Gaussian Pseudo-Maximum Likelihood Estimation of Fractional Time Series Models

J. Hualde\textsuperscript{a} and P.M. Robinson\textsuperscript{b}
\textsuperscript{a}Universidad de Navarra
\textsuperscript{b}London School of Economics
February 15, 2006

Abstract

We consider the estimation of parametric fractional time series models in which not only is the memory parameter unknown, but one may not know whether it lies in the stationary/invertible region or the nonstationary region. In these circumstances a proof of consistency (which is a prerequisite for proving asymptotic normality) can be difficult owing to non-uniform convergence of the objective function over a large admissible parameter space. In particular, this is the case for the conditional sum of squares estimate, which can be expected to be asymptotically efficient under Gaussianity. Without the latter assumption, we establish consistency and asymptotic normality for this estimate in case of a quite general univariate model. For a multivariate model we establish asymptotic normality of a one-step estimate based on an initial $\sqrt{n}$-consistent estimate.
1. Introduction

Whittle estimates of parameters in univariate stationary long memory or fractional processes have been shown to be $\sqrt{n}$-consistent and asymptotically normal, for sample size $n$, by, e.g. Fox and Taqqu (1986), Dahlhaus (1989), Giraitis and Surgailis (1990), Hosoya (1996). In many time series, for example macroeconomic ones, the possibility of non-stationarity has to be taken seriously. Unit root models have been heavily used, but they occupy a similarly specialized position relative to fractionally non-stationary processes as short memory ones do relative to fractional stationary ones. Consider the model

$$x_t = \Delta^{-\delta_0} \{ u_t 1(t > 0) \}, \quad t = 0, \pm 1, \ldots,$$

(1.1)

where

$$u_t = \theta(L; \varphi_0) \varepsilon_t, \quad t = 0, \pm 1, \ldots,$$

(1.2)

The various quantities in (1.1) and (1.2) are defined as follows: $L$ is the lag operator, $\Delta = 1 - L$ is the difference operator;

$$(1 - L)^{-\zeta} = \sum_{j=0}^{\infty} a_j(\zeta)L^j, \quad a_j(\zeta) = \frac{\Gamma(j + \zeta)}{\Gamma(\zeta)\Gamma(j + 1)},$$

with $\Gamma(\zeta) = \infty$ for $\zeta = 0, -1, \ldots$, and the convention $\Gamma(0)/\Gamma(0) = 1$; $1(\cdot)$ is the indicator function; $\delta_0$ is an unknown real number and $\varphi_0$ is an unknown $p \times 1$ vector;

$$\theta(s; \varphi) = \sum_{j=0}^{\infty} \theta_j(\varphi)s^j,$$

where for all $\varphi$, $\theta_0(\varphi) = 1$, $\theta(s; \varphi) : \mathbb{R}^1 \times \mathbb{R}^p$ is continuous in $s$ and $|\theta(s; \varphi)| \neq 0$, $|s| = 1$; $\varepsilon_t$ is a zero-mean unobservable white noise sequence. More precise conditions will be imposed below. The role of $\theta$ in (1.2) is to permit parametric short memory autocorrelation. In an important special case of our model, $\theta(s; \varphi)$ is a rational function of $s$, whose denominator and numerator are polynomials in $s$ of degrees $p_1$ and $p_2$ respectively, so that $u_t$ is an autoregressive moving average process, denoted ARMA$(p_1, p_2)$, and $x_t$ can be called a fractionally autoregressive integrated moving average process, denoted FARIMA$(p_1, \delta_0, p_2)$. We allow for the simplest case FARIMA$(0, \delta_0, 0)$ by taking $\varphi_0$ to be empty.

Due to the truncation in (1.1), $x_t$ is actually non-stationary for all $\delta_0$. However, for $\delta_0 < 1/2$, $\Delta^{-\delta_0} u_t$ is stationary and $x_t$ has an “asymptotic stationarity” property. For $\delta_0 \geq 1/2$, $x_t$ is non-stationary in a more substantial sense, in particular the variance of $x_t$ diverges as $t \to \infty$; here the truncation in (1.1) is needed to avoid explosion. In case $\delta_0 = 1$ we have a unit root process.

We wish to estimate $\tau_0 = (\delta_0, \varphi_0)'$ from observations $x_t$, $t = 1, \ldots, n$. For any admissible $\tau = (\delta, \varphi)'$, define

$$\varepsilon_t(\tau) = \Delta^\delta \theta^{-1}(L; \varphi)x_t, \quad t \geq 1,$$

(1.3)

noting that (1.1) implies $x_t = 0, t \leq 0$. Define as an estimate of $\tau_0$

$$\hat{\tau} = \arg \min_{\tau \in T} R_n(\tau),$$

(1.4)

where

$$R_n(\tau) = \frac{1}{n} \sum_{t=1}^{n} \varepsilon_t^2(\tau).$$

(1.5)
and $T = [\nabla_1, \nabla_2] \times \Psi$, where $\nabla_1 < \nabla_2$, and $\Psi$ is a compact subset of $\mathbb{R}^p$.

The estimate $\hat{\tau}$ is sometimes termed “conditional sum of squares” (though “truncated sum of squares” might be more suitable). It has the anticipated advantage of having the same limit distribution as the maximum likelihood estimate of $\tau_0$ under Gaussianity, and thereby being asymptotically efficient (though we do not assume Gaussianity). It was employed by Box and Jenkins (1971) in estimation of non-fractional ARMA models (when $\delta_0$ is a given integer), by Li and McLeod (1986) in stationary FARIMA models, where $0 < \delta_0 < 1/2$, and by Beran (1995), Tanaka (1999) in nonstationary FARIMA models, allowing $\delta_0 \geq 1/2$. Though these authors considered the limit distribution of $\hat{\tau}$ (normal, at $\sqrt{n}$ rate), none gave a rigorous proof. For the stationary case $0 < \delta_0 < 1/2$, $\hat{\tau}$ would have the same limit distribution as that of the Whittle estimate for which Fox and Taqqu (1986) rigorously established consistency and asymptotic normality under Gaussianity. However, there is a particular technical difficulty which arises with $\hat{\tau}$ when nonstationary values $\delta_0 \geq 1/2$ are to be entertained, in particular when the set $T$ in (1.4) entails a $\delta$-interval, containing $\delta_0$, of length greater than $1/2$. A familiar approach to proving consistency of $\hat{\tau}$ (a necessary preliminary for establishing its limit distribution) involves proving uniform convergence in probability of $R_n(\tau)$. This is not possible when $T$ entails a $\delta$-interval exceeding $1/2$. Velasco and Robinson (2000) established consistency, and thence asymptotic normality, of an alternative estimate of $\tau_0$, under an alternative definition of fractional nonstationarity and using tapering and “skipping” of Fourier frequencies, achieving the $\sqrt{n}$ rate but with an inflated variance.

A proof of consistency and asymptotic normality of $\hat{\tau}$ seems desirable, especially because it has an approximate maximum likelihood interpretation under Gaussianity. The following section sets down detailed regularity conditions and a formal statement of asymptotic properties. Section 3 provides asymptotically normal estimates in a multivariate extension of (1.1), (1.2). Further extension and discussion appear in Section 4. Section 5 contains the main proof details, and useful lemmas are stated and proved in Section 6.

2. Consistency and asymptotic normality

Our first three assumptions will suffice for consistency of $\hat{\tau}$.

Assumption 1

(i) $|\theta (s; \varphi)| \neq |\theta (s; \varphi_0)|$,

for all $\varphi \neq \varphi_0$, $\varphi \in \Psi$, on a set $S \subset \{ s : |s| = 1 \}$ of positive measure;

(ii) for all $\varphi$, $\theta (e^{i\lambda}; \varphi)$ is differentiable in $\lambda$ with derivative in $Lip(\varsigma)$, $\varsigma > 1/2$;

(iii) for all $\lambda$, $\theta (e^{i\lambda}; \varphi)$ is continuous in $\varphi$;

(iv) for all $\varphi$

\[ |\theta (s; \varphi)| \neq 0, \ |s| = 1. \]
Condition (i) provides identification while (ii) and (iv) ensure that $u_t$ is an $I(0)$ process (with spectrum that is bounded and bounded away from zero at all frequencies). Further, by (ii) the derivative of $\theta(e^{it}; \varphi)$ has Fourier coefficients $j \theta_j(\varphi) = O(\Delta^{-1})$ as $j \to \infty$, for all $\varphi$, by Zygmund (1977), so that, by compactness of $\Psi$ and continuity of $\theta_j(\varphi)$ in $\varphi$ for all $j$,

$$\sup_{\varphi \in \Psi} |\theta_j(\varphi)| = O(j^{-\alpha}) \quad \text{as} \; j \to \infty,$$

where $\alpha = 1 + \delta$, noting that by (ii), $\alpha > 3/2$. Also, defining

$$\theta^{-1}(s; \varphi) = \phi(s, \varphi) = \sum_{j=0}^{\infty} \phi_j(\varphi) s^j,$$

$\phi_0(\varphi) = 1$ for all $\varphi$, and (ii), (iii) and (iv) imply that

$$\sup_{\varphi \in \Psi} |\phi_j(\varphi)| = O(j^{-\alpha}) \quad \text{as} \; j \to \infty.$$

Finally, (ii) also implies that

$$\inf_{|s|=1} \phi(s; \varphi) > 0. \quad \text{ (2.1)}$$

Assumption 1 is easily satisfied by standard parameterizations of stationary and invertible ARMA processes in which autoregressive and moving average orders are not both over-specified.

**Assumption 2.** The $\varepsilon_t$ in (1.2) are stationary and ergodic with finite fourth moment,

$$E(\varepsilon_t | \mathcal{F}_{t-1}) = 0, \quad E(\varepsilon_t^2 | \mathcal{F}_{t-1}) = \sigma_0^2 \quad \text{ (2.2)}$$

almost surely, where $\mathcal{F}_t$ is the $\sigma$-field of events generated by $\varepsilon_s$, $s \leq t$, and conditional (on $\mathcal{F}_{t-1}$) third and fourth moments and cross-moments of elements of $\varepsilon_t$ equal the corresponding unconditional moments.

**Assumption 3.** For all $\delta \in [\gamma_1, \gamma_2]$, $\varphi \in \Psi$ such that $\delta_0 - \delta > 1/2$, there exists a random variable $D(\tau) > 0$ a.s. such that

$$\frac{1}{n^{2(\delta_0 - \delta)}} \sum_{t=1}^{n} \varepsilon_t^2(\tau) \Rightarrow D(\tau), \quad \text{ (2.3)}$$

where “$\Rightarrow$” denotes a suitable notion of weak convergence.

If $\varepsilon_t$ in Assumption 2 is an independent and identically distributed sequence, (2.3) mainly follows from Marinucci and Robinson (2000) and the continuous mapping theorem. Thus, although unprimitive, we believe that this is a reasonable and not very stringent condition.

**Theorem 1.** Let (1.1), (1.2) and Assumptions 1-3 hold. Then as $n \to \infty$

$$\hat{\tau} \to_p \tau_0. \quad \text{ (2.4)}$$

Next, we derive the asymptotic distribution of $\hat{\tau}$, which requires an additional regularity condition.

