$] starts with a value of $8/\pi^2$ and decreases to $6/\pi^2$ as $s \to \infty$. That is, for a very large period s, the asymptotic variance of $\hat{\mathbf{d}}_s$ is the same as the variance of $\hat{\mathbf{d}}$ from an ARFIMA(0, d, 0) model.

In order to illustrate the normality of the estimates, Figure 2 shows quantile plots for the estimated parameters $\hat{\mathbf{d}}$, see panel (a), and $\hat{\mathbf{d}}_s$, see panel (b). These estimates are based on 1000 repetitions of a SARFIMA (0, d, 0) $\times$ (0, d_s, 0)$_s$ process with $s = 6$, $d = 0.2$, $d_s = 0.2$ and sample size of 600 observations. The chi-square test for normality for these samples are $\hat{\mathbf{d}}$: $\chi^2_{31} = 26.11$ with p-value equaling 0.72 and $\hat{\mathbf{d}}_s$: $\chi^2_{31} = 30.85$ with p-value equaling 0.47. Thus, the hypothesis of normality of both samples is not rejected at the 5% significance level.

Finally, Figure 3 shows a simulated SARFIMA (2, d, 1) $\times$ (1, d_s, 0)$_s$ process with $s = 48$, sample size $n = 1440$, $d = 0.1$, $d_s = 0.2$, $\phi_1 = 0.6$, $\phi_2 = 0.2$, $\Phi_1 = 0.6$, and $\sigma_\epsilon = 1$. Panel (a) displays the series, panel (b) shows the ACF and the spectral density is shown in panel (c). Observe that the time series shows a seasonal pattern, which is reflected in the ACF, with periodicity $s = 48$. The estimated spectral density shows peaks at the origin, at periodicity 24 and 48.

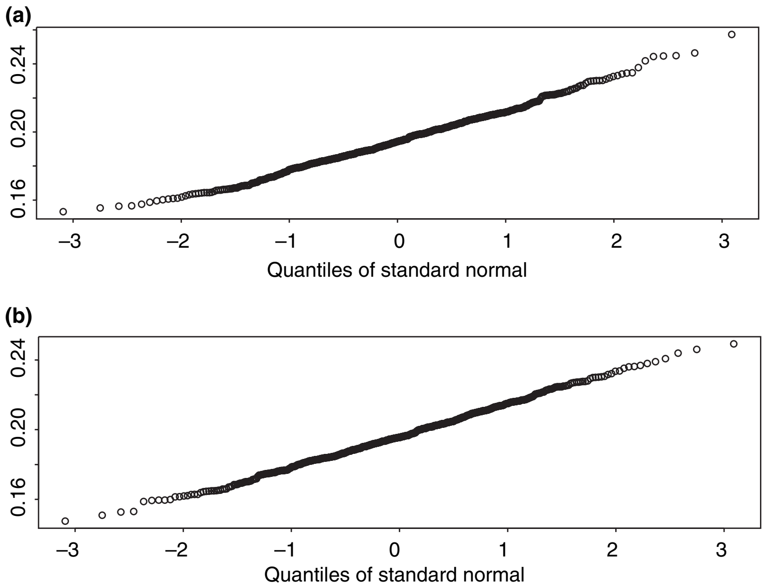

FIGURE 2. Quantile plots of (a) $\hat{\mathbf{d}}$ ($d = 0.2$) and (b) $\hat{\mathbf{d}}_s$ ($d_s = 0.2$).

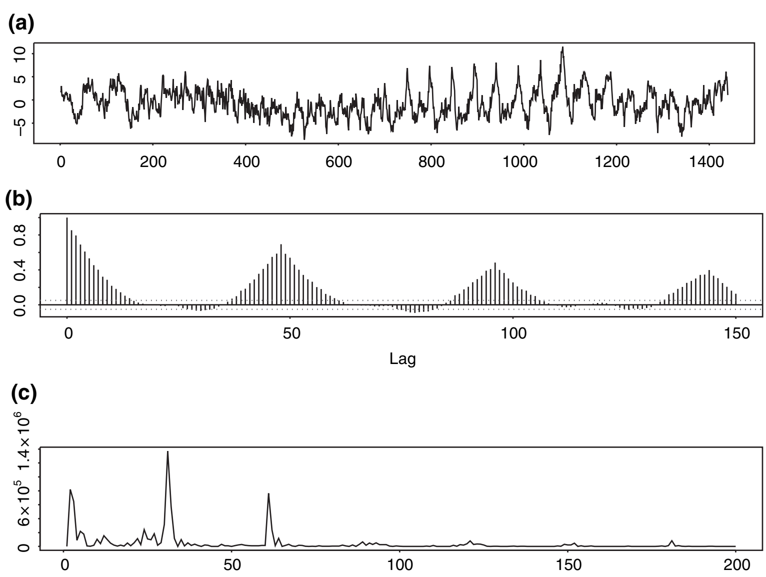

FIGURE 3. Simulated SARFIMA(2, d, 1) × (1, d_s, 0)$_s$ process: (a) series, (b) ACF and (c) spectral density.

4. AN INTERNET TRAFFIC EXAMPLE

As an illustration of the MLE discussed in the previous sections, consider the http requests to a World Wide Web server at the University of Saskatchewan between 1 June 1995 to 31 December 1995. Since it has been documented that internet traffic data display certain amount of long-memory behaviour (see, e.g. Beran, 1994), this data set serves as a good example to illustrate the long-memory seasonal modelling methodology developed in this paper and to test the effectiveness of the proposed scheme. The original data set consists of time stamps of 1-second resolution. In this paper, the data have been aggregated every 30 minutes, i.e. each point represents the total number of requests sent to the Sakastchewan's server within a 30-minute interval. To make the data more Gaussian and to stabilize their variances, we apply a logarithmic transformation.

