doi:10.1111/j.1467-9892.2008.00576.x

IDENTIFICATION OF PERSISTENT CYCLES IN NON-GAUSSIAN LONG-MEMORY TIME SERIES

By Mohamed Boutahar

University of Méditerranée

First Version received April 2007

Abstract. Asymptotic distribution is derived for the least squares estimates (LSE) in the unstable AR(p) process driven by a non-Gaussian long-memory disturbance. The characteristic polynomial of the autoregressive process is assumed to have pairs of complex roots on the unit circle. In order to describe the limiting distribution of the LSE, two limit theorems involving long-memory processes are established in this article. The first theorem gives the limiting distribution of the weighted sum,

$$\sum_{k=1}^{n} c_{n,k}\varepsilon_k, \quad \text{where } \varepsilon_k = \sum_{j\le k} b_{k-j}u_j$$

is a non-Gaussian long-memory moving-average process and $(c_{n,k}, 1 \le k \le n)$ is a given sequence of weights; the second theorem is a functional central limit theorem for the sine and cosine Fourier transforms

$$\sum_{k=1}^{[nt]} \sin(k\theta)\varepsilon_k \quad \text{and} \quad \sum_{k=1}^{[nt]} \cos(k\theta)\varepsilon_k, \quad \text{where } \theta \in]\,0,\pi[\quad \text{and} \quad t \in [0,1].$$

Keywords. Autoregressive process; Brownian motion; cycles; functional central limit theorem; least squares estimates; long memory.

1. INTRODUCTION

Consider the univariate autoregressive model

$$\phi(B)y_t = \varepsilon_t, \tag{1}$$

where y_t is the tth observation on the dependent variable, $y_t = 0$ if $t \le 0$, $\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$ is the characteristic polynomial, B is the backward shift operator, i.e. $By_t = y_{t-1}$, and the disturbance process (ε_t) is given by

$$\varepsilon_t = \sum_{j\le t} b_{t-j}u_j, \tag{2}$$

where (u_j) is a sequence of independent and identically distributed (i.i.d.) random variables (not necessarily Gaussian) with zero mean and variance 1, (b_j) is a sequence which decays hyperbolically, i.e.

$$b_j = j^{H-\frac{3}{2}}L_1(j), \quad 0 < H < 1, \quad \sum_{j=0}^{\infty} b_j^2 < \infty, \tag{3}$$

and $L_1(\cdot)$ is a slowly varying function, bounded on every finite interval. For example, (ε_t) can be either a Gaussian fractional noise or a stationary and invertible autoregressive fractionally integrated moving average process (see Hosking, 1996).

The unknown parameter $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_p)'$ is estimated by the least squares estimate (LSE):

$$\hat{\boldsymbol{\phi}}_n = \left(\sum_{k=1}^{n} \mathbf{y}_{k-1}\mathbf{y}'_{k-1}\right)^{-1} \sum_{k=1}^{n} \mathbf{y}_{k-1} y_k, \tag{4}$$

where $\mathbf{y}_k = (y_k, \ldots, y_{k-p+1})'$. The least squares error satisfies

$$\hat{\boldsymbol{\phi}}_n - \boldsymbol{\phi} = \left(\sum_{k=1}^{n} \mathbf{y}_{k-1}\mathbf{y}'_{k-1}\right)^{-1} \sum_{k=1}^{n} \mathbf{y}_{k-1} \varepsilon_k. \tag{5}$$

If (ε_t) is a Gaussian long-memory process satisfying eqns (2) and (3) with $1/2 < H < 1$, then we can summarize the results, established in the literature, describing the behaviour of the LSE and compare them with the results obtained in the short-memory setup (i.e. (ε_t) is assumed to be an i.i.d. or a martingale difference sequence) as follows.

The behaviour of the estimation error depends on that of the matrix

$$M_n = \sum_{k=1}^{n} \mathbf{y}_{k-1}\mathbf{y}'_{k-1}$$

and the vector

$$V_n = \sum_{k=1}^{n} \mathbf{y}_{k-1}\varepsilon_k;$$

the normalizations needed for these quantities and the limiting distributions obtained depend on the characteristic polynomial $\phi(z)$, more precisely on the location of its roots:

1. Stable roots (i.e. $\phi(z) = 0$ implies that $|z| > 1$): In this case, $\hat{\boldsymbol{\phi}}_n - \boldsymbol{\phi}$ converges in probability to a nonzero limit, hence the LSE is inconsistent (see Chan and Terrin, 1995, Thm 3.1); this result differs from the one obtained when (ε_t) has short memory. Recall that under the short-memory assumption, the martingale transform V_n satisfies the assumptions of the central limit theorem, hence $V_n/\sqrt{n}$ converges in distribution to a Gaussian vector, the matrix M_n is normalized by n to obtain a deterministic limit; therefore the LSE is asymptotically normal.

2. Roots equal to 1 (i.e. $\phi(z) = (1-z)^a$): The normalizations of M_n and V_n are hyperbolic (e.g. if $a = 1$ then they are n^{2H+1} and n^{2H} for M_n and V_n, respectively; see Chan and Terrin, 1995, Thm 4.1), the limit of M_n is a stochastic integral of functionals of fractional Brownian motion with respect to Lebesgue measure and that of V_n is a multiple Wiener–Itô integral; the LSE is consistent with a rate of convergence equal to $O_p(n^{-1})$. In the case of short memory, the normalizations of M_n and V_n are polynomial (if $a = 1$ then they are n^2 and n for M_n and V_n, respectively; see Dickey and Fuller, 1979, for i.i.d disturbance and Chan and Wei, 1988, if (ε_t) is a martingale difference sequence), the limit of M_n (resp. of V_n) is a stochastic integral of functionals of Brownian motion with respect to the Lebesgue measure (resp. with respect to Brownian motion); the LSE is consistent with rate of convergence equal to $O_p(n^{-1})$.

The main difference between short and long memory in the normalization used and the limiting distribution obtained can be explained by using the following two results:

(i) If (ε_t) is an i.i.d. or a martingale difference sequence with respect to an increasing sequence of σ-algebras $F = (F_n)$ then we have the functional central limit theorem (FCLT; see Billinsgley, 1968; Hall and Heyde, 1980):

$$\frac{1}{\sqrt{n}}\sum_{k=1}^{[nt]} \varepsilon_k \Longrightarrow B(t), \tag{6}$$

(ii) If (ε_t) satisfies eqns (2) and (3), then the functional non-central limit theorem holds (see Taqqu, 1975):

$$\frac{1}{n^H}\sum_{k=1}^{[nt]} \varepsilon_k \Longrightarrow B_H(t), \tag{7}$$

$X_n \Longrightarrow X$ denotes the weak convergence of a sequence of random elements X_n in D to a random element X in D, and $D = D[0,1]$ is the space of random functions that are right-continuous and have left limits, endowed with the Skorohod topology, $B(t)$ is a Brownian motion and $B_H(t)$ is a fractional Brownian motion.

3. Roots equal to -1 or complex-conjugate unit roots (i.e. $\phi(z) = (1 + z)^b$ or $\phi(z) = \prod_{m=1}^{l}(1 - 2\cos\theta_m z + z^2)^{d_m}$): The normalizations of M_n and V_n are polynomial, the limit of M_n (resp. of V_n) is a stochastic integral of functionals of Brownian motion with respect to the Lebesgue measure (resp. with respect to Brownian motion); the LSE is consistent with a rate of convergence equal to $O_p(n^{-1})$. The same results are obtained in the short-memory setup (see Chan and Wei, 1988; Chan and Terrin, 1995).

4. Explosive roots (i.e. $\phi(z) = 0$ implies that $|z| < 1$): As in the short-memory setup, the normalizations of M_n and V_n are exponential and the limits are a mixture of normal distributions; the LSE is consistent with a rate of convergence equal to $O_p(\rho^n)$ for some $\rho < 1$ (see Boutahar, 2002).