**Assumption 4**
\(\tau_0 \in \text{int } T;\)

(ii) for all \(\lambda, \theta (e^{i\lambda}; \varphi)\) is twice continuously differentiable in \(\varphi\) on a closed neighbourhood \(N_\epsilon (\varphi_0)\) of radius \(0 < \epsilon < 1/2\) about \(\varphi_0;\)

(iii) the matrix
\[
A = \begin{pmatrix}
\pi^2/6 & -\sum_{j=1}^\infty b_j (\varphi_0)/j \\
-\sum_{j=1}^\infty b_j (\varphi_0)/j & \sum_{j=1}^\infty b_j (\varphi_0) b_j (\varphi_0)
\end{pmatrix}
\]
is non-singular, where
\[
b_j (\varphi_0) = \sum_{k=0}^{j-1} \theta_k (\varphi_0) \frac{\partial \phi_{j-k} (\varphi_0)}{\partial \varphi}.
\]

By compactness of \(N_\epsilon (\varphi_0)\) and continuity of \(\partial \phi_j (\varphi) / \partial \varphi_i, \partial^2 \phi_j (\varphi) / \partial \varphi_i \partial \varphi_j,\) for all \(j, l = 1, ..., p,\) with \(\varphi_i\) the \(i\)-th element of \(\varphi,\) Assumptions 1(ii), (iv) and 4(ii) imply that, as \(j \to \infty\)
\[
\sup_{\varphi \in N_\epsilon (\varphi_0)} \left| \frac{\partial \phi_j (\varphi)}{\partial \varphi_i} \right| = O (j^{-\alpha}), \quad \sup_{\varphi \in N_\epsilon (\varphi_0)} \left| \frac{\partial^2 \phi_j (\varphi)}{\partial \varphi_i \partial \varphi_j} \right| = O (j^{-\alpha}),
\]
which again is satisfied in the ARMA case.

**Theorem 2.** Let (1.1), (1.2) and Assumptions 1-4 hold. Then as \(n \to \infty\)
\[
n^{\frac{3}{2}} (\hat{\tau} - \tau_0) \to_d N \left( 0, A^{-1} \right),
\]

(2.5)

3. Multivariate extension

When observations on several related time series are available joint modelling can achieve efficiency gains. We consider a vector \(x_t = (x_{t1}, ..., x_{tr})'\) given by
\[
x_t = \Lambda_0^{-1} \{ u_t | (t > 0) \}, \quad t = 0, \pm 1, ...
\]
where \(u_t = (u_{t1}, ..., u_{tr})',\)
\[
u_t = \Theta (L; \varphi_0) \varepsilon_t, \quad t = 0, \pm 1, ...
\]
in which \(\varepsilon_t = (\varepsilon_{t1}, ..., \varepsilon_{tr})', \varphi_0\) is (as in the univariate case) a \(p \times 1\) vector of short-memory parameters,
\[
\Theta(s; \varphi) = \sum_{j=0}^\infty \Theta_j (\varphi) s^j, \quad \Theta_0 (\varphi) = I_p \quad \text{for all } \varphi,
\]
and
\[
\Lambda_0 = \text{diag} \left( \Delta^{b_{01}}, ..., \Delta^{b_{0r}} \right),
\]
where the memory parameters \( \delta_{0i} \) are unknown real numbers. In general, they can all be distinct but for the sake of parsimony we allow for the possibility that they are known to lie in a set of dimension \( q < r \). For example, perhaps as a consequence of pre-testing, we might believe some or all the \( \delta_{0i} \) are equal, and imposing this restriction in the estimation could further improve efficiency. We introduce known functions \( \delta_i = \delta_i(\delta) \), \( i = 1, ..., r \), of \( q \times 1 \) vector \( \delta \), such that for some \( \delta_0 \) we have \( \delta_{0i} = \delta_i(\delta_0) \), \( i = 1, ..., r \). We denote \( \tau = (\delta', \varphi')' \) and define (cf. (1.3))

\[
\varepsilon_t(\tau) = \Theta^{-1}(L; \varphi)A(\delta) x_t, \ t \geq 1,
\]

where \( A(\delta) = \text{diag}\left( \Delta^{\delta_1}, ..., \Delta^{\delta_r} \right) \). Gaussian likelihood considerations suggest the multivariate analogue to (1.5)

\[
R_n(\tau) = \det \{ \Sigma_n(\tau) \}, \tag{3.3}
\]

where

\[
\Sigma_n(\tau) = \frac{1}{n} \sum_{t=1}^{n} \varepsilon_t(\tau)\varepsilon_t'(\tau),
\]

assuming that no prior restrictions link \( \tau_0 \) with the covariance matrix of \( \varepsilon_t \). Unfortunately our consistency proof for the univariate case does not straightforwardly extend to an estimate minimizing (3.3), at least if \( q > 1 \). Also (3.3) is liable to pose a more severe computational challenge than (1.5) since \( p \) is liable to be larger in the multivariate case and \( q \) may exceed 1; it may be difficult to locate an approximate minimum of (3.3) as a preliminary to iteration. We avoid both these problems by taking a single Newton step from an initial \( \sqrt{n} \)-consistent estimate \( \tilde{\tau} \). Defining

\[
H_n(\tau) = \frac{1}{n} \sum_{t=1}^{n} \left( \frac{\partial \varepsilon_t(\tau)}{\partial \tau'} \right)' \Sigma_n^{-1}(\tau) \frac{\partial \varepsilon_t(\tau)}{\partial \tau'}, \\
h_n(\tau) = \frac{1}{n} \sum_{t=1}^{n} \left( \frac{\partial \varepsilon_t(\tau)}{\partial \tau'} \right)' \Sigma_n^{-1}(\tau) \varepsilon_t(\tau),
\]

we consider the estimate

\[
\hat{\tau} = \tilde{\tau} - H_n^{-1}(\tilde{\tau})h_n(\tilde{\tau}). \tag{3.4}
\]

We collect together all the requirements for asymptotic normality of \( \hat{\tau} \) in the sample.

**Assumption 5**

(i) For all \( \varphi, \Theta(e^{i\lambda}; \varphi) \) is differentiable in \( \lambda \) with derivative in Lip(\( \varsigma \)), \( \varsigma > 1/2 \);

(ii) for all \( \varphi \)

\[\det \{ \Theta(s; \varphi) \} \neq 0, \ |s| = 1;\]

(iii) the \( \varepsilon_t \) in (3.2) are stationary and ergodic with finite fourth moment,

\[E(\varepsilon_t | F_{t-1}) = 0, \quad E(\varepsilon_t \varepsilon_t' | F_{t-1}) = \Sigma_0\]

almost surely, where \( \Sigma_0 \) is positive definite, \( F_t \) is the \( \sigma \)-field of events generated by \( \varepsilon_s, s \leq t \), and conditional (on \( F_{t-1} \)) third and fourth moments and cross-moments of elements of \( \varepsilon_t \) equal the corresponding unconditional moments;
(iv) for all $\lambda$, $\Theta(e^{i\lambda}; \varphi)$ is twice continuously differentiable in $\varphi$ on a closed neighbourhood $N_\epsilon (\varphi_0)$ of radius $0 < \epsilon < 1/2$ about $\varphi_0$;

(v) the matrix $B$ having $(i,j)$th element

$$
\sum_{k=1}^{\infty} \text{tr} \left\{ d_k^{(i)}(\varphi_0) \right\} ^t \Sigma_0^{-1} d_k^{(j)}(\varphi_0) \Sigma_0 \right\}
$$

is non-singular, where

$$
d_k^{(i)}(\varphi_0) = - \frac{\partial \delta_k(\delta_0)}{\partial \delta_i} \sum_{l=1}^{r} \sum_{m=0}^{k-l} \Phi_k^{(i)}(\varphi_0) \Theta_{k-l-m}(\varphi_0), \quad 1 \leq i \leq r,
$$

$$
= \sum_{l=1}^{r} \frac{\partial \Phi_l(\varphi_0)}{\partial \varphi_i} \Theta_{k-l}(\varphi_0), \quad r+1 \leq i \leq r+p,
$$

the $\Phi_j(\varphi)$ being coefficients in the expansion

$$
\Theta^{-1}(s; \varphi) = \Phi(s, \varphi) = \sum_{j=0}^{\infty} \Phi_j(\varphi) s^j,
$$

where $\Phi_k^{(i)}(\varphi_0)$ is an $r \times r$ matrix whose $i$-th column is the $i$-th column of $\Phi_k(\varphi_0)$ and whose other elements are all zero;

(vi) $\delta_i(\delta)$ is twice continuously differentiable in $\delta$, for $i = 1, ..., r$;

(vii) $\bar{\tau}$ is a $\sqrt{n}$-consistent estimate of $\tau_0$.

The components of Assumption 5 are mostly natural extensions of ones in Assumptions 1, 2 and 4, and require no additional discussion. The important exception is (vii). When $\Theta(s; \varphi)$ is a diagonal matrix (as in the simplest case $\Theta(s; \varphi) \equiv I_r$, when $x_i$ is a FARIMA$(0, \delta_0i, 0)$ for $i = 1, ..., r$) then $\bar{\tau}$ can be obtained by first carrying out $r$ univariate fits following the approach of Section 2, and then if necessary reducing the dimensionality in a common-sense way: for example if some of the $\delta_0i$ are a priori equal then the common memory parameter might be estimated by the arithmetic mean of estimates from the relevant univariate fits. Notice that in the diagonal–$\Theta$ case with no cross-equation parameter restrictions the efficiency achievement afforded by $\bar{\tau}$ is due solely to cross-correlation in $\epsilon_i$, i.e. non-diagonality of $\Sigma_0$.

When $\Theta(s; \varphi)$ is not diagonal it is less clear how to use the $\sqrt{n}$-consistent outcome of Theorem 2 to form $\bar{\tau}$. We can infer that $u_i$ has spectral density matrix $(2\pi)^{-1} \Theta(e^{i\lambda}; \varphi_0) \Sigma_0 \Theta(e^{-i\lambda}; \varphi_0)'$. From the $i$-th diagonal element of this (the power spectrum of $u_{it}$) we can deduce a form for the Wold representation of $u_{it}$, corresponding to (1.2). However, starting from innovations $\epsilon_i$ in (3.2) satisfying (iii) of Assumption 5, it does not follow in general that the innovations in the Wold representation of $u_{it}$ will satisfy a condition analogous to (2.2) of Assumption 2, indeed it does not help if we simply strengthen Assumption 5 such that the $\epsilon_i$ are independent and identically distributed. However, ((2.2)) certainly holds if $\epsilon_i$ is Gaussian, which motivates our estimation approach from an efficiency perspective.
Notice that if \( \mathbf{u}_t \) is a vector ARMA process with non-diagonal \( \Theta \), in general all \( r \) univariate AR operators are identical, and of possibly high degree; the formation of \( \hat{\tau} \) is liable to be affected by a lack of parsimony, or some ambiguity.

