Semiparametric Estimation for Stationary Processes whose Spectra have an Unknown Pole

by

Javier Hidalgo*
London School of Economics and Political Science

Contents:

Abstract

We consider the estimation of the location of the pole and memory parameter, λ^0 and α respectively, of covariance stationary linear processes whose spectral density function $f(\lambda)$ satisfies $f(\lambda) \sim C|\lambda - \lambda^0|^{-\alpha}$ in a neighbourhood of λ^0. We define a consistent estimator of λ^0 and derive its limit distribution Z_{λ^0}. As in related optimization problems, when the true parameter value can lie on the boundary of the parameter space, we show that Z_{λ^0} is distributed as a normal random variable when $\lambda^0 \in (0, \pi)$, whereas for $\lambda^0 = 0$ or π, Z_{λ^0} is a mixture of discrete and continuous random variables with weights equal to 1/2. More specifically, when $\lambda^0 = 0$, Z_{λ^0} is distributed as a normal random variable truncated at zero. Moreover, we describe and examine a two-step estimator of the memory parameter α, showing that neither its limit distribution nor its rate of convergence is affected by the estimation of λ^0. Thus, we reinforce and extend previous results with respect to the estimation of α when λ^0 is assumed to be known a priori. A small Monte Carlo study is included to illustrate the finite sample performance of our estimators.

Keywords: Spectral density estimation, long memory processes, Gaussian Processes.

JEL No.: C14, C22.

Contact details:
Javier Hidalgo
Department of Economics
London School of Economics
email: f.j.hidalgo@lse.ac.uk

1. INTRODUCTION

Given a covariance stationary process x_t, observed at times $t = 1, 2, .., n$, the search for cyclical components, and their estimation and testing, is of undoubted interest. This is motivated by the observed periodic behaviour exhibited in many time series and manifested by sharp peaks of the spectral density estimate.

A well known model capable of generating such a periodic behaviour is the regression model

$$x_t = \mu + \rho_1 \cos\left(\lambda^0 t\right) + \rho_2 \sin\left(\lambda^0 t\right) + \varepsilon_t, \tag{1.1}$$

where ρ_1 and ρ_2 are zero mean uncorrelated random variables with the same variance and $\{\varepsilon_t\}$ is a stationary sequence of random variables independent of ρ_1 and ρ_2. Model (1.1) has enjoyed extensive use and different techniques have been proposed for the estimation of the frequency, amplitude and phase (see Whittle (1952), Grenander and Rosenblatt (1957), Hannan (1961, 1973) and Chen (1988)). Extensions to a model with more than one periodic component have been examined by Quinn (1989) and Kavalieris and Hannan (1994), whose interest was also in testing the number of sinusoidal/cosinusoidal components. See also Quinn and Hannan (2001).

A second statistical model capable of exhibiting peaks in its spectral density function is the autoregressive $AR\left(2\right)$ process

$$\left(1 - a_1 L - a_2 L^2\right) x_t = \varepsilon_t \tag{1.2}$$

when the roots of the polynomial $\left(1 - a_1 L - a_2 L^2\right)$ are complex, with λ^0 identified as the $arc\cos\left(\frac{a_1}{2\sqrt{-a_2}}\right)$. Models (1.1) and (1.2) represent two extreme situations explaining cyclical behaviour of the data and the peakness of the spectral density function. Model (1.2) possesses a continuous spectral density function whereas model (1.1) has a spectral distribution function with a jump at the frequency λ^0. The cyclical component of the data remains constant or invariant with time in model (1.1), whereas the cyclical pattern of model (1.2) fades out with time fairly quickly.

Between these two extreme situations, there exists a class of intermediate models in which the spectral density function of x_t exhibits a pole at the frequency λ^0. For that purpose, define the spectral density function of x_t as the function $f\left(\lambda\right)$ which satisfies the relationship

$$\gamma\left(j\right) = Cov\left(x_t, x_{t+j}\right) = \int_{-\pi}^{\pi} f\left(\lambda\right) \cos\left(j\lambda\right) d\lambda, \quad j = 0, 1, 2, ... \tag{1.3}$$

We say that $f\left(\lambda\right)$ has a pole at λ^0 if

$$f\left(\lambda\right) \sim C \left|\lambda - \lambda^0\right|^{-\alpha} \quad \text{as} \quad \lambda \to \lambda^0, \tag{1.4}$$

where $C \in (0, \infty)$, $\alpha \in (0, 1)$ is the memory parameter and "$\sim$" means that the ratio of the left- and right-hand sides tends to 1. One of the main objectives of this paper is the estimation of λ^0.

One model capable of generating such a cyclical behaviour in the data has been proposed by Andel (1986) and Gray et al. (1989), and defined as

$$\left(1 - 2\left(\cos\lambda^0\right) L + L^2\right)^d x_t = \varepsilon_t, \tag{1.5}$$

where L is the backshift operator, $d = \alpha/2$ for $\lambda^0 \in (0, \pi)$ whereas for $\lambda^0 = 0$ or π, $d = \alpha/4$. The model (1.5) was coined by Gray et al. (1989) as the Gegenbauer model, who extended it to the $GARMA$ model where the innovations $\{\varepsilon_t\}$ follow an autoregressive moving average ($ARMA$) process, and later by Giraitis

and Leipus (1995) allowing for more than one pole or cyclical component. The $GARMA$ process is characterized by having the spectral density function

$$f(\lambda) = \frac{\sigma^2}{2\pi} \left| 1 - 2\left(\cos\lambda^0\right) e^{i\lambda} + e^{i2\lambda} \right|^{-2d} \left| \frac{a\left(e^{i\lambda};\theta\right)}{b\left(e^{i\lambda};\theta\right)} \right|^2, \qquad -\pi < \lambda \leq \pi, \quad (1.6)$$

where $\sigma^2 > 0$, and $a(\cdot)$ and $b(\cdot)$ are polynomials of finite degree, all of whose zeroes lie outside the unit circle. When $\lambda^0 = 0$, we have the more familiar $FARIMA$ model, apparently originated by Adenstedt (1974), and studied by Granger and Joyeux (1980) and Hosking (1981). $GARMA$ models are characterized by a stronger and more persistent cyclical behaviour than $ARMA$ models, i.e. (1.2), but unlike model (1.1), their amplitude does not remain constant over time.

If the location of the pole λ^0 is known, then under some regularity conditions and a correct specification of the model, Whittle estimates of the parameters α (or d), θ and σ^2, for example model (1.6), are known to be $n^{1/2}$-consistent and asymptotically normal. In the case of Gaussianity or linear processes, this was shown by Fox and Taqqu (1986), Dahlhaus (1989), Giraitis and Surgailis (1990) when $\lambda^0 = 0$ and generalized by Giraitis and Leipus (1995) and Hosoya (1997) for λ^0 different than 0. All these papers assume that $f(\lambda)$ is fully specified by a set of parameters $\left(\alpha, \theta', \sigma^2\right)'$.

Although knowledge of λ^0 can be realistic in some time series data, with non-seasonal data that knowledge of λ^0 is not so clear. An example of the latter is when the practitioner is interested in estimating cycles in macroeconomic or geophysical data. Recently, Giraitis et al. (2001) have shown that Whittle estimates of $\left(\alpha, \theta', \sigma^2\right)'$ are asymptotically the same irrespective of whether or not λ^0 is known. In addition, they prove that the estimate of λ^0 is n-consistent although its limit distribution remains an open problem.

However, if the ultimate interest is only the estimation of the memory parameter α, one possible criticism of the parametric approach is that an incorrect specification of the model leads to inconsistent estimates of α. One source of misspecification is the choice of a wrong value of λ^0. If that were the case, Whittle estimates of α would be inconsistent, and it would possibly estimate the value *zero*. The latter might happen even if a semiparametric approach was adopted, see Section 3. Thus, we might conclude that the data is short memory instead of long memory dependent, which could have some adverse effects on the statistical inferences of relevant statistics such as the serial covariance, see Hannan (1976) or Taqqu (1975).

The main objectives of this paper are twofold. First, under mild conditions, to provide a consistent estimator of λ^0 and characterize its limit distribution. In particular we show that the limiting distribution of the estimator of λ^0 depends on whether $\lambda^0 \in (0, \pi)$ or $\lambda^0 = 0$ or π. The second objective is to investigate the consequences that the lack of knowledge of λ^0 might have on the estimation of α.

Some earlier related work has been completed by Müller and Prewitt (1992) and Yajima (1996). In the former, they estimate the peak of the spectral density $f(\lambda)$ in a model, like that in (1.2), whose spectral density function is continuous in $[0, \pi]$. Looking at $\arg\sup_\lambda \widehat{f}(\lambda)$, where $\widehat{f}(\lambda)$ is a smoothed nonparametric estimate of $f(\lambda)$, they show its consistency and the limit distribution to be a normal random variable when $\lambda^0 \in (0, \pi)$. Yajima (1996) considers the estimation of λ^0 in a model with spectral density possessing a pole at λ^0. Based on the maximum of the periodogram of the data, he gave consistency and an upper rate of convergence for the estimate of λ^0. Unfortunately, the limit distribution was not provided which is required for statistical inferences. In addition, his results rely on the assumption that the data is Gaussian which is not required in the present paper. Finally, it

should be mentioned that Giraitis and Leipus (1995) prove the consistency of λ^0 in a model like (1.6).

The paper is organized as follows. In Section 2, we describe a semiparametric estimator of α when λ^0 is known and the estimator $\widehat{\lambda}^0$ of λ^0. In Section 3, we discuss the statistical properties of $\widehat{\lambda}^0$ and we show that the asymptotic properties of a two-step estimator of α remain the same irrespective of whether λ^0 is known or estimated. The finite sample behaviour of the estimators of λ^0 and α are analyzed in Section 4 through a Monte-Carlo study. Section 5 provides the proofs of the results given in Section 3 which apply a series of lemmas given in Section 6. Finally, Section 7 contains a summary.

2. REGULARITY CONDITIONS AND THE ESTIMATORS OF THE POLE AND MEMORY PARAMETER

Let $\{x_t\}$ be a covariance stationary linear process observed at times $t = 1, 2, ..., n$, with spectral density $f(\lambda)$ satisfying (1.4). When λ^0 is known, under the semiparametric specification (1.4), several estimators of the memory parameter α have been proposed and examined. In this paper, to estimate α we shall use a modification of the log-periodogram estimator, see Robinson (1995a), which we now describe. Consider the average periodogram spectral density estimator of $f(\lambda)$,

$$\ddot{f}_\ell = \ddot{f}(\lambda_\ell) = \frac{1}{2k_1 + 1} \sum_{|j| \leq k_1} I_{\ell+j}, \qquad (2.1)$$

where $I_\ell = I(\lambda_\ell)$ denotes the periodogram of x_t, that is

$$I_\ell = \left| (2\pi n)^{-1/2} \sum_{t=1}^n x_t e^{it\lambda_\ell} \right|^2 \qquad \ell = 1, ..., [n/2] \qquad (2.2)$$

and $I_0 = 0$, where $\lambda_\ell = (2\pi\ell)/n$ for $\ell = 1, 2, ..., [n/2]$, $[z]$ denotes the integer part of z and $k_1 = k_1(n)$ is a positive number which increases slowly with n, that is $k_1^{-1} + n^{-1}k_1 \to 0$. (Observe that the definition of I_ℓ entails sample-mean correction.)

