TRIMMING AND TAPERING SEMI PARAMETRIC ESTIMATES IN ASYMMETRIC LONG MEMORY TIME SERIES

By J. Arteche and C. Velasco

University of the Basque Country and Universidad Carlos III de Madrid

First Version received November 2003

Abstract. This paper considers semi-parametric frequency domain inference for seasonal or cyclical time series with asymmetric long memory properties. It is shown that tapering the data reduces the bias caused by the asymmetry of the spectral density at the cyclical frequency. We provide a joint treatment of different tapering schemes and of the log-periodogram regression and Gaussian semi-parametric estimates of the memory parameters. Tapering allows for a less restrictive trimming of frequencies for the analysis of the asymptotic properties of both estimates when allowing for asymmetries. Simple rules for inference are feasible thanks to tapering and their validity in finite samples is investigated in a simulation exercise and for an empirical example.

Keywords. Seasonality; cycles; periodogram; long range dependence; asymptotic normality.

1. INTRODUCTION

A time series x_t, $t = 0, \pm 1, \ldots,$ with spectral density function $f(\lambda)$ and lag j autocovariance γ_j such that

$$\gamma_j = \int_{-\pi}^{\pi} \cos(j\lambda) f(\lambda) \mathrm{d}\lambda \quad j = 0, \pm 1, \ \pm 2, \ldots,$$

has standard long memory or long range dependence if

$$f(\lambda) \sim C|\lambda|^{-2d} \quad \text{as } \lambda \to 0, \tag{1}$$

where $0 < C < \infty$ and $1/2 < d < 1/2$ is a real parameter known as memory parameter. Similarly, if x_t has long memory then

$$\gamma_j \sim K j^{2d-1} \quad \text{as } j \to \infty, \tag{2}$$

where K is a finite constant. Although standard long memory is the most popular case, there exist processes which similarly show strong persistence at some frequency $\omega \in (0, \pi]$ such that the spectral density satisfies

$$f(\omega + \lambda) \sim C|\lambda|^{-2d} \quad \text{as } \lambda \to 0. \tag{3}$$

A time series x_t with such a spectral density displays cycles of period $2\pi/\omega$, more persistent the larger d is. The condition $d < 1/2$ entails stationarity and $d > \quad 1/2$

is usually required for invertibility. If $d > 0$, $f(\lambda)$ diverges at ω and, more precisely, we say that x_t has persistence or positive memory; if $d < 0$, $f(\lambda)$ has a zero at ω and we say that x_t has antipersistence or negative memory and if d 0, x_t has short memory. Arteche and Robinson (1999) called such property Seasonal or Cyclical Long Memory (SCLM) and review some issues related to SCLM processes such as parametric models, estimation and statistical inference. More recently Ould Haye and Viano (2003) discuss the effects of seasonal long memory on some limit theorems. Oppenheim *et al.* (2000) and Lildholdt (2002) show that the seasonal long memory that has been empirically found in many macroeconomic time series can be explained by cross sectional aggregation and structural changes, providing ways of generating parametric seasonal long memory models.

The autocovariances of SCLM processes show an asymptotic slow decay typical of long memory but with oscillations that depend on the frequency ω, where the spectral pole or zero occurs such that

$$\gamma_j \sim K \cos(j\omega) j^{2d-1} \quad \text{as } j \to \infty \tag{4}$$

(see e.g. Chung, 1996 or Gray *et al.*, 1989).

The extension of the concept of long memory from ω 0 to any ω between 0 and π broadens the scope for modelling because (3) can be extended to

$$\begin{aligned} f(\omega + \lambda) &\sim C_1 \lambda^{-2d_1} \quad \text{as} \quad \lambda \to 0^+ \\ f(\omega \quad \lambda) &\sim C_2 \lambda^{-2d_2} \quad \text{as} \quad \lambda \to 0^+, \end{aligned} \tag{5}$$

where $\omega \in (0, \pi)$,

$$0 < C_i < \infty, \qquad |d_i| < \frac{1}{2}, \qquad i \quad 1, 2 \tag{6}$$

and we permit

$$d_1 \ / \ d_2 \quad \text{and/or} \quad C_1 \ / \ C_2. \tag{7}$$

Since the spectrum is symmetric about zero and π, the possibility (7) is excluded for ω $0, \pi$, but for $\omega \in (0, \pi)$ any values of C_i and d_i satisfying (6) are possible. Clearly (5) nests (3) as a special case. This class of processes has been introduced by Arteche and Robinson (2000) and called Seasonal or Cyclical Asymmetric Long Memory (SCALM).

The spectral asymmetry involves a different persistence of the cycles of period just shorter and just larger than $2\pi/\omega$. To shed some light in this concept consider quarterly data with ω $\pi/2$ and $d_1 > d_2$. This implies that the cycles of period slightly shorter than four are more persistent than the corresponding cycles of period just larger, so that in the long run the cycles tend to be slightly shorter than four periods. In this case the ACF displays asymmetric cycles so correlation before multiple lags of four decays slower than correlation for lags after four such that summer observations are more related with long horizon future spring observations than with the corresponding future fall (see Arteche and Robinson, 2000). The opposite situation arises when $d_2 > d_1$. The asymmetric long memory entails this asymmetric asymptotic behaviour in the (cyclical) persistence of the series.

Nowadays there exist several well established estimators of the memory parameters. Time domain parametric techniques, such as exact maximum likelihood, require an entire knowledge of the autocovariances and have a difficult implementation for SCLM processes due to the complicated functional form of the autocovariances. Moreover, explicit form of these autocovariances may not be available in some cases. Semiparametric techniques that only consider asymptotic behaviour of the autocovariances are only valid (if they are at all) for estimation of the largest memory parameter if there exist more than one. Thus frequency domain techniques, more closely related with our definition (3) of SCLM or (5) of SCALM, seem more appropriate and permit the estimation of different memory parameters corresponding to different spectral poles. Since we only aim to estimate d, we avoid fully parametric methods that, although more efficient under a complete and correct specification of the model, suffer inconsistency under misspecification of $f(\lambda)$, even if f is only misspecified at frequencies far from ω.

This paper focuses on two widely extended semiparametric techniques: the variant of Robinson (1995a) of the log periodogram regression first introduced by Geweke and Porter Hudak (1983) and the more efficient Gaussian semiparametric estimator of Robinson (1995b). Both have been considered under SCALM by Arteche and Robinson (2000). When d_1 / d_2, trimming of frequencies close to ω seems necessary to avoid the distorting influence of the periodogram at the other side of the spectral pole under investigation, the larger the difference between d_1 and d_2, the stronger the trimming needed. Recent work by Velasco (1999a, 1999b) regarding semiparametric estimation for nonstationary long memory series suggests that this trimming can be reduced by tapering the data and using the tapered periodogram instead of the raw periodogram.

The properties of both estimators depend strongly on the normalized periodogram $I_{ij}/C_i\lambda_j^{-2d_i}$, where $I_{1j} \quad I(\omega + \lambda_j)$, $I_{2j} \quad I(\omega \quad \lambda_j)$ are periodogram ordinates $I(\lambda) \quad |W(\lambda)|^2$, where $W(\lambda)$ is the Discrete Fourier Transform (DFT),

$$W(\lambda) : \quad (2\pi n)^{-1} \sum_{t \quad 1}^{n} x_t \mathrm{e}^{\mathrm{i}t\lambda}$$

and $\lambda_j \quad 2\pi j/n$ are the Fourier frequencies with n the sample size. Its behaviour is discussed in Section 2, considering j both fixed and tending to ∞. It is shown that tapering the data may help to reduce the bias caused by the asymmetry of the spectral density. Section 3 focuses on the tapered log periodogram regression and Section 4 pays attention to the Gaussian semiparametric estimator. Section 5 shows the finite sample behaviour of both tapered and untapered estimates and Section 6 applies both to a growth rate series of the monthly US industrial production. Technical details are placed in the Appendix.

2. BIAS OF THE PERIODOGRAM

The periodogram is the basic tool to estimate the spectral density function. Hannan (1973) showed that the periodograms evaluated at Fourier frequencies close to a fixed frequency λ are asymptotically independent and identically distributed as $(f(\lambda)/2)\chi_2^2$, where χ_2^2 is the chi square distribution with two degrees of freedom. However his assumptions rule out the possibility of long range dependence. Yajima (1989) allowed for the possibility of long memory and gave the joint asymptotic distribution of the periodogram evaluated at a set of fixed frequencies not depending on n, so that Fourier frequencies are not considered. These results have led some authors (e.g. Geweke and Porter Hudak, 1983, based their proof on Hannan's theorem) to conclude that the log periodogram estimator proposed by Geweke and Porter Hudak is asymptotically normal with variance $\pi^2/6$. However, the log periodogram and the Gaussian semiparametric estimators, are based on Fourier frequencies $\lambda_j \quad 2\pi j/n$. Consequently these frequencies do change with n, so that Yajima's result can not be applied. On the other hand, for $d < 0$ Hannan (1973) stated that the periodogram evaluated at a finite number of Fourier frequencies close to the origin converges in probability to zero. However, when we normalize with the spectral density the remainder is divided by a quantity which approaches zero, and therefore need not be negligible. These facts have been noted in Hurvich and Beltrao (1993) and Robinson (1995a), who considered the asymptotic distribution of the periodogram normalized by the spectral density of weakly stationary long memory time series at Fourier frequencies, $\lambda_j \quad 2\pi j/n$, where j is fixed and $n \to \infty$. They proved that in this context the normalized periodograms are not asymptotically identically distributed. In fact $\lim_{n\to\infty} E[I(\lambda_j)/f(\lambda_j)]$ depends on j and d and is typically greater than 1, implying positive asymptotic relative bias in the periodogram as estimate of $f(\lambda)$. In this section we extend these results to stationary SCALM processes with spectral density as in (5). We focus first on the relative bias of the periodogram at frequencies $\omega + \lambda_j$ with fixed j.

THEOREM 1. *Let x_t be a real valued stationary process with spectral density* (5). *Let j be fixed and denote*

$$L_j(d_1, d_2) : \quad E\left[\frac{I_{1j}}{C_1 \lambda_j^{-2d_1}}\right].$$

Then:

(a) *If $d_2 < d_1$*

$$\lim_{n\to\infty} L_j(d_1, d_2) \quad |2\pi j|^{2d_1} \int_0^\infty \psi_j(\lambda; d_1)\mathrm{d}\lambda,$$

(b) *if $d_2 \quad d_1 \quad d$*

$$\lim_{n\to\infty} L_j(d_1, d_2) \quad |2\pi j|^{2d} \int_0^\infty \psi_j(\lambda; d)\mathrm{d}\lambda + \frac{C_2}{C_1}|2\pi j|^{2d} \int_{-\infty}^0 \psi_j(\lambda; d)\mathrm{d}\lambda$$

(c) *and if* $d_1 < d_2$

$$\lim_{n\to\infty} n^{2(d_1-d_2)} L_j(d_1, d_2) = \frac{C_2}{C_1}|2\pi j|^{2d_1}\int_{-\infty}^{0}\psi_j(\lambda; d_2)\mathrm{d}\lambda,$$

where

$$\psi_j(\lambda; d) := \frac{2}{\pi}\frac{\sin^2\frac{\lambda}{2}}{(2\pi j - \lambda)^2}|\lambda|^{-2d}.$$

Theorem 1 focuses on the behaviour of the normalized periodogram at Fourier frequencies just after ω. A similar result is obtained for frequencies just before the spectral pole/zero. In particular, the asymptotic relative bias evaluated at those frequencies diverges as $n \to \infty$, when $d_1 > d_2$.

When $d_1 = d_2$ and $C_1 = C_2$ our Theorem 1 corresponds to that in Hurvich and Beltrao (1993). However, when $d_2 < d_1$, the asymptotic relative bias, although depending on d_1 and j, reduces with respect to that obtained by Hurvich and Beltrao at zero frequency. Finally, when $d_1 < d_2$ the asymptotic relative bias of the periodogram increases without limit as $n \to \infty$. Hurvich and Beltrao (1993) and Hurvich and Ray (1995) suggest tapering the data to reduce the bias of the periodogram as estimate of the spectral density at Fourier frequencies close to a spectral pole/zero. Based on this bias reduction Arteche and Robinson (2000) pointed out the possibility of avoiding the trimming in the log periodogram and Gaussian semiparametric estimation under SCALM by using a tapered periodogram

$$I^{\mathrm{T}}(\lambda) = |W^{\mathrm{T}}(\lambda)|^2 := \left|\left(2\pi\sum_{t=1}^{n}\left|h_t^{\mathrm{T}}\right|^2\right)^{-1/2}\sum_{t=1}^{n}h_t^{\mathrm{T}}x_t\mathrm{e}^{\mathrm{i}t\lambda}\right|^2, \tag{8}$$

where $\{h_t^{\mathrm{T}}\}_{t=1}^{n}$ is a sequence of constants (the taper) such that $\sum\left|h_t^{\mathrm{T}}\right|^2 = bn$ and $0 < b < \infty$, $I^{\mathrm{T}}(\lambda)$ and $W^{\mathrm{T}}(\lambda)$ are the tapered periodogram and DFT respectively at frequency λ and denote $I_{1j}^{\mathrm{T}} = I^{\mathrm{T}}(\omega + \lambda_j)$ and $I_{2j}^{\mathrm{T}} = I^{\mathrm{T}}(\omega - \lambda_j)$.

There are several alternative tapering schemes with desirable properties to control leakage from remote frequencies. Following Velasco (1999a, 1999b) we may consider a general class of tapers of type I and orders $p = 1, 2, \ldots$ denoted as $\{h_t^{(\mathrm{I},p)}\}$, whose DFT satisfies

$$D^{(\mathrm{I},p)}(\lambda) := \sum_{t=1}^{n}h_t^{(\mathrm{I},p)}\mathrm{e}^{\mathrm{i}t\lambda} = \frac{a(\lambda)}{n^{p-1}}\left(\frac{\sin[n\lambda/2p]}{\sin[\lambda/2]}\right)^p, \tag{9}$$

where $a(\lambda)$ is a complex function whose modulus is positive and bounded. Some examples of tapers, which satisfy (9) are the triangular Barlett window ($p = 2$), Parzen window ($p = 4$) or the Zhurbenko (1979) taper for integer p.