An alternative approach could involve first estimating the \( \delta_{0t} \) by some semiparametric approach, using these estimates to form differenced \( \mathbf{x}_t \) and then estimating \( \varphi_0 \) from these proxies for \( \mathbf{u}_t \). This initial estimate will be less-than-\( \sqrt{n} \)-consistent, but its rate can be calculated given a rate for the bandwidth used in the semiparametric estimation. One can then calculate the (finite) number of iterations of form (3.4) needed to produce an estimate satisfying (2.5), following Theorem 5 and the discussion on p.539 of Robinson (1988).

**Theorem 3.** Let (3.1), (3.2) and Assumption 5 hold. Then as \( n \to \infty \)

\[
n^{1/2} (\hat{\tau} - \tau_0) \to_d N (0, B^{-1}).
\] (3.5)

4. Further comments and extensions

Our univariate and multivariate structures cover a wide range of parametric models for stationary and nonstationary time series, with memory parameters allowed to lie in a set that can be arbitrarily large. Unit root series are a special case, but unlike in the bulk of the large literature on these models we do not have to assume knowledge that memory parameters are 1. As the nondiagonal structure of \( \mathbf{A} \) and \( \mathbf{B} \) suggests, there is efficiency loss in estimating \( \varphi_0 \) if memory parameters are unknown, but on the other hand if these are misspecified \( \varphi_0 \) will in general be inconsistently estimated. Our limit distribution theory can be used to test hypotheses on the memory and other parameters, after straightforwardly forming consistent estimates of \( \mathbf{A} \) or \( \mathbf{B} \).

Our multivariate system (3.1), (3.2) does not cover fractionally cointegrated systems (such as the parametric ones considered by Robinson and Hualde, 2003) because \( \Sigma_0 \) is required to be positive definite. On the other hand our theory for univariate estimation should cover estimation of individual memory parameters of observations, so long as Assumption 2, in particular, can be reconciled with the full system specification. Moreover, again on an individual basis, it should be possible to derive analogous properties of estimates of memory parameters of cointegrating errors based on residuals that use simple estimates of cointegrating vectors, such as least squares. In a more standard regression setting, for example with determinist regressors such as polynomial functions of time, it should be possible to extend our theory for univariate and multivariate models to residual-based estimates of memory parameters of errors.

Nonstationary fractional series can be defined in many ways. Our definition ((1.1) and (3.1)) is a leading one in the literature, and has been termed “Type II”. Another popular one (“Type I”) was used by Velasco and Robinson (2000) for an alternate type of estimate. Their estimate is generally less efficient than \( \hat{\tau} \) due to the tapering required to handle nonstationarity, and (cf. Robinson, 2005) it seems likely that the asymptotic theory derived in this paper for \( \hat{\tau} \) can also be established in a “Type I” setting.

5. Proofs

Proof of Theorem 1
Define $T_1 = [\delta_0 - \gamma, \gamma_2] \times \Psi$ for some $1/3 < \gamma < 1/2$, $T_2 = [\gamma_1, \delta_0 - \gamma) \times \Psi$ when $\gamma_1 < \delta_0 - \gamma$ or $T_2$ is the empty set otherwise. Defining

$$S_n (\tau) = R_n (\tau) - R_n (\tau_0),$$

(2.4) follows if for $\tau \in T_1$ there is a deterministic function $U (\tau)$ (not depending on $n$), such that

$$S_n (\tau) = U (\tau) - T_n (\tau),$$

where

$$\inf_{\tau \in T_1} U (\tau) > \epsilon$$

for some $\epsilon > 0$ and $\epsilon > 0$, where $T_1 = \{ \tau : ||\tau - \tau_0|| > \epsilon \}$,

$$\sup_{\tau \in T_1} |T_n (\tau)| = o_p (1)$$

and also

$$\Pr \left( \inf_{\tau \in T_2} S_n (\tau) \leq 0 \right) \to 0 \text{ as } n \to \infty.$$

(5.3)

Since $x_t = 0$, $t \leq 0$, for $\tau \in T_1$ we set

$$\varepsilon_t (\tau) = \zeta_t (\tau) + \xi_t (\tau),$$

where

$$\zeta_t (\tau) = \Delta_{\delta - \delta_0} \phi (L; \varphi) u_t,$$

$$\xi_t (\tau) = - \Delta_{\delta - \delta_0} \phi (L; \varphi) \{ u_{t \tau} (t \leq 0) \},$$

and

$$U (\tau) = E \zeta_t^2 (\tau) - \sigma^2_0,$$

$$T_n (\tau) = R_n (\tau_0) - \sigma^2_0 - \{ R_n (\tau) - E \zeta_t^2 (\tau) \}.$$

We may write

$$U (\tau) = \sigma^2_0 \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{g (\lambda)}{g_0 (\lambda)} d\lambda - 1 \right),$$

where

$$g (\lambda) = |1 - e^{i\lambda} |^{2(\delta - \delta_0)} \phi (e^{i\lambda}; \varphi) |^2,$$

$$g_0 (\lambda) = g (\lambda)|_{\tau = \tau_0}.$$

For all $\tau$

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \log (g (\lambda)/g_0 (\lambda)) d\lambda = 0,$$

so by Jensen’s inequality

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{g (\lambda)}{g_0 (\lambda)} d\lambda \geq 1.$$  

(5.4)
Under Assumption 1(i), we have strict inequality in (5.4) for all \( \tau \neq \tau_0 \), so that by continuity in \( \tau \) of the left side of (5.4), (5.1) holds.

Next, noting that by Lemma 1
\[
eq \sum_{j=0}^{t-1} c_j(\tau) u_{t-j},
\]
where
\[
c_j(\tau) = \sum_{k=0}^{j} \phi_k(\varphi) a_{j-k}(\delta_0 - \delta),
\]
because by Lemma 2
\[
\frac{1}{n} \sum_{t=1}^{n} \left( \varepsilon_t^2(\tau_0) - a_0^2 \right) \rightarrow_p 0, \text{ as } n \rightarrow \infty,
\]
(5.2) would hold on showing that
\[
\sup_{T_1} \left| \frac{1}{n} \sum_{t=1}^{n} \left( \sum_{j=0}^{t-1} c_j u_{t-j} \right)^2 - E \left( \sum_{j=0}^{t-1} c_j u_{t-j} \right)^2 \right| = o_p(1),
\]
(5.6)
\[
\sup_{T_1} \left| \frac{1}{n} \sum_{t=1}^{n} \sum_{j=0}^{t-1} \sum_{k=t}^{\infty} c_j c_k \gamma(j-k) \right| = o_p(1),
\]
(5.7)
\[
\sup_{T_1} \left| \frac{1}{n} \sum_{t=1}^{n} \sum_{j=t}^{\infty} \sum_{k=t}^{\infty} c_j c_k \gamma(j-k) \right| = o_p(1),
\]
(5.8)
where \( \gamma(k) = E(u_i u_{i-k}) \), and throughout \( c_j = c_j(\tau) \). We first deal with (5.6). The term whose modulus is taken is
\[
\frac{1}{n} \sum_{j=0}^{n-1} \sum_{t=1}^{n-j} (u_t^2 - E u_t^2) + \frac{2}{n} \sum_{j=0}^{n-2} \sum_{k=j+1}^{n-1} c_j c_k \sum_{i=k-j+1}^{n-j} (u_i u_{i-(k-j)} - E(u_i u_{i-(k-j)})) = (a) + (b).\]
(5.9)
First,
\[
\sup_{T_1} |(a)| = O_p \left( n^{-\frac{1}{2}} \sum_{j=1}^{\infty} j^{2\gamma-2} \right) = O_p \left( n^{-\frac{1}{2}} \right),
\]
by Lemma 1. Next, by summation by parts, (b) is equal to
\[
\frac{2c_{n-1}}{n} \sum_{j=0}^{n-2} c_j \sum_{k=j+1}^{n-1} \sum_{t=k-j+1}^{n-j} (u_i u_{i-(k-j)} - E(u_i u_{i-(k-j)}))
\]
\[
- \frac{2c_{n-2}}{n} \sum_{j=0}^{n-2} c_j \sum_{k=j+1}^{n-1} \sum_{r=j+1}^{n-j} (c_{k+1} - c_k) \sum_{r=j+1}^{n-j} (u_i u_{i-(r-j)} - E(u_i u_{i-(r-j)}))
\]
\[
= (b_1) + (b_2).
\]
Since
\[
\text{Var} \left( \sum_{k=j+1}^{n-1} \sum_{l=k-j}^{n-j} u_l u_{(k-j)} \right) = O \left( n^2 \right),
\]
we have
\[
E \sup_{T_1} |(b_1)| \leq Kn^{\varphi-2} \sum_{j=1}^{n} j^{\varphi-1} \left\{ \text{Var} \left( \sum_{k=j+1}^{n-1} \sum_{l=k-j}^{n-j} u_l u_{(k-j)} \right) \right\}^{\frac{1}{2}} \leq Kn^{2\varphi-1},
\]
by Lemma 1, where \( K \) throughout denotes a generic finite constant. Similarly,
\[
E \sup_{T_1} |(b_2)| \leq Kn^{-1} \sum_{j=1}^{n} j^{\varphi-1} \sum_{k=j+1}^{n} k^{\max(\varphi-2, \alpha)} \left\{ \text{Var} \left( \sum_{r=1+j+1}^{n} \sum_{l=r-j+1}^{n-j} u_l u_{(r-j)} \right) \right\}^{\frac{1}{2}},
\]
by Lemma 1. Noting that
\[
\text{Var} \left( \sum_{r=1+j+1}^{n} \sum_{l=r-j+1}^{n-j} u_l u_{(r-j)} \right) \leq K (k-j) (n-j),
\]
if \( \varphi - 2 > -\alpha \)
\[
E \sup_{T_1} |(b_2)| \leq Kn^{-\frac{1}{2}} \sum_{j=1}^{n} j^{\varphi-1} \sum_{k=j+1}^{n} k^{\varphi-2} (k-j)^{\frac{1}{2}} \leq Kn^{-\frac{1}{2}} \sum_{j=1}^{n} j^{\varphi-1} \sum_{k=1}^{n} (k+j)^{\varphi-2} k^{\frac{1}{2}}.
\]
For \( \alpha \in (1/6, \varphi) \) (where we assumed \( \varphi > 1/3 \)), the right side of (5.10) is bounded by
\[
Kn^{-\frac{1}{2}} \sum_{j=1}^{n} j^{\varphi-1-\alpha} \sum_{k=1}^{n} k^{\varphi-3/2+\alpha},
\]
because for such \( \alpha \)
\[(k+j)^{\varphi-2} \leq j^{-\alpha} k^{\varphi-2+\alpha}.
\]
Thus, (5.11) is bounded by
\[
Kn^{\varphi-1+\alpha} \sum_{j=1}^{n} j^{\varphi-1-\alpha} \leq Kn^{2\varphi-1}.
\]
Alternatively, if \( \varphi - 2 \leq -\alpha \),
\[
E \sup_{T_1} |(b_2)| \leq Kn^{-\frac{1}{2}} \sum_{j=1}^{n} j^{\varphi-1} \sum_{k=1}^{n} (k+j)^{-\alpha} k^{\frac{1}{2}} \leq Kn^{-\frac{1}{2}} \sum_{j=1}^{n} j^{\varphi-1} \sum_{k=1}^{n} k^{\frac{1}{2} - \alpha} \leq Kn^{\varphi-\frac{1}{2}},
\]
11
since $\alpha > 3/2$, to complete the proof of (5.6).

