GAUSSIAN ESTIMATION OF PARAMETRIC SPECTRAL DENSITY WITH UNKNOWN POLE[*]

by

L Giraitis, J Hidalgo and P M Robinson
London School of Economics and Political Science

Contents:

Abstract

We consider a parametric spectral density with power-law behaviour about a fractional pole at the unknown frequency ω. The case of known ω, especially $\omega = 0$, is standard in the long memory literature. When ω is unknown, asymptotic distribution theory for estimates of parameters, including the (long) memory parameter, is significantly harder. We study a form of Gaussian estimate. We establish n- consistency of the estimate of ω, and discuss its (non-standard) limiting distributional behaviour. For the remaining parameter estimates, we establish $\sqrt{n}$ - consistency and asymptotic normality.

Keywords: long range dependence; unknown pole

JEL classification: C22.

Contact address: Professor P M Robinson, Department of Economics, London School of Economics & Political Science, Houghton Street, London WC2A 2AE.
E-mail: pm.robinson@lse.ac.uk/

1   Introduction

Cyclic behaviour in covariance stationary time series is typically manifested in a pronounced peakedness in spectral density estimates. If the spectral density of a series x_t, $t = 0, \pm1, \pm2, \ldots$, is estimated over the Nyqvist band $(-\pi, \pi]$, a peak at frequency $\omega \in (0, \pi]$ corresponds to a cycle of $2\pi/\omega$. Various statistical models for such a phenomenon have been proposed. Two long-standing ones are of autoregressive (AR) and cosinusoidal type. For example the $AR(2)$ model

$$(1 - \phi_1 L - \phi_2 L^2)x_t = \varepsilon_t, \quad t = 0, \pm1, \ldots, \tag{1.1}$$

where L is the lag operator and ε_t is a sequence of uncorrelated and homoscedastic zero-mean random variables, has spectrum with peak at

$$\omega = \text{arc } \cos(\frac{\phi_1(\phi_2 - 1)}{4\phi_2})$$

when the zeroes of $1 - \phi_1 z - \phi_2 z^2$ are complex and $|\phi_1(\phi_2 - 1)/4\phi_2| < 1$. On the other hand, a very simple cosinusoidal model is

$$x_t = \alpha \cos(\omega t) + \beta \sin(\omega t), \tag{1.2}$$

where α and β are uncorrelated random variables with zero means and the same variance. Whereas (1.1) implies that x_t has a spectral density that is analytic even at a peak, x_t given by (1.2) has a spectral distribution function that jumps at ω. The model (1.2) cannot describe real data, unlike the nonstationary-in-the-mean modification

$$x_t = \alpha \cos(\omega t) + \beta \sin(\omega t) + u_t, \tag{1.3}$$

where α and β are interpreted as fixed unknown constants and u_t is an unobservable covariance stationary process with smooth spectral density. However, (1.1) and (1.2) constitute mathematically radically different descriptions of cyclic behaviour within the stationary class, and help to motivate and place in perspective the model studied in the current paper.

An intermediate possibility between (1.1) and (1.2) is that x_t has a spectral density with a pole at ω. Denote by $f(\lambda)$ the spectral density of x_t, satisfying

$$\gamma_j \overset{def}{=} E(x_0 x_j) = \int_{-\pi}^{\pi} f(\lambda) \cos(j\lambda)d\lambda, \quad j = 0, \pm1, \ldots.$$

We are implicitly assuming here, as in (1.1) and (1.2), that x_t has zero mean, but we will later indicate how our results apply in case x_t has unknown mean. We say that $f(\lambda)$ has a (fractional) pole of order $\alpha \in (0, 1)$ at ω if

$$f(\lambda) \sim C|\lambda - \omega|^{-\alpha} \quad \text{as} \quad \lambda \to \omega, \tag{1.4}$$

where $C \in (0, \infty)$ and $' \sim'$ indicates that the ratio of left and right hand sides tends to 1. In case $\omega = 0$, (1.4) is a familiar description of long-memory time series, so that for

$\omega \in (0, \pi)$ (1.4) can be said to denote long-memory at a non-zero frequency. Parametric models for $f(\lambda)$, specifying f for all λ as a given function of unknown parameters and λ, and satisfying (1.4) for some $\omega \in (0, \pi)$, have been proposed by Hosking (1981), Andel (1986), Gray et al. (1989) and others. The simplest of these modifies (1.1) to

$$a(L; \omega, d)x_t = \varepsilon_t, \tag{1.5}$$

where

$$a(z; \omega, d) = (1 - 2z\cos\omega + z^2)^d, \tag{1.6}$$

and was termed a Gegenbauer model by Gray et al. (1989). Writing $\sigma^2 = Var(\varepsilon_t)$, we deduce that, for $\lambda \in (-\pi, \pi]$,

$$f(\lambda) = \frac{\sigma^2}{2\pi}|a(e^{i\lambda}; \omega, d)|^{-2} = \frac{\sigma^2}{2\pi}\left|4\sin(\frac{\lambda+\omega}{2})\sin(\frac{\lambda-\omega}{2})\right|^{-2d}. \tag{1.7}$$

When $d > 0$, $f(\lambda)$ has a pole at $\lambda = \omega$ (and, when $\omega \neq 0$, at $-\omega$, as anticipated from symmetry). Moreover, when $\omega \in (0, \pi)$, (1.7) satisfies (1.4) with $\alpha = 2d$ and $C = (\sigma^2/2\pi)|2\sin\omega|^{-2d}$. When $\omega = 0$ or π, (1.7) satisfies (1.4) with $\alpha = 4d$ and $C = \sigma^2/2\pi$, noting in case $\omega = \pi$ that $|\sin\frac{1}{2}(\lambda+\omega)| = |\sin\frac{1}{2}(\lambda-\omega)|$. Correspondingly, x_t is covariance stationary for $d < 1/2$ when $\omega \in (0, \pi)$ and for $d < 1/4$ when $\omega = 0$ or π. When $\omega = 0$, x_t is a standard long-memory model, ARIMA $(0,2d,0)$, though in this case the usual notation replaces $2d$ by d. Note that a zero in $f(\lambda)$ occurs in (1.4) when $\alpha < 0$ or in (1.7) when $d < 0$, but we do not pursue this case. Hosking (1981), Gray et al. (1989) extended (1.5) by replacing ε_t by a covariance stationary and invertible autoregressive moving average (ARMA) process, when (1.4) is satisfied in the same way. Robinson (1994), Giraitis and Leipus (1995) considered more general models, providing spectral poles at several frequencies. They are motivated in part by seasonal processes, for which pole location is known. Hosoya (1997) established $\sqrt{n}$-consistency and asymptotic normality of Gaussian estimates of the remaining parameters in such models, extending work for the case of a known pole at zero frequency by Fox and Taqqu (1986), Dahlhaus (1989) and Giraitis and Surgailis (1990). In applications of non-seasonal, single-pole (in $(0, \pi]$) models such as (1.5), however, it is likely that ω is unknown, as when investigating the length of a cycle in geophysical or macroeconomic time series.

The present paper proposes estimates based on a Gaussian objective function in case of parametric models satisfying (1.4) with unknown ω and studies their asymptotic statistical properties. The theory seems significantly harder than for the case of known ω. The following section sets down the basic model with examples, and describes the parameter estimates, which are prompted by an approximate Gaussian likelihood, though we nowhere assume Gaussianity. In Section 3 the regularity conditions and limiting behaviour of the parameter estimates are presented, with discussion, along with the main steps of the proof and a small numerical example, the remaining details appearing in Sections 4 - 7. Our main findings are that the estimates of ω are n-consistent (like those proposed by Hannan (1973a) for (1.3)), while the estimates of the remaining parameters are $\sqrt{n}$-consistent and have the same normal distribution as when ω is known.

2 Model and parameter estimates

We parameterize $f(\lambda)$ in the first place by writing

$$f(\lambda) = \frac{\sigma_0^2}{2\pi} k(\lambda; \theta_0, \omega_0), \quad -\pi < \lambda \leq \pi, \tag{2.1}$$

where σ_0^2 is an unknown positive scalar, θ_0 is an unknown $p-$dimensional column vector, assumed to be in a compact set $\Theta \subset \mathbf{R}^p$, and $k(\lambda; \theta, \omega)$ is a known function of λ, θ, ω, such that, for $\theta \in \Theta$, $\omega \in \Pi = [0, \pi]$,

$$k(\lambda; \theta, \omega) > 0, \quad -\pi < \lambda \leq \pi,$$

$$\int_{-\pi}^{\pi} \log k(\lambda; \theta, \omega) d\lambda = 0. \tag{2.2}$$

The zero subscripted quantities in (2.1) denote true values, θ and ω denoting any admissible values (so that α_0, ω_0, should replace α, ω in (1.4), for example). Following the discussion of Section 1, we subdivide θ as $\theta = (\tau', \alpha)'$ and correspondingly write $\Theta = \Theta_\tau \times \Theta_\alpha$ for compact sets Θ_τ, Θ_α, $\Theta_\alpha \subset [0, 1]$. The $(p - 1)$-vector τ is empty when $p = 1$, in which case a particular $k(\lambda; \theta, \omega)$ can be deduced from (1.7). For $p > 1$, τ essentially describes short-range dependence, for example it can contain the coefficients of a stationary and invertible ARMA, so that we have the more general Gegenbauer model of Gray et al. (1989), with

$$k(\lambda; \theta, \omega) = \left| \frac{a(e^{i\lambda}; \omega, \alpha/2) b(e^{i\lambda}; \tau)}{c(e^{i\lambda}; \tau)} \right|^{-2}, \quad -\pi < \lambda \leq \pi, \tag{2.3}$$

where a is given by (1.6), and with τ_j the j th element of τ,

$$b(z; \tau) = 1 - \sum_{j=1}^{p_1} \tau_j z^j 1_{\{1 \leq p_1 \leq p-1\}}, \quad c(z; \tau) = 1 - \sum_{j=p_1+1}^{p-1} \tau_j z^{j-p_1} 1_{\{0 \leq p_1 \leq p-2\}}, \tag{2.4}$$

where $1_{\{.\}}$ is the indicator function, $0 \leq p_1 \leq p - 1$, all zeroes of $b(z; \tau)$ and of $c(z; \tau)$ are outside the unit circle, and $b(z; \tau)$ and $c(z; \tau)$ have no zeroes in common. More generally, we can consider models

$$k(\lambda; \theta, \omega) = \left| \frac{h(\lambda; \tau)}{a(e^{i\lambda}; \omega, \alpha/2)} \right|^2 \quad -\pi < \lambda \leq \pi, \tag{2.5}$$

where $h(\lambda; \tau)$ is bounded and bounded away from zero. Condition (2.2) indicates that for a process with spectral density $(\sigma^2/2\pi) k(\lambda; \theta, \omega)$, the free parameter σ^2 (functionally independent of θ and ω) is the variance of the one-step-ahead best linear predictor (see Hannan, 1970, pp 157-163). In view of (2.2), which is satisfied by (2.3) and (2.5), we might consider, following Hannan (1973b), estimating $(\theta_0', \omega_0)'$ by

$$\begin{pmatrix} \tilde{\theta} \\ \tilde{\omega} \end{pmatrix} = \arg \min_{\Theta \times \Pi} S(\theta, \omega),$$

3

where

$$S(\theta, \omega) = \tilde{n}^{-1} \sum_{j=0}^{\tilde{n}} \frac{I(\lambda_j)}{k(\lambda_j; \theta, \omega)}, \qquad (2.6)$$

$\tilde{n} = [n/2]$ and we introduce the periodogram

$$I(\lambda_j) = (2\pi n)^{-1} \left| \sum_{t=1}^{n} x_t e^{it\lambda_j} \right|^2,$$

evaluated at the Fourier frequencies $\lambda_j = 2\pi j/n$.

Under the conditions of Hannan (1973b), the objective function $S(\theta, \omega)$ approximates a Gaussian log-likelihood in the sense that $(\tilde{\theta}', \tilde{\omega})'$ has the same limit distribution as a Gaussian maximum likelihood estimate (though Hannan did not assume Gaussianity). However, Hannan's (1973b) conditions are not all satisfied due to the pole ω in $k(\lambda; \theta, \omega)$. In case the true pole ω_0 is known (for example it is taken for granted that $\omega_0 = 0$, as in standard long-memory models) then it is already known that the conditions of Hannan (1973b) can be suitably relaxed, as shown by Fox and Taqqu (1986), Dahlhaus (1989), Giraitis and Surgailis (1990) when $\omega_0 = 0$, and by Hosoya (1997) when $\omega_0 \in (0, \pi]$, (though strictly these authors consider different approximations to the log-likelihood function from (2.6)). Indeed the consistency proof of Hannan (without rates of convergence) still holds in case of known or unknown pole, the latter case having been considered by Giraitis and Leipus (1995). For limit distribution theory with a known pole ω_0, the smoothness conditions of Hannan can be relaxed due to the "compensation" to $I(\lambda)$ afforded by the reciprocal of $k(\lambda; \theta_0, \omega_0)$ near $\lambda = \omega_0$. When ω_0 is unknown, derivation of limit distribution theory, and even rates of convergence, is significantly more difficult. An attempt at this was made, in the context of an alternative type of approximation to the Gaussian likelihood, by Chung (1996a, b); we shall briefly comment on Chung's treatment subsequently. Although $k(\lambda; \theta, \omega)^{-1} = 0$ at $\lambda = \omega$, the derivative $(\partial/\partial\omega)k^{-1}(\lambda; \theta, \omega)$ is not well behaved near $\lambda = \omega$ in case of (2.3) and (2.5). For this reason we have chosen to study instead the estimate

$$\begin{pmatrix} \hat{\theta} \\ \hat{\omega} \end{pmatrix} = \arg \min_{\Theta \times Q} S(\theta, \lambda_q)$$

where $Q = Q_n = \{q : q = 0, 1, \ldots, \tilde{n}\}$. Thus minimization with respect to ω is now carried out over a discrete set. We could in fact consider a finer set, with spacing δ/n for arbitrary $\delta > 0$, but it is essential that the intervals not be $o(n^{-1})$ as $n \to \infty$. We can estimate σ_0^2 by

$$\hat{\sigma}^2 = S(\hat{\theta}, \hat{\omega}).$$

3   Regularity conditions and asymptotic properties

We introduce first the following assumptions, denoting by ∇_β the vector of partial derivatives $\partial/\partial\beta$, for a column vector or scalar β.

4

ASSUMPTION A.1. x_t has spectral density $f(\lambda)$ given by (2.1), (2.5), that is

$$f(\lambda) = \frac{\sigma_0^2}{2\pi} \left| 4 \sin(\frac{\lambda + \omega_0}{2}) \sin(\frac{\lambda - \omega_0}{2}) \right|^{-\alpha_0} |h(\lambda; \tau_0)|^2, \quad -\pi < \lambda \le \pi, \qquad (3.1)$$

where τ_0 denotes the true value of τ, and for $-\pi < \lambda \le \pi$, $\tau \in \Theta_\tau$, $h(\lambda; \tau)$ is even in λ and bounded away from zero, and the derivatives $\nabla_\tau h(\lambda; \tau), \nabla_\lambda h(\lambda; \tau), \nabla_\tau \nabla_\lambda h(\lambda; \tau), \nabla_\tau \nabla'_\tau h(\lambda; \tau)$ are continuous.

ASSUMPTION A.2. (2.2) holds for all $\theta \in \Theta, \omega \in [0, \pi]$.

ASSUMPTION A.3.

$$\inf_{(\theta', \omega) \in \Theta \times \Pi} (2\pi)^{-1} \int_{-\pi}^{\pi} \frac{k(\lambda; \theta_0, \omega_0)}{k(\lambda; \theta, \omega)} d\lambda = 1 \qquad (3.2)$$

and the set

$$\{\lambda : k(\lambda; \theta, \omega) \ne k(\lambda; \theta_0, \omega_0)\}, \quad (\theta', \omega) \ne (\theta_0', \omega_0) \qquad (3.3)$$

has positive Lebesgue measure. Also, the matrix

$$\Omega = (w_{ij})_{i,j=1,\dots,p} = \frac{1}{4\pi} \int_{-\pi}^{\pi} \nabla_\theta \log k(\lambda; \theta_0, \omega_0) \nabla'_\theta \log k(\lambda; \theta_0, \omega_0) d\lambda \qquad (3.4)$$

is positive definite.

ASSUMPTION A.4. θ_0 is an interior point of Θ and $\omega_0 \in [0, \pi]$, such that $0 < \alpha_0 < 1$ for $0 < \omega_0 < \pi$ and $0 < \alpha_0 < 1/2$ for $\omega_0 = 0, \pi$.

ASSUMPTION A.5. We have

$$x_t = \sum_{j=0}^{\infty} \phi_j \varepsilon_{t-j}, \quad \sum_{j=0}^{\infty} \phi_j^2 < \infty, \quad \phi_0 = 1, \qquad (3.5)$$

where $\{\varepsilon_t\}$ is ergodic and

$$E[\varepsilon_t | F_{t-1}] = 0, \quad E[\varepsilon_t^2 | F_{t-1}] = \sigma_0^2, \quad a.s.,$$

$$E[\varepsilon_t^i | F_{t-1}] = \mu_i \quad a.s., \quad i = 3, 4,$$

such that μ_3 and μ_4 are non-stochastic, and F_t is the σ- algebra generated by $\varepsilon_s, s \le t$; also, for some $\eta > 0$,

$$\max_t E|\varepsilon_t|^{4+\eta} < \infty. \qquad (3.6)$$

ASSUMPTION A.6. Uniformly in $\lambda \in [0, \pi] \backslash \{\omega_0\}$ the function $\phi(\lambda) := \sum_{j=0}^{\infty} \phi_j e^{ij\lambda}$ has the property

$$|(d/d\lambda)\phi(\lambda)| = O(|\phi(\lambda)||\lambda - \omega_0|^{-1}).$$

Assumption A.1 covers a wide range of short memory spectral densities h, including both invertible ARMA ones and Bloomfield (1973) ones, though it also permits the modelling of processes with autocovariances that decay much more slowly than exponentially.