This class of tapers provides interesting insight in the behaviour of the periodogram of time series with spectral densities displaying peaks or troughs but have the undesirable property of introducing some extra dependence among adjacent periodogram ordinates. This implies that the design of many frequency domain memory estimates requires to skip some periodogram ordinates or either adapt inference to that dependence. The use of a restricted set of Fourier frequencies generally leads to a loss of efficiency (Velasco, 1999a). To reduce the size of such set of omitted frequencies, Hurvich *et al.* (2002) and Hurvich and Chen (2000) propose alternative complex data tapers. Setting

$$h_t^{(2,s)} = h_{t,n}^{(2,s)} := \left(1 - \mathrm{e}^{2\mathrm{i}\pi t/n}\right)^{s-1}, \tag{10}$$

the tapered periodogram and DFT with a type II taper $\{h_t^{(2,s)}\}$ of order $s = 1,2,\ldots$, is obtained by

$$I^{(2,s)}(\lambda) = |W^{(2,s)}(\lambda)|^2 := \left| \left(2\pi \sum_{t=1}^{n} |h_t^{(2,s)}|^2 \right)^{-1/2} \sum_{t=1}^{n} h_t^{(2,s)} x_t \mathrm{e}^{\mathrm{i}t\lambda} \right|^2.$$

It can be shown that $\sum_{t=1}^{n} |h_t^{(2,s)}|^2 = n a_s$, where $a_s := \binom{2(s-1)}{s-1}$. Here the order s is equivalent to $s-1$ as set by Hurvich *et al.* (2002) but equivalent to the order p of Velasco (1999a, 1999b) or Hurvich and Chen (2000), so both tapers of orders $p = s = 1$ are equivalent to the usual DFT and periodogram. For higher orders they share similar properties since both classes satisfy the inequality, $p = s$,

$$|D^{(v,p)}(\lambda)| \leq C \frac{n}{(1+n|\lambda|)^p} \leq C \min\{n, n^{1-p}|\lambda|^{-p}\}, \tag{11}$$

$v = 1, 2$. The main difference is that tapers of type I are not exactly ortogonal, that is, for any Fourier frequencies λ_j, λ_k and $p > 1$

$$A_{jk}^{(v,p)} := \int_{-\pi}^{\pi} E_{jk}^{(v,s)}(\lambda) \mathrm{d}\lambda$$

with

$$E_{jk}^{(v,p)}(\lambda) := \left(2\pi \sum_{t=1}^{n} \left| h_t^{(v,p)} \right|^2 \right)^{-1} D^{(v,p)}(\omega + \lambda_j - \lambda) \overline{D^{(v,p)}}(\omega + \lambda_k - \lambda), \quad v = 1, 2,$$

for $\overline{D^{(v,p)}}$ the complex conjugate of $D^{(v,p)}$, and $\lim_{n\to\infty} A_{jk}^{(1,p)}$ is not null, whereas

$$A_{jk}^{(2,s)} = \int_{-\pi}^{\pi} E_{jk}^{(2,s)}(\lambda) \mathrm{d}\lambda = 0$$

for all positive integer s and $|j - k| \geq s \bmod n$. When $|j - k| < s$ we obtain (cf. eqn 7 in Hurvich *et al.*, 2002) that

$$A_{jk}^{(2,s)} = (a_s)^{-1} (-1)^{j-k} \binom{2(s-1)}{s-1+j-k}.$$

For data tapers of type I, we can achieve for $j, k = 1, \ldots, n/2$, $j \neq k$,

$$A_{jk}^{(1,p)} = \int_{-\pi}^{\pi} E_{jk}^{(1,p)}(\lambda) d\lambda = O(|j-k|^{-p}), \tag{12}$$

which shows that if the Fourier frequencies λ_j and λ_k are sufficiently apart, the kernels $D^{(1,p)}(\omega + \lambda_j \quad \lambda)$ and $\overline{D^{(1,p)}}(\omega + \lambda_k \quad \lambda)$ are almost, but not exactly, orthogonal.

Theorem 2 investigates the bias of the tapered periodogram at Fourier frequencies for fixed j.

Theorem 2. *Let x_t have spectral density* (5) *and let $I^{(v,p)}(\lambda)$ be the tapered periodogram* (8) *with a taper satisfying* (9) *for $v \quad 1$, or* (10) *for $v \quad 2$. Let j be fixed and denote*

$$L_j^{(v,p)}(d_1, d_2) : \quad E\left[\frac{I_{1j}^{(v,p)}}{C_1 \lambda_j^{-2d_1}}\right].$$

Then:

(a) *If $d_2 < d_1$*

$$\lim_{n\to\infty} L_j^{(v,p)}(d_1, d_2) \quad |2\pi j|^{2d_1} \int_0^{\infty} \psi_j^{(v,p)}(\lambda; d_1)\mathrm{d}\lambda.$$

(b) *If $d_2 \quad d_1 \quad d$*

$$\lim_{n\to\infty} L_j^{(v,p)}(d_1, d_2) \quad |2\pi j|^{2d} \int_0^{\infty} \psi_j^{(v,p)}(\lambda; d)\mathrm{d}\lambda + \frac{C_2}{C_1}|2\pi j|^{2d} \int_{-\infty}^{0} \psi_j^{(v,p)}(\lambda; d)\mathrm{d}\lambda.$$

(c) *If $d_1 < d_2$*

$$\lim_{n\to\infty} n^{2(d_1 - d_2)} L_j^{(v,p)}(d_1, d_2) \quad \frac{C_2}{C_1}|2\pi j|^{2d_1} \int_{-\infty}^{0} \psi_j^{(v,p)}(\lambda; d_2)\mathrm{d}\lambda,$$

where

$$\psi_j^{(1,p)}(\lambda; d) : \quad \frac{2^{2p-1}|a(0)|^2}{\pi b} \frac{\sin^{2p}\left(\frac{2\pi j - \lambda}{2p}\right)}{(2\pi j \quad \lambda)^{2p}} |\lambda|^{-2d},$$

and

$$\psi_j^{(2,p)}(\lambda; d) : \quad \frac{2}{\pi}|\lambda|^{-2d} \sin^2\frac{\lambda}{2} a_p^{-1} \left(\sum_{k \quad 0}^{p-1} \binom{p \quad 1}{k} \frac{(\quad 1)^k}{(2\pi(j+k) \quad \lambda)}\right)^2.$$

The bias of the periodogram is reduced with an adequate taper but when $d_2 > d_1$ it is still growing with n. Consider for example the cosine bell or Hanning taper as suggested by Hurvich and Ray (1995) with weights $h_t^{\cos} \quad 0.5(1 \quad \cos(2\pi t/n))$. In this case $\sum_{t \quad 1}^{n} (h_t^{\cos})^2 \quad 3n/8$ and

$$D^{\cos}(\lambda) \quad \frac{1}{2}D(\lambda) \quad \frac{1}{4}D\left(\lambda \quad \frac{2\pi}{n}\right) \quad \frac{1}{4}D\left(\lambda + \frac{2\pi}{n}\right),$$

where $D(\lambda) \quad \sum_1^n \exp(\mathrm{i}t\lambda)$ is the Dirichlet Kernel. Note that the cosine bell does not satisfy condition (9) but shares some asymptotic properties with type I tapers with p 1 and 3. The asymptotic relative bias of the periodogram in this case is the one given in Theorem 2 with

$$\psi_j^{\cos}(\lambda; d) \quad \frac{16}{3\pi}\sin^2\frac{\lambda}{2}|\lambda|^{-2d}\left\{\frac{0.5}{2\pi j \quad \lambda} \quad \frac{0.25}{2\pi(j+1) \quad \lambda} \quad \frac{0.25}{2\pi(j \quad 1) \quad \lambda}\right\}^2,$$

which shows that tapers of type II are a complex generalization of the cosine bell (which would be of order, s 3).

It is also interesting to analyse the covariances between discrete Fourier transforms at frequencies $\omega + \lambda_j$ for increasing j. For the raw series see Arteche and Robinson (2000). Consider the following assumptions:

Assumption 1. *The spectral density $f()$ satisfies for $\beta \in (0, 2]$ and at frequency $\omega \in (0, \pi)$*

$$f(\omega + \lambda) \sim C_1\lambda^{-2d_1}(1 + O(\lambda^\beta))$$
$$f(\omega \quad \lambda) \sim C_2\lambda^{-2d_2}(1 + O(\lambda^\beta))$$

as $\lambda \to 0^+$ for $0 < C_i < \infty$, $(1/2) < d_i < (1/2)$, i 1, 2.

Assumption 2. $f(\lambda) \quad \alpha(\lambda)\overline{\alpha}(\lambda)/2\pi$ *and in a neighbourhood* $(\ \varepsilon, 0) \cup (0, \varepsilon)$ *of* ω $\alpha()$ *is differentiable and as* $\lambda \to 0^+$,

$$\left|\frac{d}{d\lambda}\alpha(\omega + \lambda)\right| \quad O(\lambda^{-1-d_1}),$$

$$\left|\frac{d}{d\lambda}\alpha(\omega \quad \lambda)\right| \quad O(\lambda^{-1-d_2}).$$

Then we obtain Theorem 3, which is valid for both types of tapers since they satisfy (11).

Theorem 3. *Let Assumptions 1 and 2 hold and let k $k(n)$ and j $j(n)$ be two sequences of positive integers such that $j > k$, η: j k, $j/n \to 0$ as $n \to \infty$, and consider tapers of orders $p \geq 2$. Then as $n \to \infty$, v 1, 2,*

(a) $EW^{(v,p)}(\omega + \lambda_j)\overline{W^{(v,p)}(\omega + \lambda_j)} \quad f(\omega + \lambda_j) + O(\vartheta_{jj}(d^*) + \lambda_j^{-2d_1}j^{-1})$.
(b) $EW^{(v,p)}(\omega + \lambda_j)W^{(v,p)}(\omega + \lambda_j) \quad O(\vartheta_{jj}(d^*) + \lambda_j^{-2d_1}j^{-p})$.
(c) $EW^{(v,p)}(\omega + \lambda_j)\overline{W^{(v,p)}(\omega + \lambda_k)} \quad f_{jk}A_{jk}^{(v,p)} + O(\vartheta_{jk}(d^*) + \lambda_j^{-d_1}\lambda_k^{-d_1}(jk)^{-1/2}\eta^{1-p})$.
(d) $EW^{(v,p)}(\omega + \lambda_j)W^{(v,p)}(\omega + \lambda_k) \quad O(\vartheta_{jk}(d^*) + \lambda_j^{-d_1}\lambda_k^{-d_1}(jk)^{-p/2})$,

where $\vartheta_{jk}(d): \quad (jk)^{\frac{1}{2}-p}(\lambda_j\lambda_k)^{-d}$, d^*: $\max\{d_1, d_2\}$ *and* $f_{jk}: \quad \alpha_j\alpha_k/2\pi$, α_i: $\alpha(\omega + \lambda_i)$, i j, k.

Note that for tapers of type I $f_{jk}A_{jk}^{(1,p)} \quad O((\lambda_j\lambda_k)^{-d_1}\eta^{-p})$, cf. (12), whereas for type II tapers $A_{jk}^{(2,p)}$ 0 if j $k \geq p$.

The relative bias of the periodogram is crucial for the properties of frequency domain estimators. It affects not only the memory parameters estimates as those considered in Sections 3–7, but also the estimation of an unknown pole. In many applications the location and number of spectral poles can be known in advance, but for many other cases Hidalgo and Soulier (2003) considered the estimation of the location of the pole $\omega \in (0, \pi)$ by means of maximizing the (usual) periodogram at Fourier frequencies. Using the tapered periodogram $I^{(v,p)}$ we set the pole estimate as

$$\hat{\omega}_n^{(v,p)} = \frac{2\pi}{n} \arg \max_{1 \leq k \leq [n/2]} I^{(v,p)}(\lambda_k).$$

Under slightly stronger conditions than those imposed in this paper (for example only one spectral singularity is permitted) Hidalgo and Soulier (2003) show the almost rate n consistency of their estimate of ω, $\omega_n^{(v,1)}$. Under spectral asymmetry the situation is different. Using the results in Theorem 3 and Lemmas 1 and 2 in the Appendix it can be shown that as $n \to \infty$, $p > 1$,

$$\frac{n}{v_n}\left(\hat{\omega}_n^{(v,p)} - \omega\right) \xrightarrow{p} 0$$

for v_n a positive nondecreasing sequence such that

$$\lim_{n \to \infty} \frac{n^{\delta(d^* - d_*)}}{v_n^{\delta(d^* - d_* + p - 0.5) - 1} \log^{\delta} n} + \lim_{n \to \infty} \frac{\log n}{v_n^{2d^*}} = 0, \tag{13}$$

where $d_* := \min\{d_1, d_2\} > 0$ and 2δ is a positive integer such that Assumption 5 holds with $E\varepsilon_t^{2\delta}$ finite for some $\delta \geq 2$. For the symmetric situation in Hidalgo and Soulier (2003) the first condition in (13) is redundant since in that case $d^* = d_* = 0$, $p = 1$ and $\delta = 4$. Under spectral asymmetry, condition (13) imposes a slower rate of convergence of $\hat{\omega}_n^{(v,p)}$ the higher the difference $(d^* - d_*)$. This distorting effect can be reduced using a taper of higher order p.

3. LOG-PERIODOGRAM REGRESSION

The basic log periodogram regression estimates, $\hat{d}_i$ for $i = 1, 2$, in the version suggested by Robinson (1995a), are obtained by least squares in the regression

$$\log I_{ij}^{(v,p)} = a + d(-2 \log \lambda_j) + u_j, \quad j = 1, 2, \ldots, m, \tag{14}$$

where m goes to infinity but at a slower rate than n, in such a way that the band of frequencies used in the estimation degenerates to zero as the sample size increases. Clearly the properties of $\hat{d}_i$ are closely related to those of $I_{ij}^{(v,p)} / C_i \lambda_j^{-2d_i}$. These variables are not asymptotically independent nor identically distributed. To avoid the negative influence of these undesirable characteristics, Künsch (1986) and Robinson (1995a) suggested omitting l frequencies close to the spectral pole/zero, although this trimming seems not necessary (Hurvich *et al.*, 1998), at least in the symmetric case $\omega \quad 0$. However, under SCALM, Theorems 1 and 3 suggest that trimming is unavoidable due to the $n^{2(d^{*}-d_i)}$ term, at least for the estimation of the lower memory parameter. Tapering can reduce the bias of the periodogram at frequencies close to ω, so that a less restrictive trimming is sufficient.

The correlation found among periodogram ordinates can be taken into account in different ways when the log periodogram regression is designed. A first alternative is the use of only asymptotically uncorrelated periodogram ordinates. For type II tapers such approach would imply to neglect $p \quad 1$ frequencies of every p in the log periodogram regression. To alleviate the efficiency loss incurred following such policy, Hurvich *et al.* (2002) used a pooling of periodogram ordinates as proposed by Robinson (1995a), where the periodogram $I_{ij}^{(2,p)}$ is replaced by

$$\tilde{I}_{ij}^{(2,p)} : \quad \sum_{k \quad (M+p-1)(j-1)+1}^{(M+p-1)(j-1)+M} I_{ik}^{(2,p)}$$

for $j \quad 1,\ldots,\tilde{K}$, $\tilde{K} : \quad [m/(M+p \quad 1)]$. Up to $p \quad 1$ frequencies are dropped of every $M+p \quad 1$ and the efficiency is of magnitude $M/(M+p \quad 1)$ compared to a $1/p$ efficiency, when no pooling is employed. However for data tapers of type I ($p > 1$) there is not clear cut, because the correlation among periodograms is never zero. Nevertheless, it dies out very fast in $|j \quad k|$ for both types of tapers (for type II the leading term is zero for $|j \quad k| \geq p$) so we can consider the use of all frequencies in the regression, that is

$$\check{I}_{ij}^{(2,p)} : \quad \sum_{k \quad M(j-1)+1}^{Mj} I_{ik}^{(2,p)}$$

for $j \quad 1,\ldots,K$, $K : \quad [m/M]$, where the correlation among adjacent $\log \check{I}_{ij}^{(2,p)}$ should appear in the asymptotic variance of the log periodogram estimates (see definition of $\Omega_M^{(v,p)}$ in Theorem 4 below). For tapers of type II, if $M \geq p \quad 1$, the correlation is due at most from the previous and next log periodograms, but if p is large compared to M, many more pooled periodograms might correlate, but at most a fixed number of them. For tapers of type I all pooled periodograms display correlation.