Next, the term whose modulus is taken in (5.7) is

$$
\frac{1}{n} \sum_{t=1}^{n} \int_{-\pi}^{\pi} f(\lambda) \sum_{j=0}^{t-1} \sum_{k=1}^{\infty} c_j c_k e^{ikj\lambda} d\lambda,
$$

(5.12)

where $f(\lambda)$ denotes the spectral density of $u_t$. By the Cauchy inequality, (5.12) is bounded by

$$
Kn^{-1} \sum_{t=1}^{n} \left\{ \int_{-\pi}^{\pi} \left| \sum_{j=0}^{t-1} c_j e^{ij\lambda} \right|^2 d\lambda \right\}^{\frac{1}{2}} \leq Kn^{-1} \sum_{t=1}^{n} \left\{ \sum_{j=0}^{t-1} c_j^2 \sum_{k=t}^{\infty} c_k^2 \right\}^{\frac{1}{2}},
$$

so the left side of (5.7) is bounded by

$$
Kn^{-1} \sum_{t=1}^{n} \left\{ \int_{-\pi}^{\pi} \sum_{j=t}^{\infty} c_j^2 \sum_{k=t}^{\infty} c_k^2 d\lambda \right\}^{\frac{1}{2}} \leq Kn^{-1} \sum_{t=1}^{n} t^{\nu - \frac{1}{2}} \leq Kn^{\nu - \frac{1}{2}},
$$

by Lemma 1, to establish (5.7).

Finally, by a similar reasoning, the term whose modulus is taken in (5.8) is bounded by

$$
Kn^{-1} \sum_{t=1}^{n} \left\{ \int_{-\pi}^{\pi} \sum_{j=t}^{\infty} c_j^2 \sum_{k=t}^{\infty} c_k^2 d\lambda \right\}^{\frac{1}{2}} \leq Kn^{-1} \sum_{t=1}^{n} t^{\nu - 1} \leq Kn^{\nu - 1},
$$

to conclude the proof of (5.8).

Next, by (5.5), (5.3) holds on establishing

$$
\Pr \left( \inf_{T_{2n}} S'_{n}(\tau) \leq \varepsilon \right) \rightarrow 0 \text{ as } n \rightarrow \infty,
$$

(5.13)

for an arbitrarily small $\varepsilon > 0$, where

$$
S'_{n}(\tau) = \frac{1}{n} \sum_{t=1}^{n} \left\{ x_{t}(\tau) - \sigma_{t}^{2} \right\}.
$$

(5.13) holds on showing

$$
\Pr \left( \inf_{\varphi \in \Psi} S'_{n}(\tau) \leq \varepsilon \right) \rightarrow 0 \text{ as } n \rightarrow \infty,
$$

(5.14)

$$
\Pr \left( \inf_{\varphi \in \Psi} S'_{n}(\tau) \leq \varepsilon \right) \rightarrow 0 \text{ as } n \rightarrow \infty.
$$

(5.15)
Consider first the proof of (5.14). Denote for any process \( \zeta_t \),

\[
w_\zeta(\lambda) = n^{-\frac{1}{2}} \sum_{t=1}^{n} \zeta_t e^{i \lambda t}, \quad I_\zeta(\lambda) = |w_\zeta(\lambda)|^2,
\]

the finite Fourier transform and periodogram respectively, and \( \lambda_j = 2\pi j/n \). For \( V(\tau) \) satisfying (6.7) in Lemma 3, setting \( \tau^* = (\delta, \varphi_0) \),

\[
\sum_{t=1}^{n} \varepsilon_t^2(\tau) = \sum_{j=1}^{n} I_{\varepsilon(t)}(\lambda_j) = \sum_{j=1}^{n} \left| \xi(e^{i \lambda_j}; \varphi) \right|^2 I_{\varepsilon(\tau^*)}(\lambda_j) + V(\tau),
\]

where

\[
\xi(s; \varphi) = \frac{\theta(s; \varphi_0)}{\theta(s; \varphi)} = \sum_{j=0}^{\infty} \xi_j(\varphi) s^j.
\]

Then

\[
\inf_{\delta_0 - 1/2 \leq \delta < \delta_0 - \sqrt{\nu}} \frac{1}{n} \sum_{t=1}^{n} \varepsilon_t^2(\tau) \geq \inf_{\lambda \in [-\pi, \pi]} \left| \xi(e^{i \lambda}; \varphi) \right|^2 \inf_{\delta_0 - 1/2 \leq \delta < \delta_0 - \sqrt{\nu}} \frac{1}{n} \sum_{t=1}^{n} \varepsilon_t^2(\tau^*) \\
- \sup_{\delta_0 - 1/2 \leq \delta < \delta_0 - \sqrt{\nu}} \frac{1}{n} |V(\tau)|.
\]

Assumption 1 implies (see (2.1)) that there exists \( \epsilon > 0 \) such that

\[
\inf_{\lambda \in [-\pi, \pi]} \left| \xi(e^{i \lambda}; \varphi) \right|^2 > \epsilon.
\]

Thus

\[
\inf_{\delta_0 - 1/2 \leq \delta < \delta_0 - \sqrt{\nu}} \frac{1}{n} \sum_{t=1}^{n} \varepsilon_t^2(\tau) \geq \epsilon \inf_{\lambda \in [-\pi, \pi]} \frac{1}{n} \sum_{t=1}^{n} \left( \sum_{j=0}^{t-1} a_j \varepsilon_{t-j} \right)^2 \\
- \sup_{\delta_0 - 1/2 \leq \delta < \delta_0 - \sqrt{\nu}} \frac{1}{n} |W(\delta)|,
\]

where \( a_j = a_j(\delta_0 - \delta) \), and by Lemma 2

\[
W(\delta) = \epsilon \sum_{t=1}^{n} \nu_t^2(\delta) + 2\epsilon \sum_{t=1}^{n} \nu_t(\delta) \sum_{j=0}^{t-1} a_j \varepsilon_{t-j}.
\]

By Lemma 2 and (6.10) in Lemma 3

\[
\sup_{\delta_0 - 1/2 \leq \delta < \delta_0 - \sqrt{\nu}} \frac{1}{n} |W(\delta)| = o_p(1),
\]
and also by Lemma 3
\[
\sup_{\delta_0 - 1/2 \leq \delta < \delta_0 - \nabla} \frac{1}{n} |V(\tau)| = o_p(1). \tag{5.18}
\]
Next, note that for \(\delta_0 - 1/2 \leq \delta < \delta_0 - \nabla\)
\[
\frac{\partial a^j}{\partial \delta} = -2 (\psi(j + \delta_0 - \delta) - \psi(\delta_0 - \delta)) a^2_j < 0, \tag{5.19}
\]
where we introduce the digamma function \(\psi(x) = \frac{d \log \Gamma(x)}{dx}\).

From (5.19) and the fact that \(\psi(x)\) is strictly increasing in \(x > 0\),
\[
\inf_{\delta_0 - 1/2 \leq \delta < \delta_0 - \nabla} n^{-1} \left( \sum_{t=1}^{n} \sum_{j=0}^{t-1} a_j \xi_{t-j} \right)^2 \geq n^{-1} \sum_{t=1}^{n} \sum_{j=0}^{t-1} a_j^2 (\nabla) \xi^2_{t-j} - \sup_{\delta_0 - 1/2 \leq \delta < \delta_0 - \nabla} n^{-1} \sum_{t=1}^{n} \sum_{j \neq k} a_j a_k \xi_{t-j} \xi_{t-k}.
\]

By a similar analysis to that of (b) in (5.9)
\[
\sup_{\delta_0 - 1/2 \leq \delta < \delta_0 - \nabla} \left| n^{-1} \sum_{t=1}^{n} \sum_{j \neq k} a_j a_k \xi_{t-j} \xi_{t-k} \right| = O_p(1),
\]
or equivalently
\[
\Pr \left( \sup_{\delta_0 - 1/2 \leq \delta < \delta_0 - \nabla} \left| n^{-1} \sum_{t=1}^{n} \sum_{j \neq k} a_j a_k \xi_{t-j} \xi_{t-k} \right| > K \right) \rightarrow 0, \tag{5.20}
\]
as \(n \rightarrow \infty\). Then, noting (5.16), (5.17), (5.18), (5.20), we deduce (5.14) if
\[
\Pr \left( \frac{1}{n} \sum_{t=1}^{n} \sum_{j=0}^{t-1} a_j^2 (\nabla) \xi^2_{t-j} \leq K \right) \rightarrow 0 \text{ as } n \rightarrow \infty. \tag{5.21}
\]

Now
\[
\frac{1}{n} \sum_{t=1}^{n} \sum_{j=0}^{t-1} a_j^2 (\nabla) \xi^2_{t-j} = a_0^2 \Gamma(1 - 2\nabla) + \frac{1}{n} \sum_{t=1}^{n} \sum_{j=0}^{t-1} a_j^2 (\nabla) (\xi^2_{t-j} - \sigma^2_0) - \frac{1}{n} \sum_{t=1}^{n} \sum_{j=t}^{\infty} a_j^2 (\nabla).
\]
The third term on the right is clearly \(O(n^{2\nabla-1})\), whereas, as in the treatment of (a) in (5.9), the second is \(O_p(n^{-1/2})\), so that (5.21) holds as \(\Gamma(1 - 2\nabla) / \Gamma^2(1 - \nabla)\) can be made arbitrarily large on setting \(\nabla\) arbitrarily close to (but smaller than) 1/2.
Next, (5.15) holds if

$$\Pr \left( \inf_{\varphi \in \Psi} S'_n (\tau) \leq \varepsilon \right) \to 0 \text{ as } n \to \infty, \quad (5.22)$$

$$\Pr \left( \inf_{\varphi \in \Psi} S'_n (\tau) \leq \varepsilon \right) \to 0 \text{ as } n \to \infty, \quad (5.23)$$

for \( \eta \in (0, 1/4) \) to be described subsequently. We first deal with (5.22). Using a similar decomposition to (5.16), and noting that \( \eta < 1/4 \),

$$\sup_{\delta_0 - 1/2 - \eta < \delta < \delta_0 - 1/2} \frac{1}{n} |W(\delta)| = o_p(1), \quad (5.24)$$

by Lemma 2, whereas

$$\sup_{\delta_0 - 1/2 - \eta < \delta < \delta_0 - 1/2} \frac{1}{n} |V(\tau)| = o_p(1), \quad (5.25)$$

by Lemma 3. Denoting

$$f_n (\delta) = n^{-1} \sum_{t=1}^{t-1} \left( \sum_{j=0}^{t-1} a_j \varepsilon_{t-j} \right)^2,$$

by (5.24), (5.25), it follows that (5.22) holds if for any \( \varkappa > 0 \) there exists an integer \( n_0 \) such that

$$\Pr \left( \inf_{\delta_0 - 1/2 - \eta < \delta < \delta_0 - 1/2} f_n (\delta) > K \right) > 1 - \varkappa, \quad (5.26)$$

for all \( n \geq n_0 \). (5.26) holds on showing that there exist integers \( n_1, n_2 \), such that

$$\Pr \left( \inf_{\delta_0 - 1/2 - \eta < \delta < \delta_0 - 1/2} f_n (\delta) \geq f_n (\delta_0 - 1/2) \right) > 1 - \frac{\varkappa}{2}, \quad (5.27)$$

for any \( n \geq n_1 \), and

$$\Pr (f_n (\delta_0 - 1/2) > K) > 1 - \frac{\varkappa}{2}, \quad (5.28)$$

for any \( n \geq n_2 \). This is because

$$\Pr \left( \inf_{\delta_0 - 1/2 - \eta < \delta < \delta_0 - 1/2} f_n (\delta) \geq f_n (\delta_0 - 1/2) \right)$$

$$= \Pr \left( \inf_{\delta_0 - 1/2 - \eta < \delta < \delta_0 - 1/2} f_n (\delta) \geq f_n (\delta_0 - 1/2), f_n (\delta_0 - 1/2) \leq K \right)$$

$$+ \Pr \left( \inf_{\delta_0 - 1/2 - \eta < \delta < \delta_0 - 1/2} f_n (\delta) \geq f_n (\delta_0 - 1/2), f_n (\delta_0 - 1/2) > K \right)$$

$$\leq \Pr (f_n (\delta_0 - 1/2) \leq K) + \Pr \left( \inf_{\delta_0 - 1/2 - \eta < \delta < \delta_0 - 1/2} f_n (\delta) > K \right),$$

15
so that, setting \( n_0 = \max (n_1, n_2) \), and noting that (5.28) implies that
\[
\Pr (f_n (\delta_0 - 1/2) \leq K) < \frac{\kappa}{2},
\]
for any \( n \geq n_2 \), we have
\[
\Pr \left( \inf_{\delta_0 - 1/2 - \eta < \delta < \delta_0 - 1/2} f_n (\delta) > K \right) \geq \Pr \left( \inf_{\delta_0 - 1/2 - \chi < \delta < \delta_0 - 1/2} f_n (\delta) \geq f_n (\delta_0 - 1/2) \right) - \frac{\kappa}{2}
\]
for any \( n \geq n_0 \).

First, we show (5.27). Denoting by \( \Omega \) the sample space, for any \( n \geq 1 \), and arbitrarily small positive constants \( \epsilon, \chi \), define the families of sets
\[
A_n = \left\{ \omega \in \Omega : \left. \frac{\partial f_n (\delta)}{\partial \delta} \right|_{\delta = \delta_0 - 1/2} < -\epsilon \right\},
\]
\[
B_n (\chi) = \left\{ \omega \in \Omega : \inf_{\delta_0 - 1/2 - \chi < \delta < \delta_0 - 1/2} f_n (\delta) \geq f_n (\delta_0 - 1/2) \right\},
\]
and denote \( f_n (\delta) = f_n (\delta; \omega) \), whenever necessary for clarity. For particular \( n \) and \( \omega \), say \( \overline{\pi} \) and \( \overline{\omega} \) respectively, there exists a real number \( \eta_{\pi, \omega}(\overline{\omega}) \) such that
\[
\left. \frac{\partial f_{\pi, \omega} (\delta; \overline{\omega})}{\partial \delta} \right|_{\delta = \delta_0 - 1/2} < -\epsilon \Rightarrow \inf_{\delta_0 - 1/2 - \eta_{\pi, \omega} < \delta < \delta_0 - 1/2} f_{\pi, \omega} (\delta; \overline{\omega}) \geq f_{\pi, \omega} (\delta_0 - 1/2; \overline{\omega}).
\]
Then assuming there is a real number \( \eta \in (0, 1/4) \) such that
\[
\inf_{\omega \in \Omega, n \geq 1} \eta_n (\omega) > \eta,
\]
we have
\[
A_n \subseteq B_n (\eta),
\]
for all \( n \), so that
\[
\Pr (A_n) \leq \Pr (B_n (\eta)).
\]
Thus, (5.27) follows if
\[
\Pr (A_n) \geq 1 - \frac{\kappa}{2}, \tag{5.29}
\]
for any $n \geq n_1$. Now,

$$
\frac{\partial f_n (\delta)}{\partial \delta} \bigg|_{\delta = \delta_0 - 1/2} = -\frac{2\sigma_0^2}{n} \sum_{t=1}^{n} \sum_{j=0}^{t-1} (\psi (j + 1/2) - \psi (1/2)) a_j^2 (1/2) \\
- \frac{2}{n} \sum_{t=1}^{n} \sum_{j=0}^{t-1} (\psi (j + 1/2) - \psi (1/2)) a_j^2 (1/2) (\varepsilon_{t-j} - \sigma_0^2) \\
- \frac{2}{n} \sum_{t=2}^{n} \sum_{j=0}^{t-2} \sum_{k=j+1}^{t-1} (\psi (j + 1/2) - \psi (1/2)) a_j (1/2) a_k (1/2) \varepsilon_{t-j} \varepsilon_{t-k} \\
- \frac{2}{n} \sum_{t=2}^{n} \sum_{j=0}^{t-2} \sum_{k=j+1}^{t-1} (\psi (k + 1/2) - \psi (1/2)) a_j (1/2) a_k (1/2) \varepsilon_{t-j} \varepsilon_{t-k} \\
= (c) + (d) + (e) + (f)
$$

First, note that since $\psi (x)$ is an increasing function of $x$ for all $x > 0$, $(c)$ is nonpositive and

$$
\Pr (A_n) = \Pr ((c) + (d) + (e) + (f) < -\epsilon) \geq \Pr (- (c) > \epsilon + |(d)| + |(e)| + |(f)|).
$$

Now

$$(d) = - \frac{2}{n} \sum_{t=1}^{n} \sum_{j=1}^{n-t} (\varepsilon_t^2 - \sigma_0^2) \psi (j + 1/2) a_j^2 (1/2) + \frac{2\psi (1/2)}{n} \sum_{t=1}^{n} \varepsilon_t^2 - \sigma_0^2 \sum_{j=1}^{n-t} a_j^2 (1/2) = (d_1) + (d_2).
$$

Obviously, $E (d_1) = 0$, and

$$
\text{Var} ((d_1)) = O \left( n^{-2} \sum_{t=1}^{n} \left( \sum_{j=1}^{n-t} \psi (j + 1/2) a_j^2 (1/2) \right)^2 \right) = O \left( n^{-1} \log^4 n \right),
$$

so that

$$
|(d_1)| = O_p \left( n^{-1/2} \log^2 n \right), \quad (5.30)
$$

and similarly

$$
|(d_2)| = O_p \left( n^{-1/2} \log n \right). \quad (5.31)
$$

Next, noting that for any finite $c > 0$, $a_{j+1} (c) \psi (j + 1 + c) - a_j (c) \psi (j + c)$ is

$$
(a_{j+1} (c) - a_j (c)) \psi (j + 1 + c) + a_j (c) (\psi (j + 1 + c) - \psi (j + c)) = O \left( j^{c-2} \log j \right), \quad (5.32)
$$

since

$$
|\psi (j + 1 + c) - \psi (j + c)| \leq K |\psi' (j + c)| \leq K (j + 1)^{-1},
$$

and

$$
|\psi (j + c)| \leq K \log (j + 1), \quad (5.33)
$$
by a simple modification of the analysis of expression (b) in (5.9) to account for the presence of the digamma functions and evaluating at \( c = 1/2 \), it is straightforward to show that
\[
|\epsilon| + |(f)| = O_p(\log n).
\]
Thus, noting (5.30), (5.31), and denoting
\[
y_n = \epsilon + |(d)| + |(e)| + |(f)|,
\]
there exists a positive finite constant \( K_1 \) and integer \( n'_1 \) such that
\[
\Pr(y_n > K_1 \log n) < \frac{\alpha}{2},
\]
for any \( n \geq n'_1 \), and also for any \( \epsilon > 0 \), there exists an integer \( n''_1(\epsilon) \) such that
\[
\Pr(y_n > K_1 \log n) \geq \Pr(y_n > \epsilon \log^2 n),
\]
for any \( n \geq n''_1(\epsilon) \). Bounding sums from below by integrals, there exist strictly positive constants \( K'_2, K''_2 \), which do not depend on \( n \), such that for all \( n > 1 \)
\[
\frac{-\langle c \rangle}{2\sigma_0^2 \log^2 n} = \frac{1}{\sigma_0^2 n \log^2 n} \sum_{j=1}^{n} (n-j) a_j^2(\frac{1}{2}) \sum_{i=1}^{j} \frac{1}{2i-1} \geq \frac{K'_2}{n \log^2 n} \int_1^n \frac{n-x}{x} \left( \int_1^x \frac{1}{2y-1} dy \right) dx
\]
\[
= \frac{K'_2}{n \log^2 n} \int_1^n \frac{n-x}{x} \log (2x-1) dx \geq \frac{K'_2}{n \log^2 n} \int_1^n \frac{n-x}{x} \log x dx
\]
\[
= \frac{K'_2}{\log^2 n} \int_1^n \frac{\log x}{x} dx = \frac{K'_2}{\log^2 n} \int_1^n \log x dx \geq \frac{K''_2}{2},
\]
so that denoting \( K_2 = K''_2/2 \)
\[
\inf_{n>1} \frac{-\langle c \rangle}{\log^2 n} > K_2.
\]
Then
\[
\Pr(-\langle c \rangle > y_n) \geq \Pr(2\sigma_0^2 K_2 \log^2 n > y_n),
\]
for all \( n \). Now
\[
\Pr(2\sigma_0^2 K_2 \log^2 n > y_n) = 1 - \Pr(y_n \geq 2\sigma_0^2 K_2 \log^2 n) \geq 1 - \Pr(y_n > K_1 \log n),
\]
for \( n \geq n''_1(2\sigma_0^2 K_2) \), and
\[
1 - \Pr(y_n > K_1 \log n) \geq 1 - \frac{\epsilon}{2},
\]
for \( n \geq n_1 \). Hence, defining \( n_1 = \max(n_1' (2\sigma_0^2 K_2)) \), (5.29) holds, and the proof of (5.27) is completed.
Next we show (5.28). Clearly

\[
f_n(\delta_0 - 1/2) = \frac{\sigma_0^2}{n} \sum_{t=1}^{n} \sum_{j=0}^{t-1} a_j^2 (1/2) + \frac{1}{n} \sum_{t=1}^{n} \sum_{j=0}^{t-1} a_j^2 (1/2) (\varepsilon_{t-j}^2 - \sigma_0^2)
\]

\[
+ \frac{2}{n} \sum_{t=2}^{n} \sum_{j=0}^{t-2} \sum_{k=j+1}^{t-1} a_j (1/2) a_k (1/2) \varepsilon_{t-j} \varepsilon_{t-k} = (g) + (h) + (i),
\]

where \(g\) is obviously nonnegative and

\[\Pr(f_n(\delta_0 - 1/2) > K) \geq \Pr((g) > K + |(h)| + |(i)|).\]

Clearly, \(E((h)) = 0\), and

\[\text{Var}((h)) = O \left( n^{-2} \sum_{t=1}^{n} \left( \sum_{j=0}^{t-1} a_j^2 (1/2) \right)^2 \right) = O \left( n^{-1} \log^2 n \right),\]

so that

\[|(h)| = O_p \left( n^{-1/2} \log n \right),\]

whereas by previous results \((i) = O_p(1)\). Denoting

\[z_n = K + |(h)| + |(i)|,\]

there exists a positive finite constant \(K_3\) and integer \(n'_2\) such that

\[\Pr(z_n > K_3) < \frac{\alpha}{2},\]

for any \(n \geq n'_2\). Also, for any \(\epsilon > 0\), there exists an integer \(n''_2(\epsilon)\) such that

\[\Pr(z_n > K_3) \geq \Pr(z_n > \epsilon \log n),\]

for any \(n \geq n''_2(\epsilon)\). By a very similar analysis to that of (5.34), there exists a constant \(K_4\) such that

\[\inf_{n > 1} \frac{1}{n \log n} \sum_{t=1}^{n} \sum_{j=0}^{t-1} a_j^2 (1/2) > K_4,\]

so that

\[\Pr((g) > z_n) \geq \Pr(\sigma_0^2 K_4 \log n > z_n),\]

for all \(n\). Now

\[\Pr(\sigma_0^2 K_4 \log n > z_n) = 1 - \Pr(z_n \geq \sigma_0^2 K_4 \log n) \geq 1 - \Pr(z_n > K_3),\]

for any \(n \geq n''_2(\sigma_0^2 K_4)\) and

\[1 - \Pr(z_n > K_3) \geq 1 - \frac{\alpha}{2},\]
for any \( n \geq n'_2 \). Hence, defining \( n_2 = \max (n'_2, n''_2 (\sigma^2 \beta K)) \), the proof of (5.28) is completed. Finally, regarding (5.23),

\[
\Pr \left( \inf_{\varphi \in \Psi} S'_n (\tau) \leq \varepsilon \right) \leq \Pr \left( n^{2\eta} \inf_{\varphi \in \Psi} \frac{1}{n^{2(\delta_0 - \delta)}} \sum_{t=1}^n \varepsilon_t^2 (\tau) \leq \varepsilon + \sigma_0^2 \right).
\]

By Assumption 3 and the continuous mapping theorem,

\[
\inf_{\varphi \in \Psi} \frac{1}{n^{2(\delta_0 - \delta)}} \sum_{t=1}^n \varepsilon_t^2 (\tau) \Rightarrow \inf_{\varphi \in \Psi} D (\tau) > 0 \text{ a.s.,}
\]

as the infimum is taken over a compact set. Thus, it is clear that

\[
\Pr \left( n^{2\eta} \inf_{\varphi \in \Psi} \frac{1}{n^{2(\delta_0 - \delta)}} \sum_{t=1}^n \varepsilon_t^2 (\tau) \leq \varepsilon + \sigma_0^2 \right) \to 0,
\]

as \( n \to \infty \), to conclude (5.23), and thus complete the proof.

**Proof of Theorem 2**

By the mean value theorem, (2.5) follows if we prove that

\[
\frac{\sqrt{n} \partial R_n (\tau_0)}{2} \to_d N (0, \sigma_0^4 A), \tag{5.35}
\]

\[
\frac{1}{2} \frac{\partial^2 R_n (\tau)}{\partial \tau \partial \tau'} \to_p \sigma_0^2 A, \tag{5.36}
\]

where \(|\tau - \tau_0| \leq |\tilde{\tau} - \tau_0| \). First, (5.35) holds if

\[
\frac{\sqrt{n} \partial R_n (\tau_0)}{2} - \frac{1}{\sqrt{n}} \sum_{t=1}^n \varepsilon_t \sum_{j=1}^\infty m_j (\varphi_0) \varepsilon_{t-j} = o_p (1), \tag{5.37}
\]

\[
\frac{1}{\sqrt{n}} \sum_{t=2}^n \varepsilon_t \sum_{j=1}^\infty m_j (\varphi_0) \varepsilon_{t-j} \to_d N (0, \sigma_0^4 A), \tag{5.38}
\]

where

\[ m_j (\varphi_0) = \left( \frac{-j^{-1} \ b'_j (\varphi_0) }{s_1 + s_2} \right). \]

By Lemma 2, the left side of (5.37) is

\[
\left( r_1 + r_2 + r_3 \right) = \left( \begin{array}{c} r_1 \\ r_2 \\ r_3 \\ s_1 + s_2 \end{array} \right),
\]

\[
\begin{array}{c}
r_1 \\
r_2 \\
r_3 \\
s_1 + s_2
\end{array}
\]

20
where

\[
   r_1 = \frac{1}{\sqrt{n}} \sum_{t=2}^{n} \varepsilon_t \sum_{j=1}^{t-1} \frac{1}{j} \varepsilon_{t-j}, \quad r_2 = \frac{1}{\sqrt{n}} \sum_{t=2}^{n} \varepsilon_t \sum_{j=1}^{t-1} \sum_{k=t-j}^{\infty} \phi_k(\varphi_0) u_{t-j-k},
\]

\[
   r_3 = -\frac{1}{\sqrt{n}} \sum_{t=2}^{n} v_t(\delta_0) \sum_{j=1}^{t-1} \sum_{k=0}^{t-j-1} \phi_k(\varphi_0) u_{t-j-k}, \quad s_1 = \frac{1}{\sqrt{n}} \sum_{t=2}^{n} \varepsilon_t \sum_{j=t}^{\infty} \frac{\partial \phi_j(\varphi_0)}{\partial \varphi} u_{t-j},
\]

\[
   s_2 = \frac{1}{\sqrt{n}} \sum_{t=2}^{n} v_t(\delta_0) \sum_{j=1}^{t-1} \frac{\partial \phi_j(\varphi_0)}{\partial \varphi} u_{t-j}.
\]

Clearly, \( E(r_1) = 0 \), and

\[
   \text{Var}(r_1) = \frac{1}{n} \sum_{t=2}^{n} \sum_{j=1}^{\infty} \sum_{s=2}^{\infty} \sum_{k=s}^{\infty} \frac{1}{j^k} E(\varepsilon_t \varepsilon_s) E(\varepsilon_{t-j} \varepsilon_{s-k}) = O\left( \frac{1}{n} \sum_{t=2}^{n} \sum_{j=1}^{t} \frac{1}{j^2} \right) = O\left( \frac{\log n}{n} \right),
\]

noting that for any \( t, s, j, k \)

\[
   E(\varepsilon_t \varepsilon_s \varepsilon_{t-j} \varepsilon_{s-k}) = \kappa = 0,
\]

where \( \kappa \) denotes here the fourth cumulant of \( \varepsilon_t, \varepsilon_s, \varepsilon_{t-j}, \varepsilon_{s-k} \). Thus, \( r_1 = O_p\left( n^{-1/2} \log^{1/2} n \right) \). Next, \( E(r_2) = 0 \) and

\[
   \text{Var}(r_2) = \frac{1}{n} \sum_{t=2}^{n} \sum_{j=1}^{t-1} \sum_{s=2}^{\infty} \sum_{l=1}^{s-1} \sum_{m=s-1}^{\infty} \frac{\phi_k(\varphi_0) \phi_m(\varphi_0)}{j^l} E(\varepsilon_t \varepsilon_s) E(u_{t-j-k} u_{s-l-m}),
\]

(5.39)

noting that for any \( t, j, k, s, l, m \)

\[
   E(\varepsilon_t u_{s-l-m}) E(\varepsilon_s u_{t-j-k}) = \kappa = 0,
\]

where \( \kappa \) denotes now the fourth cumulant of \( \varepsilon_t, \varepsilon_s, u_{t-j-k}, u_{s-l-m} \). The right side of (5.39) is bounded by

\[
   K \frac{1}{n} \sum_{t=2}^{n} \int_{-\infty}^{\infty} \left| \sum_{j=1}^{t-1} \sum_{k=t-j}^{\infty} \phi_k(\varphi_0) e^{(j+k)\mu} \right|^2 d\mu \leq K \frac{1}{n} \sum_{t=2}^{n} \sum_{j=1}^{t-1} \frac{(t-j)^{1-\alpha} (t-j)^{-\alpha}}{j} \sum_{i=1}^{t-1} \frac{(t-l)^{-\alpha}}{l}.
\]

Now

\[
   \sum_{t=1}^{t-1} \frac{(t-l)^{-\alpha}}{l} = \sum_{l=1}^{[t/2]} \frac{(t-l)^{-\alpha}}{l} + \sum_{l=1}^{t-1} \frac{(t-l)^{-\alpha}}{l} \leq K \left( t^{-\alpha} \log t + t^{-1} \right),
\]

where \([ \cdot ]\) denotes integer part. Then

\[
   \text{Var}(r_2) = O\left( \frac{1}{n} \sum_{t=2}^{n} \frac{1}{j} \sum_{j=1}^{t-1} \frac{1}{j} \right) = O\left( \frac{\log^2 n}{n} \right),
\]

21
and \( r_2 = O_p(\log n/n^{1/2}) \). Next, by Lemma 2

\[
  r_3 = O_p\left(n^{-\frac{1}{2}} \sum_{t=2}^{n} t^{\frac{1}{2} - \alpha} \log t\right) = O_p\left(n^{-\frac{1}{2}}\right).
\]

Also, \( E(s_1) = 0 \), and

\[
  \text{Var}(s_1) = O\left(\left\| \frac{1}{n} \sum_{t=2}^{n} \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} \frac{\partial \phi_j(\varphi_0)}{\partial \varphi} \frac{\partial \phi_k(\varphi_0)}{\partial \varphi'} E(u_{t-j}u_{t-k}) \right\| \right)
  = O\left(\frac{1}{n} \sum_{t=2}^{n} t^{1-2\alpha}\right) = O\left(n^{-1}\right),
\]

since \( \alpha > 3/2 \), \( \| \cdot \| \) denoting Euclidean norm. Finally, by Lemmas 2 and 4

\[
  s_2 = O_p\left(n^{-\frac{1}{2}} \sum_{t=1}^{n} t^{\frac{1}{2} - \alpha}\right) = O_p(n^{-\frac{1}{2}}),
\]

to conclude the proof of (5.37).

Next, (5.38) holds by the Cramer-Wold device and Theorem 1 of Brown (1971) on showing that

\[
  E\left(\varepsilon_t \sum_{j=1}^{\infty} m_j(\varphi_0) \varepsilon_{t-j} \bigg| \mathcal{F}_{t-1}\right) = 0 \quad \text{a.s.,} \quad (5.40)
\]

and

\[
  \frac{1}{n} \sum_{t=2}^{n} E\left(\varepsilon_t^2 \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} m_j(\varphi_0) m_k(\varphi_0) \varepsilon_{t-j} \varepsilon_{t-k} \bigg| \mathcal{F}_{t-1}\right)
  = -\frac{1}{n} \sum_{t=2}^{n} E\left(\varepsilon_t^2 \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} m_j(\varphi_0) m_k(\varphi_0) \varepsilon_{t-j} \varepsilon_{t-k}\right) \rightarrow_p 0, \quad (5.41)
\]

because \( E\left(\varepsilon_t^2 \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} m_j(\varphi_0) m_k(\varphi_0) \varepsilon_{t-j} \varepsilon_{t-k} \bigg| \mathcal{F}_{t-1}\right) \) has expectation \( \sigma_A^2 \), noting that the Lindeberg condition is satisfied as \( \varepsilon_t \sum_{j=1}^{\infty} m_j(\varphi_0) \varepsilon_{t-j} \) is stationary with finite variance. First, (5.40) follows as \( \varepsilon_{t-j}, j \geq 1 \) is \( \mathcal{F}_{t-1}\)-measurable, whereas the left side of (5.41) is

\[
  \frac{\sigma^2_A}{n} \sum_{t=2}^{n} \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} m_j(\varphi_0) m_k(\varphi_0) (\varepsilon_{t-j} \varepsilon_{t-k} - E(\varepsilon_{t-j} \varepsilon_{t-k})) \rightarrow_p 0,
\]

because \( \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} m_j(\varphi_0) m_k(\varphi_0) (\varepsilon_{t-j} \varepsilon_{t-k} - E(\varepsilon_{t-j} \varepsilon_{t-k})) \) is stationary ergodic with mean equal to zero.
Next, we show (5.36). Denote $N_\epsilon (\tau_0)$ an open neighbourhood of radius $\epsilon < 1/2$ about $\tau_0$, and
\[
W_n (\tau) = \frac{1}{n} \sum_{t=2}^{n} \left( \sum_{j=0}^{t-1} \sum_{k=1}^{t-1} c_{(j,k)} \frac{\partial^2 c_k}{\partial \tau \partial \tau'} \gamma (k-j) + \sum_{j=1}^{t-1} \sum_{k=1}^{t-1} \frac{\partial c_j}{\partial \tau} \frac{\partial c_k}{\partial \tau'} \gamma (k-j) \right). \tag{5.42}
\]
Then
\[
\frac{1}{2} \frac{\partial^2 R_n (\tau)}{\partial \tau \partial \tau'} - \frac{1}{2} \frac{\partial^2 R_n (\tau)}{\partial \tau \partial \tau'} = W_n (\tau) + W_n (\tau) - W_n (\tau_0) + W_n (\tau_0).
\]
As $\tau$ is consistent for $\tau_0$, (5.36) holds on showing
\[
\sup_{\tau \in N_\epsilon (\tau_0)} \left\| \frac{1}{2} \frac{\partial^2 R_n (\tau)}{\partial \tau \partial \tau'} - W_n (\tau) \right\| = o_p (1), \tag{5.43}
\]
for some $\epsilon > 0$,\n\[
W_n (\tau) - W_n (\tau_0) = o_p (1), \tag{5.44}
\]
\[
W_n (\tau_0) \rightarrow \sigma^2_0 A, \tag{5.45}
\]
as $n \rightarrow \infty$. First, choosing $\epsilon < 1/2$, the proof for (5.43) is almost identical to that for (5.6), noting the orders in Lemma 4. Next, (5.44) holds by Slutzky’s theorem as $W_n (\tau)$ is continuous at $\tau_0$ and $\tau \rightarrow_p \tau_0$. Finally, (5.45) holds by the results given in the proof of (5.38) noting that by Lemmas 2 and 4 the norm of first term of $W_n (\tau_0)$ (see (5.42)) is equal to
\[
\left\| \frac{1}{n} \sum_{t=2}^{n} E \left( u_t (\delta_0) \frac{\partial^2 \varepsilon_t (\delta_0, \varphi_0)}{\partial \tau \partial \tau'} \right) \right\| \leq \frac{1}{n} \sum_{t=2}^{n} \left( E u_t^2 (\delta_0) E \left\| \frac{\partial^2 \varepsilon_t (\delta_0, \varphi_0)}{\partial \tau \partial \tau'} \right\|^2 \right)^{1/2} \leq \frac{K}{n},
\]
by Lemma 4, to conclude the proof.

**Proof of Theorem 3**

By standard arguments (3.5) follows on showing
\[
\sqrt{n} h_0 (\tau_0) \rightarrow_d N (0, B), \tag{5.46}
\]
\[
H_n (\tau_0) \rightarrow_p B, \tag{5.47}
\]
\[
H_n (\tau) - H_n (\tau_0) \rightarrow_p 0, \tag{5.48}
\]
for $\| \tau - \tau_0 \| \leq \| \tau - \tau_0 \|$. We only show (5.46), as (5.47), (5.48) follow from similar arguments to those given in the proof of (5.36). Noting that $\partial \varepsilon_1 (\tau_0) / \partial \tau' = 0$, whereas for $t \geq 2$, $\partial \varepsilon_t (\tau_0) / \partial \tau'$ equals
\[
\sum_{j=1}^{t-1} \left( -\Phi^{(1)} (\varphi_0) \sum_{k=1}^{t-1} \frac{1}{k} u_{t-j-k} \cdots -\Phi^{(r)} (\varphi_0) \sum_{k=1}^{t-1} \frac{1}{k} u_{t-j-k} + \frac{\partial \Phi_t (\varphi_0)}{\partial \varphi_1} u_t - \cdots - \frac{\partial \Phi_t (\varphi_p)}{\partial \varphi_p} u_{t-j} \right),
\]
by similar arguments to those in the proof of Theorem 2, it can be shown that the left side of (5.46) equals

$$\frac{1}{\sqrt{n}} \sum_{t=2}^{n} \left( \sum_{j=1}^{\infty} d_{j}^{(1)} (\varphi_{0}) \varepsilon_{t-j} \cdots \sum_{j=1}^{\infty} d_{j}^{(r+p)} (\varphi_{0}) \varepsilon_{t-j} \right) \sum_{0}^{-1} \varepsilon_{t} + o_{p} (1).$$

Then by the Cramer-Wold device, (5.46) holds if for any $(r+p)$-dimensional vector $\vartheta$ (with $i$-th component $\vartheta_{i}$)

$$\frac{1}{\sqrt{n}} \sum_{t=2}^{n} \sum_{j=1}^{\infty} \varepsilon_{t-j} M_{j}^{(r+p)} (\varphi_{0}) \Sigma_{0}^{-1} \varepsilon_{t} \rightarrow_{d} N (0, \vartheta' \vartheta),$$

where

$$M_{j} (\varphi_{0}) = \sum_{k=1}^{r+p} \vartheta_{k} d_{j}^{(k)} (\varphi_{0}).$$

As in the proof of (5.38), (5.49) holds by Theorem 1 of Brown (1971) noting that

$$E \left( \sum_{j=1}^{\infty} \varepsilon_{t-j} M_{j}^{(r+p)} (\varphi_{0}) \Sigma_{0}^{-1} \varepsilon_{t} \right) = E \left( \sum_{j=1}^{\infty} \sum_{k=1}^{r+p} \varepsilon_{t-j} M_{j}^{(r+p)} (\varphi_{0}) \Sigma_{0}^{-1} M_{k} (\varphi_{0}) \varepsilon_{t-k} \right)$$

$$= \sum_{j=1}^{\infty} \sum_{k=1}^{r+p} \text{tr} \left\{ \varepsilon_{t-j} M_{j}^{(r+p)} (\varphi_{0}) \Sigma_{0}^{-1} M_{k} (\varphi_{0}) \varepsilon_{t-k} \right\}$$

$$= \vartheta' \vartheta,$$

to conclude the proof.

6. Technical Lemmas

**Lemma 1.** Under Assumption 1

$$\varepsilon_{t} (\tau) = \sum_{j=0}^{t-1} c_{j} (\tau) u_{t-j},$$

with $c_{0} (\tau) = 1$, where for any $\delta \in [\gamma_{1}, \gamma_{2}]$, as $j \rightarrow \infty$,

$$\sup_{\varphi \in \Psi} |c_{j} (\tau)| = O \left( j^{\max (\delta_{0} - \delta - 1, -\alpha)} \right),$$

$$\sup_{\varphi \in \Psi} |c_{j+1} (\tau) - c_{j} (\tau)| = O \left( j^{\max (\delta_{0} - \delta - 2, -\alpha)} \right).$$

**Proof**

Clearly,

$$c_{j} (\tau) = \sum_{k=0}^{j} \varphi_{k} (\varphi) a_{j-k},$$

24
so that for any $\delta \in [\gamma_1, \gamma_2]$, by Stirling’s approximation

\[
\sup_{\varphi \in \Psi} \left| c_j (\tau) \right| \leq K \sum_{k=0}^{j-1} (j-k)^{\delta_0 - \delta - 1} \sup_{\varphi \in \Psi} |\phi_k (\varphi)| \leq K \sum_{k=0}^{[j/2]} (j-k)^{\delta_0 - \delta - 1} \sup_{\varphi \in \Psi} |\phi_k (\varphi)| + K \sum_{k=[j/2]}^{j-1} (j-k)^{\delta_0 - \delta - 1} \sup_{\varphi \in \Psi} |\phi_k (\varphi)|. \quad (6.4)
\]

(6.4) is bounded by

\[
K j^{\delta_0 - \delta - 1} \sum_{k=1}^{\infty} k^{-\alpha} + K j^{-\alpha} \sum_{k=0}^{j-1} (j-k)^{\delta_0 - \delta - 1} = O \left( j^{\max(\delta_0 - \delta - 1, -\alpha)} \right).
\]

The proof of (6.2) is almost identical on noting

\[
c_{j+1} - c_j = \phi_{j+1} (\varphi) + \sum_{k=1}^{j} \phi_k (\varphi) (a_{j+1-k} - a_{j-k}),
\]

and Lemma D.1 of Robinson and Hualde (2003).

**Lemma 2.** Under Assumptions 1, 2

\[
\varepsilon_t (\tau^*) = \sum_{j=0}^{t-1} a_j \varepsilon_{t-j} + v_t (\delta),
\]

where $\tau^* = (\delta, \varphi_0)$, and for any $\kappa \geq 1/2$

\[
\sup_{\delta_0 - \kappa \leq \delta \leq \delta_0 - \gamma} |v_t (\delta)| = O_p \left( t^{\kappa - 1} \right),
\]

and

\[
v_t (\delta_0) = O_p \left( t^{1/2 - \alpha} \right).
\]

**Proof**

From (6.1), (6.3)

\[
\varepsilon_t (\tau^*) = \sum_{j=0}^{t-1} a_j \sum_{k=0}^{t-j-1} \phi_k (\varphi_0) u_{t-j-k} = \sum_{j=0}^{t-1} a_j \varepsilon_{t-j} + v_t (\delta),
\]

where

\[
v_t (\delta) = -\sum_{j=0}^{t-1} a_j \sum_{k=t-j}^{\infty} \phi_k (\varphi_0) u_{t-j-k}. \quad (6.5)
\]
Thus
\[ \sup_{\delta_0 - \kappa \leq \delta < \delta_0 - \sqrt{\delta}} |v_t(\delta)| \leq K \sum_{j=1}^{t} j^{\kappa-1} \left| \sum_{k=t-j}^{\infty} \phi_k(\varphi_0) u_{t-j-k} \right|. \]

Now
\[ \text{Var} \left( \sum_{k=t-j}^{\infty} \phi_k(\varphi_0) u_{t-j-k} \right) \leq K \sum_{k=t-j}^{\infty} \phi_k^2(\varphi_0) \leq K (t-j)^{1-2\alpha}. \]

Thus
\[ \sup_{\delta_0 - \kappa \leq \delta < \delta_0 - \sqrt{\delta}} |v_t(\delta)| = O_p \left( \sum_{j=1}^{t-1} j^{\kappa-1} (t-j)^{1/2-\alpha} \right) = O_p \left( t^{\kappa-1} \right), \]
as in the proof of Lemma 1, noting that \( \alpha > 3/2 \). Finally, by (6.5)
\[ v_t(\delta_0) = - \sum_{k=t}^{\infty} \phi_k(\varphi_0) u_{t-j-k} = O_p \left( t^{1/2-\alpha} \right), \]
by previous arguments.

**Lemma 3**
\[ \sum_{j=1}^{n} I_{\tau(\lambda)}(\lambda_j) = \sum_{j=1}^{n} \left| \frac{\theta(\epsilon^{i\lambda}; \varphi_0)}{\theta(\epsilon^{i\lambda}; \varphi)} \right|^2 I_{\tau(\lambda)}(\lambda_j) + V(\tau), \tag{6.6} \]
where for any real number \( \kappa \geq 1/2 \), under Assumptions 1, 2
\[ \sup_{\varphi \in \Psi} |V(\tau)| = O_p \left( \log^2 n1(\kappa = 1/2) + n^{2\kappa-1}(\kappa > 1/2) \right). \tag{6.7} \]

**Proof**
Since
\[ \varepsilon_t(\tau) = \xi(\lambda; \varphi) \varepsilon_t(\delta, \varphi_0), \]
following similar steps as in Brockwell and Davis (1990, p.346),
\[ w_{\varepsilon(\tau)}(\lambda) = \xi_{n-1}(\epsilon^{ik}\lambda; \varphi) w_{\varepsilon(\lambda)}(\lambda) + U(\lambda; \tau), \]
where \( \xi_{n-1}(\epsilon; \varphi) = \sum_{j=0}^{n-1} \xi_j(\varphi) \epsilon^j \) and
\[ U(\lambda; \tau) = -n^{-1/2} \sum_{k=1}^{n} \xi_k(\varphi) e^{ik\lambda} \sum_{t=n-k+1}^{n} \varepsilon_t(\tau^*) e^{it\lambda}. \]
so that (6.6) holds with

$$V(\tau) = \sum_{j=1}^{n} \left( |\xi_{n-1}(e^{i\lambda_j}; \varphi)|^2 - |\xi(e^{i\lambda_j}; \varphi)|^2 \right) I_{\xi(\tau^*)}(\lambda_j) + \sum_{j=1}^{n} |U(\lambda_j; \tau)|^2 + 2 \Re \left\{ \sum_{j=1}^{n} \xi_{n-1}(e^{i\lambda_j}; \varphi) w_{\xi(\tau^*)}(\lambda_j) U(-\lambda_j; \tau) \right\}. \quad (6.8)$$

The third term of (6.8) is

$$-2 \sum_{k=1}^{n-1} \sum_{l=1}^{n-1} \sum_{s=n-l+1}^{n} \xi_k(\varphi) \xi_l(\varphi) \varepsilon_l(\tau^*) \varepsilon_s(\tau^*) \sum_{j=1}^{n} e^{(k+l-s)\lambda_j}$$

$$= -2 \sum_{k=1}^{n-1} \sum_{l=1}^{n-1} \xi_k(\varphi) \xi_l(\varphi) \sum_{s=n-l+1}^{n} \varepsilon_{s+l-k}(\tau^*) \varepsilon_s(\tau^*),$$

where by Lemma 2

$$\varepsilon_s(\tau^*) = \sum_{j=0}^{s-1} a_j \varepsilon_{s-j} + v_s(\delta). \quad (6.9)$$

By summation by parts, for \( s \geq 2 \), the first term on the right of (6.9) is

$$a_{s-1} \sum_{j=0}^{s-1} \varepsilon_{s-j} - \sum_{j=0}^{s-2} (a_{j+1} - a_j) \sum_{k=0}^{j} \varepsilon_{s-k},$$

so that

$$E \sup_{\delta_0 - \kappa < \delta < \delta_0 - \vartheta} \left| \sum_{j=0}^{s-1} a_j \varepsilon_{s-j} \right| \leq K \left( \log s1(\kappa = 1/2) + s^{\kappa-1/2}1(\kappa > 1/2) \right), \quad (6.10)$$

whereas by Lemma 2

$$E \sup_{\delta_0 - \kappa < \delta < \delta_0 - \vartheta} |v_s(\delta)| \leq K s^{\kappa-1}.$$

Then

$$E \sup_{\delta_0 - \kappa < \delta < \delta_0 - \vartheta} \left| \sum_{k=1}^{n-1} \sum_{l=1}^{n-1} \xi_k(\varphi) \xi_l(\varphi) \sum_{s=n-l+1}^{n} \varepsilon_{s+l-k}(\tau^*) \varepsilon_s(\tau^*) \right|$$

$$\leq K \left( \log^2 n1(\kappa = 1/2) + n^{2\kappa-1}/1(\kappa > 1/2) \right) \sup_{\varphi \in \Psi} \sum_{k=0}^{\infty} |\xi_k(\varphi)| \sum_{l=0}^{\infty} l! |\xi_l(\varphi)|$$

$$\leq K \left( \log^2 n1(\kappa = 1/2) + n^{2\kappa-1}/1(\kappa > 1/2) \right).$$

Following similar steps to previous ones, it is immediate to show that

$$\sup_{\delta_0 - \kappa < \delta < \delta_0 - \vartheta} \sum_{j=1}^{n} |U(\lambda_j; \tau)|^2 = O_p \left( \log^2 n1(\kappa = 1/2) + n^{2\kappa-1}/1(\kappa > 1/2) \right).$$
Finally
\[\sup_{\delta_0 - \kappa \leq \delta < \delta_0 - \sqrt{\gamma}} \sum_{j=1}^{n} \left( |\xi_{n-1} (e^{i\lambda_j})|^2 - |\xi (e^{i\lambda_j})|^2 \right) I_{\tau^*} (\lambda_j) \leq \sup_{\lambda \in [-\pi, \pi]} \sup_{\varphi \in \Psi} \left| \xi_{n-1} (e^{i\lambda})|^2 - |\xi (e^{i\lambda})|^2 \right| \times \sup_{\delta_0 - \kappa \leq \delta < \delta_0 - \sqrt{\gamma}} \sum_{t=1}^{n} e_t^2 (\tau^*).
\]

By previous results
\[\sup_{\delta_0 - \kappa \leq \delta < \delta_0 - \sqrt{\gamma}} \sum_{t=1}^{n} e_t^2 (\tau^*) = O_p \left(n \log^2 n 1 (\kappa = 1/2) + n^2 \log (\kappa > 1/2)\right),
\]
and noting that
\[\sup_{\lambda \in [-\pi, \pi]} \sup_{\varphi \in \Psi} \left| \xi_{n-1} (e^{i\lambda}; \varphi)|^2 - |\xi (e^{i\lambda}; \varphi)|^2 \right| = O \left(n^{1-\alpha}\right),
\]
the first term on the right of (6.8) is of smaller order, to conclude the proof.

**Lemma 4.** Under Assumption 4, given an open neighbourhood \(N_\epsilon (\tau_0)\) of radius \(\epsilon < 1/2\) about \(\tau_0\), as \(j \to \infty\),

\[\sup_{\tau \in N_\epsilon (\tau_0)} |c_j (\tau)| = O \left(j^{\epsilon-1}\right), \quad \sup_{\tau \in N_\epsilon (\tau_0)} |c_{j+1} (\tau) - c_j (\tau)| = O \left(j^{\max(\epsilon-2,-\alpha)}\right),
\]

\[\sup_{\tau \in N_\epsilon (\tau_0)} \left| \frac{\partial c_j (\tau)}{\partial \delta} \right| = O \left(j^{\epsilon-1} \log j\right), \quad \sup_{\tau \in N_\epsilon (\tau_0)} \left| \frac{\partial}{\partial \delta} (c_{j+1} (\tau) - c_j (\tau)) \right| = O \left(j^{-\alpha} + j^{\epsilon-2} \log j\right),
\]

\[\sup_{\tau \in N_\epsilon (\tau_0)} \left| \frac{\partial^2 c_j (\tau)}{\partial \delta^2} \right| = O \left(j^{\epsilon-1} \log^2 j\right), \quad \sup_{\tau \in N_\epsilon (\tau_0)} \left| \frac{\partial^2}{\partial \delta^2} (c_{j+1} (\tau) - c_j (\tau)) \right| = O \left(j^{-\alpha} + j^{\epsilon-2} \log^2 j\right),
\]

\[\sup_{\tau \in N_\epsilon (\tau_0)} \left| \frac{\partial c_j (\tau)}{\partial \varphi} \right| = O \left(j^{\epsilon-1}\right), \quad \sup_{\tau \in N_\epsilon (\tau_0)} \left| \frac{\partial}{\partial \varphi} (c_{j+1} (\tau) - c_j (\tau)) \right| = O \left(j^{\max(\epsilon-2,-\alpha)}\right),
\]

\[\sup_{\tau \in N_\epsilon (\tau_0)} \left| \frac{\partial^2 c_j (\tau)}{\partial \varphi \partial \delta} \right| = O \left(j^{\epsilon-1} \log j\right), \quad \sup_{\tau \in N_\epsilon (\tau_0)} \left| \frac{\partial^2}{\partial \varphi \partial \delta} (c_{j+1} (\tau) - c_j (\tau)) \right| = O \left(j^{-\alpha} + j^{\epsilon-2} \log j\right).
\]

**Proof**

The proof is very similar to that of Lemma 1, noting Lemma D.1 of Robinson and Hualde (2003). The only point worth mentioning is the calculation of the order of magnitude of \(\partial (a_{j+1} (c) - a_j (c)) / \partial c\) and \(\partial^2 (a_{j+1} (c) - a_j (c)) / \partial c^2\). First, \(\partial (a_{j+1} (c) - a_j (c)) / \partial c\) is

\[\psi (j+1+c) a_{j+1} (c) - \psi (j+c) a_j (c) - \psi (c) (a_{j+1} (c) - a_j (c)) = O \left(j^{\epsilon-2} \log j\right), \quad (6.11)
\]
as in (5.32), (5.33). Second, it can be shown that
\[
\frac{\partial^2}{\partial c^2} (a_{j+1}(c) - a_j(c)) = \psi(j + 1 + c) \frac{\partial a_{j+1}(c)}{\partial c} - \psi(j + c) \frac{\partial a_j(c)}{\partial c} + o \left( j^{c-2} \log^2 j \right)
\]

by (6.11) and similar lines to those in the treatment of (5.32).

References