Figure 4(a) displays a time series of the transformed series. The trace consists of 9074 measurements aggregated in every 30 minutes. Thus, the time span of the observations is about 189 days. Panel (b) shows the autocorrelation function of the data. Note that the ACF decays very slowly and exhibits a sinusoidal behaviour. The amplitude of the harmonic component also decays as the lag increases. Figure 5(a) shows a log-var graph. It indicates a possible long-memory behaviour in the data as described in Beran (1994). The spectral density of the data is displayed in panel (b). Observe that there are two major peaks in the spectrum: one at the origin and another at frequency $\omega = 2\pi \times 189/9074 =$

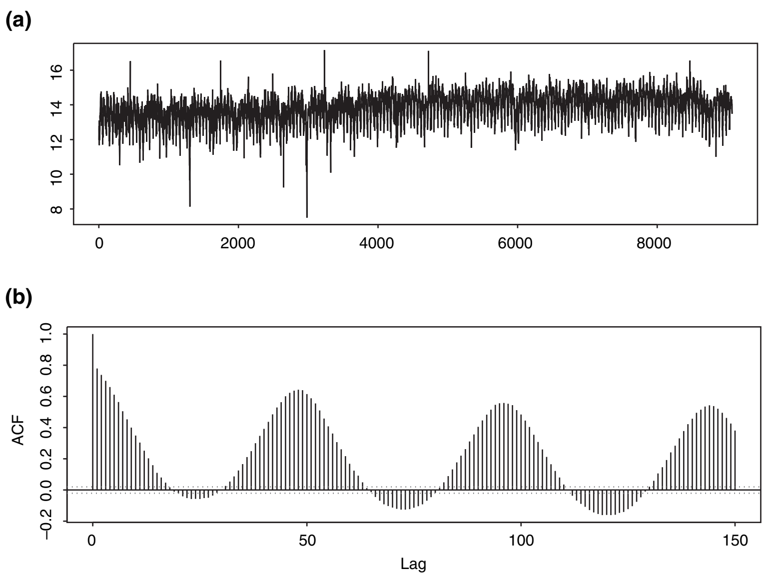

FIGURE 4. (a) Log internet traffic data; (b) autocorrelation function.

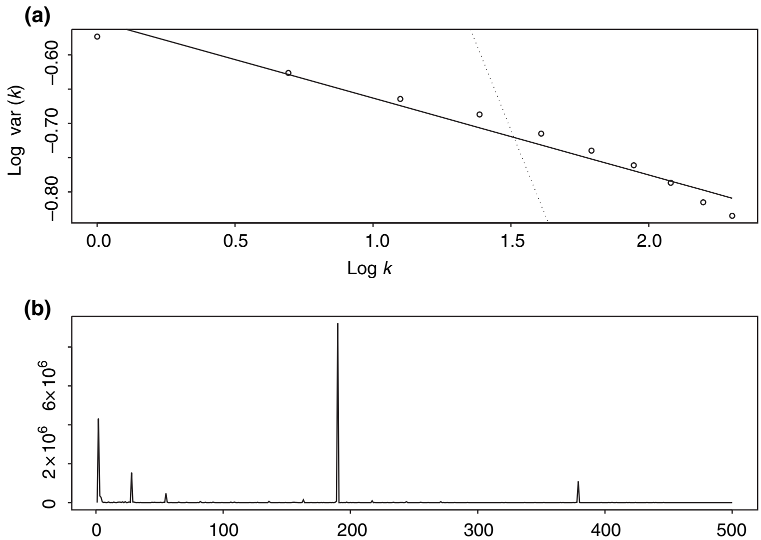

FIGURE 5. Log Internet traffic data: (a) var-log graph; (b) spectral density.

0.13 (similar to Figure 3, the third peak is an artifact of the periodogram corresponding to 2ω). These features are compatible with a seasonal long-memory process with $s = 48$, i.e. a daily seasonal pattern. Accordingly, a SARFIMA model is suggested for this time series. The model selected by Akaikes' criterion is the SARFIMA $(1, d, 1) \times (0, d_s, 0)_s$ with estimated parameters given in Table IX. Observe that the estimated values of d and d_s are small, indicating that the long-memory effect of this data set seems to be mild, albeit significant. The Student's t-values reported in Table IX are based on the numerical calculation of the inverse of the Hessian matrix, which approximates the asymptotic variance of the parameters.

The variance of the series is 0.635 and the residual variance of the model is 0.207. Thus, the model explains roughly two-thirds of the total variance of the data. The residuals and their autocorrelation function are displayed in Figure 6(a) and (b), respectively. Observe that the strong correlation structure of the data has been largely removed, suggesting the proposed model provides a reasonably good fit to the data.

TABLE IX

LOG INTERNET TRAFFIC DATA: MAXIMUM LIKELIHOOD ESTIMATION OF THE SARFIMA$(1, d, 1) \times (0, d_s, 0)_s$ MODEL

Parameter	**d**	$\mathbf{d}_s$	$\boldsymbol{\phi}$	$\boldsymbol{\theta}$
Estimate	0.076	0.148	0.917	0.583
Student's t-value	7.76	19.79	172.71	50.03

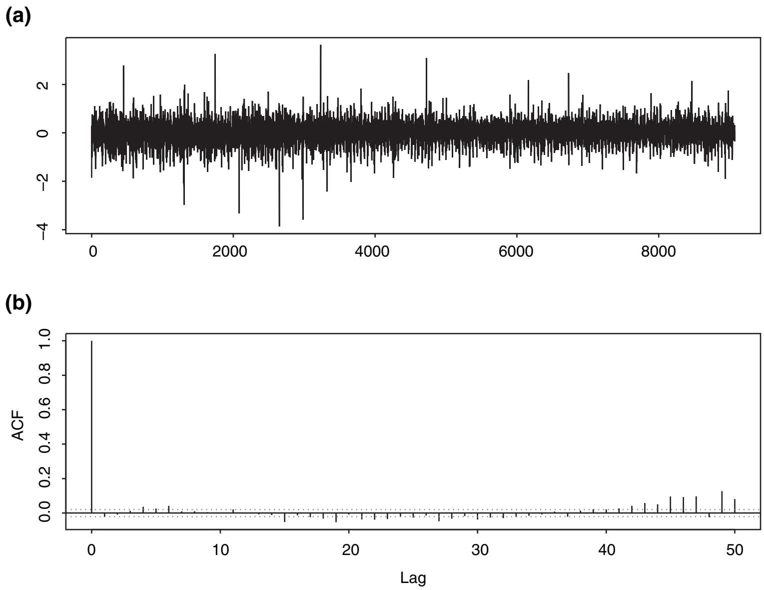

FIGURE 6. (a) Residuals of the SARFIMA model; (b) autocorrelation function of the residuals.

5. CONCLUDING REMARKS

Asymptotic properties of the MLEs of seasonal long-memory processes are derived in this paper. It is found that the MLEs are consistent, efficient and satisfy a central limit theorem. Finite sample performances of these estimates are investigated through Monte Carlo experiments and it is shown that finite sample performances of the MLE behave reasonably well even for moderate sample sizes. The exact MLE compares favourably with other estimation methods. Finally, the proposed methodology is illustrated with a disc traffic example to demonstrate its usefulness in modelling computer traffic.