Let $\psi(u)$ be a weight function in $(0, 1)$, and write

$$\widehat{\alpha}(\lambda_q) = \frac{1}{2\overline{h}_\psi k} \sum_{p=1}^k \psi_p \left(\log \widehat{f}_{q+p} + \log \widehat{f}_{q-p} \right), \qquad (2.3)$$

where $\psi_p = \psi(p/k)$, $\overline{h}_\psi = -k^{-1} \sum_{p=1}^k \psi_p \log(p/k)$, $\widehat{f}_\ell = \max\left(\ddot{f}_\ell, n^{-1}\right)$ and $k = k(n)$, a positive number which increases slowly with n, that is $k^{-1} + n^{-1}k \to 0$.

Definition 2.1. *If λ^0 is known, we define the estimator of the memory parameter α as $\widehat{\alpha}(\lambda_s)$, where λ_s is the closest Fourier frequency λ_q to λ^0.*

Remark 2.1. *The motivation to use $\widehat{f}_\ell$ instead of $\ddot{f}_\ell$ in (2.3) is due to the singular behaviour of $\log x$ at $x = 0$. Specifically for the proof of tightness, i.e. Proposition 5.4 in Section 5, we have not been able to bound some probabilities or moments for all $n \geq n_0$ as required. This problem, of course, does not appear as $n \to \infty$ as can be observed from Propositions 5.1 to 5.3 nor if our goal were to examine the behaviour of $\widehat{\alpha}(\lambda_s)$. We do not believe that this adjustment is needed in practice and have made it here only because we cannot establish Theorem 3.2 (cf. Proposition 5.4) without it, unless some additional stronger conditions were introduced. For instance, the normality of the data.*

We now define our estimator of λ^0 as $\widehat{\lambda}^0 = \lambda_{\widehat{q}} = (2\pi\widehat{q})/n$, where

$$\widehat{q} = \arg \max_{q=0,...,[n/2]} \widehat{\alpha}(\lambda_q). \qquad (2.4)$$

Note that periodicity and symmetry around zero implies that it suffices to search for the maximum in (2.4) at frequencies λ_q, with $q = 0, ..., [n/2]$. From (2.4) we could define $\widehat{\alpha}\left(\widehat{\lambda}^0\right)$ as an estimator of α, that is (2.3) evaluated at $\widehat{\lambda}^0$. However, see Section 3 below, since $\widehat{\alpha}\left(\widehat{\lambda}^0\right)$ does not have optimal properties, we will describe a two-step estimator, denoted $\check{\alpha}\left(\check{\lambda}^0\right)$, which overcomes all the adverse properties of $\widehat{\alpha}\left(\widehat{\lambda}^0\right)$, see Theorem 3.4 below.

The motivation for the estimator in (2.4) is as follows. From the proof of Theorem 3.4 part (a) below, it is easily shown that $\widehat{\alpha}(\lambda_s)$ is a consistent estimator of α. On the other hand, if λ_q is in any open set outside λ^0, that is $|\lambda_q - \lambda_s| > \delta$, for any arbitrarily small $\delta > 0$, Condition C.1 below implies that

$$f(\lambda_q) = \left|\lambda_q - \lambda^0\right|^{-\alpha} g(\lambda_q) \sim Cg(\lambda_q).$$

That is, at the frequency λ_q the spectral density function behaves as if α were equal to zero. So, from the proof of Theorem 3.4 part (a), we should expect that $\widehat{\alpha}(\lambda_q) \xrightarrow{P} 0$, implying that $\Pr\left\{\left|\widehat{\lambda}^0 - \lambda_s\right| < \delta\right\} \to 1$. That is, the estimator given in (2.4) is consistent. These heuristics will be formalized in Theorem 3.1 below. We finish this section introducing the following regularity conditions and their discussion.

C.1: There exists $\alpha \in (0, 1)$ such that

$$f(\lambda) = \begin{cases} \left|\lambda - \lambda^0\right|^{-\alpha} g(\lambda), & \text{if } 0 \leq \lambda \leq \pi, \\ \left|\lambda + \lambda^0\right|^{-\alpha} g(\lambda), & \text{if } -\pi \leq \lambda \leq 0, \end{cases}$$

where $g(\lambda)$ is a bounded symmetric and bounded away from zero function with two continuous derivatives for $0 < \lambda < \pi$.

C.2: $\{x_t\}$ is a covariance stationary linear process

$$x_t = \sum_{j=0}^{\infty} \beta_j \varepsilon_{t-j}; \qquad \sum_{j=0}^{\infty} \beta_j^2 < \infty,$$

where $\{\varepsilon_t\}$ is a zero mean iid sequence with $E\left(\varepsilon_t^2\right) = 1$ and $E\left|\varepsilon_t\right|^\ell = \mu_\ell < \infty$ for $\ell = 3, ..., 2\tau$ and some $\tau \geq 4$.

C.3: As $\lambda \to \lambda$, the function $\beta(\lambda) = \sum_{j=0}^{\infty} \beta_j e^{ij\lambda}$ satisfies
$\left|\partial\beta(\lambda)/\partial\lambda\right| = O\left(\left|\lambda - \lambda^0\right|^{-1} |\beta(\lambda)|\right).$

C.4: $k^{1+\iota}k_1^{-2} + k^{-2}k_1^3 \log k + nk_1^{-(\tau^2+2)/2\tau} \to 0$, for some $\iota > 0$ as $n \to \infty$, with $k \leq cn^{4/5}$, $0 < c < \infty$ and where τ is as in C.2.

C.5: The function $\psi(x)$ is twice continuously differentiable whose second derivative is Lipschitz of order at least $\frac{1}{2}$ in its support $(0, 1)$ and satisfies $\int_0^1 \psi(x)\,dx = 0$, $0 < h_\psi = -\int_0^1 \psi(x)(\log x)\,dx < \infty$, $0 < \overline{\psi}'' = \int_0^1 \psi''(x)(\log x)\,dx$, where $\psi''(x) = \frac{d^2}{dx^2}\psi(x)$, and $\left|x^{-2}\psi(x)\right| + \left|(1-x)^{-1}\psi(x)\right| \leq D < \infty.$

We now discuss Conditions C.1–C.5. Condition C.1 is much the same as that employed by Robinson (1995a, b). Indeed, C.1 implies that as $\lambda \to \lambda^0$,

$$f(\lambda) = C\left|\lambda - \lambda^0\right|^{-\alpha}\left(1 + C^{-1}g'\left(\lambda^0\right)\left(\lambda - \lambda^0\right) + O\left(\left|\lambda - \lambda^0\right|^2\right)\right)$$

by Taylor expansion of $g(\lambda)$ around λ^0 and where $C = g\left(\lambda^0\right)$. Observe that $g'\left(\lambda^0\right) = 0$ when $\lambda^0 = \{0, \pi\}$ by symmetry of $f(\lambda)$ obtaining then the corresponding condition used in Robinson (1995a, b). However, we prefer to state the condition in its present form since, in Theorems 3.1 and 3.2 below, some regularity

conditions on $f(\lambda)$ are needed outside any open set containing $\pm\lambda^0$. Examples of processes whose spectral density function satisfies C.1 are the $FARIMA(p, \alpha/2, q)$ and the $GARMA$ model given in (1.6). Finally, the last part of Condition C.1 is quite standard in spectral density estimation literature. C.2 is needed for the proof of tightness (see the proof of Theorems 3.2 and 3.1). It is also required to show the uniform convergence of $\widehat{f}$ although for the latter property, at the expense of stronger conditions on the rate of convergence of k_1^{-1} to zero, fewer moments of ε_t can be assumed. Obviously C.2 is satisfied if ε_t is Gaussian. Condition C.3 is the same as Robinson (1995b). Condition C.4 controls the rate of increase of k and k_1. For instance denoting $k = n^{\gamma_2}$ and $k_1 = n^{\gamma_1}$, in the Gaussian case, we can take $0 < \gamma_1 < 8/15$ and $3\gamma_1/2 < \gamma_2 \leq \min\{2\gamma_1/(1+\iota), 4/5\}$, whereas for $\tau = 4$, the bounds are $4/9 < \gamma_1 < 8/15$ and $3\gamma_1/2 < \gamma_2 \leq 4/5$. Finally, Condition C.5 characterizes the type of weight in (2.3). An example of $\psi(u)$ satisfying C.5 is $\psi(u) = -u^2 + 35u^{2.5}/6 - 29u^3/6 + 2u^3 \log u$.

It is worth mentioning that the quadratic behaviour of the weight $\psi(u)$, as $u \to 0$, guarantees that the first moment of $\xi(v)$, see Theorem 3.2 below, has a parabolic structure so that the maximum of $\xi(v)$ is easily obtainable. Obviously, other different types of weights can be used which would not prevent the consistency of the estimator of λ^0. However, for weights not having a quadratic behaviour, the asymptotic distribution of the estimate of the pole is not guaranteed to be normally distributed. We will return to this condition subsequently after Theorem 3.2.

3. STATISTICAL PROPERTIES OF THE ESTIMATORS OF THE POLE AND MEMORY PARAMETER

In this section, we prove a functional limit theorem for a process operating on increments of $\widehat{\alpha}(\lambda_q)$ near λ^0 which together with the continuous mapping theorem will allow obtaining the asymptotic distribution of $\widehat{\lambda}^0$. A similar approach was used by Eddy (1980) to estimate the mode of a probability density function and Müller (1992) for the estimation of the break point in a regression model. Apart from providing the consistency and rate of convergence of $\widehat{\lambda}^0$ to λ^0, the limit distribution will guarantee that asymptotic valid inferences around the true value of λ^0 may be implemented.

The strategy of the proof to obtain the asymptotic distribution of $\widehat{\lambda}^0$ consists of three steps, see Van der Vaart and Wellner (1996, Ch. 3). Step 1 establishes the consistency of $\widehat{\lambda}^0$ to λ^0. Step 2 establishes the rate of convergence of $\widehat{\lambda}^0$ to λ^0, and Step 3 shows that suitably rescaled versions of $\widehat{\alpha}(\lambda_q)$ converges weakly to a limit, denoted $\xi(v)$ in Theorem 3.2, in the space $\mathbb{D}[-M, M]$ for each finite $0 < M < \infty$. Note that convergence in $\mathbb{D}[-M, M]$ for each finite $0 < M < \infty$ is to be meant convergence in $\mathbb{D}(-\infty, \infty)$. See Pollard (1984). From here, the continuous mapping theorem will conclude that $\widehat{\lambda}^0$, after normalization, will converge in distribution to the $\arg\max_v \xi(v)$.

The next theorem gives the consistency and rate of convergence of $\widehat{\lambda}^0$ to λ^0, i.e. Steps 1 and 2. Theorem 3.2 justifies Step 3, whereas Corollary 3.3 examines the asymptotic distribution of $\widehat{\lambda}^0$.

Theorem 3.1. *Assuming C.1-C.5,* $\left|\widehat{\lambda}^0 - \lambda^0\right| = O_p\left(k^{1/2}n^{-1}\right)$.