5

We might call (3.1) a 'generalized' Gegenbauer model. We have already discussed Assumption A.2, but add that in view of (1.6) and (2.5), it is equivalent to

$$\int_{-\pi}^{\pi} \log h(\lambda; \tau) d\lambda = 0, \quad \tau \in \Theta_\tau.$$

Assumption A.3 is an identifiability condition; note that in view of the structure of a, it will be implied if the set $\{\lambda : h(\lambda; \tau) \neq h(\lambda; \tau_0)\}, \tau \neq \tau_0$ has positive Lebesgue measure and the matrix

$$\frac{1}{4\pi} \int_{-\pi}^{\pi} \nabla_\tau \log h(\lambda; \tau_0) \nabla'_\tau \log h(\lambda; \tau_0) d\lambda$$

is positive definite. These conditions are satisfied in the ARMA case $h(\lambda; \tau) = |c(e^{i\lambda}; \tau)/b(e^{i\lambda}; \tau)|^2$, where $b(z; \tau)$ and $c(z; \tau)$ (see (2.4)) have no zeroes in common. Note that Assumption A.4 entails $\alpha_0 > 0$, which is essential for Assumption A.3 to hold, because ω_0 is not identifiable when $\alpha_0 = 0$. In Assumption A.5, the normalization $\phi_0 = 1$ is consistent with Assumption A.2. Assumption A.5 is similar to ones used by Hannan (1973b) in Whittle estimation for short memory series, and on the other hand by Robinson (1995b) in narrow-band semiparametric Whittle estimation of long memory. However, Hannan's assumptions only extend to second moments in order to establish the Central Limit Theorem for estimators of τ_0 in case $\alpha_0 = 0$, so that $f(\lambda) = (\sigma_0^2/2\pi)h(\lambda; \tau_0)$, whereas Robinson allowed $\eta = 0$ in (3.6). Assumption A.6 is like one used by Robinson (1995b).

We also establish our results over a slightly different class than provided by Assumptions A.1, A.4:

ASSUMPTION A.1′. x_t has spectral density given by (2.1) where

$$k(\lambda; \theta, \omega) = \begin{cases} |\lambda - \omega|^{-\alpha} g(\lambda; \theta, \omega) & \text{if} \quad 0 \leq \lambda \leq \pi, \\ |\lambda + \omega|^{-\alpha} g(\lambda; \theta, \omega) & \text{if} \quad -\pi < \lambda < 0, \end{cases} \tag{3.7}$$

such that for $-\pi < \lambda \leq \pi$, $\theta \in \Theta$, $\omega \in \Pi$, $g(\lambda; \theta, \omega)$ is even in λ and bounded and bounded away from zero, the derivatives $\nabla_\theta g(\lambda; \theta, \omega), \nabla_\theta \nabla'_\theta g(\lambda; \theta, \omega)$ are continuous and bounded, $g(\lambda; \theta, \omega)$ and $\nabla_\theta g(\lambda; \theta, \omega)$ satisfy uniform Lipschitz conditions in ω of order greater than $1/2$, while, for $0 < |\lambda| < \pi, \theta \in \Theta$, $\omega \in \Pi$, $\nabla_\lambda g(\lambda; \theta, \omega)$ and $\nabla_\lambda \nabla_\theta g(\lambda; \theta, \omega)$ are bounded.

We correspondingly replace Assumption A.4 by

ASSUMPTION A.4′. θ_0 is an interior point of Θ such that $0 < \alpha_0 < 1$, and $\omega_0 \in [0, \pi]$.

Note that for $\omega \neq 0, \pi$, we can write the generalized Gegenbauer spectrum (2.5) in the form (3.7) by taking

$$g(\lambda; \theta, \omega) = \left\{ \frac{|\lambda - \omega|^\alpha}{|a(e^{i\lambda}; \omega, \alpha/2)|^2} \right\} |h(\lambda; \tau)|^2, \quad \lambda \in \Pi,$$

with the corresponding expression for $-\pi < \lambda < 0$, where the factor in braces tends to $(2 \sin \omega)^{-\alpha}$ as $\lambda \to \omega$, and is continuous but not differentiable at $\lambda = 0, \pi$ because of the inevitable evenness and periodicity of spectral densities, explaining our avoidance of

6

differentiability in λ at these frequencies. In other words, if $g(\lambda; \theta, \omega)$ were differentiable at $\lambda = 0, \pi$, then $k(\lambda; \theta, \omega)$ would not be, and the Gegenbauer model (2.3), for example, would be excluded. On the other hand, Assumption A.1' also includes models for which $g(\lambda; \theta, \omega)$ is everywhere smooth, such as when

$$k(\lambda; \theta, \omega) = c(\theta, \omega)|\lambda - \omega|^{-\alpha}, \quad \lambda \in \Pi, \tag{3.8}$$

with $c(\theta, \omega) = \exp\{(\alpha/\pi)[\omega(\log \omega - 1) + (\pi - \omega)(\log(\pi - \omega) - 1)]\}$, so that Assumption A.2 is satisfied. Moreover, under Assumption A.1' we constrain α to $(0, 1)$, not $(0, 1/2)$ at $\omega = 0, \pi$, so we avoid this type of discontinuity. Notice that from (1.7), $|a(e^{i\lambda}; \omega, \alpha/2)|^{-2} \sim (2\sin\omega)^{-\alpha}|\lambda - \omega|^{-\alpha}$ as $\lambda \to \omega$ for $\omega \neq 0, \pi$, so that $|a(e^{i\lambda}; \omega, \alpha/2)|^{-2}$ and $|\lambda - \omega|^{-\alpha}$ behave similarly around the pole. The model (3.8) was mentioned by Hosoya (1997). One notable feature of (3.8) is that besides the pole at $\lambda = \omega$, intended by the modeller, it entails lack of differentiability in λ at $\lambda = 0, \pi$, which, alongside the infinite differentiability of k at all $\lambda \neq 0, \omega, \pi$, is probably not a feature that a modeller would intend. On the other hand, models for short memory series with non-differentiable peaks or troughs at frequencies $0, \pi$, were considered by Robinson (1978).

To discuss further the distinction between A.1, A.4 and A.1', A.4', note that we can also write (2.5) as

$$k(\lambda; \theta, \omega) = \begin{cases} |\lambda - \omega|^{-\alpha}|\lambda + \omega|^{-\alpha}g_1(\lambda; \theta, \omega) & \text{if} \quad 0 \le \lambda \le \pi/2, \\ |\lambda - \omega|^{-\alpha}|\lambda + \omega - 2\pi|^{-\alpha}g_2(\lambda; \theta, \omega) & \text{if} \quad \pi/2 < \lambda \le \pi, \end{cases} \tag{3.9}$$

where

$$g_1(\lambda; \theta, \omega) = \left| \frac{(\lambda - \omega)(\lambda + \omega)}{4\sin(\frac{\lambda+\omega}{2})\sin(\frac{\lambda-\omega}{2})} \right|^{\alpha} |h(\lambda; \tau)|^2,$$

$$g_2(\lambda; \theta, \omega) = \left| \frac{(\lambda - \omega)(\lambda + \omega - 2\pi)}{4\sin(\frac{\lambda+\omega}{2})\sin(\frac{\lambda-\omega}{2})} \right|^{\alpha} |h(\lambda; \tau)|^2.$$

From (3.9) it follows that (3.7) holds with

$$g(\lambda; \theta, \omega) = \begin{cases} |\lambda + \omega|^{-\alpha}g_1(\lambda; \theta, \omega) & \text{if} \quad 0 \le \lambda \le \pi/2, \\ |\lambda + \omega - 2\pi|^{-\alpha}g_2(\lambda; \theta, \omega) & \text{if} \quad \pi/2 < \lambda \le \pi. \end{cases} \tag{3.10}$$

Although the functions g_1, g_2 satisfy Assumptions A.1' on g, (3.10) does not satisfy these assumptions, since it is unbounded at $\lambda = \omega$ for $\omega = 0, \pi$. We shall use in the proofs the fact that the Gegenbauer density k can be written in the form (3.9); the fact that in this expression g_1, g_2 satisfy Assumption A.1' on g; and the fact that if $\omega \in [\delta, \pi - \delta]$ where $\delta > 0$, i.e. when ω is separated from $0, \pi$, the function g, given by (3.10) satisfies Assumption A.1' on g.

For brevity we say that Assumptions **A** are satisfied if either A.1-A.6 hold or A.1', A.2, A.3, A.4', A.5 and A.6 hold.

Theorem 3.1 *Under Assumptions* **A**, *as* $n \to \infty$,

$$\hat{\theta} - \theta_0 = O_P(n^{-1/2}), \quad \hat{\omega} - \omega_0 = O_P(n^{-1}).$$

7

PROOF. It suffices to show that for all $\epsilon > 0$ we can choose $K = K_\epsilon$ such that

$$P\{n||\widehat{\theta} - \theta_0||^2 \geq K\} + P\{n|\widehat{\omega} - \omega_0| \geq 2\pi(K+1)\} \leq \epsilon \qquad (3.11)$$

for n sufficiently large, where $||.||$ denotes Euclidean norm. Define $q_0 = q_{0n} = \arg\min_Q |\omega_0 - \lambda_q|$ so that $|q_0 - \frac{n\omega_0}{2\pi}| \leq \frac{1}{2}$. If there are two such q, define q_0 as the smaller. Thus the left side of (3.11) is bounded by twice

$$P\{u_n(\hat{\psi}) \geq K\}, \qquad (3.12)$$

where $\psi = (\theta', \omega)'$, $\hat{\psi} = (\hat{\theta}', \hat{\omega})'$ and $u_n(\psi) = n||\theta - \theta_0||^2 + |\bar{q}|$ with $\bar{q} = q - q_0$. By a standard type of argument for proving consistency of implicitly defined extremum estimates, (3.12) is bounded by

$$P(\inf_{\psi \in \Psi_K} \{S(\psi) - S(\psi_0)\} \leq 0) = P(\inf_{\psi \in \Psi_K} \{u_n(\psi)^{-1}(S(\psi) - S(\psi_0))\} \leq 0), \qquad (3.13)$$

where $S(\psi) = S(\theta, \omega)$ and $\Psi_K = \{\psi : \theta \in \Theta, q \in Q, u_n(\psi) \geq K\}$. Now (3.13) is bounded by

$$P\left(\sup_{\Psi_K} |u_n(\psi)^{-1} U_n(\psi)| + \sup_{\Psi_K} |u_n(\psi)^{-1} V_n(\psi)| + K^{-1} \geq \inf_{\Psi_K}(u_n(\psi)^{-1} T_n(\psi))\right), \qquad (3.14)$$

where

$$T_n(\psi) = \sum_{j(q,q_0)}' \left\{\frac{k_j(\psi_0)}{k_j(\psi)} - 1\right\}, \quad V_n(\psi) = 1 - \frac{2\pi I_{q_0}}{\sigma_0^2 k_{q_0}(\psi)}, \qquad (3.15)$$

$$U_n(\psi) = \sum_{j(q_0)}' \left\{\frac{k_j(\psi_0)}{k_j(\psi)} - 1\right\}\left\{\frac{2\pi I_j}{\sigma_0^2 k_j(\psi_0)} - 1\right\}, \qquad (3.16)$$

and $I_j = I(\lambda_j)$, $k_j(\psi) = k(\lambda_j; \psi)$, $\sum_{j(q)}' = \sum_{j=0, j \neq q}^{\tilde{n}}$, $\sum_{j(q,q_0)}' = \sum_{j=0, j \neq q, q_0}^{\tilde{n}}$, so that

$$S(\psi_0) - S(\psi) = \frac{\sigma_0^2}{2\pi\tilde{n}}\left\{-U_n(\psi) + V_n(\psi) - T_n(\psi) - 1_{\{q = q_0\}}\right\}.$$

It is shown in Lemmas 6.1 and 6.2 below that as $n \to \infty, K \to \infty$,

$$E \sup_{\psi \in \Psi_K} |u_n(\psi)^{-1} U_n(\psi)| \to 0, \quad E \sup_{\psi \in \Psi_K} |u_n(\psi)^{-1} V_n(\psi)| \to 0; \qquad (3.17)$$

whereas by Lemma 7.1 below, for some $c > 0$,

$$\inf_{\psi \in \Psi_K}(u_n(\psi)^{-1} T_n(\psi)) \geq c \qquad (3.18)$$

where c does not depend on K, n. Thus, by Markov's inequality (3.14) is bounded by

$$\frac{E(\sup_{\Psi_K} |u_n(\psi)^{-1} U_n(\psi)| + \sup_{\Psi_K} |u_n(\psi)^{-1} V_n(\psi))| + K^{-1})}{\inf_{\Psi_K}(u_n(\psi)^{-1} T_n(\psi))} \to 0 \quad (n \to \infty, K \to \infty).$$

$\square$

8

Theorem 3.2 *Under Assumptions* **A**, *as* $n \to \infty$,

$$n^{1/2}(\widehat{\theta} - \theta_0) \overset{d}{\Rightarrow} N(0, \Omega^{-1}).$$

PROOF. From Theorem 3.1 and the fact that θ_0 is an interior point of Θ, for n sufficiently large

$$0 = \nabla_\theta S(\widehat{\psi}) = \nabla_\theta S(\theta_0, \widehat{\omega}) + M(\widehat{\theta} - \theta_0), \tag{3.19}$$

where the i-th row of the $p \times p$ matrix M is the i-th row of $\nabla_\theta \nabla_\theta' S(\theta, \widehat{\omega})$ evaluated at $\theta^{(i)}$, where $||\theta^{(i)} - \theta_0|| \leq ||\widehat{\theta} - \theta_0||$. Since $\{\varepsilon_t\}$ is ergodic, by a well - known argument (see e.g. Theorem 3.5.8 of Stout (1974)), the moving average sequence $\{x_t\}$ (3.5) is also ergodic. Therefore from Lemma 1 of Hannan (1973b) and Assumptions **A**, it follows that

$$M \overset{P}{\Rightarrow} \pi^{-1}\sigma_0^2 \Omega. \tag{3.20}$$

To complete the proof it remains to prove that

$$n^{1/2}\nabla_\theta S(\theta_0, \widehat{\omega}) \overset{d}{\Rightarrow} N(0, \pi^{-2}\sigma_0^4 \Omega), \tag{3.21}$$

which is shown in Lemma 5.2 below. $\qquad\square$

Theorem 3.2 indicates that the estimates of the short and long memory parameters τ_0 and α_0 have the same limit distribution when ω_0 is unknown as when it is known. In order to use the theorem in conducting inference on θ_0 we might consistently estimate Ω by

$$\hat{\Omega} = n^{-1}\sum_{j(\hat{q})}{}' \nabla_\theta \log k(\lambda_j; \hat{\theta}, \hat{\omega}) \nabla_\theta' \log k(\lambda_j; \hat{\theta}, \hat{\omega}).$$

Given Theorem 3.1, it is straightforward to show that $n^{1/2}(\widehat{\sigma}^2 - \sigma_0^2) \overset{d}{\Rightarrow} N(0, 2\sigma_0^4 + \kappa_4)$ as $n \to \infty$ (where κ_4 is the 4-th cumulant of ε_t), as in Whittle estimation of other models.

We now allow that $Ex_t = \mu$ is unknown, $\mu \in \mathbf{R}$, estimating $(\theta_0', \omega_0)'$ instead by

$$\begin{pmatrix} \hat{\theta}^* \\ \hat{\omega}^* \end{pmatrix} = \arg \min_{\Theta \times Q} S^*(\theta, \lambda_q) \quad \text{where} \quad S^*(\theta, \omega) = \tilde{n}^{-1}\sum_{j=1}^{\tilde{n}} \frac{I(\lambda_j)}{k(\lambda_j; \theta, \omega)}.$$

Theorem 3.3 *Under Assumptions* **A**, *with* $x_t = \mu + \sum_{j=0}^\infty \phi_j \varepsilon_{t-j}$ *replacing the representation for* x_t *in A.5, the results of Theorems 3.1 and 3.2 remain valid for the modified estimate* $(\hat{\theta}^{*\prime}, \hat{\omega}^*)'$.

PROOF. As is well known, for the Fourier frequencies $\lambda_j, j = 1, \ldots, \tilde{n}$, we can write $I(\lambda_j) = (2\pi n)^{-1}\left|\sum_{t=1}^n (x_t - Ex_t)e^{it\lambda_j}\right|^2$ since $I(\lambda_j)$ is invariant to location shift in x_t for such j. Hence

$$S^*(\theta, \omega) = S(\theta, \omega) - \tilde{n}^{-1}R_n(\theta, \omega), \tag{3.22}$$

9

where $R_n(\theta, \omega) = k(0; \theta, \omega)^{-1} I(0)$. Similarly to the proof of Lemma 6.2 it can be shown that

$$E \sup_{\psi \in \Psi_K} |u_n(\psi)^{-1} R_n(\theta, \omega)| \leq C K^{-\eta}$$

for some $\eta > 0$ where $C > 0$ does not depend on n, η and

$$E \sup_{\psi \notin \Psi_K} ||n^{-1} \nabla_\theta R_n(\theta, \omega)|| = o(n^{-1/2}), \quad E \sup_{\psi \notin \Psi_K} |n^{-1} R_n(\theta, \omega)| = o(n^{-1/2}) \quad (n \to \infty).$$

These relations imply that the term $\tilde{n}^{-1} R_n(\theta, \omega)$ in (3.22) is negligible. Namely, $S(\theta, \omega)$ approximates $S^*(\theta, \omega)$ and the same method of proof as in Theorems 3.1 - 3.2 can be applied. $\square$

Chung (1996a, 1996b) claims to have established the limit distribution of related estimates of both θ_0 and ω_0 in the context of the Gegenbauer model (2.3). Chung uses a conditional sum of squares method suggested by Gray et. al. (1989) in the same setting, and employed earlier by Box and Jenkins (1971) for ARMA models. This involves a time domain approximation to the Gaussian log likelihood, in which the conditional mean and variance of x_t given $x_s, 1 \leq s < t$ are replaced respectively by the conditional expectation given $x_s, -\infty < s < t$ with x_s then set to zero for $s \leq 0$, and the innovations variance. Our frequency domain approximation to the Gaussian log likelihood is proposed in part for computational reasons, because the functional form of the spectral density is typically of simple form and immediately identifiable, as in the Gegenbauer case, whereas time domain features such as AR coefficients are relatively cumbersome, while our approach can also make direct use of the fast Fourier transform. Though the limit distribution Chung states for his estimate of θ_0 is identical to that of ours, we are unable to check various details of his proofs of limit theory for estimates of θ_0 or ω_0. Perhaps, most notably, Chung claims that consistency of his estimates follows from the property that the expectation of his log likelihood approximation has zero derivative at the true parameter point. This property is insufficient, however, especially in the context of implicitly defined extremum estimates, where an initial consistency proof is an essential first step to deriving limit distribution theory, and indeed it is the proof of consistency that is the most challenging problem in the present situation due to the different rates of convergence of the estimates of θ_0 and ω_0, as we believe the proof of our Theorem 3.1 illustrates.

In the semiparametric context (1.4) there are known partial answers. Yajima (1996), in Gaussian case, and Hidalgo (1999), in the linear process case, have proposed estimators for ω_0 that are n^δ-consistent for any $\delta < 1$. Though Yajima (1996) did not obtain the asymptotic distribution, Hidalgo (1999) established asymptotic normality and suggested an estimator for θ_0 having the same rate of convergence and limiting distribution as in case of known ω_0.

Theorem 3.1 establishes n-consistency of $\hat{\omega}$. We are unable to derive its limit distribution. In fact we believe that none exists, due to the fact that $\hat{\omega}$ minimizes $S(\theta, \omega)$ not over the interval $[0, \pi]$ but over the grid Q, with mesh $2\pi/n$. If $\omega_0 \neq 0$, the Fourier frequency

10

λ_{q_0}, closest to ω_0, satisfies $|\lambda_{q_0} - \omega_0| \le 2\pi/n$ but the limit $n|\lambda_{q_0} - \omega_0|$ $(n \to \infty)$ does not exist. To explain the problem more precisely, note that

$$S(\hat{\theta}, \lambda_{\hat{q}}) - S(\theta_0, \lambda_{q_0}) = (S(\theta_0, \lambda_{\hat{q}}) - S(\theta_0, \lambda_{q_0})) + (1/2)(\hat{\theta} - \theta_0)'M(\hat{\theta} - \theta_0). \qquad (3.23)$$

By Theorem 3.2 and (3.20), the limit $\tilde{n}(\hat{\theta} - \theta_0)'M(\hat{\theta} - \theta_0) \Rightarrow (\sigma_0^2/2\pi)Z'\Omega Z$ as $n \to \infty$ exists and does not depend on λ_q, ω, where $Z \sim N(0, \Omega^{-1})$. Therefore, by (3.23), the limit distribution of $\hat{q}$ can be defined as arg $\min_q \lim_{n\to\infty} V_n(q)$ where $V_n(q) = \tilde{n}(S(\theta_0, \lambda_q) - S(\theta_0, \lambda_{q_0}))$. Write

$$V_n(q) = -\sum_{j(q_0)}{}' k_j^{-1}(\psi_0)I_j + \sum_{j(q)}{}' k_j^{-1}(\theta_0, \lambda_q)I_j = \sum_{j(q_0)}{}' d_{n,j}z_n(j) + k_{q_0}^{-1}(\theta_0, \lambda_q)I_{q_0}$$

where $d_{n,j} = \sigma_0^2(k_j^{-1}(\theta_0, \lambda_q)k_j(\theta_0, \lambda_{q_0}) - 1)$, $z_n(j) = (\sigma_0^2 k_j(\theta_0, \lambda_{q_0}))^{-1}I_j$. Note that for fixed finite j, as $n \to \infty$, $d_{n,j} \sim \tilde{d}_{n,j} = \sigma_0^2\big(|(j-q)/(j-q_0)|^{\alpha_0} - 1\big)$. If the limit $z_n(j) \Rightarrow z(j)$ exists we would expect that $V_n(q) \sim V(q) = \sum_{j=0: j\neq q_0}^{\infty} d_j z(j) + k_{q_0}^{-1}(\theta_0, \lambda_q)I_{q_0}$. However, in case $\omega_0 \neq 0$, even the limit of $Ez_n(j)$ does not exist. Indeed,

$$Ez_n(j) = \frac{k_j(\theta_0, \omega_0)}{k_j(\psi_0)}(\sigma_0^2 k_j(\theta_0, \omega_0))^{-1}EI_j.$$

By Robinson (1995a), $(\sigma_0^2 k_j(\theta_0, \omega_0))^{-1}EI_j$ has finite limit as $n \to \infty$ but

$$k_j(\theta_0, \omega_0)/k_j(\psi_0) \sim \Big|(\lambda_j - \omega_0)/(\lambda_j - \lambda_{q_0})\Big|^{\alpha_0} = \Big|1 + (2\pi n(\lambda_{q_0} - \omega_0))/(j - q_0)\Big|^{\alpha_0}$$

does not converge since as $n \to \infty$ the limit of $n(\lambda_{q_0} - \omega_0)$ does not exists.

To illustrate the finite-sample performance of our procedure, a small Monte Carlo experiment was carrried out by Dr Gilles Teyssiére on the basis of the simple Gegenbauer model (1.5), with ε_t Gaussian, $d = 0.4$ and $\omega = 0.25, 0.5, 1.0, 2.0, 2.5, 3.0$. $5,000$ replications of sample sizes $n = 64$ and 256 were generated using $3,000$ presample innovations and truncating the MA expansion at $3,000$ terms. Monte Carlo bias and standard deviation (SD) of $\hat{\omega}$ and $\hat{d} = \hat{\alpha}/2$ are reported in Tables 1 and 2. The bias in $\hat{\omega}$ seems fairly uniform across ω, whereas the bias of $\hat{d}$ is noticeably greater near $\omega = \pi/2$. Biases significantly decline as n increases, while in a more specific way the asymptotic theory, which predicts that $\text{SD}(\hat{d})$ and $\text{SD}(\hat{\omega})$ should be respectively doubled and quadrupled going from $n = 256$ to $n = 64$, is fairly well reflected by comparison of the two tables.

Table 1: $n = 64$ Table 2: $n = 256$

ω	BIAS($\hat{\omega}$)	SD($\hat{\omega}$)	BIAS($\hat{d}$)	SD($\hat{d}$)	ω	BIAS($\hat{\omega}$)	SD($\hat{\omega}$)	BIAS($\hat{d}$)	SD($\hat{d}$)
0.25	-.001	.085	-.013	.074	0.25	-.002	.019	-.006	.031
0.5	-.014	.081	-.042	.080	0.5	-.003	.020	-.002	.037
1.0	-.023	.110	-.075	.111	1.0	.001	.020	-.011	.049
1.5	-.015	.119	-.108	.138	1.5	-.002	.019	-.031	.059
2.0	.000	.107	-.085	.127	2.0	.000	.022	-.009	.055
2.5	.000	.089	-.050	.096	2.5	.002	.018	-.016	.039
3.0	-.026	.078	-.054	.079	3.0	-.004	.019	-.011	.031

4   Central Limit Theorem for weighted sums of periodograms

Throughout this section we assume that $\{x_t\}$ satisfies Assumptions **A**, so it has spectral density

$$f(\lambda) = \frac{\sigma_0^2}{2\pi} |\lambda - \omega_0|^{-\alpha} g(\lambda), \quad 0 \le \lambda \le \pi \tag{4.1}$$

where $\sigma_0^2 = E\varepsilon_0^2$ and $g(\lambda)$ is bounded away from infinity and zero and has bounded derivative $(\partial/\partial\lambda)g(\lambda)$ uniformly over $0 < \lambda < \pi$. However, whereas Assumptions **A** entail $\alpha \in (0,1)$, we allow in the current section also for $\alpha = 0$. Note that Assumption A.2 does not need to be imposed, however.

We prove now a Central Limit Theorem for the sums

$$S_n := \sum_{j(q_0)}{}' b_j \left(\frac{I_j}{f(\lambda_j)} - 1 \right) \tag{4.2}$$

with real weights $b_j \equiv b_{n,j}$, $j = 0, \dots, \tilde{n}$, the conditions on which are formulated in terms of

$$n_1(b) := n^{-1/2} \sum_{j(q_0)}{}' b_j, \quad n_2(b) := \sum_{j(q_0)}{}' b_j^2$$

and

$$n_3(b, \gamma) := \sum_{j=-1: j \neq q_0, q_0-1}^{\tilde{n}} |b_j - b_{j+1}||j - q_0|^{\gamma} \quad (\gamma > 0),$$

setting $b_{-1} = b_{\tilde{n}+1} = 0$.

Lemma 4.1 *Suppose that as $n \to \infty$*

$$n_1(b) \to v_1 < \infty, \quad n_2(b) \to v_2 < \infty \quad and \quad n_3(b, \gamma) \to 0 \tag{4.3}$$

for some $\gamma > 1/3$. Then S_n, given by (4.2), satisfies

$$S_n \overset{d}{\Rightarrow} N(0, \sigma^2), \tag{4.4}$$

where $\sigma^2 = v_2 + v_1^2 \sigma_0^{-4} \kappa_4$.

The proof of Lemma 4.1 is technical and is reserved for the next section.

We consider now a special case of the weights b_j.

Theorem 4.2 *Let*
$$b_j = n^{-1/2} h_n(\lambda_j), \quad j = 0, \ldots, \tilde{n}$$
where $h_n(\lambda), \lambda \in [0, \pi], n \geq 1$ are real valued functions. Suppose there exist $C > 0, K > 0$, $0 \leq \alpha' < 1/2$ and $0 \leq \gamma' < 3/2$, independent of n, such that for all sufficiently large n

c1) $|h_n(\lambda)| \leq C|\lambda - \omega_0|^{-\alpha'}$ *uniformly in $|\lambda - \omega_0| \geq \pi/n$,*

c2) $|(d/d\lambda)h_n(\lambda)| \leq C|\lambda - \omega_0|^{-\gamma'}$ *uniformly in $|\lambda - \omega_0| \geq K/n$, $0 < \lambda < \pi$,*

and also

c3) $\lim_{n \to \infty} h_n(\lambda) = h(\lambda)$ *exists for $\lambda \neq \omega_0$.*

Then (4.4) holds with
$$v_1 = (2\pi)^{-1} \int_0^\pi h(\lambda) d\lambda, \quad v_2 = (2\pi)^{-1} \int_0^\pi h^2(\lambda) d\lambda.$$

PROOF. By Lemma 4.1, it suffices to show that as $n \to \infty$ the weights $b_j = n^{-1/2} h_n(\lambda_j)$ satisfy (4.3). We prove first that $n_3(b, \gamma) \to 0$ for any $\gamma \in (1/3, 1/2)$. By c1), with $b_{-1} = b_{\tilde{n}+1} = 0$,

$$
\begin{aligned}
n_3(b, \gamma) &= n^{-1/2} \sum_{j=0: |j-q_0| \geq 2K}^{\tilde{n}-1} |h_n(\lambda_j) - h_n(\lambda_{j+1})| |j - q_0|^\gamma + o(1) \\
&\leq Cn^{-3/2} \sum_{j=0: |j-q_0| \geq 2K}^{n} (n/|j - q_0|)^{\gamma'} |j - q_0|^\gamma + o(1) = o(1)
\end{aligned}
$$

since by c2),

$$|h_n(\lambda_j) - h_n(\lambda_{j+1})| \leq Cn^{-1} \sup_{\xi \in [\lambda_j, \lambda_{j+1}]} |h'_n(\xi)| \leq Cn^{-1} |\lambda_j - \lambda_{q_0}|^{-\gamma'}. \tag{4.5}$$

Convergence of the sums $n_1(b) = n^{-1} \sum'_{j(q_0)} h_n(\lambda_j) \to v_1$, $n_2(b) = n^{-1} \sum'_{j(q_0)} h_n^2(\lambda_j) \to v_2$ follows by Lemma 5.3 below, because c1) - c3) imply $h_n^2(\lambda) \leq C|\lambda - \omega_0|^{-2\alpha'}$ uniformly in $|\lambda - \omega_0| \geq \pi/n$ and $|(d/d\lambda)h_n^2(\lambda)| \leq C|\lambda - \omega_0|^{-\alpha'-\gamma'}$ uniformly in $|\lambda - \omega_0| \geq K/n$. $\square$

We provide now an auxiliary lemma on the approximation of normalised periodograms. Denote $I_{j,\varepsilon} = (2\pi n)^{-1} |\sum_{t=1}^n \varepsilon_t e^{it\lambda_j}|^2$ and write

$$\xi_{j,n}^{(1)} = f^{-1}(\lambda_j) I_j - 1; \quad \xi_{j,n}^{(2)} = 2\pi I_{j,\varepsilon}/\sigma_0^2 - 1; \quad \xi_{j,n}^{(3)} = f^{-1}(\lambda_j) I_j - 2\pi I_{j,\varepsilon}/\sigma_0^2. \tag{4.6}$$

Define for $i = 1, 2, 3$,

$$Z_{j,n}^{(i)} = \begin{cases} \sum_{\ell=q_0+1}^{j} \xi_{\ell,n}^{(i)} & \text{if } j = q_0 + 1, \ldots, \tilde{n}, \\ \sum_{\ell=j+1}^{q_0-1} \xi_{\ell,n}^{(i)} & \text{if } j = -1, \ldots, q_0 - 2, \end{cases} \tag{4.7}$$

recalling that λ_{q_0} is the closest Fourier frequency to ω_0.

Lemma 4.3 *As $n \to \infty$, for $j = 1, \ldots, \tilde{n}, j \neq q_0$,*

$$E(Z_{j,n}^{(i)})^2 \leq \begin{cases} C|j - q_0| & \text{if} \quad i = 1, 2, \\ C|j - q_0|^{2p} & \text{if} \quad i = 3, \end{cases} \tag{4.8}$$

for any $p > 1/3$ where C does not depend on n and j.

PROOF. Since $\xi_{j,n}^{(1)} = \xi_{j,n}^{(2)} + \xi_{j,n}^{(3)}$, then relation (4.8) for $i = 1$ follows if it is valid for $i = 2, 3$, for any $p > 1/3$. The proof of (4.8) in case $i = 2$ is the same as that of (4.9) in Robinson (1995b). Relation (4.8) in case $i = 3$ follows since

$$E(Z_{j,n}^{(3)})^2 \leq C|j - q_0|^{2/3}(\log|j - q_0|)^{4/3} \tag{4.9}$$

which can be obtained similarly to relation (4.8) in Robinson (1995b), applying Lemma 4.1 and

$$\frac{1}{2\pi n} \int_{-\pi}^{\pi} \left|\frac{\phi(\lambda)}{\phi(\lambda_j)} - 1\right|^2 \left|\sum_{l=1}^{n} e^{il(\lambda - \lambda_j)}\right|^2 d\lambda \leq C|j - q_0|^{-1} \tag{4.10}$$

which holds uniformly over $0 \leq j \leq \tilde{n}, j \neq q_0$. Estimate (4.10) generalizes Robinson's (1995b) Lemma 3 for all $\omega_0 \in [0, \pi]$ and extends it over all Fourier frequences λ_j, $j \neq q_0$. The proof of (4.10) requires Assumption A.6. We omit the proof since, in general, it repeats the proof of Robinson's (1995b) Lemma 3. (Note that relation (4.8) in Robinson (1995b) where the peridogram I_j is normalized by the approximate spectral density $C|\lambda_j - \lambda_{q_0}|^{-\alpha_0}$, $C > 0$, due to the approximation, contains the additional term $|j - q_0|^{\beta+1}n^{-\beta}$.)
$\square$

We now consider jointly covariance stationary processes $\{y_t\}$ and $\{z_t\}$, individually satisfying Assumptions **A**, (or more precisely, Assumptions A.5 and A.6) with the same innovations sequence $\{\varepsilon_t\}$, but with possibly different memory parameters, denoted α_y, α_z. In fact we allow also for $\alpha_y = 0$ and/or $\alpha_z = 0$, in order to apply Lemma 4.4 in the proof of Lemma 4.3 with (y_t, z_t) representing (x_t, x_t), (x_t, ε_t) and $(\varepsilon_t, \varepsilon_t)$. Denote by f_y and f_z the spectral densities of y_t, z_t, respectively, and by $R_{yz}(\lambda)$ their coherency. Introduce

$$v_y(\lambda) = \left(\sum_{t=1}^{n} y_t e^{it\lambda}\right)/\{f_y(\lambda)2\pi n\}^{1/2}, \quad v_z(\lambda) = \left(\sum_{t=1}^{n} z_t e^{it\lambda}\right)/\{f_z(\lambda)2\pi n\}^{1/2}.$$

Lemma 4.4 *Let the sequences $\{y_t\}, \{z_t\}$ satisfy Assumptions **A** with $0 \leq \alpha_y, \alpha_z < 1$. Then, as $n \to \infty$, the following relations hold uniformly over $q_0 + 1 \leq k, j \leq \tilde{n}$ and $0 \leq k, j \leq q_0 - 1$ such that $|q_0 - j| > |k - q_0|$:*

$$E[v_y(\lambda_j)\overline{v_z(\lambda_j)}] = R_{yz}(\lambda_j) + O(|j - q_0|^{-1}\log|j - q_0|);$$

$$E[v_y(\lambda_j)v_z(\lambda_j)] = O(|j - q_0|^{-1}\log|j - q_0|);$$

$$\max\left(|E[v_y(\lambda_j)\overline{v_z(\lambda_k)}]|, |E[v_y(\lambda_j)v_z(\lambda_k)]|\right) = O(|k - q_0|^{-1}\log|j - q_0|).$$

14

PROOF. We can write $y_t = \sum_{j=0}^{\infty} \phi_{yj} \varepsilon_{t-j}$, $z_t = \sum_{j=0}^{\infty} \phi_{zj} \varepsilon_{t-j}$, where $\{\phi_{yj}\}, \{\phi_{zj}\}$ satisfy the conditions on $\{\phi_j\}$ implied by Assumptions **A**. Further, f_y and f_z have representations of the form (4.1), with α replaced by $\alpha_y \in [0,1)$ and $\alpha_z \in [0,1)$, respectively, while the cross-spectral density of (y_t, z_t) is of the form $(\sigma_0^2/2\pi)\phi_y(\lambda)\overline{\phi_z(\lambda)}$, where $\phi_y(\lambda) = \sum_{j=0}^{\infty} \phi_{yj} e^{ij\lambda}$, $\phi_z(\lambda) = \sum_{j=0}^{\infty} \phi_{zj} e^{ij\lambda}$, and thus has derivative in λ that is $O(|\lambda - \omega_0|^{-1-(\alpha_y+\alpha_z)/2})$ for $\lambda \in [0,\pi]\backslash\{\omega_0\}$. It follows that Assumptions A.1- A.2 of Robinson (1995a) are satisfied, and the proof follows as in Robinson's (1995a) Theorem 2. $\qquad\square$

5 Proof of Lemma 4.1

Write, in the notation of (4.6),

$$S_n \equiv \sideset{}{'}\sum_{j(q_0)} b_j \xi_{j,n}^{(1)} = \sideset{}{'}\sum_{j(q_0)} b_j \xi_{j,n}^{(2)} + \sideset{}{'}\sum_{j(q_0)} b_j \xi_{j,n}^{(3)} =: M_n + R_n.$$

(4.4) follows if we show that as $n \to \infty$

$$M_n \overset{d}{\Rightarrow} N(0,\sigma^2) \tag{5.1}$$

and

$$R_n \overset{P}{\Rightarrow} 0. \tag{5.2}$$

We begin with the proof of (5.2). By summation by parts,

$$
\begin{aligned}
E|R_n| &= E \Big| \sum_{j=-1: j\neq q_0, q_0-1}^{\tilde{n}} (b_j - b_{j+1}) Z_{j,n}^{(3)} \Big| \\
&\leq \sum_{j=-1: j\neq q_0, q_0-1}^{\tilde{n}} |b_j - b_{j+1}| (E(Z_{j,n}^{(3)})^2)^{1/2} \leq C n_3(b,\gamma).
\end{aligned}
$$

by (4.8) where $\gamma > 1/3$ is the same as in (4.3). So (5.2) follows by (4.3).

It remains to show (5.1). Since we employ a similar approach to that of Theorem 2 of Robinson (1995b) and S_n is invariant to σ_0^2, we set $\sigma_0^2 = 1$ in the proof as he did. Write M_n as

$$M_n = \sum_{j=0}^{\tilde{n}} b_j (2\pi I_{j,\varepsilon} - 1) = \sum_{t=1}^{n} z_{t,n} \tag{5.3}$$

where we can set $b_{q_0} = 0$ because M_n excludes $j = q_0$, and

$$z_{t,n} = 2\varepsilon_t \sum_{s=1}^{t-1} c_{t-s}\varepsilon_s + c_0(\varepsilon_t^2 - 1), \quad c_s = n^{-1}\sum_{j=0}^{\tilde{n}} b_j \cos(s\lambda_j).$$

$z_{t,n}$ is a triangular martingale difference array since $E[z_{t,n}|F_{t-1}] = 0$ by A.5. Thus, as in the proof of Theorem 2 of Robinson (1995b), (5.1) follows if

$$V_n = \sum_{t=1}^{n} E\left(z_{t,n}^2|F_{t-1}\right) \overset{P}{\to} \sigma^2 \quad (n \to \infty) \tag{5.4}$$

15

and

$$\sum_{t=1}^{n} E\left(z_{t,n}^2 1_{\{|z_{t,n}|>\delta\}}\right) \xrightarrow{P} 0 \quad \text{for all} \quad \delta > 0. \tag{5.5}$$

To derive (5.4) it suffices to show that

$$EV_n = n_2(b)(1 + n^{-2}) + \kappa_4 n_1^2(b) \tag{5.6}$$

and

$$Var(V_n) \to 0 \quad (n \to \infty) \tag{5.7}$$

since together with (4.3) they imply (5.4). By definition of $z_{t,n}$,

$$z_{t,n}^2 = 4\varepsilon_t^2 \sum_{1 \le r,s < t} c_{t-s} c_{t-r} \varepsilon_s \varepsilon_r + c_0^2(\varepsilon_t^2 - 1)^2 + 4\varepsilon_t(\varepsilon_t^2 - 1)c_0 \sum_{1 \le s < t} c_{t-s}\varepsilon_s.$$

By Assumption A.5, $E[\varepsilon_t^2|F_{t-1}] = 1$, $E[(\varepsilon_t^2 - 1)^2|F_{t-1}] = \mu_4 - 1$, $E[\varepsilon_t(\varepsilon_t^2 - 1)|F_{t-1}] = \mu_3$. Therefore

$$V_n = 4\sum_{t=1}^{n} \sum_{1 \le r,s < t} c_{t-s} c_{t-r} \varepsilon_s \varepsilon_r + (\mu_4 - 1)c_0^2 n + 4\mu_3 c_0 \sum_{t=1}^{n} \sum_{1 \le s < t} c_{t-s}\varepsilon_s,$$

and thus

$$EV_n = 4\sum_{t=1}^{n} \sum_{s=1}^{t-1} c_{t-s}^2 + (\mu_4 - 1)c_0^2 n. \tag{5.8}$$

Since by definition

$$nc_0^2 = n_1^2(b), \tag{5.9}$$

(5.6) follows from (5.8) if

$$i_n := 4\sum_{t=2}^{n} \sum_{s=1}^{t-1} c_{t-s}^2 = n_2(b)(1 + n^{-2}) - 2n_1^2(b) \tag{5.10}$$

which can be obtained using similar algebra as in Robinson (1995b). Indeed,

$$i_n = 4\sum_{t=1}^{n-1} \sum_{s=1}^{n-t} c_s^2 = 4n^{-2} \sum_{j,k=0}^{\tilde{n}} b_j b_k \sum_{t=1}^{n-1} \sum_{s=1}^{n-t} [\cos(s\lambda_j)\cos(s\lambda_k)]$$

$$= 4n^{-2} \sum_{j=0}^{\tilde{n}} b_j^2 \sum_{t=1}^{n-1} \sum_{s=1}^{n-t} \cos^2(s\lambda_j) + 2n^{-2} \sum_{j,k=0:j\neq k}^{\tilde{n}} b_j b_k \sum_{t=1}^{n-1} \sum_{s=1}^{n-t} [\cos(s(\lambda_j + \lambda_k)) + \cos(s(\lambda_j - \lambda_k))],$$

and applying the equalities

$$\sum_{t=1}^{n-1} \sum_{s=1}^{n-t} \cos^2(s\lambda_j) = (n-1)^2/4, \quad \sum_{t=1}^{n-1} \sum_{s=1}^{n-t} [\cos(s(\lambda_j + \lambda_k)) + \cos(s(\lambda_j - \lambda_k))] = -n$$

16

we obtain

$$
\begin{aligned}
i_n &= n^{-2}(n-1)^2 \sum_{j=0}^{\tilde{n}} b_j{}^2 - 2n^{-1} \sum_{j,k=0:j\neq k}^{\tilde{n}} b_j b_k \\
&= \sum_{j=0}^{\tilde{n}} b_j{}^2 (1 + n^{-2} - 2n^{-1}) - 2n^{-1}[(\sum_{j=0}^{\tilde{n}} b_j)^2 - \sum_{j=0}^{\tilde{n}} b_j{}^2] = n_2(b)(1 + n^{-2}) - 2n_1^2(b),
\end{aligned}
$$

establishing (5.6).

Next we show (5.7). We have

$$
V_n - EV_n = 4 \sum_{t=1}^{n} \sum_{1 \le r,s < t} c_{t-s} c_{t-r} (\varepsilon_s \varepsilon_r - E[\varepsilon_s \varepsilon_r]) + 4\mu_3 c_0 \sum_{t=1}^{n} \sum_{1 \le s < t} c_{t-s} \varepsilon_s. \tag{5.11}
$$

By (5.11) and elementary inequalities,

$$
\begin{aligned}
Var(V_n) &\le 32\Big(E(\sum_{t=1}^{n} \sum_{1 \le r,s<t} c_{t-s} c_{t-r}(\varepsilon_s \varepsilon_r - E[\varepsilon_s \varepsilon_r]))^2 + E\big(\mu_3^2 c_0^2(\sum_{t=1}^{n} \sum_{1 \le s < t} \varepsilon_s c_{t-s})^2\big)\Big) \\
&\le C\Big(\sum_{t_1,t_2=1}^{n} \big[\sum_{s=1}^{\min(t_1,t_2)} |c_{t_1-s} c_{t_2-s}|\big]^2 + c_0^2 \sum_{t_1,t_2=1}^{n} \big[\sum_{s=1}^{\min(t_1,t_2)} |c_{t_1-s} c_{t_2-s}|\big]\Big).
\end{aligned}
$$

For $1 \le s \le n - 1$, $|c_s| = o(1)n^{-p} \min(s, n-s)^{p-1}$ by Lemma 5.1 below and (4.3) for any $p \in (1/3, \min(\gamma, 1/2))$. Thus,

$$
\begin{aligned}
\sum_{s=1}^{\min(t_1,t_2)} |c_{t_1-s} c_{t_2-s}| &= o(n^{-2p}) \sum_{s=1}^{\min(t_1,t_2)} \Big\{\min(|t_1 - s|_+, |t_1 - s - n|_+)^{p-1} \\
&\quad \cdot \min(|t_2 - s|_+, |t_2 - s - n|_+)^{p-1}\Big\} = o(n^{-2p}) \min(|t_1 - t_2|_+, |t_1 - t_2 - n|_+)^{2p-1}
\end{aligned}
$$

uniformly in t_1, t_2, where $|t|_+ = \max(|t|, 1)$. By (5.9) and (4.3), $c_0^2 = O(n^{-1})$. Hence,

$$
\begin{aligned}
Var(V_n) &= o(1)\Big(n^{-4p} \sum_{t_1,t_2=1}^{n} |t_1 - t_2|_+^{2(2p-1)} + n^{-2p-1} \sum_{t_1,t_2=1}^{n} |t_1 - t_2|_+^{2p-1}\Big) \\
&= o(1)\Big(n^{-4p} \sum_{t_1=1}^{n} \sum_{u=-n}^{n} |u|_+^{4p-2} + n^{-2p-1} \sum_{t_1=1}^{n} \sum_{u=-n}^{n} |u|_+^{2p-1}\Big) = o(1)
\end{aligned}
$$

since $p > 1/3$. Thus, (5.7) is established, to complete the proof of (5.4).

To complete the proof of Lemma 4.1 we need to examine (5.5). Write

$$
z_{t,n} = 2\Big\{\varepsilon_t \sum_{1 \le s < t} \varepsilon_s c_{t-s}\Big\} + \Big\{c_0(\varepsilon_t^2 - 1)\Big\} =: X_t + Y_t.
$$

Then

$$
\begin{aligned}
E[z_{t,n}^2 1_{\{|z_{t,n}| \ge \delta\}}] &\le E[(|X_t| + |Y_t|)^2 1_{\{|X_t + Y_t| \ge \delta\}}] \\
&\le E[(2|X_t|)^2 1_{\{|X_t| \ge |Y_t|, |X_t| \ge \delta/2\}}] + E[(2|Y_t|)^2 1_{\{|Y_t| \ge |X_t|, |Y_t| \ge \delta/2\}}]. \tag{5.12}
\end{aligned}
$$

17

The first term in (5.12) is bounded by

$$4E[|X_t|^2 1_{\{|X_t| \geq \delta/2\}}] \leq 4(\delta/2)^{-2} E|X_t|^4 = CE(\varepsilon_t)^4 E\Big(\sum_{1 \leq s < t} \varepsilon_s c_{t-s}\Big)^4$$

$$\leq C\Big(\sum_{s=1}^{t-1} c_s^2\Big)^2 \leq Cn^{-4p}\Big(\sum_{s=1}^{\infty} s^{2p-2}\Big)^2 \leq Cn^{-4p}$$

applying (5.13) with $p \in (1/3, \min(\gamma, 1/2))$. The second term of (5.12) is bounded by

$$4E[|Y_t|^2 1_{\{|Y_t| \geq \delta/2\}}] \leq 4(\delta/2)^{-\eta/2} E|Y_t|^{2+\eta/2} \leq CE|c_0(\varepsilon_t^2 - 1)|^{2+\eta/2}$$
$$\leq C|c_0|^{2+\eta/2} E|\varepsilon_t|^{4+\eta} = o(n^{-1})$$

since $|c_0| \leq Cn^{-1/2}$ and by Assumption A.5, $E|\varepsilon_t|^{4+\eta} < \infty$ for some $\eta > 0$. Hence

$$\sum_{t=1}^{n} E[z_{t,n}^2 1_{\{|z_{t,n}| \geq \delta\}}] = [O(n^{-4p}) + o(n^{-1})] \sum_{t=1}^{n} 1 = o(1) \quad (n \to \infty).$$

This completes the proof of (5.5) and Lemma 4.1. $\qquad\qquad\qquad \square$

Lemma 5.1 *Under the assumptions of Lemma 4.1, for any $0 < p < 1$,*

$$|c_s| \leq Cn^{-p} \min(s, n-s)^{p-1} n_3(b, p) \qquad\qquad (5.13)$$

uniformly in $1 \leq s \leq n-1$.

PROOF. Recall that $b_{\tilde{n}+1} = b_{-1} = 0$. Summation by parts implies

$$c_s = n^{-1}\Big\{ \sum_{j=q_0+1}^{\tilde{n}} (b_j - b_{j+1}) \sum_{l=q_0+1}^{j} \cos(l\lambda_s) + \sum_{j=0}^{q_0-1} (b_j - b_{j-1}) \sum_{l=j}^{q_0-1} \cos(l\lambda_s)\Big\}.$$

Note that $|\sum_{l=k+1}^{m} \cos(l\lambda_s)| \leq |m-k|^p |\sum_{l=k+1}^{m} \cos(l\lambda_s)|^{1-p} \leq C|m-k|^p (n/\min(s, n-s))^{1-p}$ which holds because $|\sum_{l=k+1}^{m} \cos(l\lambda_s)| \leq C|\sin(\lambda_s/2)|^{-1} \leq C(n/\min(s, n-s))$ uniformly in $0 \leq k < m$ and $1 \leq s \leq n-1$. Thus

$$|c_s| \leq (n/\min(s, n-s))^{1-p} n^{-1}\Big[\sum_{j=q_0+1}^{\tilde{n}} |b_j - b_{j+1}||j - q_0|^p + \sum_{j=0}^{q_0-1} |b_j - b_{j-1}||j - q_0|^p\Big]$$

$$\equiv Cn^{-p} \min(s, n-s)^{p-1} n_3(b, p).$$

$\square$

Lemma 5.2 *Under the conditions of Theorem 3.2 the convergence (3.21) holds.*

PROOF. By Theorem 3.1, it suffices to consider $\hat{\omega} = 2\pi\hat{q}/n$ such that $|\hat{q} - q_0| \leq K$ where $K > 0$ is large enough. The ℓ-th element of $n^{1/2}\nabla_\theta S(\theta_0, \lambda_q)$ equals to $(n/\tilde{n})W_{n,q}^{(\ell)}$, where

$$W_{n,q}^{(\ell)} = n^{-1/2}\sum_{j=0}^{\tilde{n}}(\partial/\partial\theta_\ell)k_j^{-1}(\theta_0, \lambda_q)I_j,$$

and θ_ℓ is the ℓ-th element of θ. Setting $h_{n,q}^{(\ell)}(\lambda) = (\sigma_0^2/2\pi)k(\lambda;\psi_0)(\partial/\partial\theta_\ell)k^{-1}(\lambda, \theta_0, \lambda_q)$, write $W_{n,q}^{(\ell)} = v_{n,q}^{(\ell)} + e_{n,q}^{(\ell)} + r_{n,q}^{(\ell)}$ where

$$v_{n,q}^{(\ell)} = n^{-1/2}{\sum_{j(q_0)}}' h_{n,q}^{(\ell)}(\lambda_j)\Big(\frac{2\pi I_j}{\sigma_0^2 k_j(\psi_0)} - 1\Big), \quad e_{n,q}^{(\ell)} = n^{-1/2}{\sum_{j(q_0)}}' h_{n,q}^{(\ell)}(\lambda_j),$$

$r_{n,q}^{(\ell)} = n^{-1/2}(\partial/\partial\theta_\ell)k_{q_0}^{-1}(\theta_0, \lambda_q)I_{q_0}$. To complete the proof it suffices to show: 1) that as $n \to \infty$,

$$\Big(v_{n,q_0}^{(\ell)}\Big)_{\ell=1,\ldots,p} \overset{d}{\Rightarrow} \Big(Z^{(\ell)}\Big)_{\ell=1,\ldots,p}, \tag{5.14}$$

where $(Z^{(1)}, \ldots, Z^{(p)})'$ is a Gaussian vector with zero mean and covariance $E(Z^{(\ell)}Z^{(\ell')}) = (2\pi)^{-2}\sigma_0^4 w_{\ell\ell'}$; and 2) that for $k = -K, \ldots, K; \ell = 1, \ldots, p$,

$$v_{n,q_0+k}^{(\ell)} - v_{n,q_0}^{(\ell)} \overset{P}{\Rightarrow} 0, \tag{5.15}$$

$$e_{n,q_0+k}^{(\ell)} \to 0, \qquad E|r_{n,q_0+k}^{(\ell)}| \to 0. \tag{5.16}$$

(5.14) - (5.16) imply (3.21). Also, as a by - product, (5.14) - (5.15) imply that the distribution of $\hat{\theta}$ in Theorem 3.2 is the same whether ω_0 is known or estimated.

By the Cramer-Wold device, the convergence (5.14) holds if for any sequence of real numbers $a_1, \ldots, a_p, p \geq 1$, as $n \to \infty$,

$$Q_n := \sum_{\ell=1}^p a_\ell v_{n,q_0}^{(\ell)} \overset{d}{\Rightarrow} Q := \sum_{\ell=1}^p a_\ell Z^{(\ell)}. \tag{5.17}$$

Write

$$Q_n = n^{-1/2}{\sum_{j(q_0)}}' h_n(\lambda_j)(2\pi(\sigma_0^2 k_j(\psi_0))^{-1}I_j - 1), h_n(\lambda) = \sum_{\ell=1}^p a_\ell h_{n,q_0}^{(\ell)}(\lambda).$$

To derive (5.17) we apply Theorem 4.2. Note that $\lambda_{q_0} \to \omega_0$ as $n \to \infty$, so that

$$h_n(\lambda) \to h(\lambda) = \sum_{\ell=1}^p a_\ell h^{(\ell)}(\lambda), h^{(\ell)}(\lambda) = -(\sigma_0^2/2\pi)(\partial/\partial\theta_\ell)\log k(\lambda;\theta_0, \omega_0).$$

By Assumption A.2, $\int_0^\pi h^{(\ell)}(\lambda)d\lambda = 0$, $\int_0^\pi h(\lambda)d\lambda = 0$. In addition, from Assumption A.1' or A.1 it is easy to obtain that there exists $\varepsilon \in (0, 1/2)$ such that

$$|h_{n,q_0}^{(\ell)}(\lambda)| \leq C|\lambda - \omega_0|^{-\varepsilon}, \quad |(d/d\lambda)h_{n,q_0}^{(\ell)}(\lambda)| \leq C|\lambda - \omega_0|^{-1-\varepsilon} \tag{5.18}$$

uniformly in $|\lambda - \omega_0| \geq 2K/n$. Therefore, the functions $h_n, n \geq 1$ satisfy conditions c1) - c3) of Theorem 4.2 which yields $Q_n \overset{d}{\Rightarrow} Q = N(0, s^2)$, $s^2 = (2\pi)^{-1}\int_0^\pi h^2(\lambda)d\lambda$. Hence (5.14) holds.

19

To show (5.15), write

$$v_{n,q_0+k}^{(\ell)} - v_{n,q_0}^{(\ell)} = n^{-1/2}\sum_{j(q_0)}' \tilde{h}_n(\lambda_j)\big((2\pi/\sigma_0^2)k_j(\psi_0)^{-1}I_j - 1\big)$$

where for $\lambda \neq \omega_0$, $\tilde{h}_n(\lambda) = h_{n,q_0+k}^{(\ell)}(\lambda) - h_{n,q_0}^{(\ell)}(\lambda) \to 0$ as $n \to \infty$. Then the convergence (5.15) follows by the same argument as in the proof of (5.14).

Finally we show (5.16). By (5.18) and Lemma 5.3 below,

$$e_{n,q_0+k}^{(\ell)} = (n^{1/2}/2\pi)\int_0^\pi 1_{\{|\lambda-\omega_0|\geq 2K/n\}}h_{n,q_0+k}^{(\ell)}(\lambda)d\lambda + o(1).$$

Furthermore, since $\int_0^\pi h^{(\ell)}(\lambda)d\lambda = 0$,

$$
\begin{aligned}
|e_{n,q_0+k}^{(\ell)}| &\leq Cn^{1/2}\int_0^\pi 1_{\{|\lambda-\omega_0|\geq 2K/n\}}|h_{n,q_0+k}^{(\ell)}(\lambda) - h^{(\ell)}(\lambda)|d\lambda \\
&+ Cn^{1/2}\int_0^\pi 1_{\{|\lambda-\omega_0|<2K/n\}}|h^{(\ell)}(\lambda)|d\lambda + o(1).
\end{aligned}
$$

It is easy to check that under Assumptions **A** (i.e. under Assumptions A.1 or A.1′), uniformly in $|\lambda - \omega_0| \geq 2K/n$,

$$
\begin{aligned}
|h_{n,q_0+k}^{(\ell)}(\lambda) - h^{(\ell)}(\lambda)| &= \frac{\sigma_0^2}{2\pi}\Big|\frac{k(\lambda;\theta_0,\lambda_{q_0})}{k(\lambda;\theta_0,\lambda_{q_0+k})}\frac{\partial}{\partial\theta_\ell}\log k(\lambda;\theta_0,\lambda_{q_0+k}) - \frac{\partial}{\partial\theta_\ell}\log k(\lambda;\theta_0,\omega_0)\Big| \\
&\leq C\big(\lambda_k(\lambda - \lambda_{q_0})^{-1} + \lambda_k^\beta\big)\log n \leq C(n^{-1}(\lambda - \lambda_{q_0})^{-1} + n^{-\beta})\log n
\end{aligned}
$$

where $\beta \in (1/2,1)$ is the Lipschitz parameter in Assumption A.1′ (or A.1), and $|h^{(\ell)}(\lambda)| = |(\partial/\partial\theta_\ell)\log k(\lambda;\theta_0,\lambda_{q_0+k})| \leq C|\lambda - \lambda_{q_0}|^{-\varepsilon}$ with $0 < \varepsilon < 1/2$. Thus, as $n \to \infty$,

$$
\begin{aligned}
|e_{n,q_0+k}^{(\ell)}| &\leq Cn^{1/2}\Big(\log n\int_{0:|\lambda-\omega_0|\geq 2K/n}^\pi (n^{-1}(\lambda - \lambda_{q_0})^{-1} + n^{-\beta})d\lambda \\
&+ Cn^{1/2}\int_{0:|\lambda-\omega_0|<2K/n}^\pi |\lambda - \lambda_{q_0}|^{-\varepsilon}d\lambda\Big) + o(1) = o(1).
\end{aligned}
$$

It remains to estimate $E|r_{n,q_0+k}^{(\ell)}|$. Since $|(\partial/\partial\theta_\ell)k^{-1}(\lambda_{q_0};\theta_0,\lambda_{q_0+k})| \leq C|\lambda_k|^{\alpha_0}|\log\lambda_k| \leq Cn^{-\alpha_0}\log n$, then $E|r_{n,q_0+k}^{(\ell)}| \leq Cn^{-1/2-\alpha_0}\log n\, EI_{q_0} \to 0$ follows by standard arguments.
$\square$

The following lemma is used in the proofs of Lemma 5.2 and Lemma 7.1, Section 7.

Lemma 5.3 *Suppose the functions $\psi_n(\lambda)$, $n \geq 1$, satisfy Assumptions c1) - c3) of Theorem 4.2 on $h_n(\lambda)$ with $0 \leq \alpha' < 1$ and $0 \leq \gamma' < 2, \gamma' \neq 1$. Then, for large enough $K > 0$,*

$$n^{-1}\sum_{j(q_0)}' \psi_n(\lambda_j) = (2\pi)^{-1}\int_0^\pi 1_{\{|\lambda-\omega_0|\geq K/n\}}\psi_n(\lambda)d\lambda + O\big(n^{-1+\max(\alpha',\gamma'-1)}\big) \tag{5.19}$$

$$= (2\pi)^{-1}\int_0^\pi \psi(\lambda)d\lambda + o(1).$$

20

PROOF. Denote by J_n the left hand side of (5.19). By c1)

$$
\begin{aligned}
J_n &= n^{-1} \sum_{j=0:|j-q_0|\geq K}^{\tilde{n}-1} \psi_n(\lambda_j) + O(n^{\alpha'-1}) \\
&= n^{-1} \sum_{j=0:|j-q_0|\geq K}^{\tilde{n}-1} \int_j^{j+1} (\psi_n(2\pi[\lambda]/n) - \psi_n(2\pi\lambda/n)) d\lambda
\end{aligned}
$$

$$
+ n^{-1} \int_0^{\tilde{n}} 1_{\{|\lambda-q_0|\geq K\}} \psi_n(2\pi\lambda/n) d\lambda + O(n^{\alpha'-1}) =: i_{n,1} + i_{n,2}. \tag{5.20}
$$

By (4.5)

$$
i_{n,1} = O\left(n^{-2} \sum_{j=0:|j-q_0|\geq K}^{\tilde{n}} (|j-q_0|/n)^{-\gamma'}\right) = O(n^{-1+\max(0,\gamma'-1)}), \tag{5.21}
$$

whereas by c1), c3) and Lebesgue's dominated convergence theorem

$$
i_{n,2} = (2\pi)^{-1} \int_{0:|\lambda-\omega_0|\geq 2\pi K/n}^{\pi} \psi_n(\lambda) d\lambda = (2\pi)^{-1} \int_0^{\pi} \psi(\lambda) d\lambda + o(1). \tag{5.22}
$$

(5.20) - (5.22) imply (5.19). □

The following Lemma 5.4 is used in the proof of Lemma 6.1.

Lemma 5.4 *Under the assumptions of Lemma 4.3 and with $Z_{j,n}^{(2)}$ defined by (4.7),*

$$
|Z_{j,n}^{(2)}| \leq C|j-q_0|^{1/2} Z_n + \tilde{Z}_{j,n}^{(2)} \tag{5.23}
$$

where $Z_n = n^{-1/2}|\sum_{t=1}^n (\varepsilon_t^2 - \sigma_0^2)|$ and $\tilde{Z}_{j,n}^{(2)} = 2\sigma_0^{-2}|\sum_{1\leq s<t\leq n} c_{t-s}\varepsilon_t\varepsilon_s|$, which satisfy

$$
EZ_n^2 \leq C, \quad E(\tilde{Z}_{j,n}^{(2)})^4 \leq C|j-q_0|^2, \tag{5.24}
$$

uniformly in $j = 0,\ldots,\tilde{n}$ $(j \neq q_0)$.

PROOF. Set $b_l = 1_{\{q_0+1\leq l\leq j\}}$ if $j > q_0$; $b_l = 1_{\{j+1\leq l\leq q_0-1\}}$ if $j < q_0$. By (4.7), similarly to (5.3)

$$
Z_{j,n}^{(2)} = \sum_{l=0}^{\tilde{n}} b_l \xi_{l,n}^{(2)} = \sigma_0^2 \sum_{t=1}^n z_{t,n} \quad \text{where} \quad z_{t,n} = \frac{2\varepsilon_t}{\sigma_0^2} \sum_{s=1}^{t-1} c_{t-s}\varepsilon_s + c_0\left(\frac{\varepsilon_t^2}{\sigma_0^2} - 1\right).
$$

Since $c_0 = n^{-1/2} n_1(b) = n^{-1}\sum'_{j(q_0)} b_j = n^{-1}|j-q_0| \leq n^{-1/2}|j-q_0|^{1/2}$, (5.23) follows. Clearly by A.5, $EZ_n^2 \leq C$. On the other hand

$$
E(\tilde{Z}_{j,n}^{(2)})^4 = 16 \prod_{i=1}^4 \sum_{1\leq s_i<t_i\leq n} c_{t_i-s_i} E[\varepsilon_{t_1}\varepsilon_{s_1}\ldots\varepsilon_{t_4}\varepsilon_{s_4}].
$$

Since the number of equal indices in the set $\{t_1, s_1, \ldots, t_4, s_4\}$ does not exceed 4, by Assumption A.5 it follows $|E[\varepsilon_{t_1}\varepsilon_{s_1}\ldots\varepsilon_{t_4}\varepsilon_{s_4}]| \leq C$. Moreover, by A.5, the inequality

21

$E[\varepsilon_{t_1}\varepsilon_{s_1}\ldots\varepsilon_{t_4}\varepsilon_{s_4}] \neq 0$ can hold only if any t_i, s_i are repeated in $\{t_1, s_1, \ldots, t_4, s_4\}$ at least twice. Hence, applying the Cauchy inequality we obtain that

$$
\begin{aligned}
E(\tilde{Z}_{j,n}^{(2)})^4 &\leq C\prod_{i=1}^{4}\Big(\sum_{1\leq s_i<t_i\leq n}c_{t_i-s_i}^2\Big)^{1/2} = C\Big(\sum_{1\leq s<t\leq n}c_{t-s}^2\Big)^2 \\
&\leq C(n_2(b) + n_1(b)^2)^2 \leq Cn_2(b)^2 = 4|j-q_0|^2,
\end{aligned}
$$

by (5.10), which proves (5.24). $\qquad\qquad\square$

6 Proof of (3.17)

Lemma 6.1 *Under the assumptions of Theorem 3.1, there exists $\eta > 0$ such that as $n \to \infty$,*

$$
E\sup_{\psi\in\Psi_K} u_n(\psi)^{-1}|U_n(\psi)| \leq CK^{-\eta}
$$

where $C < \infty$ does not depend on n and K.

PROOF. From (3.16) we may write $U_n(\psi) = \sum'_{j(q_0)}b_j(\psi)\xi_{j,n}^{(1)}$ where

$$
b_j(\psi) = \frac{k_j(\psi_0)}{k_j(\psi)} - 1, \quad \xi_{j,n}^{(1)} = \frac{2\pi I_j}{\sigma_0^2 k_j(\psi_0)} - 1. \tag{6.1}
$$

By summation by parts

$$
|U_n(\psi)| \leq \sum_{j=-1:j\neq q_0,q_0-1}^{\tilde{n}}|b_j(\psi) - b_{j+1}(\psi)||Z_{j,n}^{(1)}|
$$

where $b_{-1}(\psi) = b_{\tilde{n}+1}(\psi) = 0$ and $Z_{j,n}^{(1)}$ is defined by (4.7). Then $|U_n(\psi)| \leq U_n^{(1)}(\psi) + U_n^{(2)}(\psi)$, where

$$
U_n^{(1)}(\psi) = \sum_{j=0:j\neq q_0,q_0-1}^{\tilde{n}-1}|b_j(\psi) - b_{j+1}(\psi)||Z_{j,n}^{(1)}|,
$$

$$
U_n^{(2)}(\psi) = 1_{\{q_0<\tilde{n}\}}|b_{\tilde{n}}(\psi)Z_{\tilde{n},n}^{(1)}| + 1_{\{q_0>0\}}|b_0(\psi)Z_{-1,n}^{(1)}|.
$$

It suffices to show that

$$
E\sup_{\psi\in\Psi_K} u_n(\psi)^{-1}|U_n^{(i)}(\psi)| \leq CK^{-\eta}, \quad i = 1, 2. \tag{6.2}
$$

We begin with $i = 2$. We have

$$
\begin{aligned}
&E\Big(\sup_{\psi\in\Psi_K} u_n(\psi)^{-1}U_n^{(2)}(\psi)\Big) \\
&\leq \sup_{\psi\in\Psi_K} u_n(\psi)^{-1}\Big[1_{\{q_0<\tilde{n}\}}|b_{\tilde{n}}(\psi)| + 1_{\{q_0>\tilde{n}\}}|b_0(\psi)|\Big]\Big(E|Z_{\tilde{n},n}^{(1)}| + E|Z_{-1,n}^{(1)}|\Big) \leq CK^{-1/2}
\end{aligned}
$$

since, by (6.18) below, $1_{\{q_0 < \tilde{n}\}} |b_{\tilde{n}}(\psi)| + 1_{\{q_0 > 0\}} |b_0(\psi)| \leq C(\|\theta - \theta_0\| + |\lambda_{\bar{q}}|^{1/2}) \leq Cn^{-1/2} u_n(\psi)^{1/2}$ and, by Lemma 4.3, $E|Z_{\tilde{n},n}^{(1)}| + E|Z_{-1,n}^{(1)}| \leq Cn^{1/2}$.

Next we show (6.2) for $i = 1$. Set $A = \{\psi : \|\theta - \theta_0\| \leq (\log n)^{-1}\}$ and $B = \{\psi : \|\theta - \theta_0\| > (\log n)^{-1}\}$. Since by Lemma 6.3 below, $|b_j(\psi) - b_{j+1}(\psi)| \leq C(\beta_{j,1}(\psi) + \beta_{j,2}(\psi))$, where

$$\beta_{j,1}(\psi) = \Big(\frac{\|\theta - \theta_0\| + |\lambda_{\bar{q}}|^{1/2}}{|j - q_0|_+} + |\bar{q}|_+^{1-\varepsilon} |j - q_0|_+^{-(2-\varepsilon)}\Big) 1_A + n^{1-\varepsilon} |j - q_0|_+^{-(2-\varepsilon)} 1_B, \quad (6.3)$$

$$\beta_{j,2}(\psi) = |\bar{q}|_+ |j - q_0|_+^{-1} |j - q|_+^{-1} 1_A + n^{1-\varepsilon} |j - q_0|_+^{-(1-\varepsilon)} |j - q|_+^{-1} 1_B \quad (6.4)$$

for $j = 0, \ldots, \tilde{n} - 1$ ($j \neq q_0, q_0 - 1$), where $0 < \varepsilon < 1/2$ is an arbitrarily small positive number, it follows that

$$E\Big(\sup_{\psi \in \Psi_K} u_n(\psi)^{-1} |U_n^{(1)}(\psi)|\Big) \leq C\Big(E \sum_{j=0: j \neq q_0, q_0-1}^{\tilde{n}-1} \sup_{\psi \in \Psi_K} u_n(\psi)^{-1} \beta_{j,1}(\psi) |Z_{j,n}^{(1)}|$$

$$+ E \sum_{j=0: j \neq q_0, q_0-1}^{\tilde{n}-1} \sup_{k=0,\ldots,\tilde{n}} \sup_{\theta: \psi=(\theta,\lambda_k) \in \Psi_K} u_n(\psi)^{-1} \beta_{j,2}(\psi) |Z_{j,n}^{(1)}|\Big) =: C(A_n + R_n). \quad (6.5)$$

Thus the proof is completed if we show that both A_n and R_n in (6.5) are bounded by $CK^{-\eta}$. First, since by Lemma 4.3, $E|Z_{j,n}^{(1)}| \leq C|j - q_0|^{1/2}$, then

$$A_n \leq C \sum_{j=0: j \neq q_0, q_0-1}^{\tilde{n}-1} \beta_j^* |j - q_0|^{1/2}, \quad (6.6)$$

where $\beta_j^* = \sup_{\psi: \psi=(\theta,\lambda_q) \in \Psi_K} \{u_n^{-1}(\psi) \beta_{j,1}(\psi)\}$. To estimate β_j^*, note that for $\psi \in \Psi_K$, $u_n(\psi) \equiv n\|\theta - \theta_0\|^2 + |\bar{q}| \geq \max(K, n\|\theta - \theta_0\|^2, |\bar{q}|)$. Then

$$u_n(\psi)^{-1}(\|\theta - \theta_0\| + |\lambda_{\bar{q}}|_+^{1/2}) \leq CK^{-1/2} n^{-1/2}, \quad (6.7)$$

and

$$u_n(\psi)^{-1} |\bar{q}|^{1-\varepsilon} \leq K^{-\varepsilon} \text{ if } \psi \in A, \quad u_n(\psi)^{-1} \leq n^{-1} \log^2 n \text{ if } \psi \in B. \quad (6.8)$$

Using (6.7)-(6.8) in the definition (6.3) of $\beta_{j,1}(\psi)$ we obtain that

$$\beta_j^* \leq C(K^{-1/2} n^{-1/2} |j - q_0|^{-1} + K^{-\varepsilon} |j - q_0|^{-2+\varepsilon})$$

so that by (6.6)

$$A_n \leq \Big(K^{-1/2} n^{-1/2} \sum_{j=0: j \neq q_0}^{\tilde{n}-1} |j - q_0|^{-1/2} + K^{-\varepsilon} \sum_{j=0: j \neq q_0}^{\tilde{n}-1} |j - q_0|^{-3/2+\varepsilon}\Big) \leq CK^{-\varepsilon}.$$

It remains to examine R_n. Denoting

$$\tilde{q} = \arg\max_{q=0,\ldots,\tilde{n}} \sum_{j=0: j \neq q_0, q_0-1}^{\tilde{n}-1} \beta_{j,q}^* |Z_{j,n}^{(1)}|, \quad \beta_{j,q}^* = \sup_{\theta: \psi=(\theta,\lambda_q) \in \Psi_K} u_n^{-1}(\psi) \beta_{j,2}(\psi)$$

23

we observe that
$$R_n \leq E\Big[\sum_{q=0}^{\tilde{n}-1} 1_{\{\tilde{q}=q\}} \sum_{j=0:j\neq q_0,q_0-1}^{\tilde{n}-1} \beta_{j,q}^* |Z_{j,n}^{(1)}|\Big].$$

Since (4.6) - (4.7) and (5.23) imply $|Z_{j,n}^{(1)}| \leq |Z_{j,n}^{(2)}| + |Z_{j,n}^{(3)}| \leq C|j-q_0|^{1/2} Z_n + \tilde{Z}_{j,n}^{(2)} + |Z_{j,n}^{(3)}|$ we obtain
$$R_n \leq C\Big(ES_{n,1}' + E\sum_{q=0}^{\tilde{n}-1}[1_{\{\tilde{q}=q\}} S_{n,2}(q)] + E\sum_{q=0}^{\tilde{n}-1}[1_{\{\tilde{q}=q\}} S_{n,3}(q)]\Big), \tag{6.9}$$

where
$$S_{n,1}' = Z_n \sup_{q=0,\ldots,\tilde{n}} B(q), \quad B(q) = \sum_{j=0:j\neq q_0,q_0-1}^{\tilde{n}-1} \beta_{j,q}^* |j-q_0|^{1/2}, \tag{6.10}$$

$$S_{n,2}(q) = \sum_{j=0:j\neq q_0,q_0-1}^{\tilde{n}-1} \beta_{j,q}^* |\tilde{Z}_{j,n}^{(2)}|, \quad S_{n,3}(q) = \sum_{j=0:j\neq q_0,q_0-1}^{\tilde{n}-1} \beta_{j,q}^* |Z_{j,n}^{(3)}|.$$

It remains to show that the terms on the right hand side of (6.9) are bounded by $CK^{-\eta}$. We first bound
$$ES_{n,1}' \leq C \max_q B(q) E|Z_n| \leq C \max_q B(q).$$

We estimate now $B(q)$. Using (6.7) - (6.8) in (6.4), we obtain
$$u_n^{-1}(\psi)\beta_{j,2}(\psi) \leq C\Big(\frac{K^{-\varepsilon}|\bar{q}|_+^\varepsilon}{|j-q_0|_+|j-q|_+} + \frac{n^{-\varepsilon}(\log n)^2}{|j-q_0|_+^{1-\varepsilon}|j-q|_+}\Big) =: \nu_j(q). \tag{6.11}$$

Thus $\beta_{j,q}^* \leq \nu_j(q)$ and
$$\begin{aligned}
B(q) &\leq C \sum_{j=0:j\neq q_0,q_0-1}^{\tilde{n}-1} \nu_j(q)|j-q_0|^{1/2} \\
&\leq C\{K^{-\varepsilon}|\bar{q}|^\varepsilon \sum_{j=0}^{\tilde{n}} |j-q_0|_+^{-1/2}|j-q|_+^{-1} + n^{-\varepsilon}(\log n)^2 \sum_{j=0}^{\tilde{n}} |j-q_0|_+^{-1/2+\varepsilon}|j-q|_+^{-1}\} \\
&\leq C\{K^{-\varepsilon}|\bar{q}|_+^{-1/2+2\varepsilon} + n^{-\varepsilon}(\log n)^2|\bar{q}|_+^{-1/2+2\varepsilon}\} \leq CK^{-\varepsilon}|\bar{q}|_+^{-1/2+2\varepsilon} \tag{6.12}
\end{aligned}$$

for sufficiently large n. Therefore
$$ES_{n,1}' \leq CK^{-\varepsilon}. \tag{6.13}$$

To bound the 2nd and 3nd terms in (6.9) set $p_q = \text{Prob}\,(\tilde{q}=q)$. By Hölder's inequality it follows that
$$\begin{aligned}
&\sum_{q=0}^{\tilde{n}}\Big(E[1_{\{\tilde{q}=q\}} S_{n,2}(q)] + E[1_{\{\tilde{q}=q\}} S_{n,3}(q)]\Big) \leq \sum_{q=0}^{\tilde{n}} p_q^{3/4}\Big((ES_{n,2}(q)^4)^{1/4} + p_q^{1/2}(ES_{n,3}(q)^2)^{1/2}\Big) \\
&\leq (\sum_{q=0}^{\tilde{n}} p_q)^{3/4}\Big(\sum_{q=0}^{\tilde{n}} ES_{n,2}(q)^4\Big)^{1/4} + (\sum_{q=0}^{\tilde{n}} p_q)^{1/2}\Big(\sum_{q=0}^{\tilde{n}} ES_{n,3}(q)^2\Big)^{1/2} \\
&= \Big(\sum_{q=0}^{\tilde{n}} ES_{n,2}(q)^4\Big)^{1/4} + \Big(\sum_{q=0}^{\tilde{n}} ES_{n,3}(q)^2\Big)^{1/2}
\end{aligned}$$

24

since $\sum_{q=0}^{\tilde{n}} p_q = 1$. To complete the proof it suffices to show that for some $\eta > 0$,

$$\sum_{q=0}^{\tilde{n}} ES_{n,2}(q)^4 \leq CK^{-\eta}, \quad \sum_{q=0}^{\tilde{n}} ES_{n,3}(q)^2 \leq CK^{-\eta}. \tag{6.14}$$

By Lemma 5.4, $E(\tilde{Z}_{j,n}^{(2)})^4 \leq C|j - q_0|^2$. Therefore

$$E|\tilde{Z}_{j_1,n}^{(2)} \dots \tilde{Z}_{j_4,n}^{(2)}| \leq (E(\tilde{Z}_{j_1,n}^{(2)})^4 \dots E(\tilde{Z}_{j_4,n}^{(2)})^4)^{1/4} \leq C(\prod_{i=1}^{4} |j_i - q_0|^2)^{1/4}$$

and

$$E(S_{n,2}(q))^4 = \sum_{j_1,\dots,j_4=0:j_l \neq q_0,q_0-1}^{\tilde{n}-1} \beta_{j_1,q}^* \dots \beta_{j_4,q}^* E|\tilde{Z}_{j_1,n}^{(2)} \dots \tilde{Z}_{j_4,n}^{(2)}| \leq CB(q)^4. \tag{6.15}$$

Hence, (6.15) and (6.12) imply

$$\sum_{q=0}^{\tilde{n}} ES_{n,2}(q)^4 \leq CK^{-4\varepsilon} \sum_{q=0}^{\tilde{n}} |q - q_0|_+^{8\varepsilon-2} \leq CK^{-4\varepsilon}.$$

By Lemma 4.3, $E(Z_{j,n}^{(3)})^2 \leq C|j - q_0|^{2p}$ with $p \in (1/3, 1/2)$. Therefore

$$E|Z_{j_1,n}^{(3)} Z_{j_2,n}^{(3)}| \leq (E(Z_{j_1,n}^{(3)})^2 E(Z_{j_2,n}^{(3)})^2)^{1/2} \leq C(|j_1 - q_0||j_2 - q_0|)^p$$

and

$$ES_{n,3}^2(q) = \sum_{j_1,j_2=0:j_1,j_2 \neq q_0,q_0-1}^{\tilde{n}-1} \beta_{j_1,q}^* \beta_{j_2,q}^* E|Z_{j_1,n}^{(3)} Z_{j_2,n}^{(3)}| \leq C\Big(\sum_{j=0:j \neq q_0,q_0-1}^{\tilde{n}-1} \beta_{j,q}^* |j - q_0|^p \Big)^2.$$

Since $\beta_{j,q}^* \leq \nu_j(q)$, this and (6.11) imply

$$\begin{aligned}
ES_{n,3}(q)^2 &\leq C\Big\{ K^{-\varepsilon}|\bar{q}|_+^\varepsilon \sum_{j=0}^{\tilde{n}} |j - q_0|_+^{-1+p}|j - q|_+^{-1} + n^{-\varepsilon}(\log n)^2 \sum_{j=0}^{\tilde{n}} |j - q_0|_+^{-1+p+\varepsilon}|j - q|_+^{-1} \Big\}^2 \\
&\leq C(K^{-\varepsilon}|\bar{q}|_+^{2\varepsilon+p-1} + n^{-\varepsilon}(\log n)^2|\bar{q}|_+^{2\varepsilon+p-1})^2 \leq CK^{-2\varepsilon}|\bar{q}|_+^{4\varepsilon+2p-2}.
\end{aligned}$$

Hence

$$\sum_{q=0}^{\tilde{n}} ES_{n,3}(q)^2 \leq CK^{-2\varepsilon} \sum_{q=0}^{\tilde{n}} |q - q_0|_+^{4\varepsilon+2p-2} \leq CK^{-2\varepsilon}$$

since $4\varepsilon + 2p - 2 < -1$ for $p \in (1/3, 1/2)$ and sufficiently small $\varepsilon > 0$. Thus (6.14) holds. $\square$

Lemma 6.2 *Let* $V_n(\psi) = 1 - 2\pi\sigma_0^{-2}k(\lambda_{q_0}; \theta, \lambda_q)^{-1}I_{q_0}$. *Under the assumptions of Theorem 3.1, there exists* $\eta > 0$ *such that*

$$E \sup_{\psi \in \Psi_K} |u_n^{-1}(\psi)V_n(\psi)| \leq CK^{-\eta} \tag{6.16}$$

where $0 < C < \infty$ *does not depend on* K, n.

PROOF. By A.1′ or A.1, $k_{q_0}(\psi)^{-1} \leq C|\lambda_{\bar{q}}|^{\alpha}$. If $|\alpha - \alpha_0| \leq (\log n)^{-1}$ and $q \neq q_0$ then $|\lambda_{\bar{q}}|^{\alpha - \alpha_0} \leq C$, and $k_{q_0}(\psi)^{-1} \leq C|\lambda_{\bar{q}}|^{\alpha_0} \leq Cu_n(\psi)^{\alpha_0} n^{-\alpha_0}$. If $|\alpha - \alpha_0| \geq (\log n)^{-1}$, then $u_n(\psi) \geq n\|\theta - \theta_0\|^2 \geq n\log^{-2} n$. Hence for $\psi \in \Psi_K$,

$$u_n(\psi)^{-1}|V_n(\psi)| \leq C(u_n(\psi)^{\alpha_0 - 1} n^{-\alpha_0} + n^{-1}\log^2 n)(I_{q_0} + 1) \leq CK^{-1+\alpha_0}(n^{-\alpha_0} I_{q_0} + 1).$$

Since by standard arguments $E[n^{-\alpha_0} I_{q_0}] \leq C$, it follows that (6.16) holds with $\eta = 1 - \alpha_0$.
□

Lemma 6.3 *Let $b_j(\psi)$ be defined by* (6.1). *Under the assumptions of Lemma 6.1, there exists $0 < \varepsilon < 1/2$, such that as $n \to \infty$, uniformly in $j = 0, \ldots, \tilde{n} - 1$ ($j \neq q_0, q_0 - 1$),*

$$|b_j(\psi) - b_{j+1}(\psi)| \leq \tag{6.17}$$

$$C \begin{cases} \frac{\|\theta - \theta_0\| + |\lambda_{\bar{q}}|^{1/2}}{|j - q_0|} + |\bar{q}|_+^{1-\varepsilon}|j - q_0|^{-(2-\varepsilon)} + |\bar{q}|_+|j - q_0|^{-1}|j - q|_+^{-1} & \text{if} \quad \|\theta - \theta_0\| \leq (\log n)^{-1}, \\ n^{1-\varepsilon}|j - q_0|^{-(2-\varepsilon)} + n^{1-\varepsilon}|j - q_0|^{-(1-\varepsilon)}|j - q|_+^{-1} & \text{if} \quad \|\theta - \theta_0\| \geq (\log n)^{-1}, \end{cases}$$

where $0 < C < \infty$ above does not depend on n, ψ, and moreover,

$$1_{\{\omega_0 < \pi\}}|b_{\tilde{n}}(\psi)| + 1_{\{\omega_0 > 0\}}|b_0(\psi)| \leq C(\|\theta - \theta_0\|^2 + |\lambda_{\bar{q}}|^{1/2}). \tag{6.18}$$

PROOF. We prove first (6.17) in case $\|\theta - \theta_0\| \leq (\log n)^{-1}$. Write $|b_j(\psi) - b_{j+1}(\psi)| = |v(\lambda_j; \psi, \psi_0) - v(\lambda_{j+1}; \psi, \psi_0)|$ where

$$v(\lambda; \psi, \psi_0) = k(\lambda; \psi_0)k^{-1}(\lambda; \psi). \tag{6.19}$$

Let $j \neq q, q \pm 1$. By the mean value theorem

$$|v(\lambda_j; \psi, \psi_0) - v(\lambda_{j+1}; \psi, \psi_0)| \leq (2\pi/n) \sup_{\xi \in (\lambda_j, \lambda_{j+1})} |(d/d\xi)v(\xi; \psi, \psi_0)|. \tag{6.20}$$

Note that
$$(\partial/\partial\xi)v(\xi; \psi, \psi_0) = v(\xi; \psi, \psi_0)|(\partial/\partial\xi)\log v(\xi; \psi, \psi_0)|. \tag{6.21}$$
By inequality (6.28) of Lemma 6.4 below,

$$\sup_{\xi \in (\lambda_j, \lambda_{j+1})} n^{-1}|(\partial/\partial\xi)\log v(\xi; \psi, \psi_0)| \leq Cn^{-1}|(\partial/\partial\xi)\log v(\lambda_j; \psi, \psi_0)|$$

$$\leq C\Big((\|\theta - \theta_0\| + |\lambda_{\bar{q}}|^{1/2})|j - q_0|^{-1} + |\bar{q}||j - q_0|^{-1}|j - q|^{-1}\Big). \tag{6.22}$$

Next we bound $v(\xi; \psi, \psi_0)$. Relation (6.29) of Lemma 6.4 below implies that for $\xi \in (\lambda_j, \lambda_{j+1})$ under Assumption A.1′ or A.1 and $0 < \omega_0 < \pi$,

$$v(\xi; \psi, \psi_0) \leq C|\lambda_{j-q}|^{\alpha}|\lambda_{j-q_0}|^{-\alpha_0} \leq C|\frac{j-q}{j-q_0}|^{\alpha_0}|\lambda_{j-q}|^{\alpha - \alpha_0} \leq C|1 - \frac{\bar{q}}{j-q_0}|^{\alpha_0}, \tag{6.23}$$

26

and under Assumption A.1 and $\omega_0 = 0, \pi$,

$$v(\xi; \psi, \psi_0) \leq C|\lambda_{j-q_0}^2 - \lambda_{\bar{q}}^2|^\alpha |\lambda_{j-q_0}^2|^{-\alpha_0} \leq C|1 - (\bar{q}/(j-q_0))^2|^{\alpha_0} \qquad (6.24)$$

because $|\alpha - \alpha_0| \leq (\log n)^{-1}$ implies that $|\lambda_{j-q}|^{\alpha-\alpha_0} \leq C$ and $|\lambda_{j-q_0}^2 - \lambda_{\bar{q}}^2|^{\alpha-\alpha_0} \leq C$.

If $|\bar{q}|/|j-q_0| \leq 2$, then (6.23)-(6.24) imply that $v(\lambda_j; \psi, \psi_0) \leq C$, and from (6.20) - (6.22) it follows that

$$|b_j(\psi) - b_{j+1}(\psi)| \leq C\Big((||\theta - \theta_0|| + |\lambda_{\bar{q}}|^{1/2})|j - q_0|^{-1} + |\bar{q}||j - q_0|^{-1}|j - q|^{-1}\Big). \qquad (6.25)$$

If $|\bar{q}|/|j-q_0| > 2$, then (6.23)-(6.24) yield $v(\lambda_j; \psi, \psi_0) \leq C(|\bar{q}|/|j-q_0|)^{1-\varepsilon}$ for small enough $\varepsilon > 0$, and using (6.20) - (6.22) we obtain that

$$|b_j(\psi) - b_{j+1}(\psi)| \leq C(|\bar{q}|/|j-q_0|)^{-(1-\varepsilon)}\Big((||\theta - \theta_0|| + |\lambda_{\bar{q}}|^{1/2})|j - q_0|^{-1} + |\bar{q}||j - q_0|^{-1}|j - q|^{-1}\Big)$$

$$\leq C|\bar{q}|^{1-\varepsilon}|j - q_0|^{-(2-\varepsilon)} \qquad (6.26)$$

since $||\theta - \theta_0|| + |\lambda_{\bar{q}}|^{1/2} \leq C$ by compactness of Θ, and $|\bar{q}|/|j-q_0| > 2$ implies $|\bar{q}|/|j-q| \leq 2$. From (6.25) - (6.26) we deduce (6.17).

It remains to prove (6.17) in case $||\theta - \theta_0|| \geq (\log n)^{-1}$. From (6.23) and (6.24) it follows that $v(\xi; \psi, \psi_0) \leq C(n/|j - q_0|)^{1-\varepsilon}$ for arbitrarily small $\varepsilon > 0$. Therefore (6.20) - (6.22) imply

$$|b_j(\psi) - b_{j+1}(\psi)| \leq C(n/|j-q_0|)^{1-\varepsilon}\Big((||\theta - \theta_0|| + |\lambda_{\bar{q}}|^{1/2})|j - q_0|^{-1} + |\bar{q}||j - q_0|^{-1}|j - q|^{-1}\Big)$$

$$\leq C(n^{1-\varepsilon}|j - q_0|^{-(2-\varepsilon)} + n^{1-\varepsilon}|j - q_0|^{-(1-\varepsilon)}|j - q|^{-1}) \qquad (6.27)$$

observing that $||\theta - \theta_0|| + |\lambda_{\bar{q}}|^{1/2}$ is bounded and $|\bar{q}||j-q_0|^{-1}|j-q|^{-1} \leq 2(|j-q_0|^{-1} + |j-q|^{-1})$. From (6.27) we deduce (6.17).

We complete the proof by showing that (6.17) holds for $j = q, q \pm 1$ but $j \neq q_0, q_0 - 1$. Note that $|b_j(\psi) - b_{j+1}(\psi)| \leq k_j(\psi_0)k_j(\psi)^{-1} + k_{j+1}(\psi_0)k_{j+1}(\psi)^{-1}$. If $||\theta - \theta_0|| \leq (\log n)^{-1}$, then similarly as in (6.23) - (6.24) it follows that $|b_j(\psi) - b_{j+1}(\psi)| \leq C$ and (6.17) holds.

If $||\theta - \theta_0|| \geq (\log n)^{-1}$, then $k(\lambda_{q\pm 1}; \psi_0)k(\lambda_{q\pm 1}; \psi_q)^{-1} \leq C|\lambda_{\bar{q}}|^{-\alpha_0} \leq (n/|\bar{q}|_+)^{1-\varepsilon}$ for small enough $\varepsilon > 0$. Therefore $|b_{n,j}(\psi) - b_{n,j+1}(\psi)| \leq C(n/|\bar{q}|_+)^{1-\varepsilon}$ which is bounded by the second term in (6.17) for $j = q, q \pm 1$ and $j \neq q_0, q_0 \pm 1$.

Finally, (6.18) follows from relation (6.30) of Lemma 6.4 below. $\qquad \square$

Lemma 6.4 *Let* $v(\lambda; \psi, \psi_0)$ *be given by (6.19) and Assumptions A.1' or A.1 be satisfied. Then uniformly in* $0 < \lambda < \pi$ *such that* $|\lambda - \lambda_q| \geq \pi/n$ *and* $|\lambda - \lambda_{q_0}| \geq \pi/n$,

$$|(\partial/\partial\lambda)\log v(\lambda; \psi, \psi_0)| \leq C\Big(\frac{||\theta - \theta_0|| + |\lambda_{\bar{q}}|^{1/2}}{|\lambda - \lambda_{q_0}|} + |\lambda_{\bar{q}}||\lambda - \lambda_{q_0}|^{-1}|\lambda - \lambda_q|^{-1}\Big) \qquad (6.28)$$

where $C \in (0, \infty)$ *does not depend on* n, ψ, *and*

$$|v(\lambda; \psi, \psi_0)| \asymp \begin{cases} |\lambda - \lambda_q|^\alpha|\lambda - \lambda_{q_0}|^{-\alpha_0} & \text{if} \quad \text{A.1 holds and } 0 < \omega_0 < \pi \text{ or A.1' holds,} \\ |(\lambda - \lambda_{q_0})^2 - \lambda_{\bar{q}}^2|^\alpha|(\lambda - \lambda_{q_0})^2|^{-\alpha_0} & \text{if} \quad \text{A.1 holds and } \omega_0 = 0, \pi \end{cases} \qquad (6.29)$$

where $a_n \asymp b_n$ means that $c_1 \leq a_n/b_n \leq c_2$ as $n \to \infty$ for some $0 < c_1, c_2 < \infty$. Moreover, for any $\delta > 0$

$$|v(\lambda; \psi, \psi_0) - 1| \leq C(||\theta - \theta_0|| + |\lambda_{\bar{q}}|^{1/2}) \quad if \quad |\lambda - \lambda_{q_0}| \geq \delta. \qquad (6.30)$$

PROOF. (6.29) follows from Assumptions A.1$'$ or A.1, taking into account (3.9).

We prove now (6.28). Let A.1$'$ hold. Then $\log k(\lambda; \psi) = -\alpha \log |\lambda - \lambda_q| + \log g(\lambda; \psi)$. Since $(d/dx) \log |x| = 1/x$ $(x \neq 0)$, it follows that $(\partial/\partial\lambda) \log k(\lambda; \psi) = -\alpha(\lambda - \lambda_q)^{-1} + (\partial/\partial\lambda) \log g(\lambda; \psi)$. Therefore

$$(\partial/\partial\lambda) \log(k(\lambda; \psi)/k(\lambda; \psi_0)) = -\alpha(\lambda - \lambda_q)^{-1} + \alpha_0(\lambda - \lambda_{q_0})^{-1} + h(\lambda; \psi) - h(\lambda; \psi_0) \quad (6.31)$$

where $h(\lambda; \psi) = (\partial/\partial\lambda) \log(g(\lambda; \theta, \lambda_q))$. (6.31) implies (6.28) noting that

$$
\begin{aligned}
|\alpha(\lambda - \lambda_q)^{-1} - \alpha_0(\lambda - \lambda_{q_0})^{-1}| &\leq |(\alpha_0 - \alpha)|\lambda - \lambda_{q_0}|^{-1}| + \alpha|\lambda_{\bar{q}}(\lambda - \lambda_q)^{-1}(\lambda - \lambda_{q_0})^{-1}| \\
&\leq ||\theta - \theta_0|| \, |\lambda - \lambda_{q_0}|^{-1} + |\lambda_{\bar{q}}(\lambda - \lambda_q)|^{-1}|\lambda - \lambda_{q_0}|^{-1}
\end{aligned}
$$

and, by Assumption A.1$'$

$$
\begin{aligned}
|h(\lambda; \psi) - h(\lambda; \psi_0)| &\leq C|h(\lambda; \theta, \lambda_q) - h(\lambda; \theta_0, \lambda_q)| + |h(\lambda; \theta_0, \lambda_q) - h(\lambda; \theta_0, \lambda_{q_0})| \\
&\leq C(||\theta - \theta_0|| + |\lambda_{\bar{q}}|^{\beta}) \leq C(||\theta - \theta_0|| + |\lambda_{\bar{q}}|^{1/2})
\end{aligned}
$$

where $\beta > 1/2$ is the Lipschitz parameter.

Suppose that Assumption A.1 holds. By (3.9)

$$\log k(\lambda; \psi) = \begin{cases} -\alpha \log |\lambda - \lambda_q| - \alpha \log |\lambda + \lambda_q| + \log g_1(\lambda; \psi) & \text{if} \quad 0 \leq \lambda < \pi/2, \\ -\alpha \log |\lambda - \lambda_q| - \alpha \log |\lambda + \lambda_q - 2\pi| + \log g_2(\lambda; \psi) & \text{if} \quad \pi/2 \leq \lambda \leq \pi. \end{cases}$$

(6.28) now follows by the same argument as in case of Assumption A.1$'$ once it is observed that $|\lambda + \lambda_q - 2\pi| \geq |\lambda - \lambda_q|$ and $|\lambda + \lambda_q| \geq |\lambda - \lambda_q|$.

Finally, we show (6.30). Let $|\lambda - \lambda_{q_0}| \geq \delta$. Then under A.1$'$ or A.1 $k^{-1}(\lambda; \theta, \omega)$ and $k(\lambda; \psi_0)$ are bounded. Therefore if $|\lambda_{\bar{q}}| \geq \delta/2$, then

$$|v(\lambda; \psi, \psi_0) - 1| = |1 - k(\lambda; \psi_0)k(\lambda; \psi)^{-1}| \leq C \leq C(\delta/2)^{-1/2}(||\theta - \theta_0|| + |\lambda_{\bar{q}}|^{1/2}).$$

If $|\lambda_{\bar{q}}| \leq \delta/2$ then $|\lambda - \lambda_q| \geq |\lambda - \lambda_{q_0}| - |\lambda_{\bar{q}}| \geq \delta/2$ and by A.1$'$ or A.1,

$$
\begin{aligned}
|v(\lambda; \phi, \phi_0) - 1| &= k(\lambda; \theta, \lambda_q)^{-1}|k(\lambda; \theta, \lambda_q) - k(\lambda; \theta_0, \lambda_{q_0})| \\
&\leq C(|k(\lambda; \theta, \lambda_q) - k(\lambda; \theta_0, \lambda_q)| + |k(\lambda; \theta_0, \lambda_q) - k(\lambda; \theta_0, \lambda_{q_0})|) \leq C(||\theta - \theta_0|| + |\lambda_{\bar{q}}|^{1/2}).
\end{aligned}
$$

$\square$

28

7   Proof of (3.18)

Lemma 7.1 *Let $T_n(\psi)$ be given by (3.15). Then under Assumptions **A**, as $n \to \infty$, there exists a constant $c > 0$ such that for sufficiently large $K > 0$,*

$$\inf_{\psi \in \Psi_K} \{u_n(\psi)^{-1} T_n(\psi)\} \geq c \tag{7.1}$$

where c does not depend on K, n and $u_n(\psi) = n\|\theta - \theta_0\|^2 + |\bar{q}|$.

PROOF. Set $L(\lambda; \psi, \psi_0) = -\log(v(\lambda; \psi, \psi_0)) + v(\lambda; \psi, \psi_0) - 1$, where $v(\lambda; \psi, \psi_0)$ is given by (6.19). L is nonnegative for all λ because $-\log x + x - 1$ is, for all $x > 0$. Rewrite

$$T_n(\psi) = \sum_{j(q_0, q)}' \left\{ \frac{k(\lambda_j; \psi_0)}{k(\lambda_j; \psi)} - 1 \right\} = \sum_{j(q_0, q)}' L(\lambda_j; \psi, \psi_0) + \sum_{j(q_0, q)}' \log v(\lambda_j; \psi, \psi_0) =: D_n(\psi) + F_n(\psi).$$

(7.1) follows if we show that as $n \to \infty$,

$$D_n(\psi) \geq c u_n(\psi) \quad (\exists c > 0) \tag{7.2}$$

since by Lemma 7.3 below, $|F_n(\psi)| \leq (c/2) u_n(\psi)$. We show that (7.2) holds in each of the following three, exhaustive, cases

$(\gamma 1)\ \|\theta - \theta_0\|^2 \leq |\lambda_{\bar{q}}| \leq \delta;$ $(\gamma 2)\ |\lambda_{\bar{q}}| + \|\theta - \theta_0\|^2 \geq \delta;$ $(\gamma 3)\ |\lambda_{\bar{q}}| \leq \|\theta - \theta_0\|^2 \leq \delta$

for some $\delta > 0$.

Case $(\gamma 1)$. Then $|\bar{q}| \geq u_n(\psi)/4\pi$ and (7.2) follows if $D_n(\psi) \geq c|\bar{q}|$. To show that note that by Lemma 7.2 below, for $\lambda \in A_q := [\lambda_{q_0} + (3/4)\lambda_{\bar{q}}, \lambda_{q_0} + (7/8)\lambda_{\bar{q}}]$ and θ, q satisfying $(\gamma 1)$,

$$\inf_{\lambda \in A_q} |v(\lambda; \psi, \psi_0) - 1| \geq c_* > 0 \tag{7.3}$$

for small enough δ. Then, since $L \geq 0$

$$\begin{aligned}
D_n(\psi) &\geq \sum_{j=0:\lambda_j \in A_q}^{\tilde{n}} L(\lambda_j, \psi, \psi_0) \\
&\geq (|\bar{q}|/8) \inf_{\lambda \in A_q} L(\lambda; \psi, \psi_0) \geq (|\bar{q}|/8) \inf_{|x-1| \geq c_*} (-\log x + x - 1) \geq c|\bar{q}|.
\end{aligned}$$

Case $(\gamma 2)$. Let $\epsilon > 0$. Under Assumptions A.1$'$ or A.1, $|(\partial/\partial\lambda)L(\lambda; \psi, \psi_0)|$ is bounded uniformly in λ, ψ satisfying $|\lambda - \lambda_q| \geq \epsilon, |\lambda - \lambda_{q_0}| \geq \epsilon$. Therefore by standard arguments (see also the proof of (5.19)),

$$D_n(\psi) \geq \sum_{j:|\lambda_j - \lambda_q| \geq \epsilon, |\lambda_j - \lambda_{q_0}| \geq \epsilon} L(\lambda_j, \psi, \psi_0) = (n/2\pi) d_\epsilon(\psi) + \Delta_n(\psi)$$

29

where

$$d_\epsilon(\psi) = \int_0^\pi 1_{\{|\lambda-\lambda_q|\geq\epsilon,|\lambda-\lambda_{q_0}|\geq\epsilon\}} L(\lambda;\psi,\psi_0)d\lambda, \quad |\Delta_n(\psi)| \leq C < \infty \tag{7.4}$$

and $C = C_*(\epsilon)$ does not depend on n and ψ. By (7.4), to prove (7.2) it suffices to show that

$$d_\epsilon(\psi) \geq c(|\lambda_{\bar{q}}| + ||\theta-\theta_0||^2). \tag{7.5}$$

Since compactness of Θ implies $|\lambda_{\bar{q}}| + ||\theta-\theta_0||^2 \leq C < \infty$ for some $C > 0$, (7.5) follows if

$$d_\epsilon(\psi) \geq c > 0 \quad (\exists c > 0) \tag{7.6}$$

uniformly in ψ satisfying $(\gamma2)$. Set $d(\psi) = \int_0^\pi L(\lambda;\psi,\theta_0,\omega_0)d\lambda$. Then

$$d_\epsilon(\psi) \geq \inf_\psi d(\psi) - \sup_\psi |d_\epsilon(\psi) - d(\psi)|$$

where $\inf_\psi$ and $\sup_\psi$ are taken over the compact set $\mathcal{I}_\delta = \{\psi = (\theta,\omega) : ||\theta-\theta_0||^2 + |\omega-\omega_0| \geq \delta/4\} \cap \Theta \times [0,\pi]$. Under Assumptions A.1$'$ or A.1, $d(\psi)$ is continuous. So it achieves its minimum, $d(\psi^*) > 0$, on $\mathcal{I}_\delta$, since $L \geq 0$ by (3.3) for $||\psi_0 - \psi^*|| \geq \delta$. On the other hand, under Assumptions A.1$'$ or A.1 we can choose $\epsilon > 0$ such that

$$\sup_\psi |d_\epsilon(\psi) - d(\psi)| < d(\psi^*)/2. \tag{7.7}$$

Indeed, if A.1$'$ holds then by (6.29) $L(\lambda;\psi,\psi_0) \leq C(1 + \log|\lambda - \lambda_q| + \log|\lambda - \lambda_{q_0}| + |\lambda - \lambda_{q_0}|^{-\alpha_0})$, and therefore

$$\sup_\psi |d_\epsilon(\psi) - d(\psi)| \leq \sup_{\psi\in\Psi_K} \int_0^\pi 1_{\{|\lambda-\lambda_q|<\epsilon \text{ or } |\lambda-\lambda_{q_0}|<\epsilon\}} L(\lambda;\psi,\psi_0)d\lambda$$

can be made arbitrarily small by choosing sufficiently small ϵ, so that (7.7) holds. Under Assumption A.1, (7.7) follows using a similar argument. Hence (7.6) holds with $c = d(\psi^*)/2$.

Case $(\gamma3)$. As above it suffices to show (7.5). Under Assumptions A.1$'$ or A.1, the Taylor expansion with respect to θ can be applied in (7.4), which leads to

$$d_\epsilon(\psi) = d_\epsilon(\theta_0, \lambda_q) + \nabla d_\epsilon(\theta_0, \lambda_q)(\theta - \theta_0) + (\theta - \theta_0)' M(\theta - \theta_0)$$

$$\geq \nabla d_\epsilon(\theta_0, \lambda_q)(\theta - \theta_0) + (\theta - \theta_0)' M(\theta - \theta_0) \tag{7.8}$$

since $d_\epsilon(\theta_0, \lambda_q) \geq 0$ by $L \geq 0$ and (7.4). The i-th element of the vector $\nabla d_\epsilon(\theta_0, \lambda_q)$ is

$$(\partial/\partial\theta_i)d_\epsilon(\theta_0, \lambda_q) = \int_0^\pi 1_{\{|\lambda-\lambda_q|\geq\epsilon, |\lambda-\lambda_{q_0}|\geq\epsilon\}}(\partial/\partial\theta_i)L(\lambda;\theta_0, \lambda_q, \psi_0)d\lambda,$$

whereas the (j, i)-th element of the $p \times p$ matrix M is

$$(\partial^2/\partial\theta_j\partial\theta_i)d_\epsilon(\theta_*^{(i)}, \lambda_q) = \int_0^\pi 1_{\{|\lambda-\lambda_q|\geq\epsilon, |\lambda-\lambda_{q_0}|\geq\epsilon\}}(\partial^2/\partial\theta_j\partial\theta_i)L(\lambda; \theta_*^{(i)}, \lambda_q, \psi_0)d\lambda$$

with $||\theta_*^{(i)} - \theta_0|| \leq ||\theta - \theta_0|| \leq \delta^{1/2}$ under $(\gamma 3)$. Now if there exist $\epsilon, \delta > 0$ such that

$$(\theta - \theta_0)'M(\theta - \theta_0) \geq c||\theta - \theta_0||^2, \tag{7.9}$$

$$|\nabla d_\epsilon(\theta_0, \lambda_q)(\theta - \theta_0)| \leq c||\theta - \theta_0||^2/2, \tag{7.10}$$

then by (7.8) $d_\epsilon(\psi) \geq (\theta - \theta_0)'M(\theta - \theta_0) - |\nabla d_\epsilon(\theta_0, \lambda_q)(\theta - \theta_0)| \geq c||\theta - \theta_0||^2/2 \geq c(||\theta - \theta_0||^2 + |\lambda_{\bar{q}}|)/4$ by $(\gamma 3)$, and so (7.5) follows. To verify (7.9), note that

$$(\theta - \theta_0)'M(\theta - \theta_0) \geq (\theta - \theta_0)'\Omega(\theta - \theta_0) - |(\theta - \theta_0)'(\Omega - M)(\theta - \theta_0)|.$$

By (3.4), $\Omega > 0$, so $(\theta - \theta_0)\Omega(\theta - \theta_0)' \geq c||\theta - \theta_0||^2$ for some $c > 0$. Under Assumptions A.1$'$ or A.1 the (j, i)-th element of $M - \Omega$ is of form

$$\int_0^\pi 1_{\{|\lambda-\lambda_q|\geq\epsilon, |\lambda-\lambda_{q_0}|\geq\epsilon\}}(\partial^2/\partial\theta_j\partial\theta_i)L(\lambda; \theta_*^{(i)}, \lambda_q, \psi_0)d\lambda - \int_0^\pi (\partial^2/\partial\theta_j\partial\theta_i)L(\lambda; \psi_0, \psi_0)d\lambda$$

$$= \int_0^\pi 1_{\{|\lambda-\lambda_q|\geq\epsilon, |\lambda-\lambda_{q_0}|\geq\epsilon\}}\left((\partial^2/\partial\theta_j\partial\theta_i)L(\lambda; \theta_*^{(i)}, \lambda_q, \psi_0) - (\partial^2/\partial\theta_j\partial\theta_i)L(\lambda; \psi_0, \psi_0)\right)d\lambda$$

$$- \int_0^\pi 1_{\{|\lambda-\lambda_q|<\epsilon \text{ or } |\lambda-\lambda_{q_0}|<\epsilon\}}(\partial^2/\partial\theta_j\partial\theta_i)L(\lambda; \psi_0, \psi_0)d\lambda,$$

where each of the two integrals can be made arbitrarily small by choosing, respectively, ϵ and δ sufficiently small. Thus, there exist $\epsilon > 0$ and $\delta > 0$ such that $||\Omega - M|| \leq c/2$ and

$$(\theta - \theta_0)'M(\theta - \theta_0) \geq c||\theta - \theta_0||^2 - ||\Omega - M|| \cdot ||\theta - \theta_0||^2 \geq (c/2)||\theta - \theta_0||^2,$$

to prove (7.9).

It remains to prove (7.10). By Assumptions A.1$'$ or A.1, the partial derivative $\nabla_\theta d_\epsilon(\theta, \lambda_q)$ satisfies a Lipschitz condition in λ of order $\beta > 1/2$. Note that $\nabla_\theta L(\lambda; \psi_0, \psi_0) = 0$ and thus $\nabla_\theta d_\epsilon(\theta_0, \lambda_{q_0}) = 0$. Therefore

$$
\begin{aligned}
|\nabla_\theta d_\epsilon(\theta_0, \lambda_q)(\theta - \theta_0)| &\leq ||\nabla d_\epsilon(\theta_0, \lambda_q) - \nabla d_\epsilon(\theta_0, \lambda_{q_0})|| \cdot ||\theta - \theta_0|| \\
&\leq C(\epsilon)|\lambda_q - \lambda_{q_0}|^\beta ||\theta - \theta_0|| \leq C(\epsilon)||\theta - \theta_0||^{1+2\beta} \leq C(\epsilon)\delta^{(2\beta-1)/2}||\theta - \theta_0||^2
\end{aligned}
$$

since $|\lambda_q - \lambda_{q_0}|^\beta \leq ||\theta - \theta_0||^{2\beta} \leq \delta^\beta$ by $(\gamma 3)$. Choosing δ such that $C(\epsilon)\delta^{(2\beta-1)/2} < c/2$ we obtain (7.10). $\square$

Lemma 7.2 *If* $(\gamma 1)$ *holds then* (7.3) *is valid.*

31

PROOF. It suffices to show that uniformly in ψ satisfying $(\gamma 1)$,

$$\sup_{\lambda \in A_q} v(\lambda; \psi, \psi_0) \leq c < 1. \tag{7.11}$$

Let A.1′ hold. Then $v(\lambda; \psi, \psi_0) = |\lambda - \lambda_q|^\alpha |\lambda - \lambda_{q_0}|^{-\alpha_0} g(\lambda; \psi_0) g(\lambda; \psi)^{-1}$ and for $\lambda \in A_q$,

$$v(\lambda; \psi, \psi_0) \leq |\lambda_{\bar{q}}/4|^\alpha |3\lambda_{\bar{q}}/4|^{-\alpha_0} g(\lambda; \psi_0)/g(\lambda; \psi). \tag{7.12}$$

Since $(\gamma 1)$ implies $|\alpha - \alpha_0|^2 \leq \|\theta - \theta_0\|^2 \leq |\lambda_{\bar{q}}| \leq \delta$, then $|\lambda_{\bar{q}}/4|^{\alpha - \alpha_0} = \exp((\alpha - \alpha_0) \log(|\lambda_{\bar{q}}/4|)) = 1 + O(|\alpha - \alpha_0| \log(|\alpha - \alpha_0|^2)) = 1 + O(\delta^{1/4})$ and

$$|\lambda_{\bar{q}}/4|^\alpha |3\lambda_{\bar{q}}/4|^{-\alpha_0} = (1/3)^{\alpha_0} |\lambda_{\bar{q}}/4|^{\alpha - \alpha_0} = (1/3)^{\alpha_0} (1 + O(\delta^{1/4})). \tag{7.13}$$

Under A.1′,

$$g(\lambda; \psi_0) g(\lambda; \psi)^{-1} = 1 + O(\|\theta - \theta_0\| + |\lambda_{\bar{q}}|^\gamma) = 1 + O(\delta^{1/2}). \tag{7.14}$$

Therefore from (7.12)-(7.14) it follows that $v(\lambda; \psi, \psi_0) \leq (1/3)^{\alpha_0} (1 + O(\delta^{1/4})) < 1$ assuming that δ is sufficiently small. Hence (7.11) holds.

Suppose that A.1 holds. Let $0 < \omega_0 < \pi$. If δ is sufficiently small then the Gegenbauer spectral density (3.1) can be written in the form (3.7), and (7.11) can be shown using the same argument as above.

Let $\omega_0 = 0$. Then by (3.9), $v(\lambda; \psi, \psi_0) = |\lambda^2 - \lambda_q^2|^\alpha |\lambda|^{-2\alpha_0} g_1(\lambda; \theta, \lambda_{q_0})/g_1(\lambda; \theta, \lambda_q)$. Since for $q_0 = 0$ we have $\bar{q} = q$, then $\lambda \in A_q = [(3/4)\lambda_q, (7/8)\lambda_q]$. Similarly to (7.14), it follows that $g_1(\lambda; \theta, \lambda_q)/g_1(\lambda; \theta_0, \lambda_{q_0}) = 1 + O(\delta^{1/2})$. Therefore

$$v(\lambda; \psi, \psi_0) \leq |\lambda_{\bar{q}}^2 - (3/4)\lambda_{\bar{q}}^2|^\alpha |(3/4)\lambda_{\bar{q}}|^{-2\alpha_0} (1 + O(\delta^{1/2}))$$

$$= (7/9)^{\alpha_0} |(7/16)\lambda_{\bar{q}}^2|^{\alpha - \alpha_0} (1 + O(\delta^{1/2})). \tag{7.15}$$

The same argument as that used to prove (7.13) implies that $|(7/16)\lambda_{\bar{q}}^2|^{\alpha - \alpha_0} = 1 + O(\delta^{1/4})$. Hence $v(\lambda; \psi, \psi_0) \leq (7/9)^{\alpha_0} (1 + O(\delta^{1/4})) < 1$ and (7.11) holds.

Let $\omega_0 = \pi$. Then $\lambda_{q_0} = \pi$ and by (3.9), $v(\lambda; \psi, \psi_0) = |(\lambda - \pi)^2 - (\lambda_q - \pi)^2|^\alpha |\lambda - \pi|^{-2\alpha_0} g_2(\lambda; \theta, \lambda_q)/g_2(\lambda; \theta_0, \lambda_{q_0})$. Using a similar argument as in case $\omega_0 = 0$ we derive that for $\lambda \in A_q$, $v(\lambda; \psi, \psi_0)$ satisfies (7.15) which proves (7.11). $\square$

Lemma 7.3 *Under the assumptions of Lemma 7.1, as $n \to \infty$,*

$$\sup_{\psi \in \Psi_K} u_n(\psi)^{-1} \Big| \sum_{j(q,q_0)}{}' \log(k_j(\psi_0)/k_j(\psi)) \Big| \leq C K^{-1/2}. \tag{7.16}$$

PROOF. Let $v(\lambda; \psi, \psi_0)$ be given by (6.19). Then by (2.2), $\int_0^{n/2} \log(v(2\pi x/n; \psi, \psi_0)) dx = 0$. Therefore

$$\Big| \sum_{j(q,q_0)}{}' \log(k_j(\psi_0)/k_j(\psi)) \Big| \leq d_{n,1}(\psi) + d_{n,2}(\psi) + d_{n,3}(\psi), \tag{7.17}$$

32

where, setting $J_q = \{x \geq 0 : x \notin [q-1, q+1] \cup [q_0-1, q_0+1]\}$, $J_q^c = [0, n/2] \setminus J_q$,

$$
\begin{aligned}
d_{n,1}(\psi) &= \int_0^{n/2} \left| \log v(2\pi[x]/n; \psi, \psi_0) - \log v(2\pi x/n; \psi, \psi_0) \right| 1_{J_q} dx; \\
d_{n,2}(\psi) &= \sum_{p \in \{\tilde{n}, q-1, q_0-1\} \setminus \{q, q_0\}} |\log v(\lambda_p; \psi, \psi_0)|, \\
d_{n,3}(\psi) &= \int_0^{n/2} \left| \log v(2\pi x/n; \psi, \psi_0) \right| 1_{J_q^c} dx.
\end{aligned}
$$

It suffices to show that

$$
\sup_{\psi \in \Psi_K} u_n(\psi)^{-1} d_{n,i} \leq CK^{-1/2} \quad (i = 1, 2, 3). \tag{7.18}
$$

Using (6.28) to bound $|(\partial/\partial \xi) \log v(\xi; \psi, \psi_0)|$, by the mean value theorem we obtain

$$
\begin{aligned}
d_{n,1}(\psi) &\leq C \int_0^{n/2} n^{-1} \sup_{\xi \in [2\pi[x]/n, 2\pi x/n]} \left| \frac{\partial}{\partial \xi} \log v(\xi; \psi, \psi_0) \right| 1_{J_q} dx. \\
&\leq C \int_0^{n/2} \left((\|\theta - \theta_0\| + |\lambda_{\bar{q}}|^{1/2}) |x - q_0|^{-1} + |\bar{q}| |x - q_0|^{-1} |x - q|^{-1} \right) 1_{J_q} dx \\
&\leq C(\|\theta - \theta_0\| + |\lambda_{\bar{q}}|^{1/2}) \log n + C \log |\bar{q}|.
\end{aligned}
$$

Thus, for $\psi \in \Psi_K$, as $n \to \infty$,

$$
u_n(\psi)^{-1} |d_{n,1}(\psi)| \leq Cn^{-1/2} \log n + CK^{-1/2} \leq 2CK^{-1/2}. \tag{7.19}
$$

We estimate now $d_{n,2}(\psi)$. 1) Suppose first that $|\alpha - \alpha_0| \geq (\log n)^{-1}$. Then it is easy to check that under Assumptions A.1 or A.1′, $|\log v(2\pi p/n, \psi, \psi_0)| \leq C \log n$. Thus, $d_{n,2}(\psi) \leq C \log n \leq Cn^{-1} \log n \, u_n(\psi) = o(1) u_n(\psi)$ and (7.18) holds. 2) If $|\alpha - \alpha_0| < (\log n)^{-1}$, then $|\lambda_j|^{\alpha - \alpha_0} \leq C$ uniformly in $j = 1, \ldots, \tilde{n}$ and A.1 or A.1′ imply that in $d_{n,2}(\psi)$, $|\log v(2\pi p/n, \psi, \psi_0)| \leq C \log |\bar{q}|$. Thus $d_{n,2}(\psi) \leq C \log |\bar{q}| \leq C \log u_n(\psi)$ and (7.18) holds.

To estimate $d_{n,3}(\psi)$ note that when A.1 and $0 < \omega_0 < \pi$ hold or A.1′ holds, (6.29) and the definition of $v(\lambda, \psi, \psi_0)$ imply $|\log(v(2\pi x/n, \psi, \psi_0)| \leq C(|\alpha_0 - \alpha| \log n + \log |x - q| + \log |x - q_0| + 1)$. Then $d_{n,3}(\psi) \leq C|\alpha_0 - \alpha| \log n + C|\bar{q}|^{1/2}$ and

$$
\sup_{\psi \in \Psi_K} u_n(\psi)^{-1} |d_{n,3}(\psi)| \leq C(n^{-1/2} \log n + K^{-1/2}). \tag{7.20}
$$

When A.1 holds and $\omega_0 = 0, \pi$, using (6.29) it can be shown that (7.20) remains valid. Then (7.17) - (7.20) prove (7.16). $\square$

REFERENCES

ANDEL, J. (1986). Long - memory time series models. *Kybernetika* **22** 105 - 123.

BLOOMFIELD, P. (1973). An exponential model for the spectrum of a scalar time series. *Biometrika* **60** 217-226.

BOX, G.E.P. AND JENKINS, G.M. (1971). *Time Series Analysis. Forecasting and Control.* Holden-Day, San Francisco.

CHUNG, C.F. (1996a). Estimating a generalized long memory process. *J. Econometrics* **73** 237-259.

CHUNG, C.F. (1996b). A generalized fractionally integrated autoregressive moving average process. *J. Time Ser. Anal.* **17** 111-140.

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

FOX, R. AND TAQQU, M.S. (1986). Large sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *Ann. Statist.* **14** 517-532.

GIRAITIS, L. AND LEIPUS, R. (1995). A generalized fractionally differencing approach in long-memory modelling. *Lithuanian Math. J.* **35** 65 − 81.

GIRAITIS, L. AND SURGAILIS, D. (1990). A central limit theorem for quadratic forms in strongly deppendent random variables and its application to asymptotical normality of Whittle's estimate. *Probab. Th. Rel. Fields* **86** 87-104.

GRAY, H.,L., ZHANG, N. AND WOODWARD, W.A. (1989). On generalized fractional processes. *J. Time Ser. Anal.* **10** 233-257.

HANNAN, E.J. (1970). *Multiple Time Series.* Wiley, New York.

HANNAN, E.J. (1973a). The estimation of frequency. *J. Appl. Probab.* **10** 513 − 519.

HANNAN, E.J. (1973b). The asymptotic theory of linear time series models. *J. Appl. Probab.* **10** 130 − 145.

HOSKING, J.R.M. (1981). Fractional differencing. *Biometrika* **68** 165-176.

HOSOYA, Y (1997). Limit theory with long-range dependence and statistical inference of related models. *Ann. Statist.* **25** 105-137.

HIDALGO, J. (1999). Semiparametric estimation of the location of the pole. *Preprint* .

ROBINSON, P.M. (1978). Alternative models for stationary stochastic processes. *Stoch. Proc. Appl.* **8** 141-152.

ROBINSON, P.M. (1994). Efficient tests of nonstationary hypotheses. *J. Amer. Statist. Assoc.* **89** 1420-1437.

34

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

Robinson, P.M. (1995b). Gaussian semiparametric estimation of long range dependence. *Ann. Statist.* **23** 1630-1661.

Stout, W. F. (1974). *Almost Sure Convergence.* Academic Press, New York.

Yajima, Y. (1996). Estimation of the frequency of unbounded spectral densities. American Statistical Association, *Proceedings of the Business and Economic Statistical Section*, 4-7.

London School of Economics, Department of Economics,
Houghton Street, London WC2A 2AE.
email: L.Giraitis@lse.ac.uk; F.J.Hidalgo@lse.ac.uk; P.M.Robinson@lse.ac.uk

35