APPENDIX

LEMMA 1. *Let $f(\omega)$ be given in eqn* (1). *Let $g(\omega) \geq 0$ be such that $g(\omega) \leq M_g$. Then, there exists a constant $K > 0$, not depending on g or f, and $0 < p < 1$ such that for $\eta \in [\frac{1}{2}, 1]$,*

$$\int_{-\pi}^{\pi} |f_\eta(\omega) - f(\omega)| g(\omega) \mathrm{d}\omega \leq K M_g |\eta - 1|^p.$$

PROOF. Consider the following decomposition

$$\begin{aligned}\int_{-\pi}^{\pi} |f_\eta(\omega) - f(\omega)| g(\omega) \mathrm{d}\omega &= \int_{-\pi}^{\pi} f(\omega) h_\eta(\omega) g(\omega) \mathrm{d}\omega \\ &= \sum_{ij} \int_{A_{ij}} f(\omega) h_\eta(\omega) g(\omega) \mathrm{d}\omega + \int_{\bar{A}} f(\omega) h_\eta(\omega) g(\omega) \mathrm{d}\omega, \end{aligned} \quad \text{(A1)}$$

where

$$h_\eta(\omega) = \left| \frac{f_\eta(\omega)}{f(\omega)} - 1 \right|, \quad A_{ij} = \{\omega : |\omega - \omega_{ij}| \le (1-\eta)^{\frac{1}{m}}\}$$

are neighbourhoods around each distinct pole ω_{ij} for $i = 1, \ldots, r$, $j = 1, \ldots, m_i$, $m = \max_i\{m_i\}$, and $A = (-\pi, \pi] \backslash \bigcup A_{ij}$. Observe that for η sufficiently close to 1, the sets A_{ij} are disjoint.

To prove the lemma, it suffices to show that each term in (A1) satisfies the following inequalities:

$$\int_{A_{ij}} f(\omega) h_\eta(\omega) g(\omega) \mathrm{d}\omega \le K M_g |\eta - 1|^p, \quad \text{(A2)}$$

for $i = 1, \ldots, r$, $j = 1, \ldots, m_i$, and

$$\int_{\bar{A}} f(\omega) h_\eta(\omega) g(\omega) \mathrm{d}\omega \le K M_g |\eta - 1|^p, \quad \text{(A3)}$$

where K is a positive constant not depending on g or f, and p is a number between 0 and 1.

To establish (A2), we first show that $h_\eta(\omega)$ is uniformly bounded for $\eta \in [\frac{1}{2}, 1]$ and $\omega \in (-\pi, \pi]$. Observe that

$$h_\eta(\omega) = \left| \prod_{i=1}^{r} p_i(\delta_i) - 1 \right|,$$

where

$$\delta_i = \prod_{j=1}^{m_i} |\omega - \omega_{ij}| \quad \text{and} \quad p_i(x) = \{\frac{x}{(1-\eta)^2 + \eta x}\}^{\alpha_i}, \quad \text{for } i = 1, \ldots, r.$$

Since $\delta_i \ge 0$, we focus on the properties of the function $p_i(x)$ for $x \in [0, \infty)$. A simple calculation gives

$$p_i'(x) = \frac{\alpha_i (1-\eta)^2 x^{\alpha_i - 1}}{[(1-\eta)^2 + \eta x]^{\alpha_i + 1}},$$

which is non-negative for $x \ge 0$. Furthermore,

$$\lim_{x \to \infty} p_i(x) = \eta^{-\alpha_i}$$

and thus

$$0 \le p_i(x) \le \eta^{-\alpha_i}, \quad \text{for all } x \ge 0.$$

Hence,

$$0 \le \prod_{i=1}^{r} p_i(\delta_i) \le \eta^{-\sum_{i=1}^{r}\alpha_i} \le 2^{\sum_{i=1}^{r}\alpha_i} \le 2^r, \quad \text{since } \eta \ge \frac{1}{2} \quad \text{and} \quad 0 \le \alpha_i < 1 \quad \text{for } i = 1, \ldots, r.$$

Note that

$$\sup_{0 \le x \le 2^r} |x-1| = \max\left\{ \sup_{0 \le x \le 1} |x-1|, \quad \sup_{1 \le x \le 2^r} |x-1| \right\} = \max\{1, 2^r - 1\} \equiv K_0.$$

Consequently, $0 \le h_\eta(\omega) \le K_0$, for any $\eta \in [\frac{1}{2}, 1]$ and $\omega \in (-\pi, \pi]$. With this bound, we obtain

$$\int_{A_{ij}} f(\omega) h_\eta(\omega) g(\omega) \mathrm{d}\omega \le K_0 \int_{A_{ij}} f(\omega) g(\omega) \mathrm{d}\omega. \tag{A4}$$

However, over A_{ij},

$$\omega^{-\alpha} \prod_{i' \ne i, j' \ne j} |\omega - \omega_{i'j'}|^{-\alpha'_i} \le K_{ij} < \infty,$$

where K_{ij} does not depend on $(\alpha, \alpha_1, \ldots, \alpha_r)$, since the only pole that the set A_{ij} contains is ω_{ij} and $(\alpha, \alpha_1, \ldots, \alpha_r)$ belong to the compact set Θ. Thus,

$$\int_{A_{ij}} f(\omega) g(\omega) \mathrm{d}\omega \le K_{ij} \int_{A_{ij}} |\omega - \omega_{ij}|^{-\alpha_i} g(\omega) \mathrm{d}\omega. \tag{A5}$$

Since $g(\omega) \le M_g$, we have

$$\int_{A_{ij}} |\omega - \omega_{ij}|^{-\alpha_i} g(\omega) \mathrm{d}\omega \le 2 M_g (1-\eta)^{\frac{1-\alpha_i}{m}}. \tag{A6}$$

Combining (A4), (A5) and (A6), we have

$$\int_{A_{ij}} f(\omega) h_\eta(\omega) g(\omega) \mathrm{d}\omega \le K M_g (1-\eta)^{\frac{1-\alpha_i}{m}},$$

where $K = 2K_0 \max_{ij}\{K_{ij}\}$. Note that this constant is finite and does not depend on f. Moreover, since $0 \le 1 - \eta < 1$,

$$(1-\eta)^{\frac{1-\alpha_i}{m}} \le (1-\eta)^p, \quad \text{for } i = 1, \ldots, r, \text{ where } 0 < p = \frac{(1 - \max \alpha_i)}{m} < 1.$$

Thus, we obtain inequality (A2).