We see that the rate of convergence of $\widehat{\lambda}^0$ to λ^0 is slower than the parametric rate n^{-1} obtained by Giraitis et al. (2001). This appears to be reasonable due to the local behaviour of our statistics. The same phenomenon occurs in other related,

although different, problems involving nonparametric statistics, see for instance Eddy (1980), Müller (1992) or Müller and Prewitt (1992).

Under Conditions C.2 and C.4, for $\tau = 4$, $\left(\widehat{\lambda}^0 - \lambda^0\right) = O_p\left(n^{\delta - 2/3}\right)$ for any arbitrarily small $\delta > 0$. However, a closer examination of these conditions and the proof of Lemma 6.3 indicate that the rate depends on the number of finite moments of the sequence ε_t in C.2. In general, with $\tau \geq 4$, $\left(\widehat{\lambda}^0 - \lambda^0\right) = O_p\left(n^{\delta - \left(2\tau^2 - 3\tau + 4\right)/2\left(\tau^2 + 2\right)}\right)$. So, the greater the number of finite moments allowed for ε_t, the faster the rate of convergence of $\widehat{\lambda}^0$ to λ^0. In the extreme case where all the moments exist, the rate of convergence of $\widehat{\lambda}^0$ becomes $n^{\delta - 1}$. This rate was obtained by Yajima (1996) in the Gaussian case and is arbitrarily close to n^{-1} obtained in Giraitis et al. (2001).

So Theorem 3.1 indicates that $\widehat{\lambda}^0 = \lambda^0 + n^{-1}\left(2\pi\left[k^{1/2}\upsilon\right]\right)$ for some $|\upsilon| \leq M < \infty$. To examine the asymptotic distribution of $\widehat{\lambda}^0$, let us introduce the following notation

$$\widehat{\xi}_n(\upsilon) = k\left(\widehat{\alpha}\left(\lambda_{s+\left[k^{1/2}\upsilon\right]}\right) - \widehat{\alpha}\left(\lambda_s\right)\right). \tag{3.1}$$

$\widehat{\xi}_n(\upsilon)$ is a random step function which is constant in the intervals $\left[i/k^{1/2}, (i+1)/k^{1/2}\right)$, $|i| \leq M$, so that $\widehat{\xi}_n(\upsilon)$ is a random element in the Skorohod space $\mathbb{D}\left[-M, M\right]$ for arbitrary $0 < M < \infty$.

We now establish our main result, i.e. the aforementioned Step 3.

Theorem 3.2. *Let $|\upsilon| \leq M$, for any arbitrary $M \in (0, \infty)$. Assuming C.1-C.5,*

$$\widehat{\xi}_n(\upsilon) \overset{weakly}{\Longrightarrow} \xi(\upsilon) \quad \text{in the space} \quad \mathbb{D}\left[-M, M\right],$$

where $\xi(\upsilon)$ is a continuous Gaussian process such that

$$E\left(\xi(\upsilon)\right) = -h_\psi^{-1}\overline{\psi}''\upsilon^2\alpha/2 \quad \text{and} \quad Cov\left(\xi(\upsilon_1), \xi(\upsilon_2)\right) = h_\psi^{-2}\varsigma\upsilon_1\upsilon_2,$$

where $\varsigma = \int_0^1 \psi'(u)^2\, du < \infty$ with $\psi'(x) = \frac{d}{dx}\psi(x)$.

The immediate consequence of Theorem 3.2 is that $\arg\max_\upsilon \xi(\upsilon)$ is a normal random variable. Indeed, because Theorem 3.2 implies that the limiting process $\xi(\upsilon)$ is Gaussian, it can be written as

$$\xi(\upsilon) = -h_\psi^{-1}\overline{\psi}''\upsilon^2\alpha/2 + h_\psi^{-1}\varsigma^{1/2}\upsilon X$$

where $X = N(0, 1)$. But $\xi(\upsilon)$ is a random parabola with fixed second derivatives and a unique maximum at

$$\upsilon^* = \left(\overline{\psi}''\alpha\right)^{-1}\varsigma^{1/2}X,$$

since by C.5, $0 < h_\psi, 0 < \overline{\psi}''$, so that $\partial^2\xi(\upsilon)/\partial\upsilon^2 = -h_\psi^{-1}\overline{\psi}''\alpha < 0$. From here we can observe the (possible) consequences of using a weight function $\psi(u)$ which does not have a parabolic structure at $u = 0$. The main implication is that if the latter were the case, $E(\xi(\upsilon))$ would not necessarily be a parabola as in Theorem 3.2. For example, it may be that $E(\xi(\upsilon)) = C|\upsilon|$, in which case not only the $\arg\max_\upsilon \xi(\upsilon)$ can be difficult to obtain, but more importantly it would no longer be a normal random variable. So, in view of the asymptotic normality achieved with a weight $\psi(u)$ satisfying condition C.5, it appears desirable to employ it. Similar issues occur when estimating the date of a break in a regression model, see for example Müller (1992).

Now we turn our attention to the asymptotic properties of $\widehat{\lambda}^0$. Note that Theorem 3.1 indicates that

$$\widehat{\lambda}^0 = \lambda_s + \frac{2\pi k^{1/2}}{n}\widehat{\upsilon}_n = \lambda^0 + \frac{2\pi k^{1/2}}{n}\widehat{\upsilon}_n + O\left(\frac{1}{n}\right), \tag{3.2}$$

where $\widehat{v}_n = \arg\max_v \widehat{\xi}_n(v)$. Then, we have the following:

Corollary 3.3. *Denote* $\Psi = \varsigma\left(\overline{\psi}''\alpha\right)^{-2}$. *Assuming C.1-C.5, as* $n \to \infty$

(a) *If* $\lambda^0 \in (0, \pi)$, *then* $\left(2\pi k^{1/2}\right)^{-1} n\left(\widehat{\lambda}^0 - \lambda^0\right) \xrightarrow{d} Z_{\lambda^0} \equiv Y = N(0, \Psi)$.

(b) *If* $\lambda^0 = 0$, *then* $\left(2\pi k^{1/2}\right)^{-1} n\widehat{\lambda}^0 \xrightarrow{d} Z_0 = Y\mathcal{I}(Y \geq 0)$, *where* $\mathcal{I}(A)$ *denotes the indicator function of the set* A.

(c) *If* $\lambda^0 = \pi$, *then* $\left(2\pi k^{1/2}\right)^{-1} n\left(\widehat{\lambda}^0 - \pi\right) \xrightarrow{d} Z_\pi = Y\mathcal{I}(Y \leq 0)$.

We now comment on the results of Corollary 3.3. First, we now see the necessity of Theorem 3.1 as it will give us the normalization needed to achieve a "proper" asymptotic distribution. Next, we observe that the limiting distribution of $\widehat{\lambda}^0$ depends on whether λ^0 is $\{0, \pi\}$ or $\lambda^0 \in (0, \pi)$. The intuition about the limiting distribution of $\widehat{\lambda}^0$ in cases (b) and (c) is as follows. As the maximization of $\widehat{\alpha}(\lambda_q)$ in (2.4) is restricted to the interval $0 \leq \lambda_q \leq \pi$, for $\lambda^0 = 0$, it implies that $\left(\widehat{\lambda}^0 - \lambda^0\right) = \widehat{\lambda}^0 \geq 0$ so that Z_0 cannot take negative values. Similarly, $\lambda^0 = \pi$ implies that $\widehat{\lambda}^0 - \pi \leq 0$ and Z_π cannot take positive values. So, the estimation of λ^0 falls into the category of a constrained optimization problem or inequality constraint estimation. Indeed, when λ^0 is an interior point of the set $[0, \pi]$ and due to the consistency of $\widehat{\lambda}^0$, we can expect that the constrained estimator, $\widehat{\lambda}^0 = \lambda_{\widehat{q}}$, coincides with the unconstrained estimator $\widetilde{\lambda}_q = \lambda_{\widetilde{q}} = (2\pi\widetilde{q})/n$ where

$$\widetilde{q} = \arg\max_{q \in \{0, \pm 1, \pm 2, \ldots\}} \widehat{\alpha}(\lambda_q), \tag{3.3}$$

whereas if $\lambda^0 = 0$, $\widehat{q} = \widetilde{q}\mathcal{I}(\widetilde{q} \geq 0)$. Similar arguments apply when $\lambda^0 = \pi$.

Once we have examined the properties of $\widehat{\lambda}^0$, we next examine the estimation of α. By Theorems 3.1 and 3.2 and the functional mapping theorem, it is easily shown that $\widehat{\alpha}\left(\widehat{\lambda}^0\right) - \widehat{\alpha}\left(\lambda^0\right) = o_p\left(k^{-1/2}\right)$. So, $k^{1/2}\left(\widehat{\alpha}\left(\widehat{\lambda}^0\right) - \alpha\right)$ and $k^{1/2}\left(\widehat{\alpha}\left(\lambda^0\right) - \alpha\right)$ have the same asymptotic distribution. However, the faster the convergence of $\widehat{\lambda}^0$ to λ^0, the slower the rate of convergence of $\widehat{\alpha}\left(\widehat{\lambda}^0\right)$ to α, and hence it becomes slower than the rate obtained when λ^0 is known. The same phenomena happens to hold in Müller (1992). Hence, to circumvent this drawback, as in Müller (1992), we propose a *two-step* procedure to estimate α. To that end, we shall use as an estimator of α that given in (2.3) where λ_q is replaced by $\widecheck{\lambda}^0 = (2\pi\widecheck{q})/n$ such that $\left|\widecheck{\lambda}^0 - \lambda^0\right| = O_p\left(k^{1/2}/n\right)$, and k and k_1 being replaced by m and m_1 respectively, satisfying

C.6: $m^{-1} + mm_1^{-2} + m_1^5 m^{-3}\log m_1 + k/m \to 0$ with $m = cn^{4/5}$, $0 < c < \infty$.

In addition, to be a bit more general regarding our choice of the weight function $\psi(u)$, we allow for the weighted function, say $w(u)$, to satisfy

C.7: $\int_0^1 w(u)\, du = 0$, $0 < h_w = -\int_0^1 w(u)(\log u)\, du < \infty$, $w(u) \sim cu^\varsigma$ as $u \to 0+$ for some $1/3 \leq \varsigma \leq 1$ and for all $0 < u_1 < u_2 < 1$,

$$|w(u_2) - w(u_1)| \leq D|u_2 - u_1|^\varsigma; \quad 0 < D < \infty.$$

So, our *two-step* estimator of α is defined as

$$\widecheck{\alpha}\left(\widecheck{\lambda}^0\right) = \frac{1}{2\overline{h}_w m}\sum_{p=1}^m w_p\left(\log\widehat{f}_{\widecheck{q}+p} + \log\widehat{f}_{\widecheck{q}-p}\right), \tag{3.4}$$

where $\widehat{f}(\lambda) = \max\left\{\ddot{f}(\lambda), n^{-1}\right\}$ and $\ddot{f}(\lambda)$ is as in (2.1) but with smoothing parameter k_1 there being replaced by m_1, $\overline{h}_w = -m^{-1}\sum_{p=1}^{m} w_p \log(p/m)$ and $w_p = w(p/m)$.