In this article we follow Ahtola and Tiao (1987a,b), Chan and Wei (1988), Chan and Terrin (1995) and Gregoir (1999) to derive the limiting distribution of LSE of AR processes with complex-conjugate unit roots, the motivation being that usually the periodogram of seasonal time series exhibits peaks at seasonal frequencies

$$\theta_k = \frac{2\pi k}{s}, \quad k = 1, \ldots, [s/2],$$

where $s = 2$,4 and 12 for semi-annual, quarterly and monthly data, respectively. However, there are also many non-seasonal time series, for example annual data with cyclical movement, which similarly produce peaks at frequencies different from seasonal time series. Peaks at frequency $\theta = 0$ are often indicative of nonstationary (resp. stationary long memory) behaviour which can be removed by applying to data the unit root $1 - B$ [resp. the fractional unit root $(1 - B)^d$, $0 < d < 0.5$] operator. Peaks at low non-null frequencies imply the existence of cycles in the time series (see Conway and Frame, 2000; Birgean and Kilian, 2002 for economic data, and Priestley, 1981; Yiou *et al.* 1996, for other kinds of data). It is well known that persistent cycles can be described by complex unit roots. For instance, Bierens (2001) has concluded that National Bureau of Economic Research business cycles of the US unemployment time series are indeed because of complex-conjugate unit roots, i.e. an appropriate non-stationary model to describe the cyclical behaviour of such series is given by

$$\prod_{m=1}^{l}(1 - 2B\cos\theta_m + B^2)y_t = \varepsilon_t, \quad \text{where } 0 < \theta_1 < \cdots < \theta_l < \pi, \tag{8}$$

and (ε_t) is a stationary process. Equation (8) generates l persistent cycles of $2\pi/\theta_m$ periods, $1 \le m \le l$. Note that vanishing cycles can also be described by complex-conjugate, but stable, roots, i.e. $\rho e^{i\theta_m}$ and $\rho e^{-i\theta_m}$ with $|\rho| < 1$, and the corresponding model is stationary.

In model (8), with $l = 1$, Ahtola and Tiao (1987a) have established the limiting distribution of the LSE by assuming that (ε_t) is an i.i.d. Gaussian process. Chan and Wei (1988) have extended the result of Ahtola and Tiao (1987a) to a more general characteristic polynomial $\phi(z)$, which can also have stable roots (i.e. $\phi(z) = 0$ implies that $|z| > 1$) and roots equal to -1 and 1. Moreover, they relaxed (ε_t) to be a martingale difference sequence. Chan and Terrin (1995) have extended the result of Chan and Wei (1988) by assuming that (ε_t) is a Gaussian long-memory process, which implies that the errors ε_t are strongly correlated in the sense that their autocorrelation function is not absolutely summable; such a model is very useful to describe time series exhibiting both cyclical and long-memory properties. In Boutahar (2002), the results of Chan and Terrin (1995) were extended to the case where the roots of $\phi(z)$ are arbitrary. Unfortunately, the normality assumption of time series is usually violated in practice (see Gil-Alana, 2003; Scherrer *et al.*, 2007; Venema *et al.*, 2006; see also Tiku *et al.*, 2000 and the references therein). The aim of this article is to remove the normality hypothesis assumed in the article of Ahtola and Tiao (1987a) and in Chan and Terrin's (1995) particular model corresponding to complex-conjugate unit roots. More precisely, we consider the multiple cycles model (1)–(3) where

$$\begin{aligned} &\phi(z) = 1 - \sum_{i=1}^{p}\phi_i z^i = \prod_{m=1}^{l}\phi_{\theta_m}(z), \quad \phi_{\theta_m}(z) = (1 - 2\cos\theta_m z + z^2)^{d_m}, \\ &p = 2\sum_{m=1}^{l} d_m, \quad \theta_m \in]0,\pi[, \quad 1 \le m \le l. \end{aligned} \tag{9}$$

In this article we study only the case when the characteristic polynomial $\phi(z)$ is unstable with complex-conjugate unit roots, i.e. an appropriate non-stationary model to identify persistent cycles in non-Gaussian long-memory time series. However, the behaviour of the LSE when $\phi(z)$ has stable roots, roots equal to -1 and 1, and explosive roots remains an open problem.

This article is organized as follows. In Section 2 we give the limiting distribution of $\sum_{k=1}^{n} c_{n,k}\varepsilon_k$ and examine the particular cases of sine and cosine Fourier transforms of $\{\varepsilon_k, 1 \leq k \leq n\}$, for which we establish a FCLT. In Section 3 we consider the unstable AR(p) model with complex-conjugate roots and study the limiting distribution of the LSE. The proofs of the results of Sections 2 and 3 are given in the Appendix.

2. CLTS FOR LONG-MEMORY PROCESSES

Many central limit theorems (CLTs) were established for short-memory processes, such as i.i.d. sequence, martingale difference sequence, and so on. Such processes are weakly dependent and usually satisfy

$$\mathrm{var}\left(\sum_{k=1}^{n} \varepsilon_k\right) \sim \mathrm{C_n},$$

for some positive constant C, and hence we need to normalize the sum $\sum_{k=1}^{n} \varepsilon_k$ by $\sqrt{n}$ to obtain a Gaussian limiting distribution (see, e.g. Doukhan *et al.*, 2003 and the references therein). For long-memory processes, the normalization and/or the limit law are usually different from the short- memory setup; in this case, we say that (ε_t) satisfies a non-central limit theorem (non-CLT). Davydov (1970) has proved a non-CLT by assuming that the process (ε_t) is linear, i.e.

$$\varepsilon_t = \sum_{j \in \mathbb{Z}} b_j u_{t-j}.$$

Taqqu (1975), Dobruhsin and Major (1979) and Giraitis and Surgailis (1985) have considered the process $\varepsilon_t = G(Y_t)$, where G is a nonlinear function and (Y_t) is a Gaussian long-memory process. They proved a non-CLT for (ε_t); they proved also a CLT when (ε_t) has short memory. Surgailis (1982) and Avram and Taqqu (1987) have extended the results of Taqqu (1975), Dobruhsin and Major (1979) to the functional of non-Gaussian processes, they proved a non-CLT for $\varepsilon_t = A_m(Y_t)$ where A_m is the mth Appell polynomial associated with the distribution of Y_0 and (Y_t) is a long-memory moving average, i.e.

$$\varepsilon_t = \sum_{j \leq k} b_j u_{t-j}.$$

Finally Ho and Hsing (1997) have generalized the results of Surgailis (1982) and Avram and Taqqu (1987) to a large class of functions G. If $\varepsilon_t = G(Y_t)$

[resp. $\varepsilon_t = A_m(Y_t)$], then the limiting distribution depends on G (resp. A_m); it can be Gaussian or non-Gaussian and expressed as a multiple Wiener–Itô integral.

In this section we establish two CLTs for the causal long-memory process given by eqns (2) and (3). It can be shown that (ε_t) satisfies

$$\mathrm{var}\left(\sum_{k=1}^{n}\varepsilon_k\right) \sim C_1 n^{2H}, \tag{10}$$

for some positive constant C_1, and $a_n \sim b_n$ means that $a_n/b_n \rightarrow 1$ as $n \rightarrow \infty$.

In Theorem 1, we consider sequences of weights $\{c_{n,k}, 1 \le k \le n\}$ such that the weighted process $(c_{n,t}\varepsilon_t)$ has a short memory in the following time-domain sense:

$$\mathrm{var}\left(\sum_{k=1}^{n}c_{n,k}\varepsilon_k\right) \sim C_2 n, \text{ for some positive constant } C_2. \tag{11}$$

The weighted sum $\sum_{k=1}^{n} c_{n,k}\varepsilon_k$ was studied by Giraitis *et al.* (1996) who assumed in eqn (2) that $1/2 < H < 1$; they proved that $n^{-H}\sum_{k=1}^{n} c_{n,k}\varepsilon_k$ is asymptotically normal with asymptotic variance

$$\mathbf{Q}_n = n^{-2H}\mathrm{var}\left(\sum_{k=1}^{n}c_{n,k}\varepsilon_k\right).$$

However, if $(c_{n,t}\varepsilon_t)$ is of short memory then $\mathbf{Q}_n \rightarrow 0$ as $n \rightarrow \infty$ and the limiting distribution of $n^{-H}\sum_{k=1}^{n} c_{n,k}\varepsilon_k$ will be degenerate. Therefore, the limiting distribution of $\sum_{k=1}^{n} c_{n,k}\varepsilon_k$ cannot be obtained from Theorem 2 of Giraitis *et al.* (1996); in Theorem 1 we resolve this problem. In Theorem 2 we examine the particular weights $c_{n,k} = \sin(k\theta)$, $c_{n,k} = \cos(k\theta)$ and prove a FCLT for the two processes

$$X_n(t) = (nL(n))^{-1/2}\sum_{k=1}^{[nt]}\sin(k\theta)\varepsilon_k \quad \text{and} \quad Y_n(t) = (nL(n))^{-1/2}\sum_{k=1}^{[nt]}\cos(k\theta)\varepsilon_k.$$

Note that the process $(\sin(t\theta)\varepsilon_t)$ is not covariance-stationary and hence Davydov's (1970) results cannot be applied to obtain the weak convergence of X_n in the Skorohod space.