To prove eqn (A3), we first establish an upper bound for $h_\eta(\omega)$ for $\omega \in \bar{A}$. Observe that over

$$\bar{A}, |\omega - \omega_{ij}| > (1-\eta)^{\frac{1}{m}}, \quad \text{for } i = 1, \ldots, r, \; j = 1, \ldots, m_i.$$

Thus,

$$\delta_i = \prod_{j=1}^{m_i} |\omega - \omega_{ij}| \ge (1-\eta)^{\frac{m_i}{m}} \ge 1 - \eta.$$

But, as shown previously, $p_i(x)$ is a nondecreasing function. Consequently, $p_i(\delta_i) \ge p_i(1-\eta) = 1$. However,

$$p_i(x) \leq \lim_{x\to\infty} p_i(x) = \eta^{-\alpha_i}.$$

Thus,

$$1 \leq \prod_{i=1}^{r} p_i(\delta_i) \leq \eta^{-\sum \alpha_i}.$$

Hence

$$h_\eta(\omega) = \left|\prod_{i=1}^{r} p_i(\delta_i) - 1\right| = \prod_{i=1}^{r} p_i(\delta_i) - 1$$

and therefore

$$0 \leq h_\eta(\omega) \leq \eta^{-\sum \alpha_i} - 1, \quad \text{for } \omega \in \bar{A}.$$

With this upper bound, we obtain

$$\begin{aligned}\int_{\bar{A}} f(\omega)h_\eta(\omega)g(\omega)\mathrm{d}\omega &\leq (\eta^{-\sum \alpha_i} - 1)\int_{\bar{A}} f(\omega)g(\omega)\mathrm{d}\omega \\ &\leq (\eta^{-\sum \alpha_i} - 1)M_g \int_{-\pi}^{\pi} f(\omega)\mathrm{d}\omega \\ &= q(\eta)(1-\eta)^p M_g \int_{-\pi}^{\pi} f(\omega)\mathrm{d}\omega,\end{aligned}$$

where

$$q(\eta) = \frac{(\eta^{-\sum \alpha_i} - 1)}{(1-\eta)^p} \quad \text{and} \quad 0 < p = \frac{(1 - \max \alpha_i)}{m} < 1.$$

Since $q(\eta)$ is a continuous function on $[\frac{1}{2}, 1)$, an application of L'Hospital's rule yields

$$\lim_{\eta\to 1} q(\eta) = 0 \quad \text{if } 0 < p < 1 \quad \text{and} \quad \lim_{\eta\to 1} q(\eta) = \sum \alpha_i \quad \text{if } p = 1.$$

Consequently, $q(\eta)$ is a bounded function for $\eta \in [\frac{1}{2}, 1]$ and there exists a constant $K_1 > 0$ such that $q(\eta) \leq K_1$ for $\eta \in [\frac{1}{2}, 1]$. Hence,

$$\int_{\bar{A}} f(\omega)h_\eta(\omega)g(\omega)\mathrm{d}\omega \leq K_1 M_g (1-\eta)^p \int_{-\pi}^{\pi} f(\omega)\mathrm{d}\omega.$$

Now, since Θ is compact and $\int_{-\pi}^{\pi} f_{\boldsymbol{\theta}}(\omega)\mathrm{d}\omega$ is a continuous function of $\boldsymbol{\theta} \in \Theta$, there exists a positive constant K_2 not depending on $f_{\boldsymbol{\theta}}$ such that $\int_{-\pi}^{\pi} f_{\boldsymbol{\theta}}(\omega)\mathrm{d}\omega \leq K_2$. Hence, the lemma follows with $K = K_1 K_2$. □

Lemma 2 shows the uniform convergence of the approximated variance-covariance matrix to T as η tends to 1.

Lemma 2.

(*a*) $|T| \leq Kn$, *where the constant* K *does not depend on* $\boldsymbol{\theta}$.
(*b*) *Let* p *be given as in Lemma 1 and choose* $\delta > 0$ *such that* $\delta > \frac{1}{p}$. *Define* $\eta = \eta(n) = 1 - n^{-\delta}$. *Then* $||T_\eta - T|| \leq Kn^{1-\delta p}$ *and consequently* $||T_\eta - T|| \to 0$ *uniformly in* $\boldsymbol{\theta}$ *as* $n \to \infty$.

Proof. For part (a), observe that

$$|T(f_{\boldsymbol{\theta}})|^2 = n\sum_{k=0}^{n-1}\left(1-\frac{k}{n}\right)\gamma_{\boldsymbol{\theta}}^2(k) \le n\sum_{k=0}^{n}\gamma_{\boldsymbol{\theta}}^2(k).$$

An application of Lemma 1 of Oppenheim *et al.* (2000) to the spectral density defined by (1) establishes that $\gamma_{\boldsymbol{\theta}}(k) \le K(\boldsymbol{\theta})k^{\tilde{\alpha}-1}$, where $\tilde{\alpha} = \max\{\alpha, \alpha_1, \ldots, \alpha_r\}$. Since Θ is compact, there exists a constant K not depending on $\boldsymbol{\theta}$ such that $\gamma_{\boldsymbol{\theta}}(k) \le Kk^{\tilde{\alpha}-1}$. Therefore,

$$\sum_{k=0}^{n}\gamma_{\boldsymbol{\theta}}^2(k) \le K\sum_{k=0}^{n}k^{2\tilde{\alpha}-2}.$$

. Hence, for

$$\tilde{\alpha} < \frac{1}{2}, \quad \sum_{k=0}^{\infty}\gamma_{\boldsymbol{\theta}}^2(k) < \infty,$$

for

$$\tilde{\alpha} = \frac{1}{2}, \quad \sum_{k=0}^{n}\gamma_{\boldsymbol{\theta}}^2(k) \le K\log(n)$$

and for

$$\tilde{\alpha} > \frac{1}{2}, \quad \sum_{k=0}^{n}\gamma_{\boldsymbol{\theta}}^2(k) < Kn.$$

Thus, we conclude that

$$\sum_{k=0}^{n}\gamma_{\boldsymbol{\theta}}^2(k) < Kn, \quad \text{for any } 0 \le \tilde{\alpha} < 1$$

and the result follows.

For part (b), let $h(\omega) = f(\omega) - f_\eta(\omega)$. Thus

$$\|T_\eta - T\| = \|T(h)\| \quad \text{and} \quad \|T_\eta - T\| \le \|T(h_+)\| + \|T(h_-)\|.$$

Since $T(h_+)$ and $T(h_-)$ are semi-definite positive, we have $\|T(h_+)\| \le \|T(h_+)^{\frac{1}{2}}\|^2 = \sup_{\|\mathbf{X}\|=1}\mathbf{X}'T(h_+)\mathbf{X}$. But, for all $\mathbf{X} \in \mathbb{R}^n$ such that $\mathbf{X}'\mathbf{X} = 1$ we have

$$\begin{aligned}\mathbf{X}'T(h_+)\mathbf{X} &= \int_{-\pi}^{\pi} h_+(\omega)\left|\sum_{j=1}^{n} \mathrm{e}^{\mathrm{i}\omega j}X_j\right|^2 \mathrm{d}\omega \\ &\le \int_{-\pi}^{\pi} |f_{\eta,\boldsymbol{\theta}}(\omega) - f_{\boldsymbol{\theta}}(\omega)|\left|\sum_{j=1}^{n} \mathrm{e}^{\mathrm{i}\omega j}X_j\right|^2 \mathrm{d}\omega.\end{aligned}$$

Now, by the Cauchy–Schwarz inequality we have

$$g(\omega) = \left|\sum_{j=1}^{n} \mathrm{e}^{\mathrm{i}\omega j}X_j\right|^2 \le n \equiv M_g$$

and then

$$\sup_{\|\mathbf{X}\|=1} \mathbf{X}'T(h_+)\mathbf{X} \le K|\eta - 1|^p n,$$

by virtue of Lemma 1. Similarly,

$$\sup_{\|\mathbf{X}\|=1} \mathbf{X}'T(h_-)\mathbf{X} \le K|\eta - 1|^p n.$$

Therefore, $||T_\eta - T|| \le Kn^{1-\delta p}$. Since, $1 - \delta p < 0$, $||T_\eta - T|| \to 0$ uniformly in $\boldsymbol{\theta}$ as $n \to \infty$. □

Lemma 3. *Let $0 < \alpha, \beta < 1$, $m = 2\delta r + 1$, $\eta = \eta(n) = 1 - n^{-\delta}$ for some δ such that $1 < \delta p$ where p is given in Lemma 2,*

$$f_\eta(\omega) = H_1(\omega)|\omega|^{-\alpha} \prod_{i=1}^{r} \left[(1-\eta)^2 + \eta \prod_{j=1}^{m_i} |\omega - \omega_{ij}| \right]^{-\alpha_i},$$

$$g_\eta(\omega) = H_2(\omega)|\omega|^{-\beta} \prod_{i=1}^{r} \left[(1-\eta)^2 + \eta \prod_{j=1}^{m_i} |\omega - \omega_{ij}| \right]^{-\alpha_i},$$

where $H_1(\omega)$, $H_2(\omega)$ and ω_{ij}, $i = 1, \ldots, r, j = 1, \ldots, m_i$ satisfy conditions (1). Then,

$$||T(f_\eta)^{-1/2} T(g_\eta)^{1/2}|| = O(n^{\max\{\frac{m}{2}(\beta-\alpha),0\}}).$$

Proof. Let

$$\xi(\omega) = \prod_{i=1}^{r} [(1-\eta)^2 + \eta \prod_{j=1}^{m_i} |\omega - \omega_{ij}|]^{-\alpha_i}.$$

Then,

$$||T(f_\eta)^{-1/2} T(g_\eta)^{1/2}||^2 \le K \sup_{\|\mathbf{X}\|=1} \frac{\int_{-\pi}^{\pi} |\omega|^{-\beta} \xi(\omega) |\sum_{j=1}^{n} e^{i\omega j} X_j|^2 d\omega}{\int_{-\pi}^{\pi} |\omega|^{-\alpha} \xi(\omega) |\sum_{j=1}^{n} e^{i\omega j} X_j|^2 d\omega}.$$

Consider the following function

$$h(\omega) = \frac{\xi(\omega) |\sum_{j=1}^{n} e^{i\omega j} X_j|^2}{\int_{-\pi}^{\pi} \xi(\omega) |\sum_{j=1}^{n} e^{i\omega j} X_j|^2 d\omega}.$$

Observe that h is non-negative and

$$\int_{-\pi}^{\pi} h(\omega) d\omega = 1.$$

Furthermore, given that

$$(1-\eta)^2 + \eta \prod_{j=1}^{m_i} |\omega - \omega_{ij}| \geq (1-\eta)^2$$

and $\alpha_i > 0$ for $i = 1, \ldots, r$ we have,

$$[(1-\eta)^2 + \eta \prod_{j=1}^{m_i} |\omega - \omega_{ij}|]^{-\alpha_i} \leq (1-\eta)^{-2\alpha_i}.$$

Hence, $\xi(\omega) \leq (1-\eta)^{-2\sum_{i=1}^{r}\alpha_i}$. Since $1 - \eta = n^{-\delta}$ we obtain

$$\xi(\omega) \leq n^{2\delta \sum_{i=1}^{r}\alpha_i}.$$

As $0 \leq \alpha_i < 1$ for $i = 1, \ldots, r$, we establish the upper bound $\xi(\omega) \leq n^{2\delta r}$. In addition,

$$\left| \sum_{j=1}^{n} \mathrm{e}^{\mathrm{i}\omega j} X_j \right|^2 \leq n$$

for all $||\mathbf{X}|| = 1$. Thus, the numerator of h is bounded by $Kn^{2\delta r+1}$. On the contrary, by a similar argument as above it can be shown that $\xi(\omega) \geq C^{-r}$ for some constant $C \geq 1$. Thus, a lower bound for the denominator is C^{-r}. Consequently, $h(\omega) \leq C^r n^{2\delta r+1}$. Thus,

$$||T(f_\eta)^{-1/2} T(g_\eta)^{1/2}||^2 \leq K \sup_{h\in\mathcal{P}} \frac{\int_{-\pi}^{\pi} |\omega|^{-\beta} h(\omega)\mathrm{d}\omega}{\int_{-\pi}^{\pi} |\omega|^{-\alpha} h(\omega)\mathrm{d}\omega},$$

where

$$\mathcal{P} = \{h : h \text{ is a probability density on } [-\pi, \pi] \text{ with } h(\omega) \leq n^m, m = 2\delta r + 1\}.$$

Note that the supremum of the above inequality is attained at the function $h^* \in \mathcal{P}$, where $h^*(\omega) = n^m$ when $|\omega| \leq \frac{1}{2n^m}$ and zero otherwise. At such an h^*,

$$||T(f_\eta)^{-1/2} T(g_\eta)^{1/2}|| \leq Kn^{\max\{\frac{m}{2}(\beta-\alpha),0\}}.$$

This completes the proof. □

Lemma 4. *Let f_j and g_j, $j = 1, \ldots, q$, be spectral densities satisfying* (1) *with parameters* $(\alpha, \alpha_1, \ldots, \alpha_r)$ *and* $(\beta, \beta_1, \ldots, \beta_r)$, *respectively. Assume that* $\beta - 2\alpha \leq 0$ *and* $\beta_i - 2\alpha_i \leq 0$, *for* $i = 1, \ldots, r$. *Consider the spectral densities* f_j^η *and* g_j^η *satisfying the corresponding approximations* (5). *Then*

$$\left| \frac{1}{n} tr\left[\prod_{j=1}^{q} T(\{4\pi^2 f_j\}^{-1}) T(g_j) \right] - \frac{1}{n} tr\left[\prod_{j=1}^{q} T(\{4\pi^2 f_j^\eta\}^{-1}) T(g_j^\eta) \right] \right| \leq Kn^{1-\delta p},$$

for all $\eta \geq 1 - n^{-\delta}$, *where K is a positive constant and* δ *and* $p \in (0,1]$ *are as in Lemma* 2.