We now comment on $\check{\alpha}\left(\check{\lambda}^0\right)$ compared to $\ddot{\alpha}\left(\check{\lambda}^0\right) = \left(\overline{h}_w m\right)^{-1}\sum_{p=1}^{m} w_p \log \widehat{f}_{\check{q}+p}$. Observe that the former is a "symmetrized" version of the latter $\ddot{\alpha}\left(\check{\lambda}^0\right)$. Assume for simplicity that λ^0 is known. As in other semiparametric estimators, for example Robinson (1995a), one source of the bias of $m^{1/2}\left(\ddot{\alpha}\left(\lambda^0\right) - \alpha\right)$ comes from the replacement of $f(\lambda)$ by $g\left(\lambda^0\right)\left|\lambda - \lambda^0\right|^{-\alpha}$, which in our case, i.e. if $\lambda^0 \neq \{0, \pi\}$, will be proportional to

$$m^{-1/2}\sum_{p=1}^{m} w_p\left(\lambda_p + O\left(\lambda_p^2\right)\right) = O\left(n^{-1}m^{3/2}\right).$$

The main reason for this behaviour is that when $\lambda^0 = \{0, \pi\}$, by symmetry we have $g'(0) = g'(\pi) = 0$, whereas for $\lambda \neq 0$ or π, $g'(\lambda)$ may not be zero so that $g^{-1}\left(\lambda^0\right)\left|\lambda - \lambda^0\right|^{\alpha}f(\lambda) = 1 + g^{-1}\left(\lambda^0\right)g'\left(\lambda^0\right)\left(\lambda - \lambda^0\right) + O\left(\left|\lambda - \lambda^0\right|^2\right)$ by a Taylor expansion of $g(\lambda)$ around λ^0. Recall the comments made on C.1. However, when the estimator $\check{\alpha}\left(\lambda^0\right)$ in (3.4) is employed, the contribution of the above approximation (Taylor expansion) to the bias of $m^{1/2}\left(\check{\alpha}\left(\lambda^0\right) - \alpha\right)$ is proportional to

$$m^{-1/2}\sum_{p=1}^{m} w_p\left(-\lambda_p + O\left(\lambda_p^2\right)\right) + m^{-1/2}\sum_{p=1}^{m} w_p\left(\lambda_p + O\left(\lambda_p^2\right)\right) = O\left(m^{5/2}n^{-2}\right).$$

Note that the latter holds true also for $\lambda^0 = \{0, \pi\}$. So, the "symmetrized" estimator $\check{\alpha}\left(\lambda^0\right)$ would have a smaller bias order and thus it would have a faster rate of convergence to α than $\ddot{\alpha}\left(\lambda^0\right)$.

Theorem 3.4. *Denote* $\Phi^2 = 2^{-1}\int_0^1 w^2(x)\,dx$ *and* $B = \left(\partial^2/\partial\lambda^2 \log g\left(\lambda^0\right)\right)\int_0^1 u^2 w(u)\,du$. *Let* $\check{\lambda}^0$ *be an estimator of* λ^0 *that* $\left|\check{\lambda}^0 - \lambda^0\right| = O_p\left(k^{1/2}/n\right)$. *Assuming C.1-C.4 with* $k_1 = n^{\gamma_1}$ *and* $k = n^{\gamma_2}$, *where* $2\tau/\left(\tau^2 + 2\right) < \gamma_1 < 8/15$, $3\gamma_1/2 < \gamma_2 < \min\left\{\frac{2\gamma_1}{1+\iota}, \frac{4}{5}\right\}$, τ *as in C.2, and C6 and C7, then*

$$(a) \quad (2m)^{1/2}\left(\check{\alpha}\left(\lambda_s\right) - \alpha\right) \xrightarrow{d} N\left(4\pi^2 c^{5/2}B/\left(2^{1/2}h_w\right), \Phi^2/h_w^2\right)$$

$$(b) \quad (2m)^{1/2}\left(\check{\alpha}\left(\check{\lambda}^0\right) - \alpha\right) \xrightarrow{d} N\left(4\pi^2 c^{5/2}B/\left(2^{1/2}h_w\right), \Phi^2/h_w^2\right).$$

Remark 3.1. *It is worth mentioning that the results of Theorem 3.4 hold true if the weight* $\psi(u)$ *employed to estimate* λ^0 *were used in* $\check{\alpha}\left(\lambda^0\right)$. *However, this weight will not guarantee an asymptotic variance smaller than 1, as is the case with the weight used in the Monte-Carlo. In fact, for the weight used in the Monte-Carlo experiment,* $h_w^{-2}\Phi^2 \sim .70$ *which is smaller than the corresponding asymptotic variance of other estimators of* α *suggested in the literature. Finally, the theorem indicates that although any preliminary estimator of* λ^0, *which satisfies* $\left|\check{\lambda}^0 - \lambda^0\right| = O_p\left(k^{1/2}/n\right)$, *is adequate for the results to follow, in practice it appears that one may use that given in (2.4) for computational simplicity.*

Theorem 3.4 provides a consistent estimator of the asymptotic variance of $\widehat{\lambda}^0$, i.e. Ψ in Corollary 3.3, by replacing α by $\check{\alpha}\left(\check{\lambda}^0\right)$. But more importantly, it indicates that the *two-step* estimator $\check{\alpha}\left(\check{\lambda}^0\right)$, apart from having the same asymptotic

distribution as $\check{\alpha}\left(\lambda^0\right)$, achieves the optimal semiparametric rate of convergence obtained by Giraitis et al. (1997) when $\lambda^0 = 0$. So, asymptotically, there is no loss by using $\widehat{\lambda}^0$ instead of λ^0. However to achieve the latter, as in other nonparametric estimates, $\check{\alpha}\left(\check{\lambda}^0\right)$ will have a bias term of the same order of magnitude as the standard deviation.

4. FINITE SAMPLE BEHAVIOUR

In this section we study via Monte-Carlo analysis the finite sample performance of the estimators $\widehat{\lambda}^0$ and $\check{\alpha}\left(\widehat{\lambda}^0\right)$. The models employed throughout the simulations are

$$(1-L)^{\alpha/2}\, x_t \;=\; \varepsilon_t, \quad t = 0, \pm 1, ..., \tag{4.1}$$

$$\left(1 - 2\cos\left(\pi/2\right)L + L^2\right)^{\alpha/2} x_t \;=\; \varepsilon_t, \quad t = 0, \pm 1, ..., \tag{4.2}$$

where $\{\varepsilon_t\}$ is a zero mean sequence of iid Gaussian random variables. Model (4.1) generates a pole at $\lambda^0 = 0$ whereas model (4.2) does at $\lambda^0 = \pi/2$. We have chosen $\alpha = 0.2, 0.4, 0.6$ and 0.8. The autocorrelation functions of (4.1) and (4.2) are given by

$$\rho_j = \frac{j - 1 + \alpha/2}{j - \alpha/2}\rho_{j-1}, \quad j = 1, 2, ...$$

and

$$\rho_{2j} = \frac{1 - j - \alpha/2}{j - \alpha/2}\rho_{2(j-1)}, \text{ and } \rho_{2j-1} = 0, \quad (j = 1, 2, ...)$$

respectively, see for example Arteche and Robinson (2000). For each combination of α and λ^0, 2500 replications of series of lengths $n = 256$ and 1024 were generated by the method of Davies and Harte (1987).

Also, we have compared the performance of $\widehat{\lambda}^0$ and $\check{\alpha}\left(\widehat{\lambda}^0\right)$ with the corresponding estimators obtained using the popular, among practitioners, log-periodogram estimator of Robinson (1995a). That is, consider $\widetilde{\lambda}^0 = \lambda_{\widetilde{q}} = (2\pi\widetilde{q})/n$ where

$$\widetilde{q} = \arg \max_{q=0,...,[n/2]} \widehat{\alpha}_{LOG}\left(\lambda_q\right), \tag{4.3}$$

$$\widehat{\alpha}_{LOG}\left(\lambda_q\right) = -\left(2\sum_{j=1}^{k}\phi_j \log j\right)^{-1}\sum_{j=1}^{k}\phi_j\left(\log I_{j+q} + \log I_{q-j}\right), \tag{4.4}$$

with $\phi_j = \log j - k^{-1}\sum_{\ell=1}^{k}\log \ell$. Moreover, we have examined the behaviour of the estimator $\widehat{\alpha}_{LOG}\left(\widetilde{\lambda}^0\right)$ of α, where $k = m$ in (4.4). For the estimation of λ^0, the chosen bandwidth parameters were, for $n = 256$ and 1024, $k = 14$ and 24 respectively, and $k_1 = k^{0.6}\log\log 2k$, whereas for the *two-step* estimators $\check{\alpha}\left(\widehat{\lambda}^0\right)$ and $\widehat{\alpha}_{LOG}\left(\widetilde{\lambda}^0\right)$ of α, we have chosen $m = n/4$ and $m_1 = m^{0.6}\log\log 2m$. The weight functions used were $\psi(u) = -u^2 + 35u^{2.5}/6 - 29u^3/6 + 2u^3\log u$ and $w(u) = u^{1/3} - 9u^{1/2}/8$, respectively.

Table 4.1 illustrates the bias and standard deviation of the estimators $\widehat{\lambda}^0$, given in (2.4) and $\widetilde{\lambda}^0$ in (4.3). More specifically, since $\widehat{\lambda}^0 = (2\pi\widehat{q})/n$ and $\widetilde{\lambda}^0 = (2\pi\widetilde{q})/n$, we have reported the bias and standard deviation of $\widehat{q}$ and $\widetilde{q}$. Table 4.2 summarizes the bias, standard deviation and mean square error of $\check{\alpha}\left(\widehat{\lambda}^0\right)$ and $\check{\alpha}\left(\lambda^0\right)$. The motivation to include $\check{\alpha}\left(\lambda^0\right)$ is to investigate the relative loss we incur by lack of

knowledge of λ^0 in small samples. Recall that Theorem 3.4 indicates that asymptotically there is no loss. Moreover, Table 4.2 illustrates the finite sample performance of the corresponding estimators of α obtained using the log-periodogram estimator in (4.4), that is $\widehat{\alpha}_{LOG}\left(\lambda^0\right)$ and $\widehat{\alpha}_{LOG}\left(\widetilde{\lambda}^0\right)$.

$$\textit{TABLES 4.1 AND 4.2 ABOUT HERE}$$

Inspection of Table 4.1 indicates better performance of $\widehat{\lambda}^0$ than $\widetilde{\lambda}^0$ across different models and sample sizes, especially for $\alpha > 0.2$. For example when $\alpha = 0.8$, the finite sample performance of $\widehat{\lambda}^0$ is clearly superior to that of $\widetilde{\lambda}^0$, being this superiority greater with the sample size. With regard to the estimators of the memory parameter α, we observe that the proposed two-step estimator $\breve{\alpha}\left(\widehat{\lambda}^0\right)$ outperforms $\widehat{\alpha}_{LOG}\left(\widetilde{\lambda}^0\right)$ and has better finite sample properties for all α and λ^0. In some cases, the performance of $\widehat{\alpha}_{LOG}\left(\widetilde{\lambda}^0\right)$ is very poor compared to that of $\breve{\alpha}\left(\widehat{\lambda}^0\right)$, especially for large values of α. Finally, when comparing their performances with the estimators obtained when the location of the pole λ^0 is known, we observe that the relative loss of efficiency of $\breve{\alpha}\left(\widehat{\lambda}^0\right)$ is smaller than that of $\widehat{\alpha}_{LOG}\left(\widehat{\lambda}^0\right)$. Moreover, as Theorem 3.4 indicates, it appears that knowledge of λ^0 is not relevant to estimate α when $\breve{\alpha}\left(\widehat{\lambda}^0\right)$ is used, although it seems not to be the case when the log-periodogram is employed. Altogether, we can conclude that $\widehat{\lambda}^0$ and $\breve{\alpha}\left(\widehat{\lambda}^0\right)$ enjoy better finite sample properties than the corresponding ones based on $\widetilde{\lambda}^0$ and $\widehat{\alpha}_{LOG}\left(\widetilde{\lambda}^0\right)$.