2.1. *A CLT for a weighted long-memory moving-average process*

Unless otherwise stated, limits are always taken as n tends to infinity in this article.

THEOREM 1. *Assume that the process* (ε_t) *is given by eqns* (2)–(3). *Let* $c_{n,k} \in \mathbb{R}^p$, $1 \le k \le n$, *be a sequence such that* $\|c_{n,k}\| < \infty$ *for all* $1 \le k \le n$,

$$(nL(n))^{-1}\operatorname{var}\left(\sum_{k=1}^{n} c_{n,k}\varepsilon_k\right) \to \mathbf{R}, \tag{12}$$

and for all $a \in \mathbb{R}^p$, $r \geq 3$,

$$\sum_{j\in\mathbb{Z}}\left(\sum_{k=1}^{n} a' c_{n,k} b_{k-j}\right)^r = o\Big((nL(n))^{r/2}\Big), \tag{13}$$

where $\mathbf{R}$ is a positive-definite matrix, with $b_i = 0$ if $i < 0$, and $L(\cdot)$ is a slowly varying function, bounded on every finite interval. Then

$$(nL(n))^{-1/2}\sum_{k=1}^{n} c_{n,k}\varepsilon_k \xrightarrow{\mathcal{L}} N(0,\mathbf{R}),$$

where $\xrightarrow{\mathcal{L}}$ denotes the convergence in distribution.

2.2. *A FCLT for the Fourier transform of long-memory moving-average process*

Let $D = D[0,1]$ be the space of random functions that are right-continuous and have left limits, endowed with the Skorohod topology. The weak convergence of a sequence of random elements X_n in D to a random element X in D is denoted by $X_n \Longrightarrow X$.

Consider the process (ε_t) given by eqns (2)–(3). For $\theta \in]0,\pi[$ and $t \in [0,1]$, let

$$X_n(t) = (nL(n))^{-1/2}\sum_{k=1}^{[nt]}\sin(k\theta)\varepsilon_k, \quad Y_n(t) = (nL(n))^{-1/2}\sum_{k=1}^{[nt]}\cos(k\theta)\varepsilon_k. \tag{14}$$

In Theorem 2 we prove that X_n converges in D to a Brownian motion B. There are two sufficient conditions for convergence in D (see Billingsley, 1968):

(i) the finite-dimensional distributions of X_n converge to the finite-dimensional distributions of B,
(ii) X_n is tight.

We prove that condition (i) holds if (ε_t) satisfies (2)–(3). However, for the tightness of X_n we impose an additional assumption, that is the white-noise (u_t) of the errors has at least a finite moment of order 4.

THEOREM 2. *Assume that the process (ε_t) is given by eqns (2)–(3) such that*

(i) $E(u_0^{2\kappa_0}) < \infty$ for some integer $\kappa_0 \geq 2$,
(ii) the spectral density of (ε_t) can be written as $f(\lambda) = |\lambda|^{1-2H}L(|\lambda|^{-1})$, where L is a slowly varying function, bounded on every finite interval. Then

$$X_n \Rightarrow K(\theta, H)B_1 \tag{15}$$

and

$$Y_n \Rightarrow K(\theta, H)B_2, \tag{16}$$

where $K(\theta, H) = \sqrt{\pi}|\theta|^{\frac{1}{2}-H}$, *and* B_1 and B_2 *are two standard Brownian motions.*

3. THE LSE IN UNSTABLE AR MODEL WITH COMPLEX-CONJUGATE ROOTS

Consider the AR(p) model (1)–(3) and (9). To study the limiting distribution of the LSE given by eqn (4) we use the same analysis as in Chan and Wei (1988) and Chan and Terrin (1995).

Let

$$x_t(m) = \phi_{\theta_m}(B)^{-1}\phi(B)y_t, \quad 1 \leq m \leq l.$$

Then there exists a nonsingular matrix $\mathbf{Q}$ (Chan and Wei, 1988, Appendix 1) such that

$$\mathbf{Q}\mathbf{y}_t = (\mathbf{x}_t'(1), \ldots, \mathbf{x}_t'(l))', \quad \text{where } \mathbf{x}_t(m) = (x_t(m), \ldots, x_{t-2d_m+1}(m))'.$$

Let $y_t(m, j) = (1 - 2\cos\theta_m B + B^2)^{d_m - j}x_t(m)$, c_j^i be the coefficient of z^i in the expansion of the polynomial $(1 - 2\cos\theta_m B + B^2)^{d_m - j}$, and

$$\mathbf{C}_m = \begin{pmatrix} 1 & c_1^1 & \cdots & \cdots & \cdots & c_{2d_m-2}^1 & 0 \\ 0 & 1 & c_1^1 & \cdots & \cdots & \cdots & c_{2d_m-2}^1 \\ 1 & c_1^2 & \cdots & c_{2d_m-2}^2 & 0 & 0 & 0 \\ 0 & 1 & c_1^1 & \cdots & c_{2d_m-2}^2 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & -2\cos\theta_m & 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & -2\cos\theta_m & 1 & 0 & \cdots & 0 \\ 1 & 0 & 0 & 0 & 0 & \cdots & \\ 0 & 1 & 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix},$$

then

$$\mathbf{C}_m\mathbf{x}_t(m) = (y_t(m,1), y_{t-1}(m,1), \ldots, y_t(m,d_m), y_{t-1}(m,d_m))'.$$

To state the limiting distribution of the LSE, we define the normalization matrix

$$\mathbf{G}_n = \text{diag}(\mathbf{L}_n(1), \ldots, \mathbf{L}_n(l)), \quad \mathbf{L}_n(m) = \text{diag}(n^{-j}\mathbf{I}_2, 1 \leq j \leq d_m)\mathbf{C}_m.$$

THEOREM 3. *Consider the time series* (1) *and assume that the characteristic polynomial* $\phi(z)$ *is given by eqn* (9). *If the disturbance process* (ε_t) *satisfies the assumptions of Theorem* 2, *then we have*

$$L^{-1}(n)\mathbf{G}_n\mathbf{Q}\sum_{k=1}^{n}\mathbf{y}_{k-1}\mathbf{y}'_{k-1}\mathbf{Q}'\mathbf{G}'_n \xrightarrow{\mathcal{L}} \operatorname{diag}(\mathbf{H}_1,\ldots,\mathbf{H}_l) \tag{17}$$

and

$$\left(\mathbf{Q}'\mathbf{G}'_n\right)^{-1}(\hat{\boldsymbol{\phi}}_n-\boldsymbol{\phi}) \xrightarrow{\mathcal{L}} \left(\left(\mathbf{H}_1^{-1}\zeta_1\right)',\ldots,\left(\mathbf{H}_l^{-1}\zeta_l\right)'\right)', \tag{18}$$

where

$$\zeta_m=(\zeta_1^{(m)},\ldots,\zeta_{2d_m}^{(m)})',\quad \mathbf{H}_m=(\sigma_{i,j}^{(m)})\textit{a } 2d_m\times 2d_m \textit{ random matrix},$$

$$\begin{aligned}\zeta_{2j}^{(m)}&=(2\sin\theta_m)^{-1}\Big\{\cos\theta_m\Big(\int_0^1 f_{m,j-1}(s)\mathrm{d}B_{2m}(s)-\int_0^1 g_{m,j-1}(s)\mathrm{d}B_{2m-1}(s)\Big)\\&\quad-\sin\theta_m\Big(\int_0^1 f_{m,j-1}(s)\mathrm{d}B_{2m-1}(s)+\int_0^1 g_{m,j-1}(s)\mathrm{d}B_{2m}(s)\Big)\Big\},\end{aligned}$$

$$\zeta_{2j-1}^{(m)}=(2\sin\theta_m)^{-1}\Big(\int_0^1 f_{m,j-1}(s)\mathrm{d}B_{2m}(s)-\int_0^1 g_{m,j-1}(s)\mathrm{d}B_{2m-1}(s)\Big),$$

$$\begin{aligned}\sigma_{2k-1,2j-1}^{(m)}&=\sigma_{2k,2j}^{(m)}\\&=(4\sin^2\theta_m)^{-1}\Big(\int_0^1 f_{m,k-1}(s)f_{m,j-1}(s)\mathrm{d}s+\int_0^1 g_{m,k-1}(s)g_{m,j-1}(s)\mathrm{d}s\Big),\end{aligned}$$

$$\begin{aligned}\sigma_{2k-1,2j}^{(m)}&=\sigma_{2j,2k-1}^{(m)}\\&=(4\sin^2\theta_m)^{-1}\Big\{\cos\theta_m\Big(\int_0^1 f_{m,k-1}(s)f_{m,j-1}(s)\mathrm{d}s+\int_0^1 g_{m,k-1}(s)g_{m,j-1}(s)\mathrm{d}s\Big)\\&\quad-\sin\theta_m\Big(\int_0^1 f_{m,j-1}(s)g_{m,k-1}(s)\mathrm{d}s-\int_0^1 g_{m,j-1}(s)f_{m,k-1}(s)\mathrm{d}s\Big)\Big\},\end{aligned}$$

$$f_{m,j}(t)=(2\sin\theta_m)^{-1}\Big(\sin\theta_m\int_0^t f_{m,j-1}(s)\mathrm{d}s-\cos\theta_m\int_0^t g_{m,j-1}(s)\mathrm{d}s\Big),$$

$$g_{m,j}(t)=(2\sin\theta_m)^{-1}\Big(\cos\theta_m\int_0^t f_{m,j-1}(s)\mathrm{d}s+\sin\theta_m\int_0^t g_{m,j-1}(s)\mathrm{d}s\Big),$$

$$f_{m,0}(t)=K(\theta_m,H)B_{2m-1}(t),\quad g_{m,0}(t)=K(\theta_m,H)B_{2m}(t),$$

$K(\theta_m, H) = \sqrt{\pi}|\theta_m|^{\frac{1}{2}-H}, 1 \leq m \leq l, B_i$ are standard Brownian motions, $i = 1, \ldots, 2l$, and B_i is independent of B_j if $i \neq j$.

Remark 1. Theorem 3 implies that the LSE $\hat{\boldsymbol{\phi}}_n$ is a consistent estimator of $\boldsymbol{\phi}$, i.e. $\hat{\boldsymbol{\phi}}_n \xrightarrow{P} \boldsymbol{\phi}$, where $\xrightarrow{P}$ denotes the convergence in probability. Moreover, the rate of convergence is equal to $O_p(n^{-1})$ and is the same as the one obtained by Ahtola and Tiao (1987a), Chan and Wei (1988) and Chan and Terrin (1995).

Remark 2. In this article we have derived the limiting distribution of the LSE in model (1), where the disturbance (ε_t) is a non-Gaussian long-memory process given by eqns (2)–(3), only when the characteristic polynomial $\phi(z)$ is unstable with complex-conjugate unit roots. However, the behaviour of the LSE when $\phi(z)$ has stable roots [i.e. $\phi(z) = 0$ implies that $|z| > 1$], roots equal to -1 and 1, and explosive roots [i.e. $\phi(z) = 0$ implies that $|z| < 1$] remains an open problem and will be treated in a future study.

APPENDIX: PROOFS

Proof of Theorem 1. We shall adapt the proof of Theorem 2 of Giraitis *et al.* (1996). Let

$$T_n = (nL(n))^{-1/2} \sum_{k=1}^{n} c_{n,k} \varepsilon_k.$$

By Cramer Wold arguments, T_n converges in distribution to T if and only if for all $v \in \mathbb{R}^p$, $v'T_n$ converges in distribution to $v'T$. To prove the last convergence it is sufficient to show that

$$E\left(\mathrm{e}^{\mathrm{i}v'T_n}\right) = \mathrm{e}^{-\frac{1}{2}\sigma_n^2(v)} + o(1),$$

uniformly on compacts $\{||v|| \leq A\}$, where $\sigma_n^2(v) = \mathrm{var}(v'T_n)$.

By using eqn (12) we have

$$\sigma_n^2(v) \longrightarrow v'\mathbf{R}v,$$

hence $\sigma_n^2(v) \leq C\lambda_{\max}(\mathbf{R})\|v\|^2$, for some positive constant C, where $\lambda_{\max}(\mathbf{R})$ is the maximum eigenvalue of the matrix $\mathbf{R}$; therefore, $\sigma_n^2(v)$ is bounded uniformly on compacts.

We consider the truncated variables

$$u_{j,N}^{+} = u_j 1_{\{|u_j| > N\}} - E(u_j 1_{\{|u_j| > N\}}), \quad u_{j,N}^{-} = u_j - u_{j,N}^{+},$$

where 1_A is the indicator function (equals *1* when condition A is satisfied and 0 otherwise), and define

$$\varepsilon_{k,N}^{+} = \sum_{j \leq k} b_{k-j} u_{j,N}^{+}, \quad T_{n,N}^{+} = (nL(n))^{-1/2} \sum_{k=1}^{n} v' c_{n,k} \varepsilon_{k,N}^{+}, \quad T_{n,N}^{-} = T_n - T_{n,N}^{+}.$$

We have

$$\sigma^2_{+,N} = E((u^+_{0,N})^2) \to 0 \quad \text{as } N \to \infty$$

and

$$E((T^+_{n,N})^2) = \sigma^2_{+,N}\sigma^2_n(v) \le C\lambda_{\max}(\mathbf{R})\|v\|^2\sigma^2_{+,N} \le q_N,$$

where q_N is independent of n and $q_N \to 0$ as $N \to \infty$. Consequently

$$E(\mathrm{e}^{\mathrm{i}v'T_n}) = E(\mathrm{e}^{\mathrm{i}v'T^-_{n,N}}) + \mathrm{d}_{n,N},$$

where

$$|d_{n,N}| = |E((\mathrm{e}^{\mathrm{i}v'T^+_{n,N}} - 1)\mathrm{e}^{\mathrm{i}v'T^-_{n,N}})| \le E(|v'T^+_{n,N}|) \le q_N^{1/2}.$$

It suffices to show that for $N < \infty$ and $r \in \mathbb{N}, r \ge 3$,

$$\operatorname*{cum}_r(v'T^-_{n,N}) = o(1), \tag{19}$$

where $\mathrm{cum}_r(\cdot)$ is the rth cumulant. Note that

$$\operatorname*{cum}_r(v'T^-_{n,N}) = v_{r,N}\sum_{j\in\mathbb{Z}} t^r_{n,j},$$

where

$$v_{r,N} = \operatorname*{cum}_r(u^-_{0,N}) \quad \text{and} \quad t_{n,j} = (nL(n))^{-1/2}\sum_{k=1}^{n} v'c_{n,k}b_{k-j},$$

hence eqn (19) follows from eqn (13). □

Proof of Theorem 2. We shall prove only eqn (15) [the proof of eqn (16) is similar]. Let

$$S_{M,N} = \sum_{k=M}^{N}\sin(k\theta)\varepsilon_k, S_n = S_{1,n},$$

then $X_n(t) = (nL(n))^{-1/2}S_{[nt]}$. To prove Theorem 2 we need the following three lemmas.

Lemma 1. *For all* $0 \le t_1 < t_2 \le 1$,

$$\mathrm{var}(X_n(t_2) - X_n(t_1)) = (nL(n))^{-1}\mathrm{var}(S_{[nt_1]+1,[nt_2]}) \sim \pi(t_2 - t_1)|\theta|^{1-2H}, \tag{20}$$

$$\mathrm{cov}(X_n(t_1), X_n(t_2)) \sim \pi t_1|\theta|^{1-2H}. \tag{21}$$

Proof. Denote by

$$\gamma(k) = \int_{-\pi}^{\pi}\mathrm{e}^{\mathrm{i}k\lambda}f(\lambda)\mathrm{d}\lambda$$

the autocovariance function of $\{\varepsilon_k\}$, and let $n_j = [nt_j]$, $j = 1,2$. Then

$$\begin{aligned}\mathrm{var}(S_{[nt_1]+1,[nt_2]}) &= \sum_{l=n_1+1}^{n_2}\sum_{j=n_1+1}^{n_2}\sin(l\theta)\sin(j\theta)\gamma(j-l)\\ &= (T_{1,n} + \bar{T}_{1,n} - T_{2,n} - \bar{T}_{2,n})/4,\end{aligned} \tag{22}$$

where

$$T_{1,n} = \sum_{j=n_1+1}^{n_2} \sum_{l=n_1+1}^{n_2} \mathrm{e}^{\mathrm{i}(j-l)\theta}\gamma(j-l), \quad T_{2,n} = \sum_{j=n_1+1}^{n_2} \sum_{l=n_1+1}^{n_2} \mathrm{e}^{\mathrm{i}(j+l)\theta}\gamma(j-l) \tag{23}$$

and $\bar{T}_{j,n}$ is the complex conjugate of $T_{j,n}$, $j = 1,2$.

Now observe that

$$\begin{aligned} T_{1,n} &= \int_{-\pi}^{\pi} \Big| \sum_{k=n_1+1}^{n_2} \mathrm{e}^{\mathrm{i}k(\lambda+\theta)} \Big|^2 |\lambda|^{1-2H} L(|\lambda|^{-1})\mathrm{d}\lambda \\ &= (n_2-n_1)\int_{(n_2-n_1)(\theta-\pi)}^{(n_2-n_1)(\pi+\theta)} \Big| \frac{\mathrm{e}^{\mathrm{i}y}-1}{(n_2-n_1)(\mathrm{e}^{\frac{\mathrm{i}y}{(n_2-n_1)}}-1)} \Big|^2 |y(n_2-n_1)^{-1}-\theta|^{1-2H} \\ &\quad \times L(|(n_2-n_1)^{-1}y-\theta|^{-1})\mathrm{d}y. \end{aligned}$$

It follows that

$$\begin{aligned} T_{1,n} &\sim (n_2-n_1)L(n_2-n_1)|\theta|^{1-2H} \int_{-\infty}^{\infty} \Big| \frac{\mathrm{e}^{\mathrm{i}y}-1}{iy} \Big|^2 \mathrm{d}y \\ &= 2\pi(n_2-n_1)L(n_2-n_1)|\theta|^{1-2H}. \end{aligned}$$

Since $(n_2 - n_1) \sim n(t_2 - t_1)$ and $L(\cdot)$ is a slowly varying function we deduce that

$$T_{1,n} \sim 2\pi n(t_2-t_1)L(n)|\theta|^{1-2H}. \tag{24}$$

The second term in eqn (23) can be written as

$$\begin{aligned} T_{2,n} &= \int_{-\pi}^{\pi} \sum_{j=n_1+1}^{n_2} \mathrm{e}^{\mathrm{i}j(\lambda+\theta)} \sum_{l=n_1+1}^{n_2} \mathrm{e}^{\mathrm{i}l(\theta-\lambda)} |\lambda|^{1-2H} L(|\lambda|^{-1})\mathrm{d}\lambda \\ &= \int_{-\pi}^{\pi} \frac{\mathrm{e}^{\mathrm{i}(n_2-n_1-1)(\theta+\lambda)}-1}{\mathrm{e}^{\mathrm{i}(\theta+\lambda)}-1} \frac{\mathrm{e}^{\mathrm{i}(n_2-n_1-1)(\theta-\lambda)}-1}{\mathrm{e}^{\mathrm{i}(\theta-\lambda)}-1} \mathrm{e}^{2\mathrm{i}(n_1+1)\theta} |\lambda|^{1-2H} L(|\lambda|^{-1})\mathrm{d}\lambda. \end{aligned}$$

Since for all $x \in \mathbb{R}$, for all $0 \le \delta \le 1$, $|\mathrm{e}^{\mathrm{i}x} - 1| \le 2^{1-\delta}|x|^{\delta}$ and for all $|x| < 2\pi$, $|\mathrm{e}^{\mathrm{i}x}-1| \ge |x|/2$, we have that

$$\begin{aligned} |T_{2,n}| &\le Cn^{2\delta}(t_2-t_1)^{2\delta} \int_{-\pi}^{\pi} |\lambda+\theta|^{\delta-1}|\lambda-\theta|^{\delta-1}|\lambda|^{1-2H}L(|\lambda|^{-1})\mathrm{d}\lambda, \\ &= o(n) \quad \text{for all } 0 < \delta < 1/2. \end{aligned} \tag{25}$$

Therefore, eqn (20) follows from eqns (22)–(25).

From eqn (20) we deduce that $\mathrm{var}(X_n(t_1)) = (nL(n))^{-1}\,\mathrm{var}(S_{1,[nt_1]}) \sim \pi t_1|\theta|^{1-2H}$; hence by writing

$$\mathrm{cov}(X_n(t_1), X_n(t_2)) = \mathrm{var}(X_n(t_1)) + \mathrm{cov}(X_n(t_1), X_n(t_2) - X_n(t_1)),$$

the result (21) holds if

$$\mathrm{cov}(X_n(t_1), X_n(t_2) - X_n(t_1)) = o(1). \tag{26}$$

Write

$$\operatorname{cov}(X_n(t_1), X_n(t_2) - X_n(t_1)) = \sum_{j=1}^{n_1} \sum_{l=n_1+1}^{n_2} \sin(l\theta)\sin(j\theta)\gamma(j-l)$$
$$= (V_{1,n} + \bar{V}_{1,n} - V_{2,n} - \bar{V}_{2,n})/4nL(n), \tag{27}$$

where

$$V_{1,n} = \sum_{j=1}^{n_1} \sum_{l=n_1+1}^{n_2} \mathrm{e}^{\mathrm{i}(j-l)\theta}\gamma(j-l), \quad V_{2,n} = \sum_{j=1}^{n_1} \sum_{l=n_1+1}^{n_2} \mathrm{e}^{\mathrm{i}(j+l)\theta}\gamma(j-l).$$

Clearly,

$$V_{1,n} = \int_{-\pi}^{\pi} \frac{\mathrm{e}^{\mathrm{i}(n_2-n_1)(\theta+\lambda)} - 1}{|\mathrm{e}^{\mathrm{i}(\theta+\lambda)} - 1|^2} \left(1 - \mathrm{e}^{\mathrm{i}(n_1+1)(\theta+\lambda)}\right) |\lambda|^{1-2H} L(|\lambda|^{-1}) \mathrm{d}\lambda.$$

Let $y = (n_1 + 1)(\theta + \lambda)$, then

$$(nL(n))^{-1} V_{1,n} \to t_1 |\theta|^{1-2H} \int_{-\infty}^{\infty} \frac{\mathrm{e}^{\mathrm{i}(t_2-t_1)y/t_1} - 1}{|y|^2} \left(1 - \mathrm{e}^{\mathrm{i}y}\right) \mathrm{d}y$$
$$= 2t_1 |\theta|^{1-2H} \int_0^{\infty} \frac{\cos((t_2-t_1)y/t_1) - \cos(t_2 y/t_1) + \cos(y) - 1}{|y|^2} \mathrm{d}y$$
$$= 0, \tag{28}$$

the last equality follows by using the formula

$$\int_0^{\infty} \frac{\cos(py) - \cos(qy)}{|y|^2} \mathrm{d}y = \frac{(q-p)\pi}{2}, \quad \text{for all } p \geq 0, \quad \text{for all } q \geq 0.$$

The term $V_{2,n}$ can be written as

$$V_{2,n} = \int_{-\pi}^{\pi} \frac{\mathrm{e}^{\mathrm{i}(n_1+1)(\theta+\lambda)} - 1}{\mathrm{e}^{\mathrm{i}(\theta+\lambda)} - 1} \frac{\mathrm{e}^{\mathrm{i}(n_2-n_1)(\theta-\lambda)} - 1}{\mathrm{e}^{\mathrm{i}(\theta-\lambda)} - 1} \mathrm{e}^{2\mathrm{i}(n_1+1)\theta} |\lambda|^{1-2H} L(|\lambda|^{-1}) \mathrm{d}\lambda,$$

hence

$$(nL(n))^{-1}|V_{2,n}| \leq C(nL(n))^{-1} n^{2\delta} t_1^{\delta} (t_2 - t_1)^{\delta}$$
$$\times \int_{-\pi}^{\pi} |\lambda + \theta|^{\delta-1} |\lambda - \theta|^{\delta-1} |\lambda|^{1-2H} L(|\lambda|^{-1}) \mathrm{d}\lambda$$
$$= o(1) \quad \text{for all } 0 < \delta < 1/2. \tag{29}$$

Consequently, eqn (26) follows from eqns (27)–(29). □

Lemma 2.

$$(nL(n))^{-1/2} S_n \xrightarrow{\mathcal{L}} N(0, \pi|\theta|^{1-2H}). \tag{30}$$

Proof. *We shall apply Theorem* 1 *with* $c_{n,k} = \sin(k\theta)$. *By choosing* $t_1 = 0, t_2 = 1$ *in eqn* (20), *we get* $\operatorname{var}(S_n) \sim \pi nL(n)|\theta|^{1-2H}$. *To obtain eqn* (30), *it remains to prove condition* (13). *As* $L(\cdot)$ *is bounded*, *it suffices to show that*

$$T_{n,r} = \sum_{j\in\mathbb{Z}}\left(\sum_{i=1}^{n}\sin(i\theta)b_{i-j}\right)^{r} = o\left(n^{r/2}\right). \tag{31}$$

By using similar arguments as used in (Zygmund, 1959, p. 187), we can easily prove that, with

$$b_i = i^{\frac{2H-3}{2}}L_1(i),$$

$$\left|\sum_{i=1}^{N}\sin(i\theta)b_i\right| = O(1) \quad \text{and} \quad \left|\sum_{i=1}^{N}\cos(i\theta)b_i\right| = O(1) \tag{32}$$

uniformly on N. Note that $T_{n,r} = T_{1,n,r} + T_{2,n,r} + T_{3,n,r}$, where

$$T_{1,n,r} = \left(\sum_{i=1}^{n}\sin(i\theta)b_i\right)^{r}, \quad T_{2,n,r} = \sum_{j=1}^{\infty}\left(\sum_{i=1}^{n}\sin(i\theta)b_{i-j}\right)^{r},$$
$$T_{3,n,r} = \sum_{j=-\infty}^{-1}\left(\sum_{i=1}^{n}\sin(i\theta)b_{i-j}\right)^{r}.$$

From eqn (32) we obtain $T_{1,n,r} = O(1)$.

$$\begin{aligned}|T_{2,n,r}| &\leq \sum_{j=1}^{\infty}\left|\sum_{i=1}^{n}\sin(i\theta)b_{i-j}\right|^{r}\\ &= \sum_{j=1}^{n}\left|\sum_{l=1}^{n-j}\sin((l+j)\theta)b_l\right|^{r}\\ &= \sum_{j=1}^{n}\left|\sin(j\theta)\sum_{l=1}^{n-j}\cos(l\theta)b_l + \cos(j\theta)\sum_{l=1}^{n-j}\sin(l\theta)b_l\right|^{r},\end{aligned}$$

hence eqn (32) implies that $|T_{2,n,r}| \leq Cn = o(n^{\frac{r}{2}})$, for all $r \geq 3$.
We have that

$$T_{3,n,r} = \sum_{j=1}^{\infty}\left(\sum_{i=1}^{n}\sin(i\theta)b_{i+j}\right)^{r}.$$

To prove that $T_{3,n,r} = o(n^{\frac{r}{2}})$ it is sufficient to show that for all $M \in \mathbb{N}$, for all $\varepsilon > 0$, there exists n_0 such that for all $n \geq n_0$

$$n^{-r/2}\left|\sum_{j=1}^{M}\left(\sum_{i=1}^{n}\sin(i\theta)b_{i+j}\right)^{r}\right| < \varepsilon. \tag{33}$$

From eqn (32) we deduce that

$$\begin{aligned}\left|\sum_{j=1}^{M}\left(\sum_{i=1}^{n}\sin(i\theta)b_{i+j}\right)^{r}\right| &= \left|\sum_{j=1}^{M}\left(\cos(j\theta)\sum_{l=j+1}^{j+n}\sin(l\theta)b_l - \sin(j\theta)\sum_{l=j+1}^{j+n}\cos(l\theta)b_l\right)^{r}\right|\\ &\leq MC,\end{aligned}$$

for some positive constant C. To obtain eqn (33) it suffices to take $n_0 = [(MC/\epsilon)^{\frac{2}{r}}] + 1.$□

LEMMA 3. *For all* $0 < M \leq N$,

$$E\left(|S_{M,N}|^{2\boldsymbol{\kappa}_0}\right) \leq C\left(\mathrm{var}(S_{M,N})\right)^{\boldsymbol{\kappa}_0}, \text{ for some positive constant } C. \tag{34}$$

PROOF. Write

$$\varepsilon_k = \sum_{j\in\mathbb{Z}} b_{k-j} u_j$$

by assuming that $b_i = 0$ if $i < 0$, then

$$\mathrm{var}(S_{M,N}) = \sum_{i\in\mathbb{Z}} \left(\sum_{k=M}^{N} \sin(k\theta) b_{i-k} \right)^2 .$$

Let $\Lambda_{\boldsymbol{\kappa}_0} = \{(i_1, \ldots, i_s) \in \mathbb{N}^s$ such that $i_1 + \cdots + i_s = \boldsymbol{\kappa}_0\}$, i.e. the set of all solutions in natural numbers of the equation $i_1 + \cdots + i_s = \boldsymbol{\kappa}_0$ (without taking into account the order of the terms). Then

$$\begin{aligned}
E\left(|S_{M,N}|^{2\boldsymbol{\kappa}_0}\right) &= \sum_{\Lambda_{\boldsymbol{\kappa}_0}} a_\lambda \sum_{k_1 \neq \cdots \neq k_s} \left(\sin(M\theta) b_{k_1-M} + \cdots + \sin(N\theta) b_{k_1-N}\right)^{2i_1} \\
&\quad \cdots \left(\sin(M\theta) b_{k_s-M} + \cdots + \sin(N\theta) b_{k_s-N}\right)^{2i_s} \\
&\leq \max_{\Lambda_{\boldsymbol{\kappa}_0}} \{a_\lambda\} \left(\sum_{i\in\mathbb{Z}} (\sin(M\theta) b_{i-M} + \cdots + \sin(N\theta) b_{i-N})^2 \right)^{\boldsymbol{\kappa}_0} \\
&= \max_{\Lambda_{\boldsymbol{\kappa}_0}} \{a_\lambda\} \left(\mathrm{var}(S_{M,N})\right)^{\boldsymbol{\kappa}_0},
\end{aligned}$$

where

$$a_\lambda = \frac{E(|u_0|^{2i_1}) E(|u_0|^{2i_2}) \cdots E(|u_0|^{2i_s}) (2\boldsymbol{\kappa}_0)!}{(2i_1!) \cdots (2i_s)!}. \qquad \square$$

We now use Lemmas 1–3 to prove Theorem 2. To prove that the finite-dimensional distributions of X_n converge to those of $K(\theta,H)B_1$ it is sufficient to show that for all integer $r \geq 1$, for all $0 \leq t_1 < \cdots < t_r \leq 1$ and for all $(\alpha_1, \ldots, \alpha_r)' \in \mathbb{R}^r$,

$$Z_n = \sum_{i=1}^{r} \alpha_i X_n(t_i) \xrightarrow{\mathcal{L}} K(\theta, H) \sum_{i=1}^{r} \alpha_i B_1(t_i). \tag{35}$$

Since

$$Z_n = (nL(n))^{-1/2} \sum_{k=1}^{[nt_r]} c_{n,k} \varepsilon_k,$$

where

$$c_{n,k} = \sin(k\theta)(\alpha_1 + \cdots + \alpha_j) \quad \text{if } [nt_{j-1}] < k \leq [nt_j], \quad 1 \leq j \leq r, \quad t_0 = 0,$$

and that it is not difficult to show the condition (13) (we omit the proof), the convergence (35) holds if

$$\operatorname{var}(Z_n) \to \operatorname{var}\left(K(\theta,H)\sum_{i=1}^{r}\alpha_i B_1(t_i)\right) = K^2(\theta,H)\sum_{1\le i,j\le r}\alpha_i\alpha_j\min(t_i,t_j). \tag{36}$$

Moreover, the convergence (36) follows from eqns (20) and (21).

To prove the tightness of X_n it suffices to show the following inequality (Billingsley, 1968, Thm 15.6)

$$E(|X_n(t)-X_n(t_1)|^{\gamma}|X_n(t_2)-X_n(t)|^{\gamma}) \le (F(t_2)-F(t_1))^{\alpha} \tag{37}$$

for some $\gamma \ge 0$, $\alpha > 1$, and F is a nondecreasing continuous function on [0,1], where $0 < t_1 < t < t_2 < 1$.