Robinson (1995a), for $p \quad 1$ and all M, and Hurvich *et al.* (2002), for $p > 1$ and large M, give explicit expressions for the expectation and variance of the pooled log periodogram $\log \tilde{I}_{ij}^{(2,p)}$, which can be used to estimate the asymptotic variance of the log periodogram memory estimate. Instead of trying a generalization of these results, in order to obtain closed expressions for the limit autocorrelations (for increasing j with n, $j/n \to 0$),

$$\sigma_{M,v,p}^{2}(k) : \quad \lim_{n\to\infty} \operatorname{Cov}\left[\log \check{I}_{ij}^{(v,p)}, \log \check{I}_{ij+|k|}^{(v,p)}\right], \quad k \quad 1,2,\ldots,$$

we use the approach of Robinson (1995a) and propose a consistent estimation of such asymptotic variance based on the residuals of (14), which takes into account the correlation across Fourier frequencies. Note that $Cov[\log \check{I}_{ij}^{(v,p)}, \log \check{I}_{ij+|k|}^{(v,p)}]$ only depends asymptotically on the frequency gap k (and on M and the taper characteristics v, p), but not on the reference frequency λ_j, neither on the scale given by $f(\lambda_j)$ or $C_i\lambda_j^{-2d_i}$ because of the log transformation. Proceeding as in Robinson (1995a, proof of Thms 3 and 4) we could calculate $\sigma_{M,v,p}^2(k)$ supposing that the real and imaginary parts of the components of $\check{I}_{ij}^{(v,p)}$ are zero mean normal variates with covariance structure given by the leading terms of the approxima tions of Theorem 3, (a) and (b), after normalization by f_{jj} and f_{jk} respectively.

Let for $i\quad 1, 2$ and $v\quad 1, 2$, $p > 1$,

$$\hat{d}_{M,i}^{(v,p)} :\quad \left(\sum_{j\quad l+1}^{\bar{K}} r_j^2\right)^{-1} \sum_{j\quad l+1}^{\bar{K}} r_j \log \check{I}_{ij}^{(v,p)},$$

where $r_j :\quad 2\log y_j\quad (K\quad l)^{-1}\sum_{k\quad l+1}^{\bar{K}}(\quad 2\log y_k)$, $y_j :\quad (2j\quad 1)\pi M/n$.

We consider now the asymptotic distribution of $\hat{d}_{M,i}^{(v,p)}$ under the following assumptions, when type II tapers are used:

Assumption 3. *x_t is a Gaussian process.*

Assumption 4. *As $n \to \infty$*

$$\frac{m^{1/2}\log m}{l} + \frac{n^{2(d^*-d_i)}}{l^{2p-1+2(d^*-d_i)}} + \frac{l(\log n)^2}{m} + \frac{m^{2\beta+1}}{n^{2\beta}} \to 0.$$

The second condition term on the left hand side of Assumption 4 establishes a trimming of l frequencies, which is smaller, the larger is p. The other terms appear also in Assumption 6 of Robinson (1995a) for untapered log periodogram inference of zero frequency long memory time series and Assumption 4 in Arteche and Robinson (2000) for SCALM series.

Theorem 4. *Under Assumptions* 1 4, *as $n \to \infty$, $p > 1$,*

$$\sqrt{m}\left(\hat{d}_{M,i}^{(2,p)}\quad d_i\right) \xrightarrow{d} N\left(0, M^{-1}\Omega_M^{(2,p)}\right),$$

where

$$\Omega_M^{(v,p)} :\quad \lim_{n\to\infty} m\left(\sum_{j\quad l+1}^{\bar{K}} r_j^2\right)^{-2} \sum_{j\quad l+1}^{\bar{K}} \sum_{k\quad -[1+(p-1)/M]}^{[1+(p-1)/M]} r_j r_{j+k}\sigma_{M,v,p}^2(k).$$

Note that the second condition on Assumption 4 imposing a lowest rate of increase on l with n appears to control leakage due to possibly different memory parameters on the two sides of the spectral peak and is always satisfied if $n^2 l^{-1-2p} \to 0$ as $n \to \infty$ because $d^*\quad d_i < 1$. Note anyway that the first condition of Assumption 4 imposes the growing rate of l to be at least of order $m^{1/2}\log m$, independently of the taper order p, but in principle m needs to grow with n only slightly faster than $l(\log n)^2$. For $\beta \quad 1$, which is usual in cyclical long memory processes as those discussed in Arteche and Robinson (1999), valid choices of the bandwidth and trimming for any $p > 1$ are $m \sim n^b$ and $l \sim n^a$ such that $2/5 < a < b < 2/3$.

For type I tapers a similar result could be possible at least if the order $p > 1$ is high enough in order to guarantee that $m^{-1} \sum_{j \quad l+1}^{\bar{K}} \sum_{k \quad \ell - j+1}^{\bar{K}-j} r_j r_{j+k} \sigma^2_{M,v,p}(k)$ converge to a positive constant and the fast decaying in $A_{jk}^{(1,p)}$ allow us to employ a truncation argument. Alternatively, an increasing gap among regressors would provide an asymptotically normal estimate, but with an infinite efficiency loss (cf. Velasco, 1999b).

For type II tapers, $v \quad 2$, it is also possible to drop $p \quad 1$ frequencies, every $M + p \quad 1$, so only asymptotically uncorrelated pooled (log)periodograms are employed, and a similar asymptotic result to Theorem 4 holds with a simplified asymptotic variance given by

$$\Omega_M^{(2,p)} : \quad \lim_{n \to \infty} \left(\frac{1}{m} \sum_{j \quad l+1}^{\tilde{K}} r_j^2 \right)^{-1} \sigma^2_{M,2,p}(0).$$

For large M, we obtain

$$\sigma^2_{M,2,p}(0) \approx \frac{1}{M + p \quad 1} \frac{\Gamma(4p \quad 3)\Gamma^4(p)}{\Gamma^4(2p \quad 1)}$$

(cf. Thm 1 of Hurvich *et al.*, 2002), whereas when no taper is applied (so $p \quad 1$, for both $v \quad 1{,}2$) and no dropping neither pooling of frequencies is used, $M \quad 1$,

$$\Omega_1^{(v,1)} \quad \lim_{n \to \infty} \left(\frac{1}{m} \sum_{j \quad l+1}^{m} r_j^2 \right)^{-1} \sigma^2_{1,v,1}(0) \quad \frac{\sigma^2_{1,v,1}(0)}{4} \quad \frac{\pi^2}{24}$$

since $\sigma^2_{M,v,1}(0) \quad \psi'(M)$, where $\psi(z) \quad \Gamma'(z)/\Gamma(z)$ is the digamma function.

We now propose a consistent feasible estimate of $\Omega_M^{(v,p)}$ in the lines suggested by Robinson (1995a) by means of the observed residuals, $j \quad l+1, \ldots, K$,

$$\hat{u}_{M,j}^{(v,p)} : \quad \log \check{I}_{ij}^{(v,p)} \quad \hat{C}_{M,i}^{(v,p)} + 2\hat{d}_{M,i}^{(v,p)} \log y_j,$$

where $\hat{C}_{M,i}^{(v,p)}$ are the ordinary least squares (OLS) estimates of the intercept C_i in the log periodogram regression. Then, setting the sample residual autocovari ances

$$\hat{\sigma}^2_{M,v,p}(k) : \quad \frac{1}{K \quad |k|} \sum_{j \quad l+1}^{\bar{K}-|k|} \hat{u}_{M,j}^{(v,p)} \hat{u}_{M,j+|k|}^{(v,p)},$$

we can estimate the asymptotic variance of the log periodogram regression estimate by means of

$$\hat{\Omega}_M^{(v,p)} : \quad m \left(\sum_{j \quad l+1}^{\bar{K}} r_j^2 \right)^{-2} \sum_{j \quad l+1}^{\bar{K}} \sum_{k \quad -\ell}^{\ell} r_j r_{j+k} \hat{\sigma}^2_{M,v,p}(k),$$

where ℓ is a fixed integer such that $\ell \geq [1 + (p \quad 1)/M]$, when $v \quad 2$, fixed with n. The consistency of such estimate for type II tapers is established in Theorem 5.

THEOREM 5. *Under Assumptions* 1 4, *as* $n \to \infty$,

$$\hat{\Omega}_M^{(2,p)} \to_p \Omega_M^{(2,p)}.$$

Also two sided estimates of $\Omega_M^{(v,p)}$ could be justified. For type I tapers, $v \quad 1$, it could be chosen an increasing lag number ℓ, such that $\ell^{-1} + \ell m^{-1} \to 0$ as $n \to \infty$, being able to take into account asymptotically the correlation among all the tapered periodograms displayed in the asymptotic variance.

4. GAUSSIAN SEMIPARAMETRIC ESTIMATION

In the SCALM case, the Gaussian semiparametric estimates of d_1 and d_2 are

$$\tilde{d}_i : \quad \arg\min_{\Theta} R_i(d), \quad i \quad 1, 2, \quad v \quad 1, 2,$$

where

$$R_i(d) : \quad \log \tilde{C}_i(d) \quad \frac{2d}{m \quad l} \sum_{j \quad l+1}^{m} \log \lambda_j, \quad \tilde{C}_i(d) \quad \frac{1}{m \quad l} \sum_{j \quad l+1}^{m} \lambda_j^{2d} I_{ij},$$

and $\Theta \quad [\Delta_1, \Delta_2]$, where $\quad 0.5 < \Delta_1 < \Delta_2 < 0.5$ and $d_i \in \Theta$. Under asymmetric long memory a strong trimming is needed as pointed out by Arteche and Robinson (2000). Again this trimming can be reduced by means of tapering. The tapered Gaussian semiparametric estimates are

$$\tilde{d}_i^{(v,p)} : \quad \arg\min_{\Theta} R_i^{(v,p)}(d), \quad i \quad 1, 2,$$

where $R_i^{(v,p)}(d)$ is defined in terms of the periodogram $I_{ij}^{(v,p)}$ with a v type taper of order p.

The consistency of these estimators need the following assumptions.

ASSUMPTION 5. $x_t \quad Ex_1 \quad \sum_{j \quad 0}^{\infty} \alpha_j \varepsilon_{t-j}$ *and* $\sum_{j \quad 0}^{\infty} \alpha_j^2 < \infty$, *where* $E[\varepsilon_t | F_{t-1}] \quad 0$, $E[\varepsilon_t^2 | F_{t-1}] \quad 1$ *for* $t \quad 0, \pm 1, \pm 2, \ldots$, F_t *is the* σ *field generated by* ε_s, $s \leq t$, *and there exists a random variable* ε *such that* $E\varepsilon^2 < \infty$ *and for all* $\eta > 0$ *and some* $\kappa < 1$, $P(|\varepsilon_t| > \eta) \leq \kappa P(|\varepsilon| > \eta)$.

ASSUMPTION 6. *For* $i \quad 1, 2$,

$$\frac{n^{2(d^* - d_i)}}{l^{2p-1+2(d^*-d_i)}} \log m + \frac{l}{m} + \frac{m}{n} \to 0$$

as $n \to \infty$.

Assumption 6 relaxes significantly the trimming needed for consistency of the untapered estimate in Arteche and Robinson (2000). Considering $l \sim n^a$, Assumption 6 implies that $a > 2(d^* - d_i)[2p - 1 + 2(d^* - d_i)]^{-1}$, which for $p > 1$ significantly relaxes the condition $a > 2(d^* - d_i)[1 + 2(d^* - d_i)]^{-1}$ required in Arteche and Robinson (2000) for the untapered estimate. If in addition $m \sim n^b$ valid choices of a and b are $2/5 < a < b < 1$ for any $p > 1$.

Theorem 6. *Under Assumptions* 1, 2, 5 *and* 6, $\tilde{d}_i^{(v,p)} \xrightarrow{p} d_i$.

For the asymptotic normality we need the following additional assumptions.

Assumption 7. *Assumption* 5 *holds and*

$$E(\varepsilon_t^3 | F_{t-1}) = \mu_3 \text{ and } E(\varepsilon_t^4 | F_{t-1}) = \mu_4, \; t = 0, \pm 1, \ldots$$

for finite constants μ_3 *and* μ_4.

Assumption 8. *For* $i = 1, 2$,

$$\frac{n^{2(d^*-d_i)}}{l^{2p-1+2(d^*-d_i)}} \log m + \frac{l^2 \log^2 m}{m} + \frac{m^{2\beta+1}}{n^{2\beta}} (\log m)^2 \to 0$$

as $n \to \infty$.

Theorem 7. *Under Assumptions* 1, 2, 7 *and* 8

$$\sqrt{m}(\tilde{d}_i^{(v,p)} - d_i) \xrightarrow{d} N\left(0, \frac{1}{4}\Phi^{(v,p)}\right),$$

where

$$\Phi^{(v,p)} := \lim_{n\to\infty} n \left(\sum_1^n \left| h_t^{(v,p)} \right|^2 \right)^{-2} \sum_1^n \left| h_t^{(v,p)} \right|^4.$$

Remark. The $\Phi^{(v,p)}$ factor arises due to the correlation introduced among periodogram ordinates by tapering. It can be substituted by the quantity whose limit is evaluated in the definition, with n equal to the sample size or any other larger integer, to estimate the asymptotic variance of $\tilde{d}_i^{(v,p)}$ and to construct confidence intervals and Wald tests. Compared with the log periodogram estimate, asymptotic inference based on the Gaussian semi parametric estimate is simpler and more efficient (for the same bandwidth m) since the asymptotic variance has an explicit expression, which can be easily computed, permitting a more straightforward inference and can be showed to be smaller for $p = 1$ and any M (cf. Robinson 1995a, 1995b). Gaussian estimation also avoids dropping of frequencies and the choice of pooling, but trimming is necessary for both estimates in the presence of possible asymmetry. Nevertheless Gaussian estimation requires numerical optimization, whereas the log periodogram regression can be performed by OLS and asymptotic variance estimation can be based on standard subroutines for autocorrelation robust standard errors.

5. FINITE SAMPLE BEHAVIOUR

To illustrate the finite sample properties of the memory estimates we generate Gaussian SCALM processes with spectral density

$$f(\lambda) \quad \begin{cases} \frac{1}{2\pi}\left|1+e^{i2\lambda}\right|^{-2d_1} & \text{if } \frac{\pi}{2} < \lambda \leq \pi, \\ \frac{1}{2\pi}\left|1+e^{i2\lambda}\right|^{-2d_2} & \text{if } 0 \leq \lambda \leq \frac{\pi}{2}, \end{cases}$$

which shows a pole or zero (depending on the values of d_1 and d_2) at $\pi/2$. We take $d_1, d_2 \quad \{ \; 0.4, \; 0.2, 0, 0.2, 0.4\}$, which correspond to positive memory or

TABLE I

BIAS OF ESTIMATES OF d_1, N = 256, M = 32

				$l \quad 0$		
$d_1 \backslash d_2$		−0.4	−0.2	0	0.2	0.4
−0.4	$\tilde{d}_1(\hat{d}_1)$	0.0251 (0.0292)	0.0313 (0.0369)	0.0489 (0.0552)	0.0977 (0.1066)	0.2083 (0.2189)
	$\tilde{d}_1^{(1,2)}(\hat{d}_1^{(1,2)})$	0.0170 (0.0095)	0.0192 (0.0121)	0.0312 (0.0257)	0.1000 (0.0905)	0.3205 (0.2203)
−0.2	$\tilde{d}_1(\hat{d}_1)$	−0.0037 (0.0049)	−0.0011 (0.0074)	0.0066 (0.0167)	0.0312 (0.0417)	0.1011 (0.1171)
	$\tilde{d}_1^{(1,2)}(\hat{d}_1^{(1,2)})$	−0.0096 (0.0003)	−0.0087 (0.0013)	−0.0040 (0.0070)	0.0298 (0.0371)	0.1818 (0.1402)
0	$\tilde{d}_1(\hat{d}_1)$	−0.0154 (−0.0071)	−0.0145 (−0.0056)	−0.0117 (−0.0019)	−0.0019 (0.0088)	0.0325 (0.0453)
	$\tilde{d}_1^{(1,2)}(\hat{d}_1^{(1,2)})$	−0.0158 (−0.0047)	−0.0155 (−0.0043)	−0.0139 (−0.0019)	−0.0016 (0.0107)	0.0807 (0.0725)
0.2	$\tilde{d}_1(\hat{d}_1)$	−0.0209 (−0.0119)	−0.0206 (−0.0115)	−0.0197 (−0.0104)	−0.0165 (−0.0067)	−0.0030 (0.0078)
	$\tilde{d}_1^{(1,2)}(\hat{d}_1^{(1,2)})$	−0.0151 (−0.0034)	−0.0149 (−0.0031)	−0.0144 (−0.0024)	−0.0105 (0.0014)	0.0229 (0.0289)
0.4	$\tilde{d}_1(\hat{d}_1)$	−0.0284 (−0.0115)	−0.0284 (−0.0113)	−0.0283 (−0.0111)	−0.0276 (−0.0106)	−0.0241 (−0.0062)
	$\tilde{d}_1^{(1,2)}(\hat{d}_1^{(1,2)})$	−0.0186 (0.0070)	−0.0186 (0.0071)	−0.0184 (0.0074)	−0.0175 (0.0086)	−0.0100 (0.0165)
				$l \quad 1$		
−0.4	−0.2	0	0.2	0.4		
−0.4	$\tilde{d}_1(\hat{d}_1)$	0.0308 (0.0293)	0.0338 (0.0331)	0.0411 (0.0410)	0.0616 (0.0636)	0.1220 (0.1352)
	$\tilde{d}_1^{(1,2)}(\hat{d}_1^{(1,2)})$	0.0294 (0.0133)	0.0300 (0.0141)	0.0313 (0.0158)	0.0346 (0.0194)	0.0480 (0.0343)
−0.2	$\tilde{d}_1(\hat{d}_1)$	0.0020 (0.0113)	0.0038 (0.0129)	0.0078 (0.0177)	0.0191 (0.0291)	0.0552 (0.0709)
	$\tilde{d}_1^{(1,2)}(\hat{d}_1^{(1,2)})$	−0.0023 (0.0039)	−0.0020 (0.0046)	−0.0013 (0.0059)	0.0005 (0.0076)	0.0080 (0.0147)
0	$\tilde{d}_1(\hat{d}_1)$	−0.0100 (−0.0001)	−0.0092 (0.0006)	−0.0074 (0.0024)	−0.0024 (0.0079)	0.0156 (0.0279)
	$\tilde{d}_1^{(1,2)}(\hat{d}_1^{(1,2)})$	−0.0127 (−0.0030)	−0.0124 (−0.0026)	−0.0120 (−0.0020)	−0.0112 (−0.0008)	−0.0076 (0.0017)
0.2	$\tilde{d}_1(\hat{d}_1)$	−0.0177 (−0.0074)	−0.0173 (−0.0070)	−0.0166 (−0.0062)	−0.0146 (−0.0040)	−0.0067 (0.0030)
	$\tilde{d}_1^{(1,2)}(\hat{d}_1^{(1,2)})$	−0.0195 (−0.0086)	−0.0194 (−0.0086)	−0.0192 (−0.0083)	−0.0188 (−0.0075)	−0.0171 (−0.0062)
0.4	$\tilde{d}_1(\hat{d}_1)$	−0.0325 (−0.0113)	−0.0324 (−0.0109)	−0.0323 (−0.0107)	−0.0319 (−0.0099)	−0.0297 (−0.0078)
	$\tilde{d}_1^{(1,2)}(\hat{d}_1^{(1,2)})$	−0.0428 (−0.0123)	−0.0428 (−0.0122)	−0.0427 (−0.0120)	−0.0425 (−0.0117)	−0.0419 (−0.0108)
				$l \quad 2$		
−0.4	−0.2	0	0.2	0.4		
−0.4	$\tilde{d}_1(\hat{d}_1)$	0.0354 (0.0240)	0.0369 (0.0264)	0.0408 (0.0320)	0.0524 (0.0449)	0.0902 (0.0940)
	$\tilde{d}_1^{(1,2)}(\hat{d}_1^{(1,2)})$	0.0405 (0.0150)	0.0408 (0.0154)	0.0414 (0.0161)	0.0429 (0.0183)	0.0521 (0.0299)
−0.2	$\tilde{d}_1(\hat{d}_1)$	0.0021 (0.0081)	0.0032 (0.0098)	0.0057 (0.0129)	0.0134 (0.0207)	0.0389 (0.0516)
	$\tilde{d}_1^{(1,2)}(\hat{d}_1^{(1,2)})$	0.0015 (0.0048)	0.0017 (0.0053)	0.0020 (0.0062)	0.0029 (0.0076)	0.0085 (0.0134)
0	$\tilde{d}_1(\hat{d}_1)$	−0.0117 (−0.0034)	−0.0111 (−0.0028)	−0.0096 (−0.0010)	−0.0053 (0.0028)	0.0092 (0.0184)
	$\tilde{d}_1^{(1,2)}(\hat{d}_1^{(1,2)})$	−0.0123 (−0.0031)	−0.0122 (−0.0028)	−0.0119 (−0.0023)	−0.0114 (−0.0014)	−0.0086 (0.0017)
0.2	$\tilde{d}_1(\hat{d}_1)$	−0.0214 (−0.0124)	−0.0210 (−0.0120)	−0.0202 (−0.0111)	−0.0180 (−0.0088)	−0.0104 (−0.0023)
	$\tilde{d}_1^{(1,2)}(\hat{d}_1^{(1,2)})$	−0.0234 (−0.0108)	−0.0233 (−0.0108)	−0.0232 (−0.0105)	−0.0229 (−0.0097)	−0.0215 (−0.0080)
0.4	$\tilde{d}_1(\hat{d}_1)$	−0.0427 (−0.0184)	−0.0425 (−0.0181)	−0.0422 (−0.0176)	−0.0413 (−0.0164)	−0.0385 (−0.0134)
	$\tilde{d}_1^{(1,2)}(\hat{d}_1^{(1,2)})$	−0.0550 (−0.0173)	−0.0550 (−0.0172)	−0.0549 (−0.0171)	−0.0549 (−0.0166)	−0.0543 (−0.0153)

persistence (0.2, 0.4), short memory (0) and negative memory or antipersistence (0.2, 0.4). We consider a sample size n 256 and bandwidth m 32. The effect of the trimming on the bias and the mean square error (MSE) on both tapered and untapered estimates is assessed by considering l 0, 1, 2. The number of replications was 1000. For a description of the simulating technique see Arteche and Robinson (2000).

We just consider a representative type I taper of order p 2 (triangular Barlett). Tables I and II show the bias and MSE of the two semi parametric estimates of d_1 for the different situations considered. The untrimmed estimates are highly biased and with high MSE in those cases, where $d_2 > d_1$.

When the first frequency is omitted both bias and MSE decrease significantly in the extreme cases $d_2 > d_1$, especially, when the tapered periodogram is used. The decrease in bias compensates the increase in variance caused by tapering such that the MSE is lower for the tapered estimates. A stronger trimming (l 2) is only beneficial for the bias of the untapered estimates, when $d_2 > d_1$ and the MSE only decreases for d_2 0.4 and d_1 0.4. For the tapered estimates, the MSE grows in

TABLE II

MSE OF ESTIMATES OF d_1, N = 256, M = 32

				l 0		
$d_1\backslash d_2$		−0.4	−0.2	0	0.2	0.4
−0.4	$\tilde{d}_1$ ($\hat{d}_1$)	0.0094 (0.0176)	0.0099 (0.0185)	0.0120 (0.0219)	0.0216 (0.0315)	0.0613 (0.0722)
	$\tilde{d}_1^{(1,2)}$ ($\hat{d}_1^{(1,2)}$)	0.0120 (0.0263)	0.0122 (0.0267)	0.0134 (0.0268)	0.0257 (0.0332)	0.1282 (0.0743)
−0.2	$\tilde{d}_1$ ($\hat{d}_1$)	0.0116 (0.0185)	0.0116 (0.0190)	0.0119 (0.0191)	0.0134 (0.0219)	0.0250 (0.0355)
	$\tilde{d}_1^{(1,2)}$ ($\hat{d}_1^{(1,2)}$)	0.0178 (0.0261)	0.0178 (0.0262)	0.0180 (0.0254)	0.0188 (0.0270)	0.0549 (0.0454)
0	$\tilde{d}_1$ ($\hat{d}_1$)	0.0121 (0.0195)	0.0122 (0.0193)	0.0122 (0.0192)	0.0122 (0.0195)	0.0140 (0.0223)
	$\tilde{d}_1^{(1,2)}$ ($\hat{d}_1^{(1,2)}$)	0.0189 (0.0260)	0.0189 (0.0259)	0.0190 (0.0253)	0.0185 (0.0248)	0.0250 (0.0304)
0.2	$\tilde{d}_1$ ($\hat{d}_1$)	0.0124 (0.0193)	0.0124 (0.0195)	0.0125 (0.0199)	0.0125 (0.0200)	0.0126 (0.0199)
	$\tilde{d}_1^{(1,2)}$ ($\hat{d}_1^{(1,2)}$)	0.0182 (0.0245)	0.0182 (0.0246)	0.0182 (0.0245)	0.0180 (0.0244)	0.0179 (0.0254)
0.4	$\tilde{d}_1$ ($\hat{d}_1$)	0.0103 (0.0186)	0.0104 (0.0186)	0.0104 (0.0188)	0.0104 (0.0192)	0.0103 (0.0193)
	$\tilde{d}_1^{(1,2)}$ ($\hat{d}_1^{(1,2)}$)	0.0121 (0.0239)	0.0121 (0.0239)	0.0121 (0.0239)	0.0121 (0.0239)	0.0118 (0.0249)
				l 1		
−0.4	−0.2	0	0.2	0.4		
−0.4	$\tilde{d}_1$ ($\hat{d}_1$)	0.0124 (0.0258)	0.0126 (0.0258)	0.0133 (0.0268)	0.0165 (0.0301)	0.0319 (0.0469)
	$\tilde{d}_1^{(1,2)}$ ($\hat{d}_1^{(1,2)}$)	0.0164 (0.0375)	0.0165 (0.0374)	0.0167 (0.0374)	0.0172 (0.0384)	0.0192 (0.0404)
−0.2	$\tilde{d}_1$ ($\hat{d}_1$)	0.0160 (0.0259)	0.0159 (0.0261)	0.0159 (0.0258)	0.0163 (0.0271)	0.0203 (0.0317)
	$\tilde{d}_1^{(1,2)}$ ($\hat{d}_1^{(1,2)}$)	0.0242 (0.0381)	0.0243 (0.0378)	0.0244 (0.0374)	0.0245 (0.0377)	0.0249 (0.0387)
0	$\tilde{d}_1$ ($\hat{d}_1$)	0.0167 (0.0263)	0.0166 (0.0265)	0.0165 (0.0266)	0.0165 (0.0262)	0.0172 (0.0264)
	$\tilde{d}_1^{(1,2)}$ ($\hat{d}_1^{(1,2)}$)	0.0265 (0.0378)	0.0265 (0.0377)	0.0265 (0.0375)	0.0266 (0.0374)	0.0267 (0.0382)
0.2	$\tilde{d}_1$ ($\hat{d}_1$)	0.0169 (0.0264)	0.0168 (0.0265)	0.0168 (0.0265)	0.0167 (0.0259)	0.0167 (0.0260)
	$\tilde{d}_1^{(1,2)}$ ($\hat{d}_1^{(1,2)}$)	0.0260 (0.0375)	0.0260 (0.0377)	0.0261 (0.0377)	0.0261 (0.0376)	0.0261 (0.0379)
0.4	$\tilde{d}_1$ ($\hat{d}_1$)	0.0135 (0.0263)	0.0134 (0.0261)	0.0134 (0.0259)	0.0133 (0.0257)	0.0131 (0.0256)
	$\tilde{d}_1^{(1,2)}$ ($\hat{d}_1^{(1,2)}$)	0.0199 (0.0370)	0.0199 (0.0370)	0.0199 (0.0370)	0.0199 (0.0371)	0.0198 (0.0374)
				l 2		
−0.4	−0.2	0	0.2	0.4		
−0.4	$\tilde{d}_1$ ($\hat{d}_1$)	0.0157 (0.0359)	0.0158 (0.0353)	0.0161 (0.0347)	0.0178 (0.0372)	0.0258 (0.0449)
	$\tilde{d}_1^{(1,2)}$ ($\hat{d}_1^{(1,2)}$)	0.0218 (0.0509)	0.0218 (0.0508)	0.0220 (0.0509)	0.0223 (0.0512)	0.0242 (0.0535)
−0.2	$\tilde{d}_1$ ($\hat{d}_1$)	0.0214 (0.0359)	0.0213 (0.0354)	0.0212 (0.0349)	0.0212 (0.0357)	0.0228 (0.0366)
	$\tilde{d}_1^{(1,2)}$ ($\hat{d}_1^{(1,2)}$)	0.0317 (0.0516)	0.0318 (0.0512)	0.0319 (0.0507)	0.0321 (0.0507)	0.0327 (0.0519)
0	$\tilde{d}_1$ ($\hat{d}_1$)	0.0230 (0.0360)	0.0229 (0.0361)	0.0227 (0.0357)	0.0224 (0.0353)	0.0222 (0.0345)
	$\tilde{d}_1^{(1,2)}$ ($\hat{d}_1^{(1,2)}$)	0.0351 (0.0505)	0.0352 (0.0503)	0.0353 (0.0502)	0.0354 (0.0501)	0.0356 (0.0509)
0.2	$\tilde{d}_1$ ($\hat{d}_1$)	0.0231 (0.0359)	0.0230 (0.0361)	0.0229 (0.0358)	0.0226 (0.0349)	0.0222 (0.0347)
	$\tilde{d}_1^{(1,2)}$ ($\hat{d}_1^{(1,2)}$)	0.0345 (0.0502)	0.0345 (0.0503)	0.0345 (0.0502)	0.0347 (0.0501)	0.0348 (0.0505)
0.4	$\tilde{d}_1$ ($\hat{d}_1$)	0.0182 (0.0355)	0.0181 (0.0352)	0.0180 (0.0350)	0.0179 (0.0345)	0.0174 (0.0343)
	$\tilde{d}_1^{(1,2)}$ ($\hat{d}_1^{(1,2)}$)	0.0263 (0.0501)	0.0263 (0.0501)	0.0263 (0.0501)	0.0264 (0.0501)	0.0263 (0.0502)

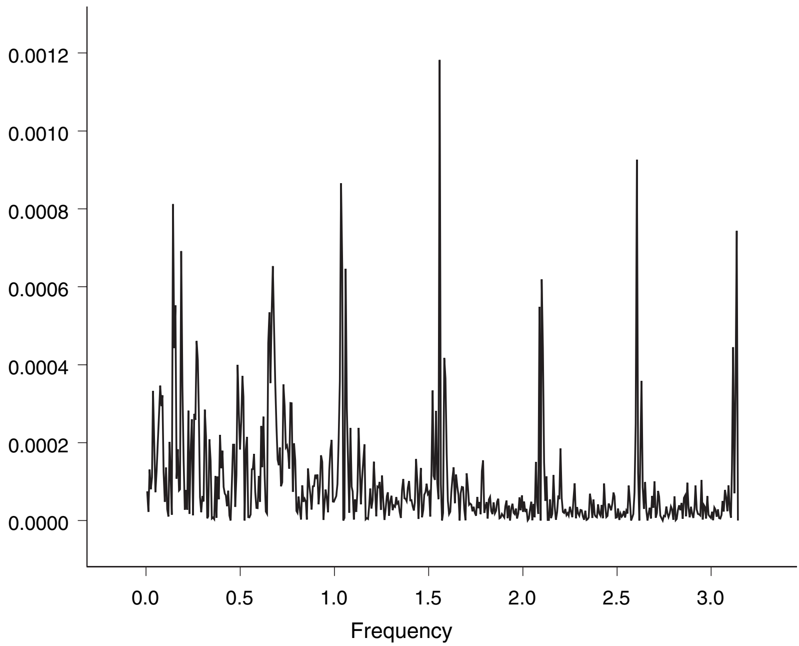

FIGURE 1. Periodogram of the growth rate of US Industrial Production Index.

every case, when l increases from 1 to 2 and only the bias of the tapered log periodogram estimate decreases in the cases (d_1, d_2) (0.4, 0.2), (0.4, 0.4), (0.2, 0.4), whereas the bias of the tapered Gaussian semi parametric estimates increases in every situation. Overall, the main contribution of the tapering is a significant reduction of the bias of the estimates of the lower memory parameter, when combined with $l \geq p$ 1. This lower bound on trimming is natural because tapering introduces extra correlation with the closest p 1 periodograms, possibly at the other side of the spectral singularity. This bias reduction is achieved with a less stringent trimming than in the untapered case and it seems that in practice it is not necessary to make it sample size dependent, but can be decided only in function of the data taper used.

6. EMPIRICAL EXAMPLE

We consider the growth rate of the monthly US Industrial Production Index with base 100 in 1997 from 1919:10 to 2003:9 (n 1008). The deterministic seasonality of the series have been subtracted by means of seasonal dummies and the strong

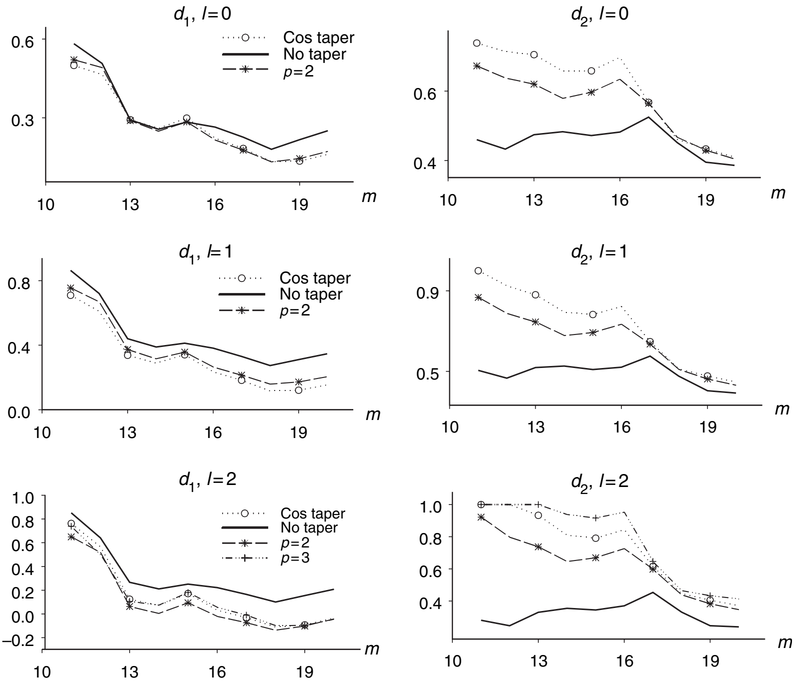

FIGURE 2. Gaussian semi-parametric memory estimates at $\pi/3$. Estimates of d_1 (left column) and d_2 (right column) are calculated for $m = 11, \ldots, 30$ and for $l = 0$ (first row), $l = 1$ (second row) and $l = 2$ (third row). No taper, cosine taper and Zhurbenko tapers of orders $p = 2$, 3 are used.

seasonality that appears in the periodogram in Figure 1 is free of different seasonal means. Since n is a multiple of 12, this extraction of the deterministic seasonality does not affect the periodogram at nonseasonal frequencies λ_j / $\pi j/6$, j 0, 1, ..., 6, which are the Fourier frequencies used in the estimation of the memory parameters (Arteche, 2002, pp. 275 6).

We focus on the analysis at frequency $\pi/3$, where the possible spectral asymmetry is finally more evident. Results for other seasonal frequencies are available upon request. Figure 2 shows Gaussian semi parametric estimates of the memory parameters d_1 and d_2 at frequency $\pi/3$ for a grid of bandwidths from m 11 to m 20. We consider untapered and tapered versions for the cosine and type I tapers of orders p 2 and p 3. The maximum m analysed is 20 in order to avoid distorting influences of possible spectral poles at neighbouring seasonal frequencies $\pi/6$ and $\pi/2$. The effect of the trimming is analysed by considering l 0, 1, 2. Although the theory shown in this paper is only valid for d_1, d_2 < 0.5, we consider also the possibility of nonstationarities appealing to Velasco (1999b).

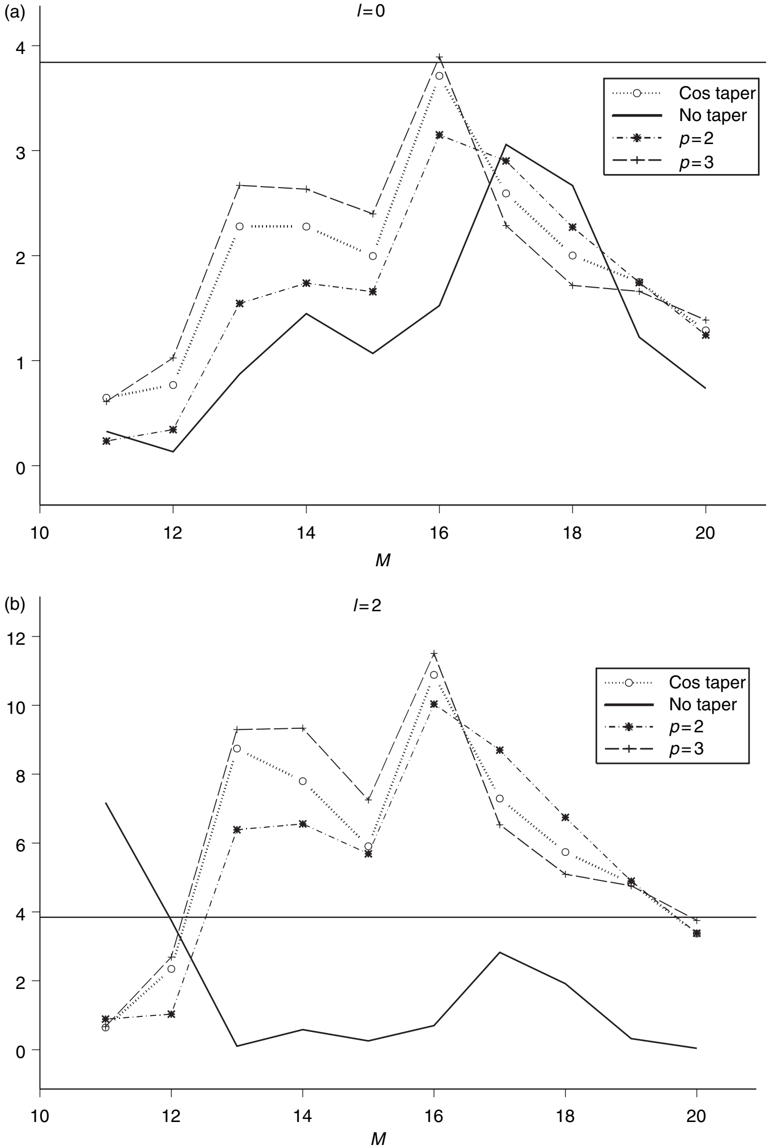

FIGURE 3. Wald tests for spectral symmetry at $\pi/3$. Wald test statistics are calculated for $m = 11, \ldots, 30$, $l = 0$ in (a) and $l = 2$ in (b), and compared with the asymptotic 5% critical value of a χ_1^2 distribution (horizontal line at 3.84). No taper, cosine taper and Zhurbenko tapers of orders $p = 2,\ 3$ are used.

The memory estimates for $l = 0$ are clearly positive and similar at both sides of the seasonal peak (especially the untapered ones), suggesting symmetric and positive long memory at $\pi/3$. However, the differences between both estimates increase with the trimming ($l > 0$) and the tapering ($p > 1$).

Figure 3 displays Wald tests of the null hypothesis of spectral symmetry $d_1 = d_2$ based on the asymptotic distribution of the Gaussian semi parametric estimates with $l = 0$ and $l = 2$ in Theorem 7. Considering the asymptotic independence of $\tilde{d}_1^{(v,p)}$ and $\tilde{d}_2^{(v,p)}$, the Wald statistics are

$$\frac{2m(\tilde{d}_1^{(1,p)} - \tilde{d}_2^{(1,p)})^2}{\hat{\Phi}_n^{(1,p)}},$$

where

$$\hat{\Phi}_n^{(1,p)} = n\left(\sum_1^n \left|h_t^{(1,p)}\right|^2\right)^{-2} \sum_1^n \left|h_t^{(1,p)}\right|^4,$$

which for $n = 1008$ and the tapers of order $p = 1$ (no taper), 2, 3 and the cosine taper is 1, 1.8, 2.24 and 1.94. The Wald statistics have an asymptotic χ_1^2 distribution under the null of spectral symmetry, $d_1 = d_2$. The untapered Wald test does not reject the null hypothesis for any of the bandwidths and trimming considered (except for $m = 11$ and $l = 2$). The situation changes with the tapered and trimmed test statistics. The larger difference between right and left estimates compensates the increase in variance so that the hypothesis of symmetry is mostly rejected.

APPENDIX

Proof of Theorems 1 and 2. Since $d_1, d_2 < 0.5$, we can write the expectation as

$$E\left[\frac{I_{1j}^{(v,p)}}{C_1 \lambda_j^{-2d_1}}\right] = \int_{-\pi}^{\pi} g_{n,j}^{(v,p)}(\lambda)\mathrm{d}\lambda, \tag{15}$$

where

$$g_{n,j}^{(v,p)}(\lambda) := K_n^{(v,p)}(\omega + \lambda_j - \lambda)\frac{f(\lambda)}{C_1 \lambda_j^{-2d_1}}$$

and

$$K_n^{(v,p)}(\lambda) := \left(2\pi \sum |h_t^{(v,p)}|^2\right)^{-1} \left|D^{(v,p)}(\lambda)\right|^2.$$

The integral (15) can be decomposed into

$$\left\{\int_{-\pi}^{\omega - n^{-\alpha}} + \int_{\omega - n^{-\alpha}}^{\omega + n^{-\alpha}} + \int_{\omega + n^{-\alpha}}^{\pi}\right\} g_{n,j}^{(v,p)}(\lambda)\mathrm{d}\lambda.$$

for some $\alpha \in (0, 0.5]$. The integral over $[\ \ \pi, \omega\ \ n^{\ \ \alpha}]$ can be written

$$\int_{\ \ \pi\ \ \omega}^{\ \ n^{\ \ \alpha}} K_n^{(v,p)}(\lambda_j \quad \lambda)\frac{f(\omega+\lambda)}{C_1\lambda_j^{\ \ 2d_1}}\mathrm{d}\lambda. \tag{16}$$

Since $K_n^{(v,p)}(\lambda) \leq \text{const.}\min(n, n^{1\ \ 2p}|\lambda|^{\ \ 2p}|)$, we have that for a sufficiently large n (16) is bounded by

$$\text{const.}n^{1\ \ 2p}|\lambda_j + n^{\ \ \alpha}|^{\ \ 2p}\lambda_j^{2d_1}\int_{\ \ \pi}^{\pi} f(\lambda)\mathrm{d}\lambda \quad O(n^{1\ \ 2p}n^{2p\alpha}n^{\ \ 2d_1}) \quad o(1)$$

for $\ \ 0.5 < d_1, d_2 < 0.5$ and $\alpha < 1 + (d_1 \quad 0.5)/p$. Similarly, the integral between $\omega + n^{\ \ \alpha}$ and π is

$$\int_{n^{\ \ \alpha}}^{\pi\ \ \omega} K_n^{(v,p)}(\lambda_j \quad \lambda)\frac{f(\omega+\lambda)}{C_1\lambda_j^{\ \ 2d_1}}\mathrm{d}\lambda \leq \text{const.}n^{1\ \ 2p}|\lambda_j \quad n^{\ \ \alpha}|^{\ \ 2p}\lambda_j^{2d_1} \quad o(1)$$

for a large enough n, such that $\lambda_j \quad n^{\ \ \alpha} < 0$ and the same α as before. Thus as $n \to \infty$ and j fixed

$$E\left[\frac{I_{1j}^{(v,p)}}{C_1\lambda_j^{\ \ 2d_1}}\right] \quad \int_{\omega\ \ n^{\ \ \alpha}}^{\omega+n^{\ \ \alpha}} g_{n,j}^{(v,p)}(\lambda)\mathrm{d}\lambda + o(1). \tag{17}$$