5. AUXILIARY RESULTS AND PROOFS

We begin with the proof of Theorem 3.2.

5.1. **Proof of Theorem 3.2.** Let $t = -\left[\upsilon k^{1/2}\right]$. We examine the case $t > 0$, that for $t < 0$ is similarly handled. First, since C.5 and Lemma 6.10 imply that $\left|\overline{h}_{\psi} - h_{\psi}\right| = O\left(k^{-1}\right)$, we have by the definition of $\widehat{\alpha}\left(\lambda_q\right)$ that

$$\widehat{\xi}_n\left(t/k^{1/2}\right) = 2^{-1}h_{\psi}^{-1}\left(\sum_{i=1}^{6}\widehat{\xi}_n^{(i)}\left(t\right)\right)\left(1 + O_p\left(k^{-1}\right)\right)$$

after observing that $\widehat{\xi}_n\left(t/k^{1/2}\right) = \widehat{\xi}_n\left(\upsilon\right)$, and where

$$\widehat{\xi}_n^{(1)}\left(t\right) = -\alpha\sum_{p=1}^{k}\psi_p\log\left(\left|p-t\right|_+/p\right); \ \widehat{\xi}_n^{(2)}\left(t\right) = -\alpha\sum_{p=1}^{k}\psi_p\log\left(\left(p+t\right)/p\right)$$

$$\widehat{\xi}_n^{(3)}\left(t\right) = \sum_{p=1}^{k}\psi_p\log\left(\frac{\widetilde{f}_{p+s-t}\lambda_{|p-t|_+}^{\alpha}}{\widetilde{f}_{p+s}\lambda_p^{\alpha}}\right); \ \widehat{\xi}_n^{(4)}\left(t\right) = \sum_{p=1}^{k}\psi_p\log\left(\frac{\widetilde{f}_{s-p-t}\lambda_{p+t}^{\alpha}}{\widetilde{f}_{s-p}\lambda_p^{\alpha}}\right)$$

$$\widehat{\xi}_n^{(5)}\left(t\right) = \sum_{p=1}^{k}\psi_p\log\left(\frac{\widetilde{f}_{p+s-t}^{-1}\widehat{f}_{p+s-t}}{\widetilde{f}_{p+s}^{-1}\widehat{f}_{p+s}}\right); \ \widehat{\xi}_n^{(6)}\left(t\right) = \sum_{p=1}^{k}\psi_p\log\left(\frac{\widetilde{f}_{s-p-t}^{-1}\widehat{f}_{s-p-t}}{\widetilde{f}_{s-p}^{-1}\widehat{f}_{s-p}}\right),$$

where $\widetilde{f}_{\ell} = (2k_1 + 1)^{-1}\sum_{j=-k_1}^{k_1}f_{(j+\ell)\mathcal{I}(j+\ell\neq s)+(s+1)\mathcal{I}(j+\ell=s)}$ and $|q|_+ = \max\left(|q|, 1\right)$.

We examine the behaviour of $\widehat{\xi}_n^{(i)}\left(t\right)$, for $i = 1, ..., 6$, in four propositions. Specifically, Propositions 5.1 and 5.2 deal with the limiting bias of $\sum_{i=1}^{4}\widehat{\xi}_n^{(i)}\left(t\right)$, although for the proof of Proposition 5.1, we will allow $t < \rho k$ for $0 < \rho < 1$. Proposition

5.3 examines the finite dimensional limiting distribution of $\widehat{\xi}_n^{(5)}(t) + \widehat{\xi}_n^{(6)}(t)$ and Proposition 5.4 its tightness. Propositions 5.1 to 5.4 imply Theorem 3.2.

Proposition 5.1. $\widehat{\xi}_n^{(1)}(t) + \widehat{\xi}_n^{(2)}(t) = -\overline{\psi}''\alpha\frac{t^2}{k} + O\left(\frac{t}{k} + \frac{t^{5/2}}{k^{3/2}}\right).$

Proof. We only examine $\widehat{\xi}_n^{(1)}(t)$, being $\widehat{\xi}_n^{(2)}(t)$ identically handled. Assume $\rho < 1/2$ first, so that $0 < t < k/2$. Then

$$\widehat{\xi}_n^{(1)}(t) = -\alpha \sum_{p=1}^{t} \psi_p \log\left(|p-t|_+ /p\right) - \alpha \sum_{p=t+1}^{k} \psi_p \log\left((p-t)/p\right), \qquad (5.1)$$

whose first term on the right is $O\left(k^{-2}t^3\right)$ because $|\psi_p| \leq Dp^2/k^2$ by C.5 and the integrability of $|u^2 \log((1-u)/u)|$. Next, the second term on the right of (5.1) is

$$-\alpha \sum_{p=t+1}^{2t} \psi_p \log\left((p-t)/k\right) - \alpha \sum_{p=t+1}^{k-t} \left(\psi_{p+t} - \psi_p\right) \log(p/k) + \alpha \sum_{p=k-t+1}^{k} \psi_p \log(p/k).$$

Proceeding as with the first term on the right of (5.1), the first term of the last displayed expression is $O\left(k^{-2}t^3 \log(k/t)\right)$, whereas the last term is $O\left(k^{-2}t^3\right)$ by Taylor expansion of $\psi(u)$ around $u = 1$, noting that C.5 implies that $\psi(1) = 0$ and that $\sum_{p=k-t+1}^{k} |\log(p/k)| = O\left(t^2/k\right)$ by Taylor expansion of $\log(x)$ around $x = 1$. Finally, the second term of the last displayed expression is

$$-\alpha\frac{t}{k} \sum_{p=t+1}^{k-t} \psi_p' \log(p/k) - \frac{\alpha}{2}\frac{t^2}{k}\frac{1}{k} \sum_{p=t+1}^{k-t} \psi_p'' \log(p/k) + O\left(\frac{t^{5/2}}{k^{3/2}}\right) \qquad (5.2)$$

by integrability of $|\log u|$ and that $\psi''(u)$ is Lipchitz continuous of order $1/2$ by C.5. By Lemma 6.10 and C.5, the first term of (5.2) is

$$-\alpha t \int_0^1 \psi'(u)(\log u)\,du + O\left(\frac{t}{k}\right) + O\left(\frac{t}{k}\left(\left\{\sum_{p=1}^{t} + \sum_{p=k-t+1}^{k}\right\} |\psi_p' \log(p/k)|\right)\right)$$

$$= \alpha t \int_0^1 \psi(u) u^{-1}\,du + O\left(k^{-1}t\left(1 + t^2 k^{-1} \log(k/t)\right)\right)$$

noting that C.5 implies that $\psi(u)(\log u)|_0^1 = 0$ and $|\psi_p'| \leq Dp/k$ and then proceeding as above. On the other hand, the second term of (5.2) is

$$-\frac{\alpha}{2}\frac{t^2}{k} \int_0^1 \psi''(u)(\log u)\,du + O\left(\frac{t^3}{k^2} \log\left(\frac{k}{t}\right)\right) = -\alpha\frac{\overline{\psi}''t^2}{2k} + O\left(\frac{t^3}{k^2} \log\left(\frac{k}{t}\right)\right),$$

so that, because $\left|k^{-1/2}t^{1/2} \log(k/t)\right| \leq D$ for $0 < t \leq k$, we conclude that

$$\widehat{\xi}_n^{(1)}(t) = \alpha t \int_0^1 \psi(u) u^{-1}\,du - \alpha\frac{\overline{\psi}''t^2}{2k} + O\left(\frac{t}{k} + \frac{t^{5/2}}{k^{3/2}}\right).$$

Now, when $1/2 \leq \rho < 1$, so that $k/2 \leq t < k$, the proof is identical since in this case the left side of (5.1) is $-\alpha \sum_{p=1}^{t} \psi_p \log\left(|t-p|_+ /k\right) - \alpha \sum_{p=1}^{k-t} \left(\psi_{p+t} - \psi_p\right) \log(p/k) + \alpha \sum_{p=k-t+1}^{k} \psi_p \log(p/k)$. Then proceed as above. Proceeding similarly, $\widehat{\xi}_n^{(2)}(t) = -\alpha t \int_0^1 \psi(u) u^{-1}\,du - \alpha\frac{\overline{\psi}''t^2}{2k} + O\left(\frac{t}{k} + \frac{t^{5/2}}{k^{3/2}}\right)$. From here the conclusion is obvious.□

Proposition 5.2. $\widehat{\xi}_n^{(3)}(t) + \widehat{\xi}_n^{(4)}(t) = o(1).$

Proof. We only examine $\widehat{\xi}_n^{(3)}(t)$, $\widehat{\xi}_n^{(4)}(t)$ is similar. By definition, $\widehat{\xi}_n^{(3)}(t)$ is

$$-\sum_{p=1}^{2k_1}\psi_p a_p - \sum_{p=2k_1+1}^{k}\psi_p a_p - \sum_{p=1}^{k}\psi_p \widetilde{g}_p, \qquad (5.3)$$

where $a_p = \log\left(f^{-1}_{(p+s)\mathcal{I}(p\neq 0)+(s+1)\mathcal{I}(p=0)}\widetilde{f}_{p+s}\right) - \log\left(f^{-1}_{(p+s-t)\mathcal{I}(p\neq t)+(s+1)\mathcal{I}(p=t)}\widetilde{f}_{p+s-t}\right)$ and $\widetilde{g}_p = \log\left(\lambda_p^{\alpha} f_{(p+s)\mathcal{I}(p\neq 0)+(s+1)\mathcal{I}(p=0)}\right) - \log\left(\lambda_{|p-t|_+}^{\alpha} f_{(p+s-t)\mathcal{I}(p\neq t)+(s+1)\mathcal{I}(p=t)}\right)$.