Let $K(t,t_1,t_2) = E(|X_n(t) - X_n(t_1)|^{\kappa_0}|X_n(t_2) - X_n(t)|^{\kappa_0})$. Using Cauchy–Schwarz inequality

$$\begin{aligned} K(t,t_1,t_2) &= (nL(n))^{-\kappa_0}E\left(|S_{[nt_1]+1,[nt]}|^{\kappa_0}|S_{[nt]+1,[nt_2]}|^{\kappa_0}\right) \\ &\le (nL(n))^{-\kappa_0}E\left(|S_{[nt_1]+1,[nt]}|^{2\kappa_0}\right)^{\frac{1}{2}}E\left(|S_{[nt]+1,[nt_2]}|^{2\kappa_0}\right)^{\frac{1}{2}}. \end{aligned}$$

Combining eqns (34) and (20), we have

$$\begin{aligned} K(t,t_1,t_2) &\le C\left((nL(n))^{-1}\operatorname{var}(S_{[nt_1]+1,[nt]})\right)^{\frac{\kappa_0}{2}}\left((nL(n))^{-1}\operatorname{var}(S_{[nt]+1,[nt_2]})\right)^{\frac{\kappa_0}{2}} \\ &\le C_1(t-t_1)^{\frac{\kappa_0}{2}}(t_2-t)^{\frac{\kappa_0}{2}} \\ &\le \left(C_1^{\frac{1}{\kappa_0}}t_2 - C_1^{\frac{1}{\kappa_0}}t_1\right)^{\kappa_0}. \end{aligned}$$

Consequently eqn (37) holds with $\gamma = \alpha = \kappa_0$ and $F(t) = C_1^{\frac{1}{\kappa_0}}t$. □

Proof of Theorem 3. For the proof of Theorem 3 we need Lemmas 4 and 5.

Lemma 4. *Let $\theta_i \in]0,\pi[$ such that $\theta_i \ne \theta_j$ if $i \ne j$ for $i,j = 1,2,\ldots,l$ and define*

$$Y_n(t_1,\ldots,t_{2l}) = (nL(n))^{-1/2} \times \left(\sum_{k=1}^{[nt_1]}\sin(k\theta_1)\varepsilon_k, \sum_{k=1}^{[nt_2]}\cos(k\theta_1)\varepsilon_k,\ldots,\sum_{k=1}^{[nt_{2l-1}]}\sin(k\theta_l)\varepsilon_k,\sum_{k=1}^{[nt_{2l}]}\cos(k\theta_l)\varepsilon_k\right).$$

Then

$$Y_n \Longrightarrow (K(\theta_1,H)B_1, K(\theta_1,H)B_2,\ldots,K(\theta_l,H)B_{2l-1},K(\theta_l,H)B_{2l}).$$

Proof. By using Theorem 2 it is sufficient to prove the asymptotic independence. The proof of which is easy and hence omitted. □

Let

$$S_t(m,j) = \sum_{k=1}^{t}\cos(k\theta_m)y_k(m,j),\ T_t(m,j) = \sum_{k=1}^{t}\sin(k\theta_m)y_k(m,j). \tag{38}$$

LEMMA 5. *For all* $0 < t_1, t_2 \le 1$, *the following convergences hold*

$$n^{-j-1/2}L^{-1/2}(n)\big(S_{[nt_1]}(m,j), T_{[nt_2]}(m,j)\big) \Longrightarrow (f_{m,j}(t_1), g_{m,j}(t_2)), \tag{39}$$

$$n^{-(j+k)}L^{-1}(n)\sum_{t=1}^{n} y_t(m,k)y_t(m,j) \xrightarrow{\mathcal{L}} \sigma^{(m)}_{2k,2j}, \tag{40}$$

$$n^{-(j+k)}L^{-1}(n)\sum_{t=1}^{n} y_{t-1}(m,k)y_t(m,j) \xrightarrow{\mathcal{L}} \sigma^{(m)}_{2k-1,2j}, \tag{41}$$

$$n^{-j}L^{-1}(n)\sum_{t=1}^{n} y_{t-1}(m,j)\varepsilon_t \xrightarrow{\mathcal{L}} \zeta^{(m)}_{2j}, \tag{42}$$

$$n^{-j}L^{-1}(n)\sum_{t=1}^{n} y_{t-2}(m,j)\varepsilon_t \xrightarrow{\mathcal{L}} \zeta^{(m)}_{2j-1}. \tag{43}$$

PROOF. Let

$$S_n = S_n(m,0) = \sum_{k=1}^{n}\cos(k\theta_m)\varepsilon_k, \quad T_n = T_n(m,0) = \sum_{k=1}^{n}\sin(k\theta_m)\varepsilon_k.$$

Then from eqn (20) we deduce that

$$S_n = O(n) \quad \text{and} \quad T_n = O(n). \tag{44}$$

Using eqn (44) it is easy to prove that all the results of Lemmas 3.3.1–3.3.6 of Chan and Wei (1988) hold and details are omitted.

By applying our Theorem 2, the continuous mapping theorem and a similar arguments as used in the proofs of Theorem 3.3.4 and the Lemma 3.3.7 of Chan and Wei (1988) the convergences (39)–(43) follow immediately. □

We now use the Lemmas 4–5 to prove Theorem 3. From eqns (39)–(43) it is easy to show that

$$\mathbf{L}_n(m)\sum_{t=1}^{n}\mathbf{x}_{t-1}(m)\mathbf{x}'_{t-1}(m)\mathbf{L}'_n(m) \xrightarrow{\mathcal{L}} \mathbf{H}_m,$$

where $\mathbf{H}_m$ is nonsingular almost surely, and

$$\big(\mathbf{L}'_n(m)\big)^{-1}\left(\sum_{t=1}^{n}\mathbf{x}_{t-1}(m)\mathbf{x}'_{t-1}(m)\right)^{-1}\sum_{t=1}^{n}\mathbf{x}_{t-1}(m)\varepsilon_t \xrightarrow{\mathcal{L}} \mathbf{H}_m^{-1}\zeta_m.$$

In view of the preceding two convergences and Lemma 5, to prove eqns (17) and (18) we only need to establish that the off-diagonal submatrices of

$$\mathbf{G}_n\mathbf{Q}\sum_{k=1}^{n}\mathbf{y}_{k-1}\mathbf{y}'_{k-1}\mathbf{Q}'\mathbf{G}'_n/L(n)$$

converge to zero in probability. Typical elements of these submatrices are

$$(L(n)n^{r+s})^{-1}\sum_{t=1}^{n}y_t(h,s)y(j,r)=\left(L(n)n^{r+s}\sin\theta_h\sin\theta_j\right)^{-1}\sum_{t=1}^{n}\{(S_t(h,s-1)\sin(t+1)\theta_h - T_t(h,s-1)\cos(t+1)\theta_h)\times(S_t(j,r-1)\cos(t+1)\theta_j - T_t(j,r-1)\cos(t+1)\theta_j)\}. \quad (45)$$

Let us, for example, examine the term

$$n^{-s-r}\sum_{t=1}^{n}S_t(h,s-1)S_t(j,r-1)\sin(t+1)\theta_h\cos(t+1)\theta_j$$

which can be written as a sum of four terms taking the form

$$n^{-s-r}\sum_{t=1}^{n}S_t(h,s-1)S_t(j,r-1)\mathrm{e}^{\mathrm{i}(t+1)\theta},$$

with $\theta = \pm(\theta_h \pm \theta_j)$.

We shall apply Theorem 2.1 of Chan and Wei (1988) to the sequence of random variables $X_n = S_n(h,s-1)S_n(j,r-1)$. Since $E(S_n^2(h,k)) = O(n^{2k+1})$, it follows that

$$\begin{aligned}E|X_n| &\le \left(E(S_n^2(h,s-1)\right)^{1/2}\left(S_n^2(j,r-1)\right)^{1/2}\\ &= O(n^{\frac{2(s-1)+1}{2}})O(n^{\frac{2(r-1)+1}{2}})\\ &= O(n^{s+r-1}).\end{aligned}$$

We need to prove that

$$|X_n - X_m| \le A_1(n,m)B_1(n,m) + A_2(n,m)B_2(n,m),$$

for some random variables $A_i(n,m)$ and $B_i(n,m)$ such that

$$E(A_i^2(n,m)) \le Cn^{\gamma_i}, E(B_i^2(n,m)) \le Cn^{\delta_i}(n-m)$$

for $n \ge m$ and some positive constants $C,\gamma_i,\delta_i,i = 1,2$. We have that

$$\begin{aligned}|X_n - X_m| &= |S_n(h,s-1)S_n(j,r-1) - S_m(h,s-1)S_m(j,r-1)|\\ &\le |S_n(h,s-1)||S_n(j,r-1) - S_m(j,r-1)|\\ &\quad + |S_m(j,r-1)||S_n(h,s-1) - S_m(h,s-1)|.\end{aligned}$$

Let

$$A_1(n,m) = |S_n(h,s-1)|, \quad B_1(n,m) = |S_n(j,r-1) - S_m(j,r-1)|, \quad n \ge m.$$

Then $E(A_1^2(n,m)) = O(n^{\gamma_1})$ with $\gamma_1 = 2s-1$.

$$\begin{aligned}B_1(n,m) &\le \left(\sum_{k=m+1}^{n}\cos^2 k\theta_j\right)^{1/2}\left(\sum_{k=m+1}^{n}y_k^2(j,r-2)\right)^{1/2}\\ &\le (n-m)^{1/2}\left\{\frac{4}{\sin^2\theta_j}\sum_{t=m+1}^{n}S_t^2(j,r-2)+T_t^2(j,r-2)\right\}^{1/2}\end{aligned}$$

Since $E(T_t^2(j,k)) = O(t^{2k+1})$, it follows that

$$E(B_1^2(n,m)) = O\left((n-m)\sum_{t=m+1}^{n} t^{2r-3}\right) = O\big((n-m)n^{\delta_1}\big), \quad \text{with } \delta_1 = 2(r-1).$$

Likewise, if we define

$$A_2(n,m) = |S_m(j,r-1)|, \quad B_2(n,m) = |S_n(h,s-1) - S_m(h,s-1)|, \quad n \geq m,$$

then we can prove that

$$E(A_2^2(n,m)) = O(n^{\gamma_2}) \quad \text{and} \quad E(B_2^2(n,m)) = O(n-m)n^{\delta_2}), \quad \text{with } \gamma_2 = \gamma_1, \delta_2 = \delta_1.$$

Therefore (see the remark of Theorem 2.1 of Chan and Wei, 1988), if we put $\alpha = s + r - 1$ then $\gamma_i + \delta_i < 2\alpha$ for $i = 1,2$ and hence Theorem 2.1 of Chan and Wei (1988) implies that

$$n^{-r-s}\sum_{t=1}^{n} S_t(h,s-1)S_t(j,r-1)\mathrm{e}^{\mathrm{i}(t+1)\theta} = o_p(1).$$

Since $L(\cdot)$ is bounded, it follows that

$$\big(L(n)n^{r+s}\sin\theta_h\sin\theta_j\big)^{-1}\sum_{t=1}^{n} S_t(h,s-1)S_t(j,r-1)\sin(t+1)\theta_h\cos(t+1)\theta_j = o_p(1).$$

Likewise, the remaining terms in the right-hand side of eqn (45) are $o_p(1)$. Consequently

$$(L(n)n^{r+s})^{-1}\sum_{t=1}^{n} y_t(h,s)y(j,r) = o_p(1).$$

This completes the proof of Theorem 3. □

ACKNOWLEDGEMENTS

I thank the associate editor and the referee for their constructive comments.

NOTE

Corresponding author: Mohamed Boutahar, Luminy Faculty of Sciences, University of Méditerranée, Marseille, France. E-mail: boutahar@univmed.fr

REFERENCES

AHTOLA, J. and TIAO, G. C. (1987a) Distribution of the least squares parameters of autoregressive process with complex roots on the unit circle. *Journal of Time Series Analysis* 8, 1–14.

AHTOLA, J. and TIAO, G. C. (1987b) A note on asymptotic inference in autoregressive models with roots on the unit circle. *Journal of Time Series Analysis* 8, 15–18.

AVRAM, F. and TAQQU, M. S. (1987) Noncentral limit theorem and Appell polynomials. *Annals of Probability* 15, 767–75.

BIERENS, H. J. (2001) Complex unit roots and business cycles: Are they real? *Econometric Theory* 17, 962–83.

BILLINGSLEY, P. (1968) *Convergence of Probability Measures*. New York: Wiley.

BIRGEAN, I. and KILIAN, L. (2002) Data-driven nonparametric spectral density estimators for economic times series: a Monte Carlo study. *Econometric Reviews* 21, 449–76.

BOUTAHAR, M. (2002) General autoregressive models with long memory noise. *Statistical Inference for Stochastic Processes* 5, 321–33.

CHAN, N. H. and TERRIN, N. (1995) Inference for unstable long-memory processes with applications to fractional unit root autoregression. *Annals of Statistics* 23, 1662–83.

CHAN, N. H. and WEI, C. Z. (1988) Limiting distribution of least squares estimates of unstable autoregressive processes. *Annals of Statistics* 16, 367–401.

CONWAY, P. and FRAME, D. (2000) A spectral *analysis of New Zealand output gaps using* Fourier transform and Wavelet techniques. *Reserve Bank of New Zealand, Discussion Paper* 2000/06.

DAVYDOV, YU. A. (1970) The invariance principle for stationary processes. *Theory and Probability Applications* 15, 487–98.

DICKEY, D. A. and FULLER, W. A. (1979) Distribution of the estimators for autoregressive time series with a unit root. *Journal of American Statistical Association* 74, 427–31.

DOBRUSHIN, R. L. and MAJOR, P. (1979) Non-central limits theorem for non-linear functionals of Gaussian fields. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 50, 27–52.

DOUKHAN, P., OPPENHEIM, G. and TAQQU, M. S. (2003) *Theory and Applications of Long-range Dependence*. Boston, MA: Birkhäuser Boston, Inc.

GIL-ALANA, L. A. (2003) Long memory in financial time series data with Non-Gaussian disturbances. *International Journal of Theory and Applied Finance* 6, 119–34.

GIRAITIS, L. and SURGAILIS, D. (1985) CLT and other limit theorems for functionals of Gaussian sequences. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 70, 191–212.

GIRAITIS, L., KOUL, H. L. and SURGAILIS, D. (1996) Asymptotic normality of regression estimators with long memory errors. *Statistic Probability Letters* 29, 317–35.

GREGOIR, S. (1999) Efficient tests for the presence of a couple of complex conjugate unit roots in real time series. Paper Presented at the European Meeting of the Econometric Society 1999, Santiago de Compostela, Spain.

HALL, P. and HEYDE, C. C. (1980) *Martingales Limit Theory and its Application*. New York: Academic Press.

HO, H. C. and HSING, T. (1997) Limit theorems for functionals of moving average. *Annals of Probability* 25, 1636–69.

HOSKING, J. R. M. (1996) Asymptotic distribution of the sample mean, autocavariances, autocorrelations of long-memory time series. *Journal of Econometrics* 73, 261–84.

PRIESTLEY, M. B. (1981) *Spectral Analysis and Time Series*. New York: Academic Press.

SCHERRER, A., LARRIEU, N., OWEZARSKI, P., BORGNAT, P. and ABRY, P. (2007) Non Gaussian and long memory statistical characterizations for Internet Traffic with Anomalies. *IEEE, Transaction on dependable and secure computing* 4, 56–70, January–March.

SURGALIS, D. (1982) Zones of attraction of self-similar multiple integrals. *Lithuanian Mathematics Journal* 22, 327–40.

TAQQU, M. S. (1975) Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 31, 287–302.

TIKU, M. L., WONG, W. K., VAUGHAN, D. C. and BIAN, G. (2000) Times series models in non–normal situations: symmetric innovations. *Journal of Time Series Analysis* 21, 571–96.

VENEMA, V., BACHNER, S., RUST, H. W. and SIMMER, C. (2006) Statistical characteristics of surrogate data based on geophysical measurements. *Nonlinear processes in Geophysics* 13, 449–66.

YIOU, P., BAERT, A. and LOUTRE, H. F. (1996) Spectral analysis of climate data. *Surveys in Geophysics* 17, 619–63.

ZYGMUND, A. (1959) *Trigonometric Series*. London: Cambridge University Press.