Since the behaviour of the spectral density (5) is different to the right and left of ω, we split the integral in (17) into two. First

$$\begin{aligned}\int_{\omega}^{\omega+n^{\ \ \alpha}} g_{n,j}^{(v,p)}(\lambda)\mathrm{d}\lambda \quad & \int_0^{n^{\ \ \alpha}} K_n^{(v,p)}(\lambda_j \quad \lambda)\frac{f(\omega+\lambda)}{C_1\lambda_j^{\ \ 2d_1}}\mathrm{d}\lambda \\ & \int_0^{n^{1\ \ \alpha}} \frac{1}{n}K_n^{(v,p)}\left(\frac{2\pi j \quad \lambda}{n}\right)\frac{f(\omega+\frac{\lambda}{n})}{C_1\lambda_j^{\ \ 2d_1}}\mathrm{d}\lambda \\ & \int_0^{\infty} h_{n,j}^{(v,p,1)}(\lambda)\mathrm{d}\lambda,\end{aligned}$$

where

$$h_{n,j}^{(v,p,1)}(\lambda) : \quad \frac{1}{n}K_n^{(v,p)}\left(\frac{2\pi j \quad \lambda}{n}\right)\frac{f(\omega+\frac{\lambda}{n})}{C_1\lambda_j^{\ \ 2d_1}}\chi_{[0,n^{1\ \ \alpha}]}$$

and $\chi_{[0,n^{1\ \ \alpha}]}$ is the indicator function of the interval $[0, n^{1\ \ \alpha}]$. Proceeding like in the proof of Theorem 1 in Hurvich and Beltrao (1993) we see that $h_{n,j}^{(v,p,1)}(\lambda)$ for the different tapers considered is dominated by an integrable function. Thus we can use Lebesgue's dominated convergence theorem (see for instance Temple, 1971, Thm 9.3.7) and we have that

$$\lim_{n\to\infty}\int_0^{\infty} h_{n,j}^{(v,p,1)}(\lambda)\mathrm{d}\lambda \quad \int_0^{\infty}\lim_{n\to\infty} h_{n,j}^{(v,p,1)}(\lambda)\mathrm{d}\lambda \quad \int_0^{\infty} h_{1,j}^{(v,p)}(\lambda)\mathrm{d}\lambda.$$

For each taper considered $h_{n,j}^{(v,p,1)}(\lambda)$ and $h_{1,j}^{(v,p)}(\lambda)$ have different expressions. With no tapering

$$h_{n,j}^{(v,1,1)}(\lambda) \quad \frac{\sin^2\left(\frac{2\pi j \ \ \lambda}{2}\right)}{2\pi n^2\sin^2\left(\frac{2\pi j \ \ \lambda}{2n}\right)}\frac{f\left(\omega+\frac{\lambda}{n}\right)}{C_1\lambda_j^{\ \ 2d_1}}\chi_{[0,n^{1\ \ \alpha}]}$$

and

$$h_{1,j}^{(v,1)}(\lambda) \quad \frac{2}{\pi} \frac{\sin^2\left(\frac{2\pi j \quad \lambda}{2}\right)}{(2\pi j \quad \lambda)^2} \left|\frac{\lambda}{2\pi j}\right|^{\quad 2d_1}$$

for $0 \leq \lambda < \quad \infty$, using the fact that $\sin^2((2\pi j \quad \lambda)/2) \quad \sin^2(\lambda/2)$ for j an integer. Similarly, for the type I taper

$$K_n^{(1,p)}\left(\frac{2\pi j \quad \lambda}{n}\right) : \quad \frac{\left|a\left(\frac{2\pi j \quad \lambda}{n}\right)\right|^2}{n^{2p \quad 1} 2\pi b} \frac{\sin^{2p}\left(\frac{2\pi j \quad \lambda}{2p}\right)}{\sin^{2p}\left(\frac{2\pi j \quad \lambda}{2n}\right)}$$

and

$$h_{1,j}^{(1,p)}(\lambda) : \quad |2\pi j|^{2d_1} \frac{2^{2p \quad 1}|a(0)|^2}{\pi b} \frac{\sin^{2p}\left(\frac{2\pi j \quad \lambda}{2p}\right)}{(2\pi j \quad \lambda)^{2p}} |\lambda|^{\quad 2d_1}.$$

Finally, for the type II, taper we use eqn (5) in Hurvich *et al*. (2002) such that

$$\begin{aligned} D^{(2,p)}\left(\frac{2\pi j \quad \lambda}{n}\right) &\quad \sum_{k \quad 0}^{p \quad 1} \binom{p \quad 1}{k} (\quad 1)^k \frac{\sin\left(\frac{2\pi(j+k) \quad \lambda}{2}\right)}{\sin\left(\frac{2\pi(j+k) \quad \lambda}{2n}\right)} \exp\left[\mathrm{i}\left(\frac{n+1}{2}\right)\left(\frac{2\pi(j+k) \quad \lambda}{n}\right)\right] \\ &(\quad 1)^{j+1} \sin\frac{\lambda}{2} \exp\left[\mathrm{i}\left(\frac{n+1}{2}\right)\left(\frac{2\pi j \quad \lambda}{n}\right)\right] \\ &\times \sum_{k \quad 0}^{p \quad 1} \binom{p \quad 1}{k} \sin^{\quad 1}\left(\frac{2\pi(j+k) \quad \lambda}{2n}\right) \exp(\mathrm{i}\pi k) \exp\left(\frac{\pi k}{n}\right) \end{aligned}$$

and thus $h_{1,j}^{(2,p)}$ is

$$h_{1,j}^{(2,p)} \quad \frac{2}{\pi} a_p^{\quad 1} \sin^2\left(\frac{\lambda}{2}\right) |2\pi j|^{2d_1} |\lambda|^{\quad 2d_1} \left(\sum_{k \quad 0}^{p \quad 1} \binom{p \quad 1}{k} \frac{(\quad 1)^k}{(2\pi(j+k) \quad \lambda)}\right)^2.$$

Regarding the integral over $[\omega \quad n^{\quad \alpha}, \omega]$, multiply it by $n^{2(d_1 \quad d_2)}$,

$$\begin{aligned} &n^{2(d_1 \quad d_2)} \int_{\omega \quad n^{\quad \alpha}}^{\omega} g_{n,j}^{(v,p)}(\lambda) \mathrm{d}\lambda \\ &\quad n^{2(d_1 \quad d_2)} \int_{\quad n^{\quad \alpha}}^{0} K_n^{(v,p)}(\lambda_j \quad \lambda) \frac{f(\omega+\lambda)}{C_1 \lambda_j^{\quad 2d_1}} \mathrm{d}\lambda \\ &\quad n^{\quad 2d_2} \int_{\quad n^{1 \quad \alpha}}^{0} \frac{1}{n} K_n^{(v,p)}\left(\frac{2\pi j \quad \lambda}{n}\right) \frac{f(\omega+\frac{\lambda}{n})}{C_1} |2\pi j|^{2d_1} \mathrm{d}\lambda \\ &\quad \int_{\quad \infty}^{0} h_{n,j}^{(v,p,2)}(\lambda) \mathrm{d}\lambda, \end{aligned}$$

where

$$h_{n,j}^{(v,p,2)}(\lambda) \quad |2\pi j|^{2d_1} n^{\quad 2d_2} \frac{1}{n} K_n^{(v,p)}\left(\frac{2\pi j \quad \lambda}{n}\right) \frac{f(\omega+\frac{\lambda}{n})}{C_1} \chi_{[\quad n^{1 \quad \alpha},0]}$$

and $\chi_{[\ \ n^{1}\ \ \alpha,0]}$ is the indicator function of the interval $[\ \ n^{1\ \ \alpha}, 0]$. Proceeding as before we get that as $n \to \infty$, $h_{n,j}^{(v,p,2)}(\lambda) \to h_{2,j}^{(v,p)}(\lambda)$ for $0 \leq \lambda < \infty$, where $h_{2,j}^{(v,1)}(\lambda)\ \ C_2 C_1^{\ \ 1}|2\pi j|^{2d_1}\psi_j(\lambda; d_2)$, $h_{2,j}^{(v,p)}(\lambda)\ \ C_2 C_1^{\ \ 1}|2\pi j|^{2d_1}\psi_j^{(v,p)}(\lambda; d_2)$, $v\ \ 1,2$, for no tapering and types I and II tapers respectively.

Now again

$$\lim_{n\to\infty}\int_{\ \ \infty}^{0} h_{n,j}^{(v,p,2)}(\lambda)\mathrm{d}\lambda \quad \int_{\ \ \infty}^{0} h_{2,j}^{(v,p)}(\lambda)\mathrm{d}\lambda$$

using Lebesgue's dominated convergence theorem. Thus if $d_1 > d_2$ then $\int_{\omega\ \ n\ \ \alpha}^{\omega} g_{n,j}^{(v,p)}(\lambda)\mathrm{d}\lambda \to 0$ as $n \to \infty$ and consequently a) is proved. When $d_1\ \ d_2$ we obtain the result stated in b). If $d_1 < d_2$ then $n^{2(d_1\ \ d_2)} \to 0$ as $n \to \infty$ so that the only integral with a limit different from zero is

$$n^{2(d_1\ \ d_2)}\int_{\omega\ \ n\ \ \alpha}^{\omega} g_{n,j}^{(v,p)}(\lambda)\mathrm{d}\lambda$$

and c) is proved. □

Proof of Theorem 3. The proof is similar to that of Theorem 2 in Robinson (1995a) and Theorem 8 in Velasco (1999a). We focus on the cases where the differences with existing work are more apparent, namely cases (c) and (d).

(c)

$$EW^{(v,p)}(\omega+\lambda_j)\overline{W^{(v,p)}}(\omega+\lambda_k)\quad A_{jk}^{(v,p)}f_{jk} + \int_{\ \ \pi}^{\pi} E_{jk}^{(v,p)}(\lambda)(f(\lambda)\quad f_{jk})\mathrm{d}\lambda. \tag{18}$$

Write the integral as

$$\left\{\int_{\ \ \pi}^{\omega+\frac{\lambda_k}{2}} + \int_{\omega+2\lambda_j}^{\pi}\right\}\{f(\lambda)\quad f_{jk}\}E_{jk}^{(v,p)}(\lambda)\mathrm{d}\lambda \tag{19}$$

$$+\int_{\omega+\frac{\lambda_k}{2}}^{\omega+\frac{\lambda_k+\lambda_j}{2}}\{f(\lambda)\quad f_{jk}\}E_{jk}^{(v,p)}(\lambda)\mathrm{d}\lambda \tag{20}$$

$$+\int_{\omega+\frac{\lambda_k+\lambda_j}{2}}^{\omega+2\lambda_j}\{f(\lambda)\quad f_{jk}\}E_{jk}^{(v,p)}(\lambda)\mathrm{d}\lambda. \tag{21}$$

First we decompose the integral in (19)

$$\int_{\ \ \pi}^{\omega\ \ \varepsilon} + \int_{\omega\ \ \varepsilon}^{\omega\ \ \lambda_j} + \int_{\omega\ \ \lambda_j}^{\omega\ \ \frac{\lambda_k}{2}} + \int_{\omega\ \ \frac{\lambda_k}{2}}^{\omega+\frac{\lambda_k}{2}} + \int_{\omega+2\lambda_j}^{\omega+\varepsilon} + \int_{\omega+\varepsilon}^{\pi}.$$

Now

$$\left|\int_{\ \ \pi}^{\omega\ \ \varepsilon} + \int_{\omega+\varepsilon}^{\pi}\right| \quad O(n^{1\ \ 2p}(1+\lambda_j^{\ \ d_1}\lambda_k^{\ \ d_1}))$$

$$O(\vartheta_{jk}(d_1)\lambda_k^{p\ \ \frac{1}{2}+d_1}\lambda_j^{p\ \ \frac{1}{2}+d_1}(1+\lambda_j^{\ \ d_1}\lambda_k^{\ \ d_1}))\quad O(\vartheta_{jk}(d_1)).$$

The integral over $[\omega \quad \varepsilon, \omega \quad \lambda_j]$ is bounded in absolute value by

$$\frac{1}{2\pi \sum |h_t^{(v,p)}|^2} \int_{\varepsilon}^{\lambda_j} |D^{(v,p)}(\lambda_j \quad \lambda)||\overline{D^{(v,p)}}(\lambda_k \quad \lambda)|(f(\omega+\lambda) \quad f_{jk})\,\mathrm{d}\lambda$$
$$O\left(n^{1 \quad 2p}\left[\int_{\lambda_j}^{\varepsilon} \lambda^{\quad 2d_2 \quad 2p} d\lambda + \lambda_j^{\quad d_1} \lambda_k^{\quad d_1} \int_{\lambda_j}^{\varepsilon} \lambda^{\quad 2p} d\lambda\right]\right)$$
$$O\left(n^{1 \quad 2p}\left[\lambda_j^{1 \quad 2d_2 \quad 2p} + \lambda_j^{1 \quad 2p \quad d_1} \lambda_k^{\quad d_1}\right]\right)$$
$$O\left(\theta_{jk}(d^*)\lambda_j^{d^*+p \quad 1/2} \lambda_k^{d^*+p \quad 1/2}\left[\lambda_j^{1 \quad 2d^* \quad 2p} + \lambda_j^{1 \quad 2p \quad d^*} \lambda_k^{\quad d^*}\right]\right)$$
$$O\left(\theta_{jk}(d^*)\left[\left(\frac{k}{j}\right)^{p+d^* \quad 1/2} + \left(\frac{k}{j}\right)^{p \quad 1/2}\right]\right)$$
$$O(\theta_{jk}(d^*))$$

using (11). Similarly

$$\left|\int_{\omega+2\lambda_j}^{\omega+\varepsilon}\right| \quad O(\vartheta_{jk}(d_1)).$$

Now,

$$\left|\int_{\omega \quad \lambda_j}^{\omega \quad \frac{\lambda_k}{2}}\right| \leq \frac{1}{2\pi \sum |h_t^{(v,p)}|^2} \int_{\lambda_j}^{\frac{\lambda_k}{2}} |D^{(v,p)}(\lambda_j \quad \lambda)||\overline{D^{(v,p)}}(\lambda_k \quad \lambda)|(f(\omega+\lambda) \quad f_{jk})\,\mathrm{d}\lambda$$
$$O(n^{1 \quad 2p}\lambda_j^{\quad p}[\lambda_j^{\quad 2d^*} + \lambda_k^{\quad 2d^*}]\lambda_k^{1 \quad p})$$
$$O(\vartheta_{jk}(d^*)\lambda_j^{d^* \quad \frac{1}{2}}\lambda_k^{\frac{1}{2}+d^*}[\lambda_j^{\quad 2d^*} + \lambda_k^{\quad 2d^*}]) \quad O(\vartheta_{jk}(d^*))$$

and

$$\left|\int_{\omega \quad \frac{\lambda_k}{2}}^{\omega+\frac{\lambda_k}{2}}\right| \leq \frac{1}{2\pi \sum |h_t^{(v,p)}|^2} \int_{\frac{\lambda_k}{2}}^{+\frac{\lambda_k}{2}} |D^{(v,p)}(\lambda_j \quad \lambda)||\overline{D^{(v,p)}}(\lambda_k \quad \lambda)|(f(\omega+\lambda) \quad f_{jk})\,\mathrm{d}\lambda$$
$$O(n^{1 \quad 2p}|\lambda_j \quad \frac{\lambda_k}{2}|^{\quad p}\lambda_k^{\quad p}[\lambda_k^{1 \quad 2d^*} + \lambda_k^{1 \quad d_1}\lambda_j^{\quad d_1}])$$
$$O(n^{1 \quad 2p}|\frac{\lambda_j}{2}|^{\quad p}\lambda_k^{1 \quad p}[\lambda_k^{\quad 2d^*} + \lambda_j^{\quad d^*}\lambda_k^{\quad d^*}])$$
$$O(\vartheta_{jk}(d^*)\lambda_j^{d^* \quad \frac{1}{2}}\lambda_k^{\frac{1}{2}+d^*}[\lambda_k^{\quad 2d^*} + \lambda_j^{\quad d^*}\lambda_k^{\quad d^*}]) \quad O(\vartheta_{jk}(d^*)).$$