Since by C.1 and $\left|\lambda^0 - \lambda_s\right| \leq \frac{\pi}{n}$, $D^{-1}(k_1/p)^{\alpha} < \lambda_{k_1}^{\alpha}f_{p+s} < D(k_1/p)^{\alpha}$ it implies that, for $p \leq 2k_1$, $|a_p| = O\left(\log(k_1/p)\right)$ by Lemma 6.1. Note that $a_t = O\left(\log k_1\right)$. Hence the absolute value of the first term of (5.3) is bounded by

$$D\left|\psi_t\right|\log(k_1) + D\sum_{p=1;p\neq t}^{2k_1}\left|\psi_p\right|\log\left(\frac{k_1}{p}\right) = O\left(\frac{k_1^3}{k^2}\right) = o(1)$$

by C.4 and because C.5 implies that $\left|\psi_p\right| \leq D(p/k)^2$. The absolute value of the second term of (5.3) is bounded by

$$D\sum_{p=2k_1+1}^{k_1\left[\log^{1/3}k_1\right]}\left|\psi_p a_p\right| + D\sum_{p=k_1\left[\log^{1/3}k_1\right]+1}^{k}\left|\psi_p a_p\right|$$

$$= O\left(\frac{k_1^3\log k_1}{k^2} + \frac{tk_1^2}{k^2}\left(1 + \log\left(\frac{k}{t}\right)\right)\right) = o(1),$$

where for the first term on the left we have used that by Lemma 6.1 part (a), $D^{-1} < \left|f^{-1}_{p+s}\widetilde{f}_{p+s}\right| < D$ and then C.5 and for the second term on the left that by Lemma 6.1 part (a), $\left|f^{-1}_{p+s}\widetilde{f}_{p+s} - 1\right| = O\left(p^{-2}k_1^2\right)$, which implies that $\left|\log\left(f^{-1}_{p+s}\widetilde{f}_{p+s}\right)\right| = O\left(p^{-2}k_1^2\right)$ by the mean value theorem, and then Lemma 6.4 with $\nu_p = k_1^{-2}\log\left(f^{-1}_{p+s}\widetilde{f}_{p+s}\right)$ there.

To complete the proof, it remains to show that the third term of (5.3) is $o(1)$. By C.1, $\sum_{p=1}^{k}\psi_p\widetilde{g}_p$ is

$$\sum_{p=1}^{k}\psi_p\left(\log\left(g\left(\lambda_{p+s}\right)\right) - \log\left(g\left(\lambda_{p+s-t}\right)\right)\right) \qquad (5.4)$$

$$-\alpha\sum_{p=1}^{k}\psi_p\left(\log\left(\left|\lambda_{p+s} - \lambda^0\right|\lambda_p^{-1}\right) - \log\left(\left|\lambda_{p-t+s} - \lambda^0\right|_+\lambda_{|p-t|_+}^{-1}\right)\right).$$

Denote the first and second derivatives of $\log\left(g\left(\lambda\right)\right)$ by $h\left(\lambda\right)$ and $h'\left(\lambda\right)$, respectively. The first term of (5.4) is

$$-\left(\frac{2\pi t}{n}\right)\sum_{p=1}^{k}\psi_p h\left(\lambda_{p+s}\right) - \left(\frac{2\pi t}{n}\right)^2\frac{1}{2}\sum_{p=1}^{k}\psi_p h'\left(\lambda_{p+s-\theta(p)t}\right)$$

$$= -\frac{4\pi^2 t}{n^2}h'\left(\lambda_s\right)\sum_{p=1}^{k}p\psi_p + o(1),$$

where $\theta(p) \in (0,1)$, by C.4 and because Lemma 6.10 and C.5 imply that $k^{-1}\sum_{p=1}^{k}\psi_p = O(k^{-1})$, so that

$$\sum_{p=1}^{k}\psi_p h(\lambda_{p+s}) = \sum_{p=1}^{k}\psi_p\left(h(\lambda_{p+s}) - h(\lambda_s)\right) + O(1)$$

$$= \frac{2\pi}{n}h'(\lambda_s)\sum_{p=1}^{k}p\psi_p + O\left(k^3 n^{-2} + 1\right).$$

Next, the second term of (5.4) is

$$-\alpha\left\{\sum_{p=1}^{2t} + \sum_{p=2t+1}^{k}\right\}\psi_p\left(\log\left((\lambda_{p+s} - \lambda^0)\lambda_p^{-1}\right) - \log\left(\left|\lambda_{p-t+s} - \lambda^0\right|_+ \lambda_{|p-t|_+}^{-1}\right)\right).$$

The contribution due to $\sum_{p=1}^{2t}$ is $o(1)$ by C.5 and then C.4, whereas the contribution due to $\sum_{p=2t+1}^{k}$, by Taylor expansion of $\log(x)$, is

$$-\alpha n\left(\frac{\lambda_s - \lambda^0}{2\pi}\right)\sum_{p=2t+1}^{k}\psi_p\left(\frac{1}{p} - \frac{1}{p-t}\right)$$

$$+\frac{\alpha}{2}n^2\left(\frac{\lambda_s - \lambda^0}{2\pi}\right)^2\sum_{p=2t+1}^{k}\psi_p\left\{\frac{1}{p^2(s)} - \frac{1}{p^2(s-t)}\right\}$$

$$= O\left(tk^{-1}\log k\right) + O\left(k^{-1} + tk^{-2}\log k\right) = o(1)$$

where $p(s-a)$ is an intermediate point between $p-a$ and $(p-a)+n(\lambda_s - \lambda^0)/(2\pi)$ and then because $n\left|\lambda^0 - \lambda_s\right| \leq \pi$, $\left|p^{-j}\psi_p\right| \leq Dk^{-j}$ for $j = 1, 2$, by C.5 and that $\left|p^2 p^{-2}(s)\right| + \left|(p-t)^2 p^{-2}(s-t)\right| \leq D$. So, we conclude that

$$\widehat{\xi}_n^{(3)}(t) = -\frac{4\pi^2 t}{n^2}h'(\lambda_s)\sum_{p=1}^{k}p\psi_p + o(1).$$

Similarly $\widehat{\xi}_n^{(4)}(t) = 4\pi^2 th'(\lambda_s)\sum_{p=1}^{k}p\psi_p/n^2 + o(1)$. From here the conclusion of the proposition is obvious. $\qquad\square$

Proposition 5.3. *The finite dimensional distributions of $\widehat{\xi}_n^{(5)}(t) + \widehat{\xi}_n^{(6)}(t)$ converge to those of a normal random variable.*

Proof. By Wold device, it suffices to show that, for any finite $l > 0$,

$$\sum_{i=1}^{l}\phi_i\left(\widehat{\xi}_n^{(5)}(t_i) + \widehat{\xi}_n^{(6)}(t_i)\right) \xrightarrow{d} N\left(0, \varsigma\sum_{i,j=1}^{l}\phi_i\phi_j v_i v_j\right),$$

where ϕ_i satisfies $\sum_{i=1}^{l}|\phi_i|^2 = 1$ and $v_i = \lim t_i/k^{1/2}$. Denoting $\widetilde{f}_\ell^{-1}\widehat{f}_\ell$ by $\widehat{g}_\ell$,

$$\widehat{\xi}_n^{(5)}(t) = -\sum_{p=1}^{2t}\psi_p\log\left(\widehat{g}_{p+s}/\widehat{g}_{p+s-t}\right) - \sum_{p=2t+1}^{k}\psi_p\log\left(\widehat{g}_{p+s}/\widehat{g}_{p+s-t}\right) \quad (5.5)$$

$$: \quad = b_{1t} + b_{2t}.$$

We begin by showing that $b_{1t} = o_p(1)$. In particular, we will show something stronger than needed. That is, that for each $\varepsilon > 0$ there exists n_0 such that

$$\Pr\left\{\sup_{t_1 \leq q \leq t_2}|b_{1q} - b_{1t_1}| > \varepsilon\right\} < DM^3 k^{-1/2}\varepsilon^{-1}\log^{-1}k_1, \quad (5.6)$$

for all $n \geq n_0$ and $0 \leq t_1 < t_2 \leq \left[k^{1/2}M\right]$. Because $t_2 \leq \left[k^{1/2}M\right]$,

$$
\begin{aligned}
\sup_{t_1 \leq q \leq t_2} |b_{1q} - b_{1t_1}| &\leq 2 \sup_{0 \leq |q| \leq 2\left[k^{1/2}M\right]} |\log \widehat{g}_{q+s}| \sum_{p=1}^{2\left[k^{1/2}M\right]} |\psi_p| \\
&\leq D \sup_{0 \leq |q| \leq 2\left[k^{1/2}M\right]} |\log \widehat{g}_{q+s}| \, M^3 k^{-1/2}
\end{aligned}
$$

since C.5 implies that $\sum_{p=1}^{a} |\psi_p| \leq Da^3/k^2$. Then, by Markov's inequality, the left side of (5.6) is bounded by

$$
DM^3 k^{-1/2} \varepsilon^{-1} E \sup_{0 \leq |q| \leq 2\left[k^{1/2}M\right]} |\log \widehat{g}_{q+s}| \leq DM^3 k^{-1/2} \varepsilon^{-1} \log^{-1} k_1,
$$

because Lemma 6.1 and the definition of $\widehat{f}_{q+s}$ imply that $\widehat{g}_{q+s}$ is bounded from below by $D^{-1} n^{-h}$, for some $h \geq 1$, and for $x > D^{-1} n^{-h}$

$$
|\log x - (x-1)| \leq D (x-1)^2 \log n, \tag{5.7}
$$

by C.4, $|\log n / \log k_1| \leq D$, and that by Lemma 6.3 and C.4, for some $\beta > 0$,

$$
E \left(\sup_{q:|q| \leq 2k_1} |\widehat{g}_{q+s} - 1|^\mu \right) \leq \frac{D}{\log^{1+\mu} k_1} \tag{5.8}
$$

$$
E \left(\sup_{q:2k_1 < |q|} |\widehat{g}_{q+s} - 1|^\mu \right) \leq \frac{D}{k_1^{\beta\mu}}. \tag{5.9}
$$

So, (5.6) holds true, which implies that $b_{1t} = o_p(1)$.

Next, the second term on the right of (5.5), i.e. b_{2t}. Denoting

$$
b_{3t} = \sum_{p=2t+1}^{k-t} (\psi_{p+t} - \psi_p) \log (\widehat{g}_{p+s}), \tag{5.10}
$$

we have that $b_{2t} - b_{3t}$ is

$$
\sum_{p=t+1}^{2t} \psi_{p+t} \log (\widehat{g}_{p+s}) - \sum_{p=k-t+1}^{k} \psi_p \log (\widehat{g}_{p+s}). \tag{5.11}
$$

Because by C.5, $\left\{ \sum_{p=t+1}^{2t} |\psi_{p+t}| + \sum_{p=q+1}^{2q} |\psi_{p+q}| \right\} \leq Dk^{-2}(t^3 + q^3) \leq DM^3 k^{-1/2}$, the first term of (5.11) satisfies that

$$
\Pr \left\{ \sup_{t_1 \leq q \leq t_2} \left| \sum_{p=t_1+1}^{2t_1} \psi_{p+t_1} \log (\widehat{g}_{p+s}) - \sum_{p=q+1}^{2q} \psi_{p+q} \log (\widehat{g}_{p+s}) \right| > \varepsilon \right\} \leq \frac{DM^3}{\varepsilon k^{1/2} \log k_1} \tag{5.12}
$$

by Markov's inequality and $(5.7) - (5.8)$, whereas the second term satisfies that

$$
\Pr \left\{ \sup_{t_1 \leq q \leq t_2} \left| \left(\sum_{p=k-t_1+1}^{k} - \sum_{p=k-q+1}^{k} \right) \psi_p \log (\widehat{g}_{p+s}) \right| > \varepsilon \right\} \leq \frac{DM}{\varepsilon \log k_1} \left(\frac{t_2 - t_1}{k^{1/2}} \right)
$$

by (5.7) and (5.9) and because $\sum_{p=k-q+1}^{k} - \sum_{p=k-t_1+1}^{k} = \sum_{p=k-q+1}^{k-t_1}$ and by C.5, $\sup_{t_1 \leq q \leq t_2} \sum_{p=k-q+1}^{k-t_1} |\psi_p| \leq \sum_{p=k-t_2}^{k-t_1} |\psi_p| \leq Dk^{-1}(t_2^2 - t_1^2) \leq DMk^{-1/2}(t_2 - t_1)$.