Proof. We prove the lemma for $q = 1$ since the other cases can be shown similarly. Let $f = f_1$ and $g = g_1$ and

$$\Delta_n(\omega) = \sum_{j=1}^{n} e^{i\omega j}.$$

Note that

$$\mathrm{tr}\left[T(f^{-1})T(g)\right] = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} |\Delta_n(\omega-\lambda)|^2 f^{-1}(\omega)g(\lambda)\mathrm{d}\omega\,\mathrm{d}\lambda.$$

Consequently,

$$\begin{aligned}
&|\mathrm{tr}[T(\{4\pi^2 f\}^{-1})T(g)] - \mathrm{tr}[T(\{4\pi^2 f^{\eta}\}^{-1})T(g^{\eta})]| \\
&\quad \le K\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} |\Delta_n(\omega-\lambda)|^2 |f(\omega)^{-1}g(\lambda) - \{f^{\eta}(\omega)\}^{-1}g^{\eta}(\lambda)|\mathrm{d}\omega\,\mathrm{d}\lambda \\
&\quad \le K\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} |\Delta_n(\omega-\lambda)|^2 f(\omega)^{-1}|g(\lambda) - g^{\eta}(\lambda)|\mathrm{d}\omega\,\mathrm{d}\lambda \\
&\qquad + K\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} |\Delta_n(\omega-\lambda)|^2 f(\omega)^{-1}f^{\eta}(\omega)^{-1}g^{\eta}(\lambda)|f(\omega) - f^{\eta}(\omega)|\mathrm{d}\omega\,\mathrm{d}\lambda.
\end{aligned}$$

But, $f(\omega)^{-1} \le K$ and $f^{\eta}(\omega)^{-1} \le K$, hence

$$\begin{aligned}
&|\mathrm{tr}[T(\{4\pi^2 f\}^{-1})T(g)] - \mathrm{tr}[T(\{4\pi^2 f^{\eta}\}^{-1})T(g^{\eta})]| \\
&\quad \le K\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} |\Delta_n(\omega-\lambda)|^2 |g(\lambda) - g^{\eta}(\lambda)|\mathrm{d}\omega\ \mathrm{d}\lambda \\
&\qquad + K\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} |\Delta_n(\omega-\lambda)|^2 g^{\eta}(\lambda)|f(\omega) - f^{\eta}(\omega)|\mathrm{d}\omega\ \mathrm{d}\,\lambda.
\end{aligned}$$

Observe that $|\Delta_n(\omega - \lambda)| \le n$ for all ω and λ. Therefore, by virtue of Lemma 1 we have

$$\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} |\Delta_n(\omega-\lambda)|^2 |g(\lambda) - g^{\eta}(\lambda)|\mathrm{d}\omega\,\mathrm{d}\lambda \le Kn^2|1-\eta|^p,$$

and

$$\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} |\Delta_n(\omega-\lambda)|^2 g^{\eta}(\lambda)|f(\omega) - f^{\eta}(\omega)|\mathrm{d}\omega\,\mathrm{d}\lambda \le Kn^2|1-\eta|^p \int_{-\pi}^{\pi} g^{\eta}(\lambda)\mathrm{d}\lambda.$$

Observe that g^{η} is integrable for all $\eta \in [0, 1]$. Hence,

$$\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} |\Delta_n(\omega-\lambda)|^2 g^{\eta}(\lambda)|f(\omega) - f^{\eta}(\omega)|\mathrm{d}\omega\,\mathrm{d}\lambda \le Kn^2|1-\eta|^p.$$

Therefore,

$$n^{-1}|\mathrm{tr}[T(\{4\pi^2 f\}^{-1})T(g)] - \mathrm{tr}[T(\{4\pi^2 f^{\eta}\}^{-1})T(g^{\eta})]| \le n|1-\eta|^p$$

and the result is obtained for all $\eta \ge 1 - n^{-\delta}$. □

LEMMA 5. *Let f_j and g_j, $j = 1, \ldots, q$, be spectral densities satisfying eqn* (1) *with parameters* $(\alpha, \alpha_1, \ldots, \alpha_r)$ *and* $(\beta, \beta_1, \ldots, \beta_r)$, *respectively* (*same parameters for all j*). *Assume that* $q(\beta - \alpha) < \frac{1}{2}$ *and* $q(\beta_i - \alpha_i) < \frac{1}{2}$, *for* $i = 1, \ldots, r$. *Then*

$$\lim_{n\to\infty}\frac{1}{n}\mathrm{tr}\left[\prod_{j=1}^{q}T(f_j)^{-1}T(g_j)\right]=\frac{1}{2\pi}\int_{-\pi}^{\pi}\left\{\prod_{j=1}^{q}\frac{g_j(\omega)}{f_j(\omega)}\right\}\mathrm{d}\omega. \tag{A7}$$

Proof. Proceeding analogously to the proof of Theorem 5.1 of Dahlhaus (1989), it suffices to prove that

$$\lim_{n\to\infty}\frac{1}{n}\mathrm{tr}\left[\prod_{j=1}^{k}T(\{4\pi^2 f_j\}^{-1})T(g_j)\right]=\frac{1}{2\pi}\int_{-\pi}^{\pi}\left\{\prod_{j=1}^{k}\frac{g_j(\omega)}{f_j(\omega)}\right\}\mathrm{d}\omega, \tag{A8}$$

for any integer k such that $1\le k\le 2q$. Observe that this limit is well defined if $2q(\beta-\alpha)<1$ and $2q(\beta_i-\alpha_i)<1$, for $i=1,\ldots,r$. In order to establish eqn (A8), define the following quantities:

$$a_n=\frac{1}{n}\mathrm{tr}\left[\prod_{j=1}^{k}T(\{4\pi^2 f_j\}^{-1})T(g_j)\right],\quad a=\frac{1}{2\pi}\int_{-\pi}^{\pi}\left\{\prod_{j=1}^{k}\frac{g_j(\omega)}{f_j(\omega)}\right\}\mathrm{d}\omega,$$

$$a_n^\eta=\frac{1}{n}\mathrm{tr}\left[\prod_{j=1}^{k}T(\{4\pi^2 f_j^\eta\}^{-1})T(g_j^\eta)\right],\quad a^\eta=\frac{1}{2\pi}\int_{-\pi}^{\pi}\left\{\prod_{j=1}^{k}\frac{g_j^\eta(\omega)}{f_j^\eta(\omega)}\right\}\mathrm{d}\omega.$$