Now (21) is bounded in absolute value by

$$\frac{1}{2\pi \sum |h_t^{(v,p)}|^2} \int_{\frac{\lambda_j+\lambda_k}{2}}^{2\lambda_j} |D^{(v,p)}(\lambda_j \quad \lambda)||\overline{D^{(v,p)}}(\lambda_k \quad \lambda)|\{f(\omega+\lambda) \quad f_{jk}\}\,\mathrm{d}\lambda$$
$$O\left(n^{\quad p}|\lambda_j \quad \lambda_k|^{\quad p}\int_{\frac{\lambda_j+\lambda_k}{2}}^{2\lambda_j} |D^{(v,p)}(\lambda_j \quad \lambda)|\{f(\omega+\lambda) \quad f_{jk}\}\,\mathrm{d}\lambda\right)$$

and by the mean value theorem, for $\lambda_j + \lambda_k/2 < \lambda < 2\lambda_j$,

$$
\begin{aligned}
f(\omega+\lambda) \quad \frac{1}{2\pi}\alpha_j\bar{\alpha}_k \quad \frac{1}{2\pi}\alpha(\omega+\lambda)\bar{\alpha}(\omega+\lambda) \quad \frac{1}{2\pi}\alpha_j\bar{\alpha}_k \\
&\le \frac{1}{2\pi}\left[|\bar{\alpha}(\omega+\lambda)|\left|\alpha(\omega+\lambda) \quad \alpha_j\right| + \left|\alpha_j\right||\bar{\alpha}(\omega+\lambda) \quad \bar{\alpha}_k|\right] \\
&\quad O(\left|\lambda \quad \lambda_j\right| \sup|\bar{\alpha}(\omega+\lambda)||\alpha'(\omega+\lambda)| + \left|\alpha_j\right||\lambda \quad \lambda_k| \sup|\alpha'(\omega+\lambda)|) \\
&\quad O(\lambda_j^{\ 1 \quad 2d_1}[|\lambda \quad \lambda_j| + |\lambda \quad \lambda_k|])
\end{aligned}
$$

and (21) is

$$
\begin{aligned}
&O\left(\eta^{\ \ p}\int_{\frac{\lambda_j+\lambda_k}{2}}^{2\lambda_j}|D^{(v,p)}(\lambda_j \quad \lambda)|\lambda_j^{\ 1 \quad 2d_1}[|\lambda \quad \lambda_j| + |\lambda \quad \lambda_k|]\,\mathrm{d}\lambda\right) \\
&\quad O\left(\eta^{\ \ p}\lambda_j^{\ 1 \quad 2d_1}\left[\int_{\lambda_j}^{\lambda_j}|D^{(v,p)}(\lambda)||\lambda|d\lambda + |\lambda_j \quad \lambda_k|\int_{\lambda_j}^{\lambda_j}|D^{(v,p)}(\lambda)|\,\mathrm{d}\lambda\right]\right) \\
&\quad O(\eta^{\ \ p}\lambda_j^{\ 1 \quad 2d_1}n^{\ \ 1} + \eta^{1 \quad p}\lambda_j^{1 \quad 2d_1}n^{\ \ 1}) \\
&\quad O\left(\eta^{1 \quad p}(jk)^{\ \ 1/2}\lambda_j^{\ \ d_1}\lambda_k^{\ \ d_1}\right).
\end{aligned}
$$

Finally (20) is bounded in absolute value by

$$
\begin{aligned}
&\frac{1}{2\pi\sum|h_t^{(v,p)}|^2}\int_{\frac{\lambda_k}{2}}^{\frac{\lambda_j+\lambda_k}{2}}|D^{(v,p)}(\lambda_j \quad \lambda)||\overline{D^{(v,p)}}(\lambda_k \quad \lambda)|\{f(\omega+\lambda) \quad f_{jk}\}\,\mathrm{d}\lambda \\
&\quad O\left(n^{\ \ p}|\lambda_j \quad \lambda_k|^{\ \ p}\int_{\frac{\lambda_k}{2}}^{\frac{\lambda_j+\lambda_k}{2}}|\overline{D^{(v,p)}}(\lambda_k \quad \lambda)|\{f(\omega+\lambda) \quad f_{jk}\}\,\mathrm{d}\lambda\right)
\end{aligned}
$$

and again by the mean value theorem, for $\frac{\lambda_k}{2} < \lambda < \frac{\lambda_j+\lambda_k}{2}$,

$$
f(\omega+\lambda) \quad \frac{1}{2\pi}\alpha_j\bar{\alpha}_k \quad O(\lambda_k^{\ 1 \quad 2d_1}|\lambda \quad \lambda_j| + \lambda_k^{\ 1 \quad d_1}\lambda_j^{\ \ d_1}|\lambda \quad \lambda_k|)
$$

and (20) is

$$
\begin{aligned}
&O\left(\eta^{\ \ p}\lambda_k^{\ 1 \quad d_1}\left[\eta n^{\ \ 1}\lambda_k^{\ \ d_1}\int_{\lambda_j}^{\lambda_j}|\overline{D^{(v,p)}}(\lambda)|d\lambda + \lambda_j^{\ \ d_1}\int_{\lambda_j}^{\lambda_j}|\overline{D^{(v,p)}}(\lambda)||\lambda|\,\mathrm{d}\lambda\right]\right) \\
&\quad O(\eta^{\ \ p}\lambda_k^{\ \ d_1}k^{\ \ 1}[\eta\lambda_k^{\ \ d_1} + \lambda_j^{\ \ d_1}]) \\
&\quad O\left(\eta^{1 \quad p}(jk)^{\ \ 1/2}\lambda_j^{\ \ d_1}\lambda_k^{\ \ d_1}\left[\left(\frac{j}{k}\right)^{1/2+d_1} + \frac{1}{\eta}\left(\frac{j}{k}\right)^{1/2}\right]\right) \\
&\quad O\left(\eta^{1 \quad p}(jk)^{\ \ 1/2}\lambda_j^{\ \ d_1}\lambda_k^{\ \ d_1}\right)
\end{aligned}
$$

if $k \ge j/2$ and

$$
\begin{aligned}
&O(n^{\ \ p}|\lambda_j \quad \lambda_k|^{\ \ p}[\lambda_j^{\ \ 2d_1} + \lambda_k^{\ \ 2d_1}]) \\
&\quad O\left(\lambda_j^{\ \ d_1}\lambda_k^{\ \ d_1}\left[\left(\frac{k}{j}\right)^{d_1} + \left(\frac{k}{j}\right)^{\ \ d_1}\right](j \quad k)^{\ \ p}\right) \\
&\quad O\left(\lambda_j^{\ \ d_1}\lambda_k^{\ \ d_1}\eta^{\ \ p}\left(\frac{j}{k}\right)^{|d_1|}\right)
\end{aligned}
$$

$$O\left(\lambda_j^{\ \ d_1}\lambda_k^{\ \ d_1}\eta^{1\ \ p} j^{\ \ 1}\left(\frac{j}{k}\right)^{|d_1|}\right)$$
$$O\left(\lambda_j^{\ \ d_1}\lambda_k^{\ \ d_1}\eta^{1\ \ p}(jk)^{\ \ 1/2}\left(\frac{j}{k}\right)^{|d_1|\ \ 1/2}\right)$$
$$O\left(\lambda_j^{\ \ d_1}\lambda_k^{\ \ d_1}\eta^{1\ \ p}(jk)^{\ \ 1/2}\right) \tag{22}$$

if $k < j/2$, because $\eta^{\ \ 1} < 2j^{\ \ 1}$ and $|d_1| \quad 0.5 < 0$.

(d)

$$EW^{(v,p)}(\omega+\lambda_j)W^{(v,p)}(\omega+\lambda_k) \quad \int_{\ \pi}^{\pi}\frac{1}{2\pi\sum|h_t^{(v,p)}|^2}D^{(v,p)}(\omega+\lambda_j+\lambda)D^{(v,p)}(\omega+\lambda_k \quad \lambda)\,\mathrm{d}\lambda. \tag{23}$$

Now

$$\left|\int_{\ \pi}^{\ \omega\ \ \varepsilon}+\int_{\ \omega+\varepsilon}^{\omega\ \ \varepsilon}+\int_{\omega+\varepsilon}^{\pi}\right| \quad O(n^{1\ \ 2p}) \quad O(\vartheta_{jk}(d_1)).$$

The absolute value of the integral over $[\ \ \omega \quad \varepsilon,\ \ \omega \quad 3\lambda_j/2]$ is bounded by

$$\frac{1}{2\pi\sum|h_t^{(v,p)}|^2}\int_{\ \varepsilon}^{\ \frac{3}{2}\lambda_j}|D^{(v,p)}(\lambda_j+\lambda)||D^{(v,p)}(2\omega+\lambda_k \quad \lambda)|f(\ \ \omega+\lambda)\,\mathrm{d}\lambda$$
$$O\left(n^{1\ \ 2p}\int_{\frac{\lambda_j}{2}}^{\varepsilon}\lambda^{\ \ p\ \ 2d_1}\,\mathrm{d}\lambda\right)$$
$$O(n^{1\ \ 2p}\lambda_j^{1\ \ 2d_1\ \ p}) \quad O(\vartheta_{jk}(d_1)).$$

We obtain similarly the same bound for the integral over $[\omega+3\lambda_k/2,\ \omega+\varepsilon]$. Now the integral over $[\ \ \omega \quad 3\lambda_j/2,\ \ \omega \quad \lambda_j/2]$ is bounded in absolute value by

$$\frac{1}{2\pi\sum|h_t^{(v,p)}|^2}\int_{\ \frac{3}{2}\lambda_j}^{\ \lambda_j/2}|D^{(v,p)}(\lambda_j+\lambda)||D^{(v,p)}(2\omega+\lambda_k \quad \lambda)|f(\ \ \omega+\lambda)\mathrm{d}\lambda$$
$$O\left(n^{\ \ p}\lambda_j^{\ \ 2d_1}\int_{\ \lambda_j/2}^{\lambda_j/2}|D^{(v,p)}(\lambda)|\mathrm{d}\lambda\right)$$
$$O(n^{\ \ p}\lambda_j^{\ \ 2d_1}) \quad O(j^{\ \frac{p}{2}}k^{\ \frac{p}{2}}\lambda_j^{\ \ d_1}\lambda_k^{\ \ d_1})$$

and likewise for the interval over $[\omega+\lambda_k/2, \omega+3\lambda_k/2]$. Similarly the integral over $[\ \ \omega \quad \lambda_j/2,\ \ \omega+\lambda_j/2]$ is bounded in absolute value by

$$\frac{1}{2\pi\sum|h_t^{(v,p)}|^2}\int_{\ \frac{\lambda_j}{2}}^{\lambda_j/2}|D^{(v,p)}(\lambda_j+\lambda)||D^{(v,p)}(2\omega+\lambda_k \quad \lambda)|f(\ \ \omega+\lambda)\,\mathrm{d}\lambda$$
$$O(n^{1\ \ 2p}\lambda_j^{\ \ p}\lambda_j^{1\ \ 2d^*}) \quad O(\vartheta_{jk}(d^*))$$

and similarly for the integral over $[\omega \quad \lambda_k/2,\ \omega+\lambda_k/2]$. Finally the absolute value of the integral over $[\ \ \omega+\lambda_j/2,\ \ \omega+\varepsilon]$ is bounded by

$$
\frac{1}{2\pi\sum|h_t^{(v,p)}|^2}\int_{\frac{\lambda_j}{2}}^{\varepsilon}|D^{(v,p)}(\lambda_j+\lambda)||D^{(v,p)}(2\omega+\lambda_k\quad\lambda)|f(\quad\omega+\lambda)\,\mathrm{d}\lambda
$$

$$
O\left(n^{\quad p}\left\{\max_{\frac{\lambda_j}{2}\le\lambda\le\varepsilon+\lambda_j}\frac{f(\quad\omega+\lambda)}{\lambda^{\frac{1}{2}\quad d_2}}\right\}\int_{\frac{\lambda_j}{2}}^{\varepsilon+\lambda_j}n^{1\quad p}\lambda^{\frac{1}{2}\quad d_2\quad p}\,\mathrm{d}\lambda\right)
$$

$$
O(n^{1\quad 2p}\lambda_j^{1\quad 2d_2\quad p})\qquad O(\vartheta_{jk}(d^*))
$$

and similarly for the integral over $[\omega\quad\varepsilon,\ \omega\quad\lambda_k/2]$. □

Proof of Theorem 4. Under Gaussianity the proof follows as in Robinson (1995a, Thm 3) using now our Theorem 3 with $v\quad 2$ instead of his Theorem 2, and Assumption 4 instead of his Assumption 6. For tapered data, we have also to replace the identity I_{JM} matrix in Robinson's (1995a) eqn (17) by a positive definite squared band matrix of same dimension containing the asymptotic correlations among tapered DFT (with typical element $A_{rs}^{(2,p)}$, for $|r\quad s|<p$, with $A_{rr}^{(2,p)}\quad 1$) and set $R\quad \frac{1}{2}I_2$. Then, the bound $o(m^{\quad 1/2})$ on the r.h.s. of his equation (17) is preserved in our problem because of Assumption 4, which is similar to Robinson's Assumption 6 except the second condition that takes into account the terms depending on $\vartheta_{jk}(d)$ in parts (c) and (d) of our Theorem 3. □

Proof of Theorem 5. It follows similarly as the proof of expression (5.13) in Robinson (1995a, pp. 1070 1). Note that in the nontapered case $p\quad 1$ we only need to consider $\hat\sigma^2_{M,2,1}(0)$ as in expression (28) of Robinson (1995a), but that when $p>1$ we have to estimate the correlation between adjacent log periodogram ordinates by means of $\hat\sigma^2_{M,2,p}(k)$, $k\quad \pm1,\pm2,\ldots,\pm\ell$. The consistency of $\hat\sigma^2_{M,2,p}(k)$ for $\sigma^2_{M,2,\mathrm{p}}(k)$ follows as when $k\quad 0$ from expression (5.22) of Robinson (1995a) but using now our Theorem 3, Asumption 4, and our tapering modified DFT variance covariance matrix, cf. proof of Theorem 4. □

Proof of Theorem 6. For simplicity of exposition we focus on the estimation of d_1, that of d_2 being similar. Using a similar notation as in Theorem 1 in Robinson (1995a, 1995b) we get that $\sup_{\Theta_1}|A(H)|\quad o_p(1)$ under Assumption 6 because for $p\ge 2$

$$
E\left|\frac{I_{1j}^{(v,p)}}{g_{1j}}\right|\quad 1+O\left(\frac{n^{2(d^*\quad d_1)}}{j^{2p\quad 1+2(d^*\quad d_1)}}+\frac{1}{j}+\left(\frac{j}{n}\right)^{\alpha}\right)\tag{24}
$$

$$
E|I_{1j}^{(v,p)}\quad|\alpha_j|^2 I_{\varepsilon j}^{(v,p)}|\quad O\left(\lambda_j^{\quad 2d_1}j^{\quad 1/2}+j^{1/2\quad p}\lambda_j^{\quad(d^*+d_1)}\right),\tag{25}
$$

where $g_{1j}\quad C_1\lambda_j^{\quad 2d_1}$, and from Velasco (1999b) in his proof of Theorem 5 (p. 114)

$$
\mathrm{Var}\left\{\sum_{j\quad l+1}^{r}(2\pi I_{\varepsilon j}^{(v,p)}\quad 1)\right\}\quad O(r\quad l)\tag{26}
$$

for $v\quad 1$ and similarly for $v\quad 2$ (see also Arteche, 2000 for details in SCALM). We need to show that $\lim_{n\to\infty}P(\inf_{\Theta_2}S(d)\le 0)\quad 0$. Now

$$P(\inf_{\Theta_2} S(d) \leq 0) \leq P\left(\left|\frac{1}{m \quad l}\sum_{l+1}^{m}(a_j \quad 1)\left(\frac{I_{1j}^{(v,p)}}{g_{1j}} \quad 1\right)\right| \geq 1\right). \tag{27}$$

Using the definition of a_j in pages 1638 and 1639 in Robinson (1995b), (24) and (25) we get that $P(\inf_{\Theta_2} S(d) \leq 0) \to 0$ if

$$\left|\frac{1}{m \quad l}\sum_{l+1}^{m}(a_j \quad 1)(2\pi I_{\varepsilon j}^{(v,p)} \quad 1)\right| \quad o_p(1),$$

which is proved in p 116 of Velasco (1999b) for $v \quad 1$ and the proof for $v \quad 2$ is similar and easier because of the limited correlation between $I_{\varepsilon j}^{(2,p)}$ and $I_{\varepsilon k}^{(2,p)}$. □

Proof of Theorem 7. The proof is an adaptation of Theorem 2 in Robinson (1995b). The main difference is in the proof of

$$\sum_{r \quad l+1}^{m}\left(\frac{r}{m}\right)^{1 \quad 2\delta}\frac{1}{r^2}\left|\sum_{j \quad l+1}^{r}\left(\frac{I_{1j}^{(v,p)}}{g_{1j}} \quad 1\right)\right| + \frac{1}{m}\left|\sum_{j \quad l+1}^{m}\left(\frac{I_{1j}^{(v,p)}}{g_{1j}} \quad 1\right)\right| \quad o_p((\log m)^{\quad 6})$$

By Lemma 2 and (26) we get the desired result under Assumption 8. This result guarantees that $d^2R_1^p(\bar{d}_1)/d^2d \xrightarrow{p} 4$ for $|\bar{d}_1 \quad d_1| \leq |\tilde{d}_1 \quad d_1|$. Next we have

$$\frac{dR_1^p(d_1)}{\mathrm{d}d} \quad \frac{2}{\sqrt{m}}\sum_{l+1}^{m} v_j\left(\frac{I_{1j}^{(v,p)}}{g_{1j}} \quad 2\pi I_{\varepsilon j}^{(v,p)}\right)(1 + o_p(1)) \tag{28}$$

$$+\frac{2}{\sqrt{m}}\sum_{l+1}^{m} v_j(2\pi I_{\varepsilon j}^{(v,p)} \quad 1)(1 + o_p(1)), \tag{29}$$

where $v_j \quad \log j \quad \sum_{l+1}^{m} \log k/(m \quad l)$. By Lemma 2 (28) is $o_p(1)$ under Assumption 8. Next

$$\frac{2}{\sqrt{m}}\sum_{l+1}^{m} v_j(2\pi I_{\varepsilon j}^{(v,p)} \quad 1) \quad \frac{2}{\sqrt{m}}\sum_{1}^{m} v_j(2\pi I_{\varepsilon j}^{(v,p)} \quad 1) \quad \frac{2}{\sqrt{m}}\sum_{1}^{l} v_j(2\pi I_{\varepsilon j}^{(v,p)} \quad 1),$$

which converges to $N(0, 4\Phi)$ by Lemma 6 of Velasco (1999b) and the fact that

$$\frac{2}{\sqrt{m}}\sum_{1}^{l} v_j(2\pi I_{\varepsilon j}^{(v,p)} \quad 1) \quad O_p(l^{1/2}m^{\quad 1/2}) \quad o_p(1).$$

□

Lemma 1. *Under Assumptions* 1 *and* 2, *for* $p \geq 2$ *and integer* j *such that* $j/n \to 0$ *as* $n \to \infty$,

$$\int_{\quad \pi}^{\pi}\left|\frac{\alpha(\lambda)}{\alpha_j} \quad 1\right|^2 K_n^{(v,p)}(\lambda \quad \lambda_j \quad \omega)\mathrm{d}\lambda \quad O\left(\frac{1}{j^2} + j^{1 \quad 2p}\left(\frac{n}{j}\right)^{2(d^* \quad d_1)}\right).$$

Proof. The proof generalizes that of Lemma 3 in Robinson (1995b), considering the same intervals of integration (around ω) and the bound of

$$|K_n^{(v,p)}(\lambda)| \leq \text{const.} \times \min(n, n^{1 \quad 2p}|\lambda|^{\quad 2p}).$$

As in the proof of Lemma 3 in Robinson (1995b) we get the bounds

$$\left|\int_{\pi}^{\omega \quad \delta}\right| + \left|\int_{\omega+\delta}^{\pi}\right| \quad O(j^{1 \quad 2p}).$$

Now $\left|\int_{\omega \quad \delta}^{\omega \quad \lambda_j/2}\right|$ is bounded by

$$O\left(f_j^{\quad 1}\left\{\left(\max_{\frac{\lambda_j}{2}\le\lambda\le\delta} \frac{f(\omega \quad \lambda)}{\lambda^{\frac{1}{2} \quad d_2}}\right) + |\bar{\alpha}_j|\left(\max_{\frac{\lambda_j}{2}\le\lambda\le\delta} \frac{\alpha(\omega \quad \lambda)}{\lambda^{\frac{1}{2} \quad d_2}}\right)\right\}\int_{\frac{\lambda_j}{2}}^{\delta} \lambda^{\frac{1}{2} \quad d_2} K_n^{(v,p)}(\quad \lambda \quad \lambda_j)\mathrm{d}\lambda\right]$$

$$+ \int_{\frac{\lambda_j}{2}}^{\delta} K_n^{(v,p)}(\quad \lambda \quad \lambda_j)\mathrm{d}\lambda\Bigg) \quad O\left(j^{1 \quad 2p}\left(\frac{n}{j}\right)^{2(d^* \quad d_1)}\right)$$

and similarly the integral over $(\omega + 2\lambda_j, \omega + \delta)$. The integral over $(\omega \pm \lambda_j/2)$ is

$$O\left(f_j^{\quad 1} n^{1 \quad 2p}\lambda_j^{\quad 2p}\int_{\frac{\lambda_j}{2}}^{\frac{\lambda_j}{2}} f(\lambda)\mathrm{d}\lambda\right) \quad O\left(j^{1 \quad 2p}\left(\frac{n}{j}\right)^{2(d^* \quad d_1)}\right)$$

and finally the integral over $(\omega + \lambda_j/2, \omega + 2\lambda_j)$ is, by the mean value theorem

$$O\left(\frac{1}{|\alpha_j|^2}\left|\max_{\frac{\lambda_j}{2}\le\lambda\le 2\lambda_j} \alpha'(\omega+\lambda)\right|^2\left[\int_0^{n \quad 1} \lambda^2 K_n^{(v,p)}(\lambda)\mathrm{d}\lambda + \int_{n \quad 1}^{\lambda_j} \lambda^2 K_n^{(v,p)}(\lambda)\mathrm{d}\lambda\right]\right)$$

$$O\left(\lambda_j^{\quad 2}\left[\int_0^{n \quad 1} \lambda^2 n \mathrm{d}\lambda + n^{1 \quad 2p}\int_{n \quad 1}^{\lambda_j} \lambda^{2 \quad 2p}\mathrm{d}\lambda\right]\right) \quad O(j^{\quad 2}). \qquad \square$$

Lemma 2. *Let* $0 < l < r \le m$ *such that* $n^{2(d^* \quad d_1)} l^{1 \quad 2p \quad 2(d^* \quad d_1)} \to 0$ *as* $n \to \infty$. *Under Assumptions* 1, 2 *and* 5 *and* $p \ge 2$

$$\sum_{j \quad l+1}^{r}\left(\frac{I_{1j}^{(v,p)}}{g_{1j}} \quad 2\pi I_{\varepsilon j}^{(v,p)}\right) \quad O_p\left(\frac{r^{\alpha+1}}{n^{\alpha}} + l + \log r + \left(\frac{r}{l}\right)^{1/4}(\log r)^{1/2}\right).$$

Proof. Is based on the proof of (4.8) in Robinson (1995b) in p. 1648 1651. Using similar notation (noting that in our case the DFT are of tapered data) we need to get the bound of

$$E\left\{\sum_{j \quad l+1}^{r}\left(\frac{I_{1j}^{(v,p)}}{f_j} \quad 2\pi I_{\varepsilon j}^{(v,p)}\right)\right\}^2 \quad a_1 + a_2 + b_1 + b_2,$$

where a_1, a_2, b_1, b_2 are defined as in p. 1648 of Robinson (1995b) with $u_j \quad \sqrt{2\pi}W_j^{(v,p)}|\alpha_j|^{\quad 1}$ and $v_j \quad \sqrt{2\pi}W_{\varepsilon j}^{(v,p)}$. Using Theorem 2 we get that

$$a_1 \quad O\left(\sum_{j \quad l+1}^{r}\left[\frac{1}{j} + \frac{n^{2(d^* \quad d_1)}}{j^{2p \quad 1+2(d^* \quad d_1)}}\right]\right) \quad O(\log r + l).$$

We similarly get the following bound for b_1

$$b_1 \quad O((\log r)^2 + l^2 + l^{\quad 1/2} r^{1/2}\log r)$$

taking into account that the $(A_{jk}^{(v,p)})^2$ terms cancel out, $\eta^{\ a} \quad O(1)$,

$$\sum_{l+1}^{r}\sum_{j>k}\frac{\eta^{\ a}}{(jk)^{1/2}} \quad O\left(\frac{1}{l^{1/2}}\sum_{l+1}^{r}k^{\ 1/2}\sum_{j>k}\eta^{\ a}\right)$$
$$O\left(\frac{1}{l^{1/2}}\sum_{l+1}^{r}k^{\ 1/2}\log r\right) \quad O\left(\left(\frac{r}{l}\right)^{1/2}\log r\right)$$

and similarly

$$\sum_{l+1}^{r}\sum_{j>k}\frac{\eta^{\ a}}{jk} \quad O\left(\frac{1}{l}\sum_{l+1}^{r}\frac{1}{k}\sum_{j>k}\eta^{\ a}\right) \quad O\left(\frac{(\log r)^2}{l}\right)$$

for $a > 1$. Finally, the terms involving cumulants are

$$a_2 \quad O(l),$$
$$b_2 \quad O(l^2 + l\log r + n^{\ 1/2}(\log r)^2),$$

using Lemma 1 as in Robinson (1995b). □

ACKNOWLEDGEMENTS

We are grateful to the Associate Editor and the referee for valuable comments. The first author acknowledges financial support from the University of the Basque Country 9/UPV 00038.321 13503/2001 and from the Spanish Ministerio de Ciencia y Tecnología and FEDER BEC2003 02028, and the second from the Spanish Dirección General de Enseñanza Superior (DGES), reference number BEC2001 1270 and the Spanish Ministerio de Educación y cipucia, ref. SEJ2004 04583/ECON.

NOTE

Corresponding author: J. Arteche, Departmento de Econometría y Estadística, University of the Basque Country, Avda. Lehendakari Agirre 83, 48015 Bilbao, Spain. Tel.: +34 94 6013852; Fax: +34 94 6013754; E mail: etpargoj@bs.ehu.es

REFERENCES

Arteche, J. (2000) Gaussian semiparametric estimation in seasonal or cyclical long memory time series. *Kybernetika* 36, 279 310.

Arteche, J. (2002) Semiparametric robust tests on seasonal or cyclical long memory time series. *Journal of Time Series Analysis* 23, 251 85.

Arteche, J. and Robinson, P. M. (1999) Seasonal and cyclical long memory. In *Asymptotics, Nonparametrics and Time Series* (ed. S. Ghosh). New York: Marcel Dekker, Inc., pp. 115 48.

Arteche, J. and Robinson, P. M. (2000) Semiparametric inference in seasonal and cyclical long memory processes. *Journal of Time Series Analysis* 21, 1 25.
Chung, C. F. (1996). Estimating a generalized long memory process. *Journal of Econometrics* 73, 237 59.
Geweke, J. and Porter-Hudak, S. (1983) The estimation and application of long-memory time series models. *Journal of Time Series Analysis* 4, 221 38.
Gray, H. L., Zhang, N. F. and Woodward, W. A. (1989) On generalized fractional processes. *Journal of Time Series Analysis* 10, 233 57.
Hannan, E. J. (1973) Central limit theorems for time series regression. *Z. Wahrscheinlichkeitstheorie und Verw Gebiete* 26, 157 70.
Hidalgo, J. and Soulier, P. (2004) Estimation of the location and exponent of the spectral singularity of a long memory process. *Journal of Time Series Analysis* 25, 173 197.
Hurvich, C. M. and Beltrao, K. I. (1993) Asymptotics for the low-frequency ordinates of the periodogram of long-memory time series. *Journal of Time Series Analysis* 14, 455 72.
Hurvich, C. M., Deo, R. and Brodsky, J. (1998). The mean square error of Geweke and Porter-Hudak's Estimator of the memory parameter of a long memory time series. *Journal of Time Series Analysis* 19, 19 46.
Hurvich, C. M. and Chen, W. W. (2000) An efficient taper for potentially overdifferenced long-memory time series. *Journal of Time Series Analysis* 21, 155 80.
Hurvich, C. M., Moulines, E. and Soulier, P. (2002) The FEXP estimator for potentialy non-stationary linear time series. *Stochastic Processes and their Applications* 97, 307 40.
Hurvich, C. M. and Ray, B. K. (1995) Estimation of the memory parameter for nonstationary or noninvertible fractionally integrated process. *Journal of Time Series Analysis* 16, 17 42.
Künsch, H. R. (1986) Discrimination between monotonic trends and long range dependence. *Journal of Applied Probability* 23, 1025 30.
Lildholdt, P. M. (2002) Sources of seasonal fractional integration in macroeconomic time series. CAF Working Paper 125.
Oppenheim, G., Ould Haye, M. and Viano, M. C. (2000) Long memory with seasonal effects. *Statistical Inference for Stochastic Processes* 3, 53 68.
Ould Haye, M. and Viano, M. C. (2003) Limit theorems under seasonal long-memory. In *Theory and Applications of Long-Range Dependence* (eds P. Doukhan, G. Oppenheim and M. S. Taqqu). Boston: Birkhäuser, pp. 101 10.
Robinson, P. M. (1995a) Log-periodogram regression of time series with long-range dependence. *Annals of Statistics* 23, 1048 72.
Robinson, P. M. (1995b) Gaussian semiparametric estimation of long-range dependence. *Annals of Statistics* 23, 1630 61.
Temple, G. (1971) *The Structure of Lebesgue Integration Theory*. Oxford: Oxford University Press.
Velasco, C. (1999a) Non-stationary log-periodogram regression. *Journal of Econometrics* 91, 325 71.
Velasco, C. (1999b) Gaussian semiparametric estimation of non-stationary time series. *Journal of Time Series Analysis* 20, 87 127.
Yajima, Y. (1989) A central limit theorem of Fourier transforms of strongly dependent stationary processes. *Journal of Time Series Analysis* 10, 375 83.
Zhurbenko, I. G. (1979) On the efficiency of estimates of a spectral density. *Scandinavian Journal of Statistics* 6, 49 56.