So, (5.6), (5.12) and the last inequality imply that for any $\varepsilon > 0$ there exists n_0 such that for all $n \geq n_0$,

$$
\Pr \left\{ \sup_{t_1 \leq q \leq t_2} \left| \left(\widehat{\xi}_n^{(5)}(q) - b_{3q} \right) - \left(\widehat{\xi}_n^{(5)}(t_1) - b_{3t_1} \right) \right| > \varepsilon \right\} \leq \frac{DM}{\varepsilon k^{1/2} \log k_1} \left(M^2 + (t_2 - t_1) \right). \tag{5.13}
$$

Clearly (5.13) implies that $\sup_{0 \leq t \leq 2[k^{1/2}M]} \left| \widehat{\xi}_n^{(5)}(t) - b_{3t} \right| = o_p(1)$. Proceeding similarly, we have that for any $\varepsilon > 0$ there exists n_0 such that for all $n \geq n_0$,

$$\Pr \left\{ \sup_{t_1 \leq q \leq t_2} \left| \left(\widehat{\xi}_n^{(6)}(q) - b_{4q} \right) - \left(\widehat{\xi}_n^{(6)}(t_1) - b_{4t_1} \right) \right| > \varepsilon \right\} \leq \frac{DM}{\varepsilon k^{1/2} \log k_1} \left(M^2 + (t_2 - t_1) \right),$$
(5.14)

where

$$b_{4t} = \sum_{p=3t+1}^{k} \left(\psi_{p-t} - \psi_p \right) \log \left(\widehat{g}_{s-p} \right). \tag{5.15}$$

Next, we examine b_{3t} and b_{4t}. Denoting $\vartheta_\ell = \widehat{g}_\ell - 1$ and writing

$$\widetilde{b}_t = \sum_{p=2t+1}^{k-t} \left(\psi_{p+t} - \psi_p \right) \vartheta_{p+s}, \quad \widetilde{\widetilde{b}}_t = \sum_{p=3t+1}^{k} \left(\psi_{p-t} - \psi_p \right) \vartheta_{s-p}, \tag{5.16}$$

Lemma 6.5 implies that $b_{3t} = \widetilde{b}_t + t/k^{1/2} o_p(1)$ and $b_{4t} = \widetilde{\widetilde{b}}_t + t/k^{1/2} o_p(1)$, where the $o_p(1)$ is uniformly in $t \leq \rho k$, for $\rho < 1/3$.

So, it remains to examine $\widetilde{b}_t$ and $\widetilde{\widetilde{b}}_t$. By Taylor expansion of ψ_p,

$$\widetilde{b}_t = \frac{t}{k} \sum_{p=2t+1}^{k-t} \psi_p' \vartheta_{p+s} + \frac{1}{2} \frac{t^2}{k^2} \sum_{p=2t+1}^{k-t} \psi'' \left(\frac{p}{k} + \delta \frac{t}{k} \right) \vartheta_{p+s}, \tag{5.17}$$

where $\delta = \delta(t) \in (0,1)$. The first term on the right of (5.17) is

$$\frac{t}{k} \sum_{p=2k_1+1}^{k-2k_1} \psi_p' \vartheta_{p+s} + \frac{t}{k} \left\{ \sum_{p=2t+1}^{2k_1} + \sum_{p=k-2k_1+1}^{k-t} \right\} \psi_p' \vartheta_{p+s}.$$

By Lemma 6.2, C.4 and C.5, the second term of the last displayed expression is clearly $tk_1^{1/2}/kO_p \left(1 + k^{-1} k_1^{\alpha+1/2} \mathcal{I}(\alpha \geq 1/2) \right) = t/k^{1/2} o_p(1)$ where the $o_p(1)$ does not depend on $t \leq \rho k$. On the other hand, writing $\eta_j = f_{j+s}^{-1} I_{j+s} - 1$, the first term is

$$\frac{t}{k} \sum_{p=2k_1+1}^{k-2k_1} \frac{\psi_p'}{2k_1+1} \sum_{j=-k_1}^{k_1} \eta_{p+j} + \frac{t}{k} \sum_{p=2k_1+1}^{k-2k_1} \frac{\psi_p'}{2k_1+1} \sum_{j=-k_1}^{k_1} \left(\frac{f_{p+j+s}}{\widetilde{f}_{p+s}} - 1 \right) \eta_{p+j}. \tag{5.18}$$

Since $\left| \widetilde{f}_{p+s}^{-1} f_{p+s} - 1 \right| = \widetilde{f}_{p+s}^{-1} f_{p+s} \left| 1 - f_{p+s}^{-1} \widetilde{f}_{p+s} \right| = O\left(k_1^2/p^2 \right)$ by Lemma 6.1, $\left| f_{p+s}^{-1} f_{p+j+s} - 1 \right| \leq D \frac{k_1}{p} \frac{|j|}{k_1}$ by C.3, and by an obvious extension of Robinson (1995b), $E \left| \sum_{j=-k_1}^{k_1} c_j \eta_{p+j} \right| = O\left(k_1^{1/2} \right)$ for any $|c_j| = c(j/k_1)| \leq D$, we obtain that the first absolute moment of the second term of (5.18) is by $C.5$ and then $C.4$, $tk^{-1/2} O\left(k_1^{1/2} k^{-1/2} \sum_{p=1}^{k} |\psi_p'|/p \right) = t/k^{1/2} o(1)$, where the $o(1)$ does not depend on $t \leq \rho k$.

On the other hand, after rearranging subindices, the first term of (5.18) is

$$\frac{t}{k} \sum_{p=2k_1+1}^{k-2k_1} \eta_p \left(\frac{1}{2k_1+1} \sum_{j=1}^{2k_1} \psi_{p+j}' \right)$$

$$+ \frac{t}{k} \sum_{p=k_1+1}^{2k_1} \eta_p \left(\frac{1}{2k_1+1} \sum_{j=1}^{p-k_1} \psi_{j+k_1}' \right) + \frac{t}{k} \sum_{p=k-2k_1+1}^{k-k_1} \eta_p \left(\frac{1}{2k_1+1} \sum_{j=p}^{k-k_1} \psi_{j-k_1}' \right).$$

After standard calculations and routine application of Robinson's (1995b) Theorem 2, the last two terms are $t/kO_p\left(k_1^{1/2}\right) = t/k^{1/2}o_p\left(1\right)$ by $C.4$ and $C.5$, whereas the first term of the last displayed expression is

$$\frac{t}{k}\sum_{p=2k_1+1}^{k-2k_1}\psi_p'\eta_p + \frac{t}{k}\sum_{p=2k_1+1}^{k-2k_1}\psi_p'\eta_p\left(\frac{1}{\psi_p'}\left[\frac{1}{2k_1+1}\sum_{j=1}^{2k_1}\psi_{p+j}'\right] - 1\right).$$

Because $\psi'\left(u\right)$ is continuous, $C.5$ implies that $\psi_{p+j}'/\psi_p' \to 1$ as $k_1/p \to 0$, and hence the expression inside the parenthesis converges to zero as $p \to \infty$. So, by Toeplitz's lemma, we conclude that the last displayed expression, and therefore also the first term of (5.17), is

$$\frac{t}{k^{1/2}}\left(\frac{1}{k^{1/2}}\sum_{p=2k_1+1}^{k-2k_1}\psi_p'\eta_p + o_p\left(1\right)\right).$$

Proceeding similarly as with the first term on the right of (5.17), the second term of (5.17) is $k^{-3/2}t^2\left(k^{-1/2}\sum_{p=2k_1+1}^{k-2k_1}\psi''\left(\frac{p}{k} + \delta\frac{t}{k}\right)\eta_p + o_p\left(1\right)\right)$, so that

$$\widetilde{b}_t = \frac{t}{k^{1/2}}\left(\frac{1}{k^{1/2}}\sum_{p=2k_1+1}^{k-2k_1}\psi_p'\eta_p + \frac{t}{k}O_p\left(1\right) + o_p\left(1\right)\right), \tag{5.19}$$

where the $o_p\left(1\right)$ and $O_p\left(1\right)$ are uniformly in $t \le \rho k$. Similarly, we obtain that

$$\widetilde{\widetilde{b}}_t = \frac{t}{k^{1/2}}\left(\frac{1}{k^{1/2}}\sum_{p=2k_1+1}^{k-2k_1}\psi_p'\eta_{-p} + \frac{t}{k}O_p\left(1\right) + o_p\left(1\right)\right), \tag{5.20}$$

where the $o_p\left(1\right)$ and $O_p\left(1\right)$ are uniformly in $t \le \rho k$.

Thus, (5.13), (5.14), Lemma 6.5, (5.19) and (5.20) imply that, for $t_i \le \left[Mk^{1/2}\right]$,

$$\sum_{i=1}^{l}\phi_i\left(\widehat{\xi}_n^{(5)}\left(t_i\right) + \widehat{\xi}_n^{(6)}\left(t_i\right)\right) = \sum_{i=1}^{l}\phi_i\left(\widetilde{b}_{t_i} + \widetilde{\widetilde{b}}_{t_i}\right) + \frac{t_i}{k^{1/2}}o_p\left(1\right) \overset{d}{\to} N\left(0, \varsigma\sum_{i,j=1}^{l}\phi_i\phi_j\upsilon_i\upsilon_j\right)$$

$$\tag{5.21}$$

by Robinson's (1995b) Theorem 2 and Toeplitz's Lemma since $\left|k^{-1}\sum_{\ell=t+1}^{k-t}\left(\psi_\ell'\right)^2 - \varsigma\right| = o\left(1\right)$ by Lemma 6.10. $\qquad\square$

Proposition 5.4. $\widehat{\xi}_n^{(5)}\left(t\right)$ *and* $\widehat{\xi}_n^{(6)}\left(t\right)$ *are tight.*

Proof. Write $c_t = \widehat{\xi}_n^{(5)}\left(t\right) - b_{3t}$, where b_{3t} is given in (5.10). To show that $\widehat{\xi}_n^{(5)}\left(t\right)$ is tight it suffices to show that c_t and b_{3t} are tight. Since the finite dimensional distributions of c_t converge to zero, cf. (5.13), Billingsley's (1968) Theorem 15.4 implies that c_t is tight if for each $\varepsilon > 0$ and $\nu > 0$ there exists a $\delta \in \left(0,1\right)$ such that

$$\Pr\left\{\vartheta''\left(c_t, \delta\right) \ge \varepsilon\right\} \le \nu \tag{5.22}$$

holds for all $n \ge n_0$, where

$$\vartheta''\left(c_t, \delta\right) = \sup\min\left\{\left|c_t - c_{t_1}\right|, \left|c_{t_2} - c_t\right|\right\},$$

and the supremum is over t_1, t, and t_2 satisfying $t_1 \le t \le t_2$ with $t_2 - t_1 \le \delta\left[k^{1/2}M\right]$ and $\delta \in \left(0,1\right)$. Observe that we can assume $k^{-1/2} \le \left[t_2/k^{1/2}M\right] - \left[t_1/k^{1/2}M\right]$. If $\left[t_2/k^{1/2}M\right] - \left[t_1/k^{1/2}M\right] < k^{-1/2}$, then either t_1 and t lie in the same subinterval $\left[\left(p-1\right)/M, p/M\right)$ or else t and t_2 do; in either of these cases the left side of (5.22) vanishes.

Inequalities (14.9) and (14.46) in Billingsley (1968) imply that (5.22) holds if

$$\Pr\left\{\vartheta\left(c_t,\delta\right)\geq\varepsilon\right\}\leq\nu,$$

for some $0<\delta\leq1$, where

$$\vartheta\left(c_t,\delta\right)=\sup_{\left|(t-v)/\left(k^{1/2}M\right)\right|<\delta}\left|c_t-c_v\right|.$$

(Observe that as c_t converges in probability to zero, which has continuous paths, the Skorohod's metric can be replaced by the uniform topology.) By the Corollary of Billingsley's (1968) Theorem 8.3, it suffices to show that

$$\sum_{i=1}^{r}\Pr\left\{\sup_{t_{i-1}\leq v\leq t_i}\left|c_v-c_{t_{i-1}}\right|\geq\varepsilon/3\right\}\leq\nu \qquad (5.23)$$

where $2^{-1}\delta<\left[k^{1/2}M\right]^{-1}\left(t_i-t_{i-1}\right)<\delta$ and $0=t_0<t_1<...<t_r=\left[k^{1/2}M\right]$. But this is the case since by (5.13),

$$\Pr\left\{\sup_{t_{i-1}\leq v\leq t_i}\left|c_v-c_{t_{i-1}}\right|\geq\varepsilon/3\right\}\leq\frac{DM^3\delta}{\varepsilon\log k_1}.$$

Now choose n_0 such that $DM^3\varepsilon^{-1}\log^{-1}k_1<\nu$ since $r\leq2\left[\delta^{-1}\right]$ to obtain (5.23). Proceeding similarly, but using (5.14) instead of (5.13), $\widehat{\xi}_n^{(6)}\left(t\right)-b_{4t}$ is also tight.

Next, we show the tightness condition for b_{3t}, the proof for b_{4t} is similar and it is omitted. Consider $t<q$. Then, $b_{3t}-b_{3q}$ is

$$\sum_{p=2t+1}^{k-t}\left(\psi_{p+t}-\psi_p\right)\log\widehat{g}_{p+s}-\sum_{p=2q+1}^{k-q}\left(\psi_{p+q}-\psi_p\right)\log\widehat{g}_{p+s} \qquad (5.24)$$

$$=\left\{\sum_{p=2t+1}^{2q}+\sum_{p=k-q+1}^{k-t}\right\}\left(\psi_{p+t}-\psi_p\right)\log\widehat{g}_{p+s}+\sum_{p=2q+1}^{k-q}\left(\psi_{p+t}-\psi_{p+q}\right)\log\widehat{g}_{p+s}.$$

The first term on the right of (5.24) is tight as we now show. Because by C.5,

$$\left|\psi_{p+t}-\psi_p\right|\leq Dk^{-1}t,\quad\left|\sum_{p=2t+1}^{2q}1+\sum_{p=k-q+1}^{k-t}1\right|\leq3\left(q-t\right),$$

abbreviating the first term on the right of (5.24) by $\zeta_{t,q}$, we obtain that

$$\sum_{i=1}^{r}\Pr\left\{\sup_{t_{i-1}\leq q\leq t_i}\left|\zeta_{t_{i-1},q}\right|\geq\varepsilon\right\}\leq\sum_{i=1}^{r}\Pr\left\{D\frac{t_i-t_{i-1}}{k^{1/2}}\sup_{p=2t+1,...,k}\left|\log\widehat{g}_{p+s}\right|\geq\varepsilon\right\}\leq\frac{DM}{\varepsilon\log k_1}$$

by Markov's inequality and $(5.7)-(5.8)$ with $\mu=1$ there. Then, choose n_0 such that $DM\varepsilon^{-1}\log^{-1}k_1<\nu$ to complete.

Next, Taylor expansion implies that the second term on the right of (5.24) is

$$\left(\frac{t-q}{k}\right)\sum_{p=2q+1}^{k-q}\psi'_{p+t}\log\widehat{g}_{p+s}+\frac{1}{2}\left(\frac{t-q}{k}\right)^2\sum_{p=2q+1}^{k-q}\psi''_{p+\ell}\log\widehat{g}_{p+s}, \qquad (5.25)$$

where ℓ is an intermediate point between t and q.

The second term of (5.25) is tight as we now show. Proceeding as with the proof of tightness of c_t, it suffices to show that for all ν and $\varepsilon>0$ there exists n_0 such that

$$\sum_{i=1}^{r}\Pr\left\{\sup_{t_{i-1}\leq q\leq t_i}\left|\left(\frac{q-t_{i-1}}{k}\right)^2\sum_{p=2q+1}^{k-q}\left|\psi''_{p+\ell}\log\widehat{g}_{p+s}\right|\right|\geq\varepsilon\right\}\leq\nu$$

for all $n \geq n_0$ and $0 \leq t_0 < ... < t_r \leq \left[k^{1/2} M \right]$. But by $(5.7) - (5.8)$ and that $\left| \psi''(u) \right| \leq D$ by $C.5$, the left side of the last displayed inequality is bounded by

$$D\varepsilon^{-1} k^{-1} \log^{-1} k_1 \sum_{i=1}^{r} (t_i - t_{i-1})^2 \leq DM^2 \varepsilon^{-1} \log^{-1} k_1 \sum_{i=1}^{r} \delta^2 < \nu$$

since $r \leq 2 \left[\delta^{-1} \right]$ and $2^{-1} \delta < \left[k^{1/2} M \right]^{-1} (t_i - t_{i-1}) < \delta$.

To finish the proof it remains to examine the first term of (5.25), denoted by $d_{t,q}$. Since from the proof of Proposition 5.3, the finite-dimensional distributions of $d_{t,q}$ converge to those of the limiting Gaussian process which has continuous paths, by Billingsley's (1968) Theorem 15.4, it implies that it suffices to check that

$$\Pr \left\{ \vartheta''(d_{t,q}, \delta) \geq \varepsilon \right\} \leq \nu$$

for some $n \geq n_0$. Now by Billingsley's (1968) Theorem 15.6, it suffices to check the moment condition

$$E \left| d_{t,q} d_{q,v} \right|^{\beta_1} \leq D \left| \frac{t - v}{k^{1/2}} \right|^{\beta_2} \tag{5.26}$$

for $t \leq q \leq v$ and some $\beta_1 > 0$ and $\beta_2 > 1$. Write

$$d_{t,q} = \left(\frac{q - t}{k} \right) \sum_{p=2q+1}^{k-q} \psi'_{p+t} \left(\widehat{g}_{p+s} - 1 \right) \tag{5.27}$$

$$+ \left(\frac{q - t}{k} \right) \left\{ \sum_{p=2q+1}^{3k_1} + \sum_{p=3k_1+1}^{k-q} \right\} \psi'_{p+t} \left(1 - \widehat{g}_{p+s} + \log \widehat{g}_{p+s} \right).$$

Using (5.7), the first absolute moment of the second term on the right of (5.27) is bounded by

$$D \left| \frac{q - t}{k} \right| \log n \left\{ \sum_{p=2q+1}^{3k_1} + \sum_{p=3k_1+1}^{k-q} \right\} \left| \psi'_{p+t} \right| E \left(\widehat{g}_{p+s} - 1 \right)^2$$

$$\leq D |q - t| \log n \left(\frac{k_1^{2\alpha}}{k^2} \mathcal{I}(\alpha \geq 1/2) + \frac{1}{k_1} \right)$$

$$\leq D \left| \frac{t - q}{k^{1/2}} \right|^{1+\xi} \log n \left(\frac{k^{(\xi+1)/2}}{k_1} + \frac{k_1^{2\alpha} k^{(1+\xi)/2}}{k^2} \mathcal{I}(\alpha \geq 1/2) \right) \leq D \left| \frac{t - q}{k^{1/2}} \right|^{1+\xi}$$

for some $0 < \xi < 3(1 - \alpha)$ by Lemma 6.2 and observing that $\sum_{p=2q+1}^{3k_1} \left| \psi'_{p+t} \right| \leq Dk_1^2 k^{-1} = o\left(k^{1/2} \right)$ by $C.4$ and $C.5$. So, the second term on the right of (5.27) satisfies (5.26) which follows by the Cauchy-Schwarz inequality choosing $\beta_1 = 1/2$ and $\beta_2 = 1 + \xi$, and that $(q - t)(v - q) \leq (v - t)^2$. Finally, proceeding as with the proof of $\widetilde{b}_t$ given in (5.19),

$$E \left| \left(\frac{q - t}{k} \right) \sum_{p=2q+1}^{k-q} \psi'_{p+t} \left(\widehat{g}_{p+s} - 1 \right) \right|^2 \leq D \left(\frac{q - t}{k^{1/2}} \right)^2$$

which implies that the first term on the right of (5.27) satisfies (5.26) choosing $\beta_1 = 1$ and $\beta_2 = 2$, and noting that $(q - t)(v - q) \leq (v - t)^2$. So, we conclude that b_{3t} is tight. Proceeding similarly b_{4t} is also tight, which completes the proof. $\square$

5.2. **Proof of Corollary 3.3.** (a) By Theorem 3.2, $\widehat{\xi}_n(\upsilon) \overset{weakly}{\Longrightarrow} \xi(\upsilon)$ in $\mathbb{D}[-M, M]$ for any arbitrary $M > 0$. Next, since the limiting Gaussian process $\xi(\upsilon)$ has continuous paths, i.e. it belongs to $\mathbb{C}[-M, M]$, the Skorohod's metric can be replaced by the uniform topology. On the other hand, by Eddy (1980), the $argmaximum$ is a continuous functional in the set of parabolas with fixed second derivatives in $\mathbb{C}[-M, M]$. So, by Van der Vaart and Wellner's (1996) Theorem 3.2.2., we obtain that

$$\widehat{\upsilon}_n = \arg\max_{\upsilon}\widehat{\xi}_n(\upsilon) \overset{d}{\to} \arg\max_{\upsilon}\xi(\upsilon) = \upsilon^*,$$

where $\upsilon^* = \Psi^{1/2}X$ and $X = N(0, 1)$. Observe that Theorem 3.1 shows that $\Pr\{|\widehat{\upsilon}_n| < L\} > 1 - \delta$ for n sufficiently large. This together with problem 1.3.9 in Van der Vaart and Wellner (1996, $p.27$) implies that $\widehat{\upsilon}_n$ is uniformly tight.

But by construction, $\widehat{\lambda}^0 = \lambda^0 + n^{-1}\left(2\pi k^{1/2}\right)\widehat{\upsilon}_n + