



HAL
open science

Estimation of the location and exponent of the spectral singularity of a long memory process

Javier Hidalgo, Philippe Soulier

► **To cite this version:**

Javier Hidalgo, Philippe Soulier. Estimation of the location and exponent of the spectral singularity of a long memory process. *Journal of Time Series Analysis*, 2004, 25 (1), pp.55-81. 10.1111/j.1467-9892.2004.00337.x . hal-00147620

HAL Id: hal-00147620

<https://hal.science/hal-00147620>

Submitted on 18 May 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Estimation of the location and exponent of the spectral singularity of a long memory process

Javier Hidalgo*

Philippe Soulier†

Abstract

We consider the estimation of the location of the pole and memory parameter ω_0 and d of a covariance stationary process with spectral density $f(x) = |1 - e^{i(x-\omega_0)}|^{-d}|1 - e^{i(x+\omega_0)}|^{-d}f^*(x)$. We investigate optimal rates of convergence for the estimators of ω_0 and d , and the consequence that the lack of knowledge of ω_0 has on the estimation of the memory parameter d . We present estimators which achieve the optimal rates. A small Monte-Carlo study is included to illustrate the finite sample performance of our estimators.

Key words : Long Memory, Fractionally differenced processes, Gegenbauer processes, Periodogram, Semiparametric estimation.

1 Introduction

Given a covariance stationary process X , the search for cyclical components is of undoubted interest. This is motivated by the observed periodic behaviour exhibited in many time series. A well known model capable of generating such a periodic behaviour is the regression model

$$x_t = \mu + \rho_1 \cos(\omega_0 t) + \rho_2 \sin(\omega_0 t) + \varepsilon_t, \quad (1.1)$$

where ρ_1 and ρ_2 are zero mean uncorrelated random variables with the same variance and $\{\varepsilon_t\}$ is a stationary sequence of random variables independent of ρ_1 and ρ_2 . Model (1.1) has enjoyed extensive use and different techniques have been proposed for the estimation of the frequency, amplitude and phase; see Whittle (1952), Grenander and Rosenblatt (1957), Hannan (1971,1973) and Chen (1988). A second model exhibiting peaks in its spectral density function is the autoregressive $AR(2)$ process

$$(I - a_1 B - a_2 B^2) X = \varepsilon, \quad (1.2)$$

where B is the backshift operator. When the roots of the polynomial $1 - a_1 z - a_2 z^2$ are not real (which implies that $a_2 < 0$), then the process X exhibits a periodic behaviour with frequency $\omega_0 =$

*Economic Department, London School of Economics, Houghton Street, London WC2A 2AE, U. K.

Supported by Economic and Social Research Council (ESRC) grant R000238212.

†Département de Mathématiques, Université d'Evry Val d'Essonne, 91025 Evry cedex, France.

Corresponding author: philippe.soulier@maths.univ-evry.fr

$\arccos(a_1/2\sqrt{-a_2})$. Models (1.1) and (1.2) represent two extreme situations explaining cyclical behaviour of the data. Model (1.2) possesses a continuous spectral density function whereas model (1.1) has a spectral distribution function with a jump at the frequency ω_0 . Whereas the cyclical pattern of model (1.2) fades out with time fairly quickly (in the sense that the autocorrelation of the process decays exponentially fast), the cyclical component of the data remains constant in model (1.1).

Between these two extreme situations, there exists a class of intermediate models capable of generating a cyclical behaviour in the data, stronger and more persistent than *ARMA* models, e.g. (1.2), but unlike model (1.1), their amplitude does not remain constant over time. Such a model has been proposed by Anel (1986) and extended by Gray et al. (1989,1994) who coined it the *GARMA* model. It is defined as

$$(I - e^{i\omega_0} B)^{-d/2} (I - e^{-i\omega_0} B)^{-d/2} X = \varepsilon, \quad (1.3)$$

where ε is an *ARMA* process. The spectral density function of the *GARMA* process is given by

$$f(x) = \frac{\sigma^2}{2\pi} |1 - e^{i(x-\omega_0)}|^{-d} |1 - e^{i(x+\omega_0)}|^{-d} |P(e^{i\lambda})/Q(e^{i\lambda})|^2 \quad (1.4)$$

where $\sigma^2 > 0$, and P and Q are polynomials without common roots and without roots inside the unit disk. As $1 - 2\cos(\omega_0)z + z^2$ is known as the Gegenbauer polynomial, *GARMA* processes are also referred to as Gegenbauer processes. When $\omega_0 = 0$, model (1.3) boils down to the more familiar FARIMA(p, d, q) model, originated by Adenstedt (1974) and examined by Granger and Joyeux (1980) and Hosking (1981). The coefficient d is frequently referred to as the memory parameter, or the fractional differencing coefficient. One can also sometimes find reference to the coefficient g defined as $g = d$ if $\omega_0 \in \{0, \pi\}$ and $g = d/2$ if $\omega_0 \in (0, \pi)$.

The spectral density has a singularity at ω_0 with a power law $|x - \omega_0|^{-\alpha}$, where $\alpha = 2d$ if $\omega_0 \in \{0, \pi\}$ and $\alpha = d$ if $\omega_0 \in (0, \pi)$. (Note that in both cases, $\alpha = 2g$).

These models exhibit “long range dependence”, meaning that the autocovariance function $\gamma(k) := \text{cov}(X_0, X_k)$ is not absolutely summable. This makes their probabilistic properties and the asymptotic distribution of some relevant statistics very different from those of usual “weakly dependent” processes such as *ARMA* models.

Maximum likelihood and Whittle estimators for these models have been investigated and proved to be \sqrt{n} -consistent and asymptotically normal when a parametric model is correctly specified. In the case of Gaussian or linear processes, this was shown by Fox and Taquq (1986), Dahlhaus (1989), Giraitis and Surgailis (1990) in the case $\omega_0 = 0$, Hosoya (1997) in the case $\omega_0 \neq 0$ known, and more recently Giraitis *et al.* (2001) dealt with the case ω_0 unknown. However, misspecification of the model can lead to inconsistent estimates of the coefficient d and of the location of the pole ω_0 . This has drawn attention of researchers and practitioners to semi-parametric methods. This means that the process ε in the fractional differencing equation (1.3) is not fully specified, but considered as an infinite dimensional nuisance parameter, while the parameters of interest are ω_0 and α . More precisely, we consider a covariance stationary linear process X with spectral density function

$$f(x) = |1 - e^{i(x-\omega_0)}|^{-d} |1 - e^{i(x+\omega_0)}|^{-d} f^*(x), \quad x \in [-\pi, \pi] \setminus \{\pm\omega_0\}, \quad (1.5)$$

where $0 < d < 1/2$ if $\omega_0 = 0$ or π and $0 < d < 1$ if $\omega_0 \in (0, \pi)$ and f^* is a positive and continuous function on $[-\pi, \pi]$. The difference between (1.3) and (1.5) lies in the fact that in (1.5), the so-called smooth component f^* of the spectral density f is neither constrained to be rational, as in (1.4), nor to be characterized by a finite dimensional parameter.

The main objectives of this paper are twofold. The first one is to provide, under mild conditions, a consistent semi-parametric estimator of ω_0 with the best possible rate of convergence. The method we have chosen is the method of Yajima (1996) which consists in maximizing the periodogram of the data. Yajima (1996) proved the consistency of this estimator under the assumption of Gaussianity. Theorem 1 below relaxes this assumption of Gaussianity and slightly improves the rate of convergence obtained by Yajima. The second objective is to investigate the consequences that the lack of knowledge of ω_0 might have on the estimation of α . Theorem 2 shows that a modified version of the GPH estimator of Geweke and Porter-Hudak (1983) (see also Robinson, 1995) has the same rate-optimality properties proved by Giraitis *et al.* (1997) where ω_0 was known, and it is asymptotically Gaussian with the same asymptotic variance obtained in Robinson (1995). That is, the statistical properties of this estimator of α are unaffected by the lack of knowledge of ω_0 . A short Monte-Carlo experiment confirms these theoretical results.

2 Definition and asymptotic properties of the estimators

Let X_1, \dots, X_n be n observations of the process X and let $I_n(x) = (2\pi n)^{-1} |\sum_{t=1}^n X_t e^{itx}|^2$ be the periodogram. Define $\tilde{n} = \lfloor (n-1)/2 \rfloor$ and let $x_k = 2k\pi/n$, $-\tilde{n} \leq k \leq \tilde{n}$ be the Fourier frequencies. The estimator of ω_0 is defined as

$$\hat{\omega}_n = \frac{2\pi}{n} \arg \max_{1 \leq k \leq \tilde{n}} I_n(x_k). \quad (2.1)$$

Theorem 1. *Let X be a strict sense linear process, i.e. there exists an i.i.d. sequence $(Z_t)_{t \in \mathbb{Z}}$ with zero mean and finite eighth moment and a square summable sequence $(a_t)_{t \in \mathbb{Z}}$ such that $X_t = \mu + \sum_{j \in \mathbb{Z}} a_j Z_{t-j}$ and define $a(x) = \sum_{j \in \mathbb{Z}} a_j e^{ijx}$. Assume that a is differentiable on $[-\pi, \pi] \setminus \{\omega_0\}$ and $\forall x \in [0, \pi] \setminus \{\omega_0\}$,*

$$K_1 |x - \omega_0|^{-\nu} \leq |a(x)| \leq K_2 |x - \omega_0|^{-\nu}, \quad (2.2)$$

$$|(x - \omega_0)a'(x)/a(x)| \leq K, \quad (2.3)$$

for some positive constants K, K_1, K_2 and $\nu \in (0, 1/2)$. Let v_n be a non-decreasing sequence such that $\lim_{n \rightarrow \infty} v_n^{-2\nu} \log(n) = 0$. Then $\frac{v_n}{v_n} (\hat{\omega}_n - \omega_0)$ converges in probability to zero.

Comments on assumptions (2.2) and (2.3) Assumption (2.2) ensures that there is a pole. Assumption (2.3) is an adaptation of Assumption A1 of Robinson (1995) in the case of a pole outside zero; it is needed to obtain upper bounds for the covariance of the discrete Fourier transform ordinates of X .

We now define the modified GPH estimator of the exponent of the singularity. Since model (1.5) does not allow the singularity to be non symmetric, as in Arteche and Robinson (2000), the definition of the estimator is symmetric around the pole. Recall that the exponent of the singularity is defined as $\alpha = d$ if $\omega_0 \in (0, \pi)$ and $\alpha = 2d$ if $\omega_0 \in \{0, \pi\}$. Denote $g(x) = -\log(|1 - e^{ix}|)$, $\bar{g}_m = m^{-1} \sum_{k=1}^m g(x_k)$, $s_m^2 = 2 \sum_{k=1}^m (g(x_k) - \bar{g}_m)^2$ and for $k = -m, \dots, -1, 1, \dots, m$, $\gamma_k = s_m^{-2} (g(x_k) - \bar{g}_m)$. The estimator $\hat{\alpha}_n$ is defined as

$$\hat{\alpha}_n = \sum_{1 \leq |k| \leq m} \gamma_k \log\{I_n(\hat{\omega}_n + x_k)\}. \quad (2.4)$$

(A1) f^* is an integrable function on $[-\pi, \pi]$ and there exists a neighborhood $\mathcal{V}_0 = [\omega_0 - \vartheta, \omega_0 + \vartheta]$ of ω_0 such that for all $x \in \mathcal{V}_0$,

$$(a) \quad |f^*(x) - f^*(\omega_0)| \leq K f^*(\omega_0) |x - \omega_0|^\beta,$$

for some $\beta \in (0, 1]$ and some positive constant $K > 0$ or, (b) f^* is differentiable at ω_0 and

$$|f^*(x) - f^*(\omega_0) - (x - \omega_0)f^{*'}(\omega_0)| \leq K|x - \omega_0|^\beta,$$

for some $\beta \in (1, 2]$ and some positive constant $K > 0$.

Theorem 2. *Let X be a Gaussian process whose spectral density f can be expressed as*

$$f(x) = |1 - e^{i(x-\omega_0)}|^{-d} |1 - e^{i(x+\omega_0)}|^{-d} f^*(x), \quad x \in [-\pi, \pi] \setminus \{\pm\omega_0\},$$

with $d \in (0, 1)$ if $\omega_0 \in (0, \pi)$ and $d \in (0, 1/2)$ if $\omega_0 \in \{0, \pi\}$ and f^* is an even function that satisfies (A1) and

$$\forall x \in [0, \pi] \setminus \{\omega_0\}, \quad |(x - \omega_0)f^{*'}(x)/f^*(x)| \leq K. \quad (2.5)$$

Let δ be a positive real number such that $\delta < 2\beta/(2\beta+1)$ and define $m = m(n) = \lceil n^\delta \rceil$. Then $m^{1/2}(\hat{\alpha}_n - \alpha)$ converges weakly to $\mathcal{N}(0, \pi^2/12)$ if $\omega_0 \in (0, \pi)$ and $m^{1/2}(\hat{\alpha}_n - \alpha)$ converges weakly to $\mathcal{N}(0, \pi^2/6)$ if $\omega_0 \in \{0, \pi\}$.

The first step in the proof of Theorem 2 is to obtain the distribution of the GPH estimator when the location of the pole is known. Let ω_n be the closest Fourier frequency to ω_0 , with the convention that if x_{k_1} and x_{k_2} are equidistant from ω_0 we take ω_n as the smallest of the two. Let $\tilde{\alpha}_n$ be the (infeasible) GPH estimator based on the knowledge of ω_0 :

$$\tilde{\alpha}_n = \sum_{1 \leq |k| \leq m} \gamma_k \log\{I_n(\omega_n + x_k)\}. \quad (2.6)$$

Proposition 1. *Under the assumptions of Theorem 2, $m^{1/2}(\tilde{\alpha}_n - \alpha)$ converges weakly to $\mathcal{N}(0, \pi^2/12)$ if $\omega_0 \in (0, \pi)$ and $\mathcal{N}(0, \pi^2/6)$ if $\omega_0 \in \{0, \pi\}$.*

Comments

- Assumption (2.5) corresponds to Assumption A1 of Robinson (1995) and is used to obtain covariance bounds for the discrete Fourier transform and log-periodogram ordinates of the process X (Lemmas 1 and 4 below). (A1) corresponds to A2 of Robinson (1995) and is used to control the bias of the GPH estimator.
- An asymptotic distribution for the estimator of ω_0 would obviously be of great interest, especially if the rate of convergence n could be obtained. Unfortunately, this has not been already achieved. Nevertheless, the rate of convergence of the present estimator is close to the parametric rate n obtained by Giraitis *et al.* (2001) and its empirical performance is quite good, as shown in section 3. Hidalgo (2001) proposes an alternative estimator for which he obtains an asymptotic distribution with a rate of convergence close to the parametric rate n provided the process X has enough finite moments.

- If $\omega_0 \in \{0, \pi\}$, note that $\hat{\alpha}_n$ is approximately equal to twice the GPH estimator \hat{d}_n of d whose asymptotic distribution under the same assumptions is $\mathcal{N}(0, \pi^2/24)$, cf. Robinson (1995). Hence, there is no efficiency loss incurred by estimating ω_0 . The asymptotic variance in the case $\omega_0 \in \{0, \pi\}$ is twofold the asymptotic variance in the case $\omega_0 \in (0, \pi)$, because $\epsilon_k(\omega_n) = \epsilon_{-k}(\omega_n)$ in the former case, while in the latter case, the m Fourier frequencies on the left and on the right of ω_0 can be considered “asymptotically i.i.d.”.
- The symmetrization procedure is necessary in the case $\omega_0 \in (0, \pi)$, in order to obtain the same rate of convergence as in the case $\omega_0 \in \{0, \pi\}$. Without the symmetrization, the maximum possible value of δ would be $(2/3) \wedge (2\beta/(2\beta + 1))$, ($2/3$ being the value corresponding to $\beta \geq 1$), instead of $2\beta/(2\beta + 1)$ here. The reason is that by symmetry, the first derivative of the spectral density at ω_0 is 0 when $\omega_0 \in \{0, \pi\}$, whereas it need not be so if $\omega_0 \in (0, \pi)$. For more details see Hidalgo (2001).
- Theorem 2 is proved under the assumption of Gaussianity. Using the techniques of Velasco (2000) and Hurvich *et al.* (2002), it could be proved under the weaker assumption of linearity in the strict sense. The derivations would then be extremely involved and lengthy. We prefer to give a simple proof under the assumption of Gaussianity.

Case of multiple singularities

Model (1.5) can be extended to allow for more than one spectral singularity or pole, that is

$$f(x) = \prod_{i=1}^s |1 - e^{i(x-\omega^i)}|^{-d_i} |1 - e^{i(x+\omega^i)}|^{-d_i} f^*(x),$$

where $\omega^i \neq \omega^j$ if $i \neq j$, $0 < d_i < 1$ if $\omega^i \neq 0$ and $0 < d_i < 1/2$ if $\omega^i = 0$, and f^* is a smooth function (C^2 over $[-\pi, \pi]$, say). Then the poles ω^i , $i = 1, \dots, s$ and the values d_i , $i = 1, \dots, s$, can be estimated sequentially due to the local character of our estimators as we now illustrate. Suppose for expositional purposes that $\omega^1 \neq 0$, $\omega^2 \neq 0$, $d_1 > d_2$ and $s = 2$. Then ω^1 and ω^2 can be estimated by

$$\hat{\omega}_n^1 = \frac{2\pi}{n} \arg \max_{1 \leq k \leq \tilde{n}} I_n(x_k) \quad \text{and} \quad \hat{\omega}_n^2 = \frac{2\pi}{n} \arg \max_{1 \leq k \leq \tilde{n}; |x_k - \hat{\omega}_n^1| \geq z_n/n} I_n(x_k),$$

respectively, where z_n is a non decreasing sequence such that for any positive real numbers κ and κ' ,

$$\log^\kappa(n) \ll \frac{n}{z_n} \ll n^{\kappa'}.$$

An example of such a sequence is $z_n = ne^{-\sqrt{\log(n)}}$. Then the rate of convergence of $\hat{\omega}_n^1$ to ω^1 will be unchanged and the rate of convergence of $\hat{\omega}_n^2$ to ω^2 will be $n/v_{2,n}$, where $v_{2,n}^{-2d_2} \log(n) = o(1)$.

Let us briefly explain why such a sequence z_n is needed to yield a $n/v_{2,n}$ consistent estimator of ω^2 . First the proof of Theorem 1 can be easily adapted to prove that $\hat{\omega}_n$ is still $n/v_{1,n}$ consistent with $v_{1,n}$ such that $v_{1,n}^{-d_1} \log(n) = o(1)$. Then if z_n is chosen as proposed above, with probability tending to one, $|\hat{\omega}_n^1 - \omega^1| \leq z_n/2n$. Hence $|\hat{\omega}_n^2 - \omega^2| \geq z_n/2n$ and

$$\max_{|x_k - \omega^1| \geq z_n/(2n), |x_k - \omega^2| \geq v_{2,n}/n} f(x_k) \leq \max((n/z_n)^{d_1}, (n/v_{2,n})^{d_2}) = (n/v_{2,n})^{d_2}.$$

As in the proof of Theorem 1, the final argument is that $\mathbb{E}[I_n(\omega_n^2)] \asymp n^{d_2}$, where ω_n^2 is the closest Fourier frequency to ω^2 .

If $d_1 = d_2$ then obviously $\hat{\omega}_n^1$ is no longer consistent. Nevertheless, it is not difficult to see that the pair $\{\hat{\omega}_n^1, \hat{\omega}_n^2\}$ is a n/v_n consistent estimator of the pair $\{\omega^1, \omega^2\}$, with $v_n = v_{1,n} = v_{2,n}$.

The estimators of d_1 and d_2 will be given by (2.4) but with $\hat{\omega}_n$ being replaced by $\hat{\omega}_n^1$ and $\hat{\omega}_n^2$ respectively and their rates of convergence will be the optimal ones. The proof of this result would be along the same lines as the proof of Theorem 2, with added technicalities.

The previous comments are valid only if the number of poles s is known. Clearly an important issue is to know how many poles do we have. For instance one may even doubt the actual presence of a pole. If indeed there is no pole then the value of $\hat{\omega}_n$ is spurious, but nevertheless we claim that in that case $\hat{\alpha}_n$ converges to zero, and more precisely, $\hat{\alpha}_n = O_P(m^{-\gamma})$ for any $\gamma \in (0, 1/2)$, under the assumptions of Theorem 2 with $\alpha = 0$ and assumption **(A1)** extended to the whole interval $[-\pi, \pi]$, that is, assuming $f = f^*$ is a smooth function over $[-\pi, \pi]$.

We briefly outline the arguments. Define $S_n(x_\ell) = \sum_{1 \leq |k| \leq m} \gamma_k \log\{I_n(x_\ell + x_k)/f(x_\ell + x_k)\}$ and let A be an arbitrary positive real number. Then:

$$\mathbb{P}(m^\gamma |\hat{\alpha}_n| \geq A) \leq \mathbb{P}(m^\gamma |S_n(\hat{\omega}_n)| \geq A/2) + \mathbb{P}(m^\gamma \left| \sum_{1 \leq |k| \leq m} \gamma_k \log\{f(\hat{\omega}_n + x_k)\} \right| \geq A/2).$$

Now, under the assumption on the sequence m , the last term is $o(1)$ since $\sum_{1 \leq |k| \leq m} \gamma_k = 0$ and it is assumed here that f satisfies **(A1)** since $\alpha = 0$. By applying Markov's inequality, we obtain, for any integer q :

$$\mathbb{P}(m^\gamma |S_n(\hat{\omega}_n)| \geq A/2) \leq \sum_{\ell=1}^{\tilde{n}} \mathbb{P}(m^\gamma |S_n(x_\ell)| \geq A/2) \leq \sum_{\ell=1}^{\tilde{n}} (A/2)^{-q} m^{\gamma q} \mathbb{E}[|S_n(x_\ell)|^q].$$

It can then be shown that, under the Gaussian assumption, $\mathbb{E}[|S_n(x_\ell)|^q] \leq C m^{-q/2}$, where the constant C depends only on f and q (using for instance Theorem 2.1 and following the lines of the proof of Theorem 3.1 in Soulier (2001)). Hence, we can conclude that

$$\lim_{A \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}(m^\gamma |S_n(\hat{\omega}_n)| \geq A/2) = 0,$$

which proves our claim.

It is plausible that the rate of convergence can be improved to $O_P(m^{-1/2})$, (possibly up to some logarithmic factor), but this is beyond the scope of the present paper.

An asymptotic distribution for $\hat{\alpha}_n$ could be obtained in some particular cases (such as Gaussian white noise), but there are probably many different situations, thus making it impossible to have a general result.

Minimax rates of convergence

In semiparametric estimation, rates of convergence are an important issue. Two problems are considered here. The first one is the estimation of the location of the pole ω_0 , and the second one is estimation of

the exponent of the singularity α when ω_0 is unknown.

When ω_0 is equal to zero and known, Giraitis *et al.* (1997) have proved that the best possible rate of convergence of an estimator of α under assumption **(A1)** is $n^{\beta/(2\beta+1)}$. Theorem 2 states that even if ω_0 is unknown, then $m^{1/2}(\hat{\alpha}_n - \alpha)$ is asymptotically normal at any rate m such that $m/n^{2\beta/(2\beta+1)} \rightarrow 0$. Since an extra unknown parameter cannot possibly result in an improved rate of convergence, Theorem 2 shows that the rate of convergence of $\hat{\alpha}_n$ is optimal as far as obtaining a central limit theorem is concerned.

The problem of estimating the pole is not yet satisfactorily answered. It has been conjectured that the best possible rate of convergence, even in a parametric context is n , meaning that there exists a constant $c > 0$ such that for any estimator $\tilde{\omega}_n$ based on observations X_1, \dots, X_n , it holds that $\lim_{n \rightarrow \infty} \mathbb{P}(n|\tilde{\omega}_n - \omega_0| > c) > 0$. In a parametric context, Giraitis *et al.* (2001) have defined an estimator $\hat{\omega}_n$ such that $\hat{\omega}_n - \omega_0 = O_P(1/n)$, but they have not proved that this rate of convergence is optimal. In the present semiparametric context, a lower bound for the rate of convergence is not known, and up to our best knowledge, there exists no estimator better (and simpler) than the one presented here.

There is an obvious problem of identifiability of ω_0 if the singularity α or the fractional differencing coefficient d is not bounded away from 0. If d is bounded away from zero, say $d > d_0$, Theorem 1 implies that the estimator presented here attains the rate of convergence $n \log^{-\gamma}(n)$ for any $\gamma > 1/2d_0$. We have not been able to obtain a lower bound for the rate of convergence of an estimator of ω_0 when d is bounded away from zero by a fixed constant. If d can get closer and closer to zero at a given rate, we have obtained the following result.

Theorem 3. *Let s be a positive integer and let d_n be a sequence of real numbers such that $\lim_{n \rightarrow \infty} (d_n + nd_n^s) = 0$. Denote $d(\omega_0) = 1$ if $\omega_0 \in (0, \pi)$ and $d(\omega_0) = 1/2$ if $\omega_0 \in \{0, \pi\}$. There exists a positive constant c such that,*

$$\inf_{\tilde{\omega}_n} \inf_{\omega_0 \in [0, \pi]} \inf_{d_n \leq d < d(\omega_0)} \sup_{f^*} \mathbb{P}_{\omega_0, d, f^*}(nd_n|\tilde{\omega}_n - \omega_0| \geq c) > 0,$$

where $\mathbb{P}_{\omega_0, d, f^*}$ denotes the distribution of any second order stationary process with spectral density $f(x) =$ and \sup_{f^*} means that the supremum is evaluated over all function f^* such that (2.5) holds.

Choosing $d_n = 1/n$ proves that ω_0 cannot be consistently estimated (in the minimax sense) if d is not bounded away from zero.

3 Monte-Carlo simulations

In order to investigate how well \hat{w}_n behaves and the relative performance of $\hat{\alpha}_n$ compared to the infeasible estimator $\tilde{\alpha}_n$ (with $\tilde{\alpha}_n$ defined in (2.6)) in small samples, a limited Monte-Carlo study was carried out. We report the result for the parameter g defined as $\alpha/2$ and the corresponding estimators \hat{g}_n and \tilde{g}_n instead of α , to be consistent with the literature. When $\omega_0 \in \{0, \pi\}$, the expected variance of \hat{g}_n is the usual $\pi^2/24m$, whereas when $\omega_0 \in (0, \pi)$, the expected variance of \hat{g}_n is $\pi^2/48m$ which matches that obtained by Arteché and Robinson (2000).

We have simulated 5000 replications of series of length $n = 256, 512$ and 1024 of Gaussian processes with spectral densities $|1 - e^{ix}|^{-2g}$ and $|1 - e^{i(x-\pi/2)}|^{2g}|1 - e^{i(x+\pi/2)}|^{2g}$ with $g = 0.1, 0.2, 0.3$ and 0.4 .

They were generated by the method of Davies and Harte (1987) using formulae for autocovariance given in Arteche and Robinson (2000). The results of the Monte-Carlo experiment are given in Tables 1 to 4.

Table 1 shows the performance of \hat{w}_n . As expected, the finite sample performance of \hat{w}_n becomes better as n and/or g increases. Specifically for $g = 0.1$, the standard deviation of \hat{w}_n is far larger than when $g = 0.4$. Moreover it appears that the precision of \hat{w}_n is not affected by the true value of w_0 . Note that the positive larger bias for $w_0 = 0$ compared to the case $w_0 = \pi/2$ is due to the fact that the estimator cannot take negative values, so that $\hat{w}_n - w_0 \geq 0$ for $w_0 = 0$, which implies that it is a nonnegative random variable.

A quick look at Tables 2 to 4, especially the Table for the *Mean Square Error* (M.S.E.), indicates that as predicted by the theory, \hat{g}_n and \tilde{g}_n given in (2.6) are very close, although for g small, for instance $g = 0.1$, \hat{g}_n tends to have a larger bias than \tilde{g}_n . On the other hand, for larger values of g , not only the M.S.E. of \hat{g}_n and \tilde{g}_n are similar but their bias and standard deviation are also very close. So, the Monte-Carlo experiment tends to confirm the theoretical results obtained in Theorem 2, that is, that $m^{1/2}(\hat{g}_n - \tilde{g}_n) = o_p(1)$. Note that the method employed to prove Theorem 2 is precisely to show that the latter is the case. The empirical standard deviation of \hat{g}_n and \tilde{g}_n are close to the theoretical ones, especially for larger values of g .

Remark A more erratic behaviour of \hat{w}_n and \hat{g}_n when g is small, for instance $g = 0.1$ is observed. One possible explanation is that the spectral density is flatter for $g = 0.1$ than for $g = 0.4$. So to locate and obtain the maximum becomes harder. This will translate into the fact that the estimator of g will be worst. Another possible explanation is due to the randomness of $I_n(x)$, which implies that the error $I_n(x) - f(x)$ becomes larger relative to $I_n(x)$, or $f(x)$, as g becomes smaller. That is, the ratio of the noise of $I_n(x)$, $I_n(x) - f(x)$, to its signal given by $I_n(x)$, or $f(x)$, becomes larger, so more observations are expected to be needed to obtain a more accurate estimate. One way to alleviate this problem, in small samples, could be by looking at the maximum not of $I_n(x_k)$ but of the average $\frac{1}{3} \{I_n(x_{k-1}) + I_n(x_k) + I_n(x_{k+1})\}$. This would have the effect to reduce the variability of the noise of $I_n(x_k)$, e.g. the variance of $I_n(x) - f(x)$.

4 Proofs

In all the subsequent proofs, c , C , $c(\dots)$ will denote numerical constants whose values depend only on their arguments and may change upon each appearance.

As always, the key ingredient of the proof is a bound on covariance of renormalised discrete Fourier transform ordinates. Such a bound has originally been obtained by Robinson (1995) in the case $\omega_0 = 0$, and then generalised to the case $\omega_0 \in (0, \pi)$ by Arteche and Robinson (2000) and Giraitis *et al.* (2001, Lemma 4.6). Define $d_k(\omega) = (2\pi n f(\omega + x_k))^{-1/2} \sum_{t=1}^n X_t e^{it(\omega + x_k)}$.

Lemma 1. *Let X be a second order stationary process with spectral density f that satisfies (2.2) and (2.3). Then, for $1 \leq k \leq j \leq \tilde{n}$,*

$$|\mathbb{E}[d_k(\omega_n)d_j(\omega_n)]| + |\mathbb{E}[d_k(\omega_n)\bar{d}_j(\omega_n)] - \delta_{k,j}| \leq C(d, l^*) \log(1+j)k^{-1}. \quad (4.1)$$

Proof of Theorem 1. We must prove that for any $\epsilon > 0$, $\lim_{n \rightarrow \infty} \mathbb{P}(nv_n^{-1}|\hat{\omega}_n - \omega_0| > \epsilon) = 0$. Recall that ω_n is the closest Fourier frequency to ω_0 . Since $|\omega_n - \omega_0| \leq \pi/n$, it is enough to prove that

$\lim_{n \rightarrow \infty} \mathbb{P}(nv_n^{-1}|\hat{\omega}_n - \omega_n| > \epsilon) = 0$. Define $M_n = \max_{1 \leq k \leq \tilde{n}; x_k \neq \omega_n} \frac{I_n(x_k)}{f(x_k)}$ and $\Lambda_{\epsilon, n} = \{k : 1 \leq k \leq \tilde{n}, |x_k - \omega_n| > \epsilon v_n/n\}$. Then, using (2.2) and the fact that $f(x) = |a(x)|^2 \mathbb{E}[Z_0^2]/(2\pi)$, we have

$$\begin{aligned} \mathbb{P}(nv_n^{-1}|\hat{\omega}_n - \omega_n| > \epsilon) &= \mathbb{P}(|\hat{\omega}_n - \omega_n| > \epsilon v_n/n) \\ &\leq \mathbb{P}(I_n(\omega_n) < M_n \max_{k \in \Lambda_{\epsilon, n}} f(x_k)) \\ &\leq \mathbb{P}(I_n(\omega_n) < M_n c(\omega_0, L)(\epsilon v_n/n)^{-2\nu}) \\ &\leq \mathbb{P}(I_n(\omega_n) < 2c(\omega_0, L) \log(n)(\epsilon v_n/n)^{-2\nu}) + \mathbb{P}(M_n > 2 \log(n)). \end{aligned}$$

Let us first prove that $\lim_{n \rightarrow \infty} \mathbb{P}(M_n > 2 \log(n)) = 0$. To that purpose, it suffices to show that $M_n/\log(n)$ converges in probability to 1.

Denote $I_n^Z(x) = \frac{1}{2\pi n} |\sum_{k=1}^n Z_k e^{ikx}|^2$ and $M_n^Z = 2\pi \max_{1 \leq k \leq \tilde{n}} I_n^Z(x_k)$. We can assume without loss of generality that $\mathbb{E}[Z_0^2] = 1$. By definition of M_n and M_n^Z ,

$$\left| \frac{M_n}{\log(n)} - \frac{M_n^Z}{\log(n)} \right| \leq \frac{\max_{1 \leq k \leq \tilde{n}; x_k \neq \omega_n} |I_n(x_k)/f(x_k) - 2\pi I_n^Z(x_k)|}{\log(n)}.$$

It has been proved in An, Chen and Hannan (1983) that $\limsup_n M_n^Z/\log(n) \leq 1$. Davis and Mikosch (1999) showed that the limsup is actually a limit and that it is equal to 1. Define

$$R_n = \max_{1 \leq k \leq \tilde{n}; x_k \neq \omega_n} |I_n(x_k)/f(x_k) - 2\pi I_n^Z(x_k)|.$$

To prove that $R_n/\log(n)$ tends to zero in probability, we need the following bound, which generalises Lemma 1 to higher moments and can be proved in a similar way as Lemma 11 in Hurvich *et al.* (2002): under the assumptions of Theorem 1, there exists a constant C such that for all n and $k \neq 0$ such that $0 < x_k + \omega_n < \pi$,

$$\mathbb{E}[\{I_n(\omega_n + x_k)/f(\omega_n + x_k) - 2\pi I_n^Z(\omega_n + x_k)\}^4] \leq C \log^2(k)k^{-2}.$$

Bounding the maximum by the sum and applying the Markov inequality we obtain that

$$\mathbb{P}(R_n/\log(n) > \epsilon) \leq \sum_{1 \leq k \leq \tilde{n}, x_k \neq \omega_n} \frac{\mathbb{E}[\{I_n(x_k)/f(x_k) - 2\pi I_n^Z(x_k)\}^4]}{\epsilon^4 \log^4(n)} = O(\log^{-4}(n)),$$

since the series $\log^2(k)k^{-2}$ is summable. Thus $R_n/\log(n)$ converges to zero in probability, which implies that $M_n/\log(n)$ converges to one in probability.

We must now prove that $\lim_{n \rightarrow \infty} \mathbb{P}(I_n(\omega_n) < 2c(\omega_0, L) \log(n)(\epsilon v_n/n)^{-2\nu}) = 0$.

If $\omega_0 \in \{0, \pi\}$, then $\omega_n = \omega_0$ for all n . If the process X is not centered, *i.e.* $\mu \neq 0$, then it is easily seen that $I_n(\omega_0) = n\mu + o_P(n)$. Hence $\lim_{n \rightarrow \infty} \mathbb{P}(I_n(\omega_0) < 2c(\omega_0, L) \log(n)(\epsilon v_n/n)^{-2\nu}) = 0$, because $\nu < 1/2$.

If now $\omega_0 \in (0, \pi)$ or $\omega_0 \in \{0, \pi\}$ and the process is centered, define $\chi_n = \{\mathbb{E}[I_n(\omega_n)]\}^{-1} I_n(\omega_n)$. As proved in Lemma 2 below, there exists a constant K such that $\mathbb{E}[I_n(\omega_n)] \geq Kn^{2\nu}$. Hence

$$\begin{aligned} \mathbb{P}(I_n(\omega_n) < 2C(\omega_0, L) \log(n)(\epsilon v_n/n)^{-2\nu}) &= \mathbb{P}(\chi_n \leq Cn^{2\nu} \{\mathbb{E}[I_n(\omega_n)]\}^{-1} (\epsilon v_n)^{-2\nu} \log(n)) \\ &\leq \mathbb{P}(\chi_n \leq C(\epsilon v_n)^{-2\nu} \log(n)). \end{aligned}$$

By assumption $\lim_{n \rightarrow \infty} \log(n)v_n^{-\nu} = 0$. Since it is proved in Lemma 2 that χ_n converges to a distribution with no mass at zero, the proof of Theorem 1 is concluded. \square

Lemma 2. *Under the assumption of Theorem 1 and if moreover the process X is centered in the case $\omega_0 \in \{0, \pi\}$, then $\{\mathbb{E}[I_n(\omega_n)]\}^{-1/2} I_n(\omega_n)$ converges weakly to a distribution without mass at 0.*

Proof. Define $\sigma_n^2 = \mathbb{E}[I_n(\omega_n)]$ and $\zeta_n = (2\pi n)^{-1/2} \sum_{k=1}^n X_k e^{ik\omega_n}$. If $\omega_0 \in (0, \pi)$, then, for large enough n , ζ_n is centered even if the process X is not. Thus we can assume, without loss of generality, that X is centered. Since moreover X is a linear process, the asymptotic normality of $\sigma_n^{-1} \zeta_n$ holds as soon as $\lim_{n \rightarrow \infty} \sigma_n = \infty$ (cf. Theorem 18.6.5 in Ibragimov and Linnik, 1971). Under assumption (2.2), we have:

$$\begin{aligned} \sigma_n^2 &= \mathbb{E}[I_n(\omega_n)] = \frac{1}{2\pi n} \int_{-\pi}^{\pi} \frac{\sin^2(nx/2)}{\sin^2(x/2)} f(x + \omega_n) dx \\ &\geq K' n^{-1} \int_{2\pi/n}^{\pi} \frac{\sin^2(nx/2)}{\sin^2(x/2)} (x + \omega_n - \omega_0)^{-2\nu} dx \\ &\geq K'' n^{-1} \int_{2\pi/n}^{\pi} \frac{\sin^2(nx/2)}{\sin^2(x/2)} x^{-2\nu} dx \geq K''' n^{2\nu}, \end{aligned}$$

which tends to infinity. □

Proof of Proposition 1. For $\omega \in [0, \pi]$, denote

$$\tilde{\alpha}_n(\omega) = \sum_{1 \leq |k| \leq m} \gamma_k \log\{I_n(\omega + x_k)\}.$$

Note that with this notation, $\hat{\alpha}_n = \tilde{\alpha}_n(\hat{\omega}_n)$ and $\tilde{\alpha}_n = \tilde{\alpha}_n(\omega_n)$. Define also $l^* := \log(f^*)$, $\epsilon_k(\omega) = \log\{I_n(\omega + x_k)/f(\omega + x_k)\} + \mathcal{E}$ (where $\mathcal{E} = -0.577216\dots$ is Euler's constant), $\beta_k = s_m^{-1}(g(x_k) - \bar{g}_m) = s_m \gamma_k$ and $\xi_m(\omega) = \sum_{1 \leq |k| \leq m} \beta_k \epsilon_k(\omega)$. With these notations, we obtain that for any ω ,

$$\tilde{\alpha}_n(\omega) = d \sum_{1 \leq |k| \leq m} \gamma_k \{g(x_k + \omega - \omega_0) + g(x_k + \omega + \omega_0)\} + \sum_{1 \leq |k| \leq m} \gamma_k l^*(\omega + x_k) + s_m^{-1} \xi_m(\omega). \quad (4.2)$$

Replacing ω with ω_n in (4.2) above, we get

$$\begin{aligned} s_m(\tilde{\alpha}_n(\omega_n) - \alpha) &= d \mathbf{1}_{(0, \pi)}(\omega_n) \sum_{1 \leq |k| \leq m} \beta_k \{g(x_k + \omega_n - \omega_0) - g(x_k)\} \\ &\quad + d \mathbf{1}_{(0, \pi)}(\omega_n) \sum_{1 \leq |k| \leq m} \beta_k g(x_k + \omega_n + \omega_0) + \sum_{1 \leq |k| \leq m} \beta_k l^*(\omega_n + x_k) + \xi_m(\omega_n). \end{aligned} \quad (4.3)$$

Since by definition, $\sum_{1 \leq |k| \leq m} \beta_k = 0$, (4.3) can be written as

$$\begin{aligned} s_m(\tilde{\alpha}_n(\omega_n) - \alpha) &= d \mathbf{1}_{(0, \pi)}(\omega_n) \sum_{1 \leq |k| \leq m} \beta_k \{g(x_k + \omega_n - \omega_0) - g(x_k)\} \\ &\quad + d \mathbf{1}_{(0, \pi)}(\omega_n) \sum_{1 \leq |k| \leq m} \beta_k \{g(x_k + \omega_n + \omega_0) - g(\omega_n + \omega_0)\} \\ &\quad + \sum_{1 \leq |k| \leq m} \beta_k \{l^*(\omega_n + x_k) - l^*(\omega_n)\} + \xi_m(\omega_n) \\ &=: b_{1,m}(\omega_n) + b_{2,m}(\omega_n) + B_m(l^*) + \xi_m(\omega_n). \end{aligned}$$

The terms $b_{1,m}$ and $b_{2,m}$ vanish if $\omega_0 \in \{0, \pi\}$ and are well defined if $\omega_0 \in (0, \pi)$, at least for large enough n , since $m/n \rightarrow 0$, which implies that $x_k + \omega_n \in (0, \pi)$ and $x_k + \omega_n + \omega_0 \neq 0$ (modulo 2π). The third bias term $B_m(l^*)$ depends upon the smoothness of l^* at ω_0 . The next Lemma gives bounds for the deterministic terms and some relevant quantities. Its proof relies on elementary computations and is omitted.

Lemma 3. *Assume that $\lim_{m \rightarrow \infty} (1/m + m/n) = 0$, then*

$$\max_{1 \leq |k| \leq m} |\beta_k| = O(\log(m)m^{-1/2}), \quad \lim_{m \rightarrow \infty} s_m^2/m = 2,$$

$$b_{1,m}^2(\omega_n) \leq C \log^2(m)m^{-1/2} \quad \text{and} \quad b_{2,m}^2(\omega_0) \leq \frac{Cm^5}{n^4(\omega_0 \wedge (\pi - \omega_0))^2} \mathbf{1}_{(0,\pi)}(\omega_0).$$

If **(A1)** holds, then $B_m^2(l^*) \leq Cm^{2\beta+1}n^{-2\beta}$.

For a Gaussian process, it is now a well established fact that Lemma 1 and the covariance inequality for functions of Gaussian vectors of Arcones (1994, Lemma 1) can be used to derive bounds for the bias and covariance of the renormalised log-periodogram ordinates. This has been shown in various places; see for instance Robinson (1995), Giraitis *et al.* (1997), Hurvich *et al.* (1998), Moulines and Soulier (1999) or Iouditsky *et al.* (2001). A central limit theorem for weighted sums of log-periodogram ordinates can also be derived, as in Robinson (1995), Moulines and Soulier (1999) and Soulier (2001) in a more general framework. Thus we state without proof the following Lemma.

Lemma 4. *Let X be a Gaussian process whose spectral density satisfies (2.5). Then, for $1 \leq |k| \leq j$,*

$$|\mathbb{E}[\epsilon_k(\omega_n)]| \leq C(d, l^*) \log(k+1)k^{-1}, \quad (4.4)$$

$$\left| \mathbb{E}[\epsilon_k(\omega_n)\epsilon_j(\omega_n)] - \frac{\pi^2}{6} \delta_{k,j} \right| \leq C(d, l^*) \log^2(j+1)k^{-2}. \quad (4.5)$$

If $\lim_{m \rightarrow \infty} (\log^2(n)/m + m/n) = 0$ then $\xi_m(\omega_n)$ converges weakly to $\mathcal{N}(0, \pi^2/3)$ if $\omega_0 \in \{0, \pi\}$ and $\mathcal{N}(0, \pi^2/6)$ if $\omega_0 \in (0, \pi)$.

Lemmas 3 and 4 yield Proposition 1. □

Proof of Theorem 2. Write now

$$s_m(\hat{\alpha}_n - \alpha) = s_m(\tilde{\alpha}_n - \alpha) + s_m(\hat{\alpha}_n - \tilde{\alpha}_n).$$

By Proposition 1 it is sufficient to prove that $s_m(\hat{\alpha}_n - \tilde{\alpha}_n)$ tends to zero in probability, *i.e.* for all $\epsilon > 0$,

$$\lim_{n \rightarrow \infty} \mathbb{P}(s_m|\tilde{\alpha}_n - \hat{\alpha}_n| > \epsilon) = 0.$$

For any $M > 0$ and $0 < \gamma < 1/2$, we can write,

$$\begin{aligned} \mathbb{P}(s_m|\hat{\alpha}_n - \tilde{\alpha}_n| > \epsilon) &\leq \mathbb{P}(s_m|\hat{\alpha}_n - \tilde{\alpha}_n| > \epsilon, nm^{-\gamma}|\hat{\omega}_n - \omega_n| \leq 2\pi M) \\ &\quad + \mathbb{P}(nm^{-\gamma}|\hat{\omega}_n - \omega_n| > 2\pi M) \\ &\leq \mathbb{P}\left(\max_{\theta \in [-M, M]} |S_n(\theta)| > \epsilon\right) + \mathbb{P}(nm^{-\gamma}|\hat{\omega}_n - \omega_n| > 2\pi M), \end{aligned}$$

where we have defined for any $\theta > 0$,

$$S_n(\theta) = s_m(\tilde{\alpha}_n(\omega_n + 2\pi[m^\gamma\theta]/n) - \tilde{\alpha}_n).$$

Theorem 1 implies that for any $0 < \gamma < 1/2$, if $m = m(n) = n^\delta$ for some $0 < \delta < 1$, then $\lim_{n \rightarrow \infty} \mathbb{P}(nm^{-\gamma}|\hat{\omega}_n - \omega_n| > M) = 0$. Hence, the proof of Theorem 2 is completed by the following Lemma. \square

Lemma 5. *Let X be a Gaussian process whose spectral density satisfies (2.5) and **(A1)**. If $m = [n^\delta]$ for some $0 < \delta < 2\beta/(2\beta + 1)$, then for any $0 < \gamma < 1/2$, the sequence $S_n(\theta)$ converges to zero in probability uniformly with respect to $\theta \in [-M, M]$, that is, for all $\epsilon > 0$,*

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(\max_{\theta \in [-M, M]} |S_n(\theta)| > \epsilon\right) = 0.$$

Proof. Denote $j(\theta) = [m^\gamma\theta]$, $\epsilon_k = \epsilon_k(\omega_n)$ and recall that by symmetry, $\beta_k = \beta_{-k}$ for all $k = 1, \dots, m$. Write now $S_n(\theta)$ as follows:

$$\begin{aligned} S_n(\theta) &= \sum_{1 \leq |k| \leq m} \beta_k \{\log(I_n(\omega_n + x_{k+j(\theta)})) - \log(I_n(\omega_n + x_k))\} \\ &= \beta_{j(\theta)} \{\log[I_n(\omega_n)] - \log[I_n(\omega_n - x_{j(\theta)})]\} \end{aligned} \quad (4.6)$$

$$+ d \sum_{1 \leq |k| \leq m, k \neq -j(\theta)} \beta_k \{g(x_{k+j(\theta)} + \omega_n - \omega_0) - g(x_k + \omega_n - \omega_0)\} \quad (4.7)$$

$$+ d \sum_{1 \leq |k| \leq m, k \neq -j(\theta)} \beta_k \{g(x_{k+j(\theta)} + \omega_n + \omega_0) - g(x_k + \omega_n + \omega_0)\} \quad (4.8)$$

$$\begin{aligned} &+ \sum_{1 \leq |k| \leq m, k \neq -j(\theta)} \beta_k \{I^*(x_{k+j(\theta)} + \omega_n) - I^*(x_k + \omega_n)\} + \sum_{1 \leq |k| \leq m, k \neq -j(\theta)} \beta_k \{\epsilon_{k+j(\theta)} - \epsilon_k\} \\ &=: \beta_{j(\theta)} \{\log[I_n(\omega_n)] - \log[I_n(\omega_n - x_{j(\theta)})]\} + dA_n(\theta) + dA'_n(\theta) + B_n(\theta) + T_n(\theta). \end{aligned}$$

Under the assumption of Gaussianity, it is easily seen that for any $\lambda \in [-\pi, \pi]$, $\mathbb{E}[\log^2(I_n(\lambda))] = O(\log^2(n))$. Applying Lemma 3, we obtain the following bound for the term (4.6):

$$\beta_{j(\theta)}^2 \{\mathbb{E}[\log^2(I_n(\omega_n))] + \mathbb{E}[\log^2(I_n(\omega_n - x_{j(\theta)}))]\} = O(m^{-1} \log^4(n)),$$

uniformly with respect to θ .

Consider now the term in (4.7), *i.e.* A_n . For $x \neq 0$, g is differentiable and $g'(x) = -\frac{1}{2} \cot(\frac{x}{2})$. Since moreover g is even, this implies that for all x, y in $[-\pi, \pi] \setminus \{0\}$, $|g(x) - g(y)| \leq |x - y|/(|x| \wedge |y|)$. Thus we get the bounds

$$\begin{aligned} &|g(x_k + x_j + \omega_n - \omega_0) - g(x_k + \omega_n - \omega_0)| \\ &\leq \begin{cases} j/(k - 1/2), & \text{if } 1 \leq k \leq m, \\ j/((-k - j) - 1/2), & \text{if } -m \leq k \leq -1 - j, \\ j/((-k - 1/2) \wedge (j + k - 1/2)), & \text{if } -j + 1 \leq k \leq -1. \end{cases} \end{aligned} \quad (4.9)$$

Using these bounds, the property $\sum_{k=1}^m \beta_k^2 = 1$, and applying the Hölder inequality, we obtain:

$$|A_n(\theta)| \leq j(\theta) \log^2(m) m^{-1/2} \leq \log^2(m) m^{\gamma-1/2}.$$

If $\omega_0 \in (0, \pi)$, we must also consider A'_n , whereas if $\omega_0 \in \{0, \pi\}$, $A'_n = A_n$. For n large enough, $\omega_n + x_k \in (0, \pi)$ and $\omega_n + x_{k+j(\theta)} \in (0, \pi)$ since by assumption $m/n \rightarrow 0$. Thus, there exists a constant C which depends only on the function g such that

$$|g(\omega_0 + \omega_n + x_{k+j(\theta)}) - g(\omega_0 + \omega_n + x_k)| \leq Cx_{j(\theta)}/(\omega_0 \wedge (\pi - \omega_0)).$$

The property $\sum_{k=1}^m \beta_k^2 = 1$ implies that $\sum_{1 \leq |k| \leq m} |\beta_k| \leq \sqrt{m}$, hence, if $\omega_0 \in (0, \pi)$, we obtain:

$$|A'_n(\theta)| \leq \frac{Cx_{j(\theta)}}{\omega_0 \wedge (\pi - \omega_0)} \sum_{1 \leq |k| \leq m} |\beta_k| \leq \frac{2\pi CMm^{1/2+\gamma}}{n(\omega_0 \wedge (\pi - \omega_0))}.$$

The term B_n presents no difficulty since the function l^* is smooth everywhere. Since $\gamma < 1/2$, it is easily obtained that

$$|B_n(\theta)| \leq Cm^{\beta+1/2}n^{-\beta}.$$

Thus the sequences A_n , A' and B_n converge to zero uniformly on compact sets. We now examine the properties of the process T_n . To prove that T_n converges to zero in probability uniformly with respect to $\theta \in [-M, M]$, we apply Theorem 15.6 in Billingsley (1968). It suffices to check that for all $\theta \in [-M, M]$, $T_n(\theta)$ tends to zero in probability and that for all $\theta, \theta_1, \theta_2 \in [-M, M]$ such that $\theta_1 < \theta < \theta_2$,

$$\mathbb{E}[|T_n(\theta_1) - T_n(\theta)| |T_n(\theta_2) - T_n(\theta)|] \leq C(\theta_1 - \theta_2)^2. \quad (4.10)$$

We can restrict our attention to those (θ_1, θ_2) such that $|\theta_1 - \theta_2| \geq m^{-\gamma}$, since otherwise, the left hand side of (4.10) is zero. Moreover, applying the Cauchy-Schwarz inequality, it is sufficient to check that for all $\theta_1, \theta_2 \in [-M, M]$ such that $|\theta_1 - \theta_2| \geq m^{-\gamma}$, it holds that

$$\mathbb{E}[|T_n(\theta_1) - T_n(\theta_2)|^2] \leq C(\theta_1 - \theta_2)^2. \quad (4.11)$$

Let $\theta_1, \theta_2 \in [-M, M]$ and consider the increments of $T_n(\theta_1, \theta_2) := T_n(\theta_1) - T_n(\theta_2)$. Assume that $\theta_1 < \theta_2$ and denote $j_i = j(\theta_i)$, $i = 1, 2$. Without loss of generality, it can be assumed that $j_1 < j_2$, since otherwise, $T_n(\theta_1, \theta_2) = 0$. We can then split $T_n(\theta_1, \theta_2)$ in the following terms.

$$\begin{aligned} T_n(\theta_1, \theta_2) &= \sum_{\substack{k=-m+j_1 \\ k \neq 0, k \neq j_1}}^{m+j_1} \beta_{k-j_1} \epsilon_k - \sum_{\substack{k=-m+j_2 \\ k \neq 0, k \neq j_2}}^{m+j_2} \beta_{k-j_2} \epsilon_k - \beta_{j_1} \epsilon_{j_1} + \beta_{j_2} \epsilon_{j_2} \\ &= \sum_{k=-m+j_1}^{-m+j_2-1} \beta_{k-j_1} \epsilon_k - \sum_{k=m+j_1+1}^{m+j_2} \beta_{k-j_2} \epsilon_k \end{aligned} \quad (4.12)$$

$$+ \sum_{k=-m+j_2}^{-1} (\beta_{k-j_1} - \beta_{k-j_2}) \epsilon_k + \sum_{k=1}^{j_1-1} (\beta_{k-j_1} - \beta_{k-j_2}) \epsilon_k \quad (4.13)$$

$$+ \sum_{k=j_1+1}^{j_2-1} (\beta_{k-j_1} - \beta_{k-j_2}) \epsilon_k + \sum_{k=j_2+1}^{m+j_1} (\beta_{k-j_1} - \beta_{k-j_2}) \epsilon_k \quad (4.14)$$

$$- \beta_{j_1} \epsilon_{j_1} + \beta_{j_2} \epsilon_{j_2} - \beta_{j_1-j_2} \epsilon_{j_1} - \beta_{j_2-j_1} \epsilon_{j_2} \quad (4.15)$$

We have three kind of terms to bound. The sums in (4.12) have only $j_2 - j_1$ terms. Applying Lemmas 3 and 4, we obtain:

$$\mathbb{E}\left[\left(\sum_{k=-m+j_1}^{-m+j_2-1} \beta_{k-j_1} \epsilon_k\right)^2\right] + \mathbb{E}\left[\left(\sum_{k=m+j_1+1}^{m+j_2} \beta_{k-j_2} \epsilon_k\right)^2\right] \leq Cm^{-1} \log^2(m)(j_2 - j_1)^2 \leq Cm^{2\gamma-1} \log^2(m)(\theta_1 - \theta_2)^2.$$

To bound the sums in (4.13) and (4.14), we need a bound for $\beta_{k-j_1} - \beta_{k-j_2}$. Recall that by definition, $\beta_k = s_m(g(x_k) - \bar{g}_m)$ and that $|g(x) - g(y)| \leq |x - y|/(|x| \wedge |y|)$. Thus, if for any integer $k \neq j_1, j_2$, it holds that

$$|\beta_{k-j_1} - \beta_{k-j_2}| \leq Cm^{-1/2}|j_1 - j_2|/(|k - j_1| \wedge |k - j_2|). \quad (4.16)$$

Let us evaluate for instance the second moment of the last term in (4.14), which can be expressed as

$$\begin{aligned} \mathbb{E}[\sum_{k=j_2+1}^{m+j_1} (\beta_{k-j_1} - \beta_{k-j_2})\epsilon_k]^2 &= \sum_{k=j_2+1}^{m+j_1} (\beta_{k-j_1} - \beta_{k-j_2})^2 \text{var}(\epsilon_k) \\ &\quad + 2 \sum_{j_2+1 \leq k < l \leq m+j_1} (\beta_{k-j_1} - \beta_{k-j_2})(\beta_{l-j_1} - \beta_{l-j_2}) \text{cov}(\epsilon_k, \epsilon_l). \end{aligned} \quad (4.17)$$

For all $k \in \{j_2 + 1, \dots, m + j_1\}$, it holds that $|k - j_1| \wedge |k - j_2| = k - j_2$. Using Lemmas 3, 4 and the bound in (4.16), we obtain that the first term on the right hand side of (4.17) is bounded by

$$Cm^{-1}(j_2 - j_1)^2 \sum_{k=j_2+1}^{m+j_1} (k - j_2)^{-2} \leq C(\theta_1 - \theta_2)^2 m^{2\gamma-1},$$

whereas the second term on the right of (4.17) is bounded in absolute value by

$$Cm^{-1}(j_2 - j_1)^2 \sum_{j_2+1 \leq k < l \leq m+j_1} k^{-2}(k - j_2)^{-1}(l - j_2)^{-1} \log^2(l) \leq C \log^3(m)(\theta_1 - \theta_2)^2 m^{2\gamma-1}.$$

The other sums in (4.13) and (4.14) are dealt with similarly. To complete the investigation of T_n , we apply Lemmas 3 and 4 to obtain that

$$(\beta_{j_1}^2 + \beta_{j_2}^2 + \beta_{j_1-j_2}^2)\mathbb{E}[\epsilon_{j_1}^2 + \epsilon_{j_2}^2] \leq C \log^2(m)m^{-1}.$$

Altogether, since by assumption $|\theta_2 - \theta_1| \geq m^{-\gamma}$, we have proved that, for sufficiently small $\eta > 0$, and for a constant depending only on ω_0 ,

$$\mathbb{E}[T_n^2(\theta_1, \theta_2)] \leq C(\omega_0)m^{-\eta}(\theta_1 - \theta_2)^2.$$

We have proved (4.11) and using the same techniques, we can prove that $E[T_n(\theta)] = O(m^{-\eta})$ for sufficiently small η . Hence, applying Billingsley (1968, Theorem 15.6) we conclude that T_n converges uniformly to 0 in probability. This concludes the proof of Lemma 5. \square

Proof of Theorem 3. Let $\Theta_n = [0, \pi] \times [d_n, 2d_n]$ and denote $\theta = (\lambda, d)$, $f_\theta^*(x) = \exp\{d \sum_{j=1}^{p_n-1} \alpha_j(\lambda) \cos(jx)\}$ and $f_\theta(x) = |1 - e^{i(x-\lambda)}|^{-d} |1 - e^{i(x+\lambda)}|^{-d} f_\theta^*(x)$, where $\alpha_j(\lambda) = 2 \cos(j\lambda)j^{-1}$ and p_n is a non decreasing sequence of integers such that $p_n = o(n)$, $d_n p_n = o(1)$. With these notations, we can write

$$\log(|1 - e^{i(x-\lambda)}| |1 - e^{i(x+\lambda)}|) = - \sum_{j=1}^{\infty} \alpha_j(\lambda) \cos(jx) \quad \text{and} \quad f_\theta(x) = \exp\{d \sum_{j=p_n}^{\infty} \alpha_j(\lambda) \cos(jx)\}.$$

The assumption $d_n p_n = o(1)$ ensures that f_θ satisfies (2.5) for large enough n . Denote \mathbb{E}_θ the expectation with respect to the distribution of a stationary Gaussian process with spectral density f_θ . Let q be a continuously differentiable probability density function on $[0, \pi]$ with finite information, *i.e.*

$\int_0^\pi q'^2(s)/q(s)ds < \infty$. Let q_n be a density defined on Θ_n by $q_n(\theta) = \pi d_n^{-1} q(\pi d_n^{-1}(d - d_n))q(\lambda)$. It then obviously holds that

$$\inf_{\hat{\lambda}_n} \sup_{d, \lambda, f^*} \mathbb{E}_{d, \lambda, f^*} [(\hat{\lambda}_n - \lambda)^2] \geq \inf_{\hat{\lambda}_n} \sup_{\theta \in \Theta_n} \mathbb{E}_\theta [(\hat{\lambda}_n - \lambda)^2] \geq \inf_{\hat{\lambda}_n} \int_{\Theta_n} \mathbb{E}_\theta [(\hat{\lambda}_n - \lambda)^2] q_n(\theta) d\theta.$$

For $\theta = (\lambda, d)$, denote $I_n^{(1)}(\theta)$ the Fisher information for the parameter λ when d is known. We can apply the so-called Bayesian information bound (cf. Theorem 1 in Gill and Levit 1995).

$$\inf_{\hat{\lambda}_n} \int_{\Theta_n} \mathbb{E}_\theta [(\hat{\lambda}_n - \lambda)^2] q_n(\theta) d\theta \geq \left(\int_{\Theta_n} I_n^{(1)}(\theta) q_n(\theta) d\theta + \tilde{I}(q) \right)^{-1}, \quad (4.18)$$

where $\tilde{I}(q) = \int_0^\pi \left(\frac{\partial}{\partial \lambda} q(\theta) \right)^2 \frac{1}{q(\theta)} d\theta$. We now need an asymptotic expansion of $\int_{\Theta_n} I_n^{(1)}(\theta) q_n(\theta) d\theta$. Let Σ_θ denote the covariance matrix of the Gaussian vector (X_1, \dots, X_n) . The Fisher information for λ is given by

$$I_n^{(1)}(\theta) = \frac{1}{2} \text{tr} \left\{ \left(\Sigma_\theta^{-1} \frac{\partial}{\partial \lambda} \Sigma_\theta \right)^2 \right\}$$

Let J_n denotes the $n \times n$ identity matrix. For a given function h , let $T_n(h)$ denote the Toeplitz matrix of order n defined by $T_n(h)_{j,k} = \hat{h}(j-k) = \int_{-\pi}^\pi h(x) e^{i(j-k)x} dx$. Define $h_\theta = f_\theta - 1$. With these notations, we get $\Sigma_\theta = 2\pi J_n + T_n(h_\theta)$. Let ρ_n denote the spectral radius of the matrix $T_n(h_\theta)$. As appears from the proof of Lemma 2.1 of Iouditsky *et al.* (2001), in the case $\lambda = 0$, under the assumptions of Theorem 3, and with the choice of p_n made here, $\rho_n = o(1)$. The proof of this result is still valid in the present context, since it only uses the bound $|\alpha_j(\lambda)| \leq 2/j$. Thus $\rho_n = o(1)$ uniformly with respect to λ , which implies that

$$I_n^{(1)}(\theta) = \frac{1}{8\pi^2} \text{tr} \left\{ \left(\frac{\partial}{\partial \lambda} \Sigma_\theta \right)^2 \right\} (1 + o(1)).$$

Define now $g_\theta = \log(f_\theta) = d \sum_{j=p_n}^\infty \alpha_j(\lambda) \cos(jx)$ and $k_\theta = f_\theta - 1 - g_\theta$. With these notations, it holds that $\Sigma_\theta = 2\pi J_n + T_n(g_\theta) + T_n(k_\theta)$ and

$$\frac{\partial}{\partial \lambda} \Sigma_\theta = \frac{\partial}{\partial \lambda} T_n(g_\theta) + \frac{\partial}{\partial \lambda} T_n(k_\theta) =: A_\theta + B_\theta.$$

It is easily seen that $A_\theta(j, k) = -2d\pi \sin(|j-k|\lambda) \mathbf{1}_{\{|j-k| \geq p_n\}}$, hence

$$\text{tr}(A_\theta^2) = 4d^2 \pi^2 \sum_{j=p_n}^{n-1} (n-j) \sin^2(j\lambda).$$

Integrating with respect to q_n yields,

$$\int_{\Theta_n} \text{tr}(A_\theta^2) q_n(\theta) d\theta = c(q) n^2 d_n^2 (1 + o(1)),$$

for some positive constant $c(q)$ depending only on q . Again, the proof of Lemma 2.1 of Iouditsky *et al.* (2001) can be adapted to prove that $\text{tr}(B_\theta^2) = o(\text{tr}(A_\theta^2))$. This finally yields

$$\int_{\Theta_n} I_n^{(1)}(\theta) q_n(\theta) d\theta = c(q) n^2 d_n^2 (1 + o(1)).$$

Putting this bound into (4.18), we conclude that

$$\liminf_{n \rightarrow \infty} \sup_{d, \lambda, f^*} n^2 d_n^2 \mathbb{E}_{d, \lambda, f^*} [(\hat{\lambda}_n - \lambda)^2] \geq \liminf_{n \rightarrow \infty} (c(q) + o(1))^{-1} > 0.$$

□

References

- [1] Adenstedt, R. K. (1974). On large-sample estimation for the mean of a stationary random sequence. *Annals of Statistics* **2**, 1095-1107.
- [2] An, H.-Z. Chen, Z.-G. and Hannan, E.J. (1983). The maximum of the periodogram. *Journal of Multivariate Analysis*, **13**, 383-400.
- [3] Andel, J. (1986). Long-memory time series models. *Kybernetika*, **22**, 105-123.
- [4] Arcones, M.A. (1994). Limit theorems for nonlinear functionals of a stationary Gaussian sequence of vectors. *Annals of Probability*, **15**, 2243-2274.
- [5] Arteche, J. and Robinson, P.M. (2000) Semiparametric inference in seasonal and cyclical long memory processes. *Journal of Time Series Analysis*, **21**, 1-25.
- [6] Billingsley P. (1968). *Convergence of probability measures*. New York, Wiley.
- [7] Chen, Z-G. (1988). Consistent estimate for hidden frequencies in a linear process. *Advances in Applied Probability*, **20**, 295-314.
- [8] Dahlhaus, R. (1989). Efficient parameter estimation for self-similar processes. *Annals of Statistics*, **17**, 1749-1766.
- [9] Davies, R. B. and Harte, D. S. (1987). Tests for Hurst effect. *Biometrika*, **74**, 95-101.
- [10] Davis, R. A. and Mikosch, T. (1999). The maximum of the periodogram of a non-Gaussian sequence. *Annals of Probability*, **27**, 522-536.
- [11] Fox, R. and Taqqu, M. S. (1986). Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *Annals of Statistics*, **14**, 517-532.
- [12] Geweke, J. and Porter-Hudak, S. (1983). The estimation and application of long-memory time series models. *Journal of Time Series Analysis*, **4**, 87-104.
- [13] Gill R. and Levit B. (1995). Applications of the van Trees inequality: A bayesian cramer-rao bound, *Bernoulli*, **1-2**, 59-79.
- [14] Giraitis, L., Hidalgo, J. and Robinson, P.M. (2001). Gaussian estimation of parametric spectral density with unknown pole. *Annals of Statistics*, **29**, 987-1023.
- [15] Giraitis L., Robinson P. M. and Samarov A. (1997). Rate optimal semiparametric estimation of the memory parameter of the Gaussian time series with long range dependence, *Journal of Time Series Analysis* **18**, 49-61.
- [16] Giraitis, L. and Surgailis, D. (1990). A central limit theorem for quadratic forms in strongly dependent linear variables and its application to asymptotic normality of Whittle's estimate. *Probability Theory and Related Fields*, **86**, 87-104.
- [17] Granger C. W. J. and Joyeux R. (1980). An introduction to long memory time series and fractional differencing. *J. of Time Series Analysis*, **1**, 15-30.
- [18] Gray, H.L., Zhang, N-F. and Woodward, W.A. (1989). On generalized fractional processes. *Journal of Time Series Analysis*, **10**, 233-257.

- [19] Gray, H.L., Zhang, N-F. and Woodward, W.A. (1994). On generalized fractional processes - a correction. *Journal of Time Series Analysis*, **15**, 561-562.
- [20] Grenander, U. and Rosenblatt, M. (1957). Statistical Analysis of stationary time series. John Wiley, New York.
- [21] Hannan, E.J. (1971). Non-linear time series regression. *Journal of Applied Probability*, **8**, 767-780
- [22] Hannan, E.J. (1973). The estimation of frequency. *Journal of Applied Probability*, **10**, 510-519.
- [23] Hidalgo, J. (2001). Semiparametric estimation when the location of the pole is unknown. Preprint.
- [24] Hosking J. R. M. (1981). Fractional differencing. *Biometrika*, **60**, 165-176.
- [25] Hosoya Y. (1997). A limit theory with long-range dependence and statistical inference on related models. *Annals of Statistics*, **25**, 105-137.
- [26] Hurvich C. M., Deo R. and Brodsky J. (1998). The mean squared error of Geweke and Porter-Hudak's estimator of a long-memory time-series. *Journal of Time Series Analysis* **19** (1998), 19-46.
- [27] Hurvich C. M., Moulines E. and Soulier Ph. (2002). The FEXP estimator for non Gaussian, potentially non stationary processes. *Stochastic Processes and their Applications*, **97**, 307-340.
- [28] Ibragimov I. A. and Linnik. Yu. V. (1971). Independent and stationary sequences of random variables. Wolters, Groningen, 1971.
- [29] Iouditsky, A., Moulines, E. and Soulier, Ph. (2001). Adaptive estimation of the fractional differencing coefficient. *Bernoulli*, **7**, 699-731.
- [30] Moulines, E. and Soulier, Ph. (1999). Log-periodogram regression of time series with long range dependence. *Annals of Statistics*, **27**, 1415-1439.
- [31] Robinson, P.M. (1995). Log-periodogram regression of time series with long range dependence. *Annals of Statistics*, **23**, 1048-1072.
- [32] Soulier, Ph. (2001). Moment bounds and a central limit theorem for functions of Gaussian vectors. *Statistics and Probability Letters*, **54**, 193-203.
- [33] Velasco, C. (2000). Non-Gaussian log-periodogram regression. *Econometric Theory* **16**, 44-79.
- [34] Whittle, P. (1952). The simultaneous estimation of a time series harmonic components and covariance structure. *Trab. Estadist.*, **3**, 43-57.
- [35] Yajima, Y. (1996). Estimation of the frequency of unbounded spectral densities. Proceedings of the Business and Economic Statistical Section. American Statistical Association.

		<i>g</i>								
		0.1		0.2		0.3		0.4		
$\omega_0 = 0$	n	256	29.55	(35.47)	8.73	(17.17)	3.15	(4.25)	1.81	(2.74)
		512	45.25	(60.78)	8.49	(17.22)	3.37	(4.95)	1.98	(2.38)
		1024	44.67	(79.78)	7.07	(10.57)	3.33	(3.86)	2.16	(2.18)
$\omega_0 = \frac{\pi}{2}$	n	256	0.671	(23.87)	0.032	(9.08)	0.015	(3.01)	0.005	(0.81)
		512	-0.189	(38.22)	-0.081	(7.65)	0.040	(2.16)	-0.004	(0.68)
		1024	-0.416	(43.13)	0.020	(5.61)	0.004	(1.68)	0.007	(0.44)

Table 1. Bias and Standard Deviation (in parentheses) of $n\hat{\omega}_n/(2\pi)$.

				g				
				0.1	0.2	0.3	0.4	
$\omega_0 = 0$	n	256	\hat{g}_n	m=64	-0.049	-0.038	-0.026	-0.024
				m=32	-0.054	-0.026	-0.024	-0.024
				m=16	-0.062	-0.074	-0.075	-0.081
		\tilde{g}_n	m=64	-0.002	-0.003	-0.004	-0.008	
			m=32	0.002	0.003	0.002	-0.009	
			m=16	-0.003	-0.005	-0.013	-0.035	
	512	\hat{g}_n	m=128	-0.047	-0.029	-0.024	-0.023	
			m=64	-0.056	-0.041	-0.033	-0.030	
			m=32	-0.067	-0.066	-0.062	-0.061	
	\tilde{g}_n	m=128	-0.003	-0.005	-0.006	-0.007		
		m=64	0.000	0.000	0.001	-0.001		
		m=32	-0.002	-0.001	-0.002	-0.012		
1024	\hat{g}_n	m=256	-0.030	-0.019	-0.019	-0.021		
		m=128	-0.041	-0.026	-0.039	-0.037		
		m=64	-0.055	-0.046	-0.044	-0.044		
	\tilde{g}_n	m=256	-0.002	-0.004	-0.005	-0.006		
		m=128	0.000	0.001	0.002	0.002		
		m=64	0.001	0.002	0.004	0.001		
$\omega_0 = \frac{\pi}{2}$	n	256	\hat{g}_n	m=32	-0.058	-0.041	-0.021	-0.013
				m=16	-0.070	-0.058	-0.033	-0.024
				m=8	-0.077	-0.083	-0.067	-0.061
		\tilde{g}_n	m=32	-0.001	-0.003	-0.003	-0.008	
			m=16	-0.001	0.000	-0.002	-0.015	
			m=8	-0.004	-0.009	-0.021	-0.045	
	512	\hat{g}_n	m=64	-0.049	-0.022	-0.011	-0.005	
			m=32	-0.060	-0.033	-0.013	-0.004	
			m=16	-0.073	-0.052	-0.027	-0.018	
	\tilde{g}_n	m=64	-0.001	-0.002	-0.003	-0.003		
		m=32	0.001	0.002	0.004	0.000		
		m=16	0.001	0.003	0.002	-0.011		
1024	\hat{g}_n	m=128	-0.029	-0.013	-0.009	-0.006		
		m=64	-0.041	-0.017	-0.005	-0.003		
		m=32	-0.057	-0.029	-0.008	-0.001		
\tilde{g}_n	m=128	-0.002	-0.003	-0.005	-0.005			
	m=64	0.001	0.000	0.003	0.004			
	m=32	0.002	0.003	0.006	0.003			

Table 1: Bias of the long-memory parameter estimators

			g					
				0.1	0.2	0.3	0.4	
$\omega_0 = 0$	n	256	\hat{g}_n	m=64	0.091	0.102	0.093	0.084
				m=32	0.121	0.141	0.135	0.119
				m=16	0.170	0.190	0.186	0.169
		\tilde{g}_n	m=64	0.087	0.087	0.087	0.080	
			m=32	0.132	0.131	0.126	0.109	
			m=16	0.202	0.195	0.179	0.154	
	n	512	\hat{g}_n	m=128	0.068	0.068	0.060	0.058
				m=64	0.086	0.096	0.089	0.082
				m=32	0.114	0.133	0.130	0.118
		\tilde{g}_n	m=128	0.058	0.058	0.058	0.057	
			m=64	0.086	0.086	0.086	0.078	
			m=32	0.130	0.130	0.124	0.108	
n	1024	\hat{g}_n	m=256	0.051	0.042	0.040	0.040	
			m=128	0.066	0.062	0.058	0.056	
			m=64	0.085	0.090	0.085	0.081	
	\tilde{g}_n	m=256	0.040	0.040	0.040	0.040		
		m=128	0.058	0.058	0.058	0.056		
		m=64	0.086	0.086	0.086	0.077		
$\omega_0 = \frac{\pi}{2}$	n	256	\hat{g}_n	m=32	0.107	0.110	0.102	0.089
				m=16	0.152	0.162	0.151	0.124
				m=8	0.233	0.236	0.223	0.188
		\tilde{g}_n	m=32	0.095	0.095	0.095	0.086	
			m=16	0.148	0.146	0.138	0.118	
			m=8	0.231	0.224	0.206	0.177	
	n	512	\hat{g}_n	m=64	0.078	0.070	0.064	0.060
				m=32	0.106	0.105	0.097	0.083
				m=16	0.153	0.159	0.147	0.118
		\tilde{g}_n	m=64	0.062	0.062	0.062	0.060	
			m=32	0.093	0.093	0.092	0.082	
			m=16	0.144	0.143	0.135	0.114	
n	1024	\hat{g}_n	m=128	0.054	0.044	0.042	0.041	
			m=64	0.074	0.066	0.062	0.059	
			m=32	0.102	0.100	0.093	0.081	
	\tilde{g}_n	m=128	0.041	0.041	0.041	0.041		
		m=64	0.060	0.060	0.061	0.058		
		m=32	0.090	0.090	0.090	0.081		

Table 2: Standard Deviation of the long-memory parameter estimators

			g					
				0.1	0.2	0.3	0.4	
$\omega_0 = 0$	n	256	\hat{g}_n	m=64	0.011	0.012	0.009	0.008
				m=32	0.018	0.022	0.020	0.016
				m=16	0.033	0.042	0.040	0.035
		\tilde{g}_n	m=64	0.008	0.008	0.008	0.007	
			m=32	0.018	0.017	0.016	0.012	
			m=16	0.041	0.038	0.032	0.025	
	512	\hat{g}_n	m=128	0.007	0.005	0.004	0.004	
			m=64	0.010	0.011	0.009	0.008	
			m=32	0.017	0.022	0.021	0.018	
	\tilde{g}_n	m=128	0.003	0.003	0.003	0.003		
		m=64	0.007	0.007	0.007	0.006		
		m=32	0.017	0.017	0.015	0.012		
1024	\hat{g}_n	m=256	0.003	0.002	0.002	0.002		
		m=128	0.006	0.005	0.004	0.004		
		m=64	0.010	0.010	0.009	0.008		
	\tilde{g}_n	m=256	0.002	0.002	0.002	0.002		
		m=128	0.003	0.003	0.003	0.003		
		m=64	0.007	0.007	0.007	0.006		
$\omega_0 = \frac{\pi}{2}$	n	256	\hat{g}_n	m=32	0.015	0.014	0.011	0.008
				m=16	0.028	0.030	0.024	0.016
				m=8	0.060	0.063	0.054	0.039
		\tilde{g}_n	m=32	0.009	0.009	0.009	0.007	
			m=16	0.022	0.021	0.019	0.014	
			m=8	0.054	0.050	0.043	0.034	
	512	\hat{g}_n	m=64	0.008	0.005	0.004	0.004	
			m=32	0.015	0.012	0.010	0.007	
			m=16	0.029	0.028	0.022	0.014	
	\tilde{g}_n	m=64	0.004	0.004	0.004	0.004		
		m=32	0.009	0.009	0.009	0.007		
		m=16	0.021	0.020	0.018	0.013		
1024	\hat{g}_n	m=128	0.004	0.002	0.002	0.002		
		m=64	0.007	0.005	0.004	0.003		
		m=32	0.014	0.011	0.009	0.007		
\tilde{g}_n	m=128	0.002	0.002	0.002	0.002			
	m=64	0.004	0.004	0.004	0.003			
	m=32	0.008	0.008	0.008	0.006			

Table 3: Mean Square Error of the long-memory parameter estimators