Then,

$$|a_n-a|\le|a_n-a_n^\eta|+|a_n^\eta-a^\eta|+|a^\eta-a|.$$

Now, for all $\eta\ge 1-n^{-\delta}$, with δ as in Lemma 2, we obtain from Lemma 4 that, $|a_n-a_n^\eta|\le Kn^{1-\delta p}$ and by Lemma 5.3.2 of Taniguchi and Kakizawa (2000), $\limsup_{n\to\infty}|a_n^\eta-a^\eta|=0$. Furthermore, using arguments similar to proving Lemma 1, we conclude that given $\epsilon>0$, $|a^\eta-a|<\epsilon$ for η in a neighbourhood of 1. Consequently, $\limsup_{n\to\infty}|a_n-a|\le\epsilon$. Since ϵ can be chosen arbitrarily small, the result holds as $\eta\to 1$.□

Lemma 6. *Consider the following three matrices:*

$$A_{\boldsymbol\theta}^{(1)}=T(f_{\boldsymbol\theta})^{-1}T(\nabla f_{\boldsymbol\theta})T(f_{\boldsymbol\theta})^{-1}T(\nabla f_{\boldsymbol\theta})T(f_{\boldsymbol\theta})^{-1},$$

$$A_{\boldsymbol\theta}^{(2)}=T(f_{\boldsymbol\theta})^{-1}T(\nabla^2 f_{\boldsymbol\theta})T(f_{\boldsymbol\theta})^{-1},\quad A_{\boldsymbol\theta}^{(3)}=T(f_{\boldsymbol\theta})^{-1}T(\nabla f_{\boldsymbol\theta})T(f_{\boldsymbol\theta})^{-1}.$$

Then, for some $\epsilon>0$, $i=1,2,3$ *and for all* $\mathbf{X}\in\mathbb{R}^k$,

$$(a)\ |\,\|A_{\boldsymbol\theta}^{(i)}\|\,|\le Kn^{\epsilon},\quad (b)\ |\,\|\nabla A_{\boldsymbol\theta}^{(i)}\|\,|\le Kn^{\epsilon},$$

$$(c)\ |\mathbf{X}'A_{\boldsymbol\theta}^{(i)}\mathbf{X}|\le K\mathbf{X}'\mathbf{X}n^{\epsilon},\quad (d)\ |\mathbf{1}'A_{\boldsymbol\theta}^{(i)}\mathbf{1}|\le Kn^{1-\alpha(\boldsymbol\theta+\epsilon)},$$

where $\mathbf{1}=(1,\ldots,1)'$.

PROOF. We only prove part (a) for the case $i = 1$, the other cases are derived similarly. Note that,

$$|||A_{\boldsymbol{\theta}}^{(i)}|||^2 = \sum_{j=1}^{q} |||A_{\boldsymbol{\theta}_j}^{(i)}||^2,$$

and

$$||A_{\boldsymbol{\theta}_j}^{(1)}||^2 = ||T^{-1}T_{\partial_j}T^{-1}T_{\partial_j}T^{-1}\ ||^2,$$

where $T = T(f)$ and $T_{\partial_j} = T(\partial f/\partial \boldsymbol{\theta}_j)$. Consequently,

$$||A_{\boldsymbol{\theta}_j}^{(1)}|| \leq ||T^{-1}||\ ||T^{-1}T_{\partial_j}||^2.$$

Since f is bounded from below

$$(f(\omega) > C > 0),\ ||T^{-1}|| \leq K \quad \text{so that } ||A_{\boldsymbol{\theta}_j}^{(1)}|| \leq K||T^{-1}T_{\partial_j}||^2.$$

Let $g(\omega) = \partial f/\partial \boldsymbol{\theta}_j$. Then,

$$\begin{aligned}\|T(f)^{-1}T(g)\| &\leq \|T(f)^{-1}T(g) - T(f)^{-1}T(g_\eta)\| + \|T(f_\eta)^{-1}T(g_\eta)\| \\ &\quad + \|T(f)^{-1}\|\ \|T(f_\eta)^{-1}T(g_\eta)\|\ \|T(f) - T(f_\eta)\| \\ &\leq K\|T(g) - T(g_\eta)\| + \|T(f_\eta)^{-1}T(g_\eta)\| \\ &\quad + K\|T(f_\eta)^{-1}T(g_\eta)\|\ \|T(f) - T(f_\eta)\|,\end{aligned}$$

where $\eta = \eta(n) = 1 - n^{-\delta}$, for some δ such that $\delta p > 1$ with p as in Lemma 1. Now, by Lemma 2(b), $||T(g) - T(g_\eta)||$ and $||T(f) - T(f_\eta)||$ are both bounded by $Kn^{1-\delta p}$. On the other hand, by Lemma 3, $||T(f_\eta)^{-1}T(g_\eta)||$ is bounded by $Kn^{m\delta_0}$, for some $\delta_0 > 0$. Consequently, $||T(f)^{-1}T(g)|| \leq Kn^{m\delta_0}$, and therefore by taking $\epsilon = 2m\delta_0$ the result is obtained. □

Lemma 7 is needed to show the equicontinuity of a class of quadratic forms. In turn, this property is needed to prove the central limit theorem for the MLE.

LEMMA 7. *For $i = 1,2$, let*

$$Z_n^{(i)} = \frac{1}{n}X'A_{\boldsymbol{\theta}}^{(i)}X - \frac{1}{n}\text{tr}\{A_{\boldsymbol{\theta}}^{(i)}T(f_{\boldsymbol{\theta}})\}.$$

Then, there exists a constant D such that for all $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in P$, and for all $\kappa > 0$,

$$P(|Z_n^{(i)}(\boldsymbol{\theta}_1) - Z_n^{(i)}(\boldsymbol{\theta}_2)| > \kappa|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2|) \leq 4J^2e^{-\kappa/D}.$$

PROOF. It suffices to prove that there exists a constant C which is independent of $\boldsymbol{\theta}_1,\boldsymbol{\theta}_2$ and l such that

$$\left|\text{cum}_l\left(\frac{\{Z_n^{(i)}(\boldsymbol{\theta}_1) - Z_n^{(i)}(\boldsymbol{\theta}_2)\}_{jk}}{|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2|}\right)\right| \leq l!C^l, \tag{A9}$$

where $\text{cum}_l(Z)$ denotes the l-cumulant of the random variable Z. Now, by Lemma 2(a),

$$|T(f_{\boldsymbol{\theta}_0})| \leq Kn. \tag{A10}$$

Consequently, in order to derive (5.20), it suffices to prove that

$$||(A_{\boldsymbol{\theta}_1}^{(i)} - A_{\boldsymbol{\theta}_2}^{(i)})_{jk}|| \leq |\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2| C(\delta) n^{\delta}. \qquad (A11)$$

Note that by the mean value theorem and part (b) of Lemma 6, we have

$$||(A_{\boldsymbol{\theta}_1}^{(i)} - A_{\boldsymbol{\theta}_2}^{(i)})_{jk}|| \leq |\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2| ||\nabla A_{\boldsymbol{\theta}}^{(i)}|| \leq |\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2| K n^{\delta}.$$

Combining eqns (A11) and (A10) with the arguments in Lemma 6.2 of Dahlhaus (1989), eqn (A9) is proved. □

ACKNOWLEDGEMENTS

The authors would like to express their gratitude to a referee and an Associate Editor for helpful and constructive comments, which led to a substantially improved version of this paper. This research was supported in part by Research Grant Council of Hong Kong CUHK400305, CUHK4043/02P and Fondecyt Grant 1040934.

REFERENCES

ABRAHAMS, M. and DEMPSTER, A. (1979) Research on seasonal analysis, progress report on the ASA/ Census Project on seasonal adjustment. Technical Report, Dept. Statistics, Harvard University.

ARTECHE, J., and ROBINSON, P. (2000) Semiparametric inference in seasonal and cyclical long memory processes. *Journal of Time Series Analysis* 21, 1–25.

BASAWA, I. V., FEIGIN, P. and HEYDE, C. C. (1976) Asymptotic properties of maximum likelihood estimators for stochastic processes. *Sankhyā, Series* A 38, 259–70.

BERAN, J. (1994) *Statistics for Long Memory Processes*. Chapman and Hall, New York.

BERTELLI, S. and CAPORIN, M. (2002) A note on calculating autocovariances of long-memory processes. *Journal of Time Series Analysis*, 23, 503–8.

BROCKWELL, P. J. and DAVIS, R. A. (1991) *Times Series: Theory and Methods*, 2nd edn. Springer-Verlag, New York.

CHAN, N. H. and PALMA, W. (1998) State space modeling of long memory processes. *Annals of Statistics* 26, 719–40.

CHUNG, C. F. (1996) A generalized fractionally integrated autoregressive moving average process. *Journal of Time Series Analysis* 17, 111–40.

DAHLHAUS, R. (1989) Efficient parameter estimation for self-similar processes. *Annals Statistics* 17, 1749–66.

FOX, R. and TAQQU, M. (1987) Central limit theorems for quadratic forms in random variables having long-range dependence. *Probability Theory of Related Fields* 74, 213–240.

GIRAITIS, L., HIDALGO, J. and ROBINSON, P. M. (2001) Gaussian estimation of parametric spectral density with unknown pole. *Annals Statistics* 29, 987–1023.

GIRAITIS, L. and LEIPUS, R. (1995) A generalized fractionally differencing approach in long memory modelling. *Lithuanian Mathematical Journal* 35, 53–65.

GRAY, H. L., ZHANG, N. F. and WOODWARD, W. A. (1989) On generalized fractional processes. *Journal of Time Series Analysis* 10, 233–57.

HANNAN, E.J. (1973) The asymptotic theory of linear time-series models. *Journal of Applied Probability* 10, 130–45.

HASSLER, U. (1994) (Mis) specification of long memory in seasonal time series. *Journal of Time Series Analysis* 15, 19–30.

HASSLER, U. and WOLTERS, J. (1995) Long memory in inflation rates: international evidence. *Journal of Business Economics and Statistics* 13, 37–46.

JONAS, A. J. (1979) Persistent Memory Random Processes. Unpublished Ph.D. dissertation, Dept. Statistics, Harvard University.

LEIPUS, R., and VIANO, M. C. (2000) Modelling long-memory time series with finite or infinite variance: a general approach. *Journal of the Time Series Analysis* 21, 61–74.

LING, S., and LI, W. K. (1997) On fractionally integrated autoregressive moving-average time series models with conditional heteroscedasticity. *Journal of the American Statistical Association*, 92, 1184–94.

MONTANARI, A., ROSSO, R., and TAQQU, M. (2000) A seasonal fractional ARIMA model applied to the Nile River monthly flows at Aswan. *Water Resources Research* 36, 1249–59.

OOMS, M. (1995) Flexible seasonal long memory and economic time series. Technical Report, Econometric Institute, Erasmus University.

OPPENHEIM, G., OULD HAYE, M. and VIANO, M.C. (2000) Long memory with seasonal effects. *Statistical Inference for Stochastic Processes* 3, 53–68.

OULD HAYE, M. (2002) Asymptotic behaviour of the empirical process for seasonal long-memory data. European Series in Applied and Industrial Mathematics (ESAIM): P&S 6, 293–309.

PORTER-HUDAK, S. (1990) An application of the seasonal fractionally differenced model to the monetary aggregates. *Journal of the American Statistical Association* 85, 338–44.

RAY, B. K. (1993) Long range forecasting of IBM product revenues using a seasonal fractionally differenced ARMA model. *International Journal of Forecasting* 9, 255–69.

REISEN, V., RODRIGUES, A. and PALMA, W. (2005a) Estimation of seasonal fractionally integrated processes. *Computational Statistics and Data Analysis*, forthcoming.

REISEN, V., RODRIGUES, A. and PALMA, W. (2005b) Estimating seasonal long-memory processes. *Journal of Statistical Computation and Simulation*, forthcoming.

TANIGUCHI, M. and KAKIZAWA, Y. (2000) *Asymptotic Theory of Statistical Inference for Time Series*. Springer-Verlag, New York.

VELASCO, C. and ROBINSON, P. M. (2000) Whittle pseudo-maximum likelihood estimation for nonstationary time series. *Journal of American Statistical Association* 95, 1229–43.

WOODWARD, W. A., CHENG, Q. C. and GRAY, H. L. (1998) A k-factor GARMA long memory model. *Journal of Time Series Analysis*, 19, 485–504.

YAJIMA, Y. (1985) On estimation of long-memory time series models. *Australian Journal of Statistics* 27, 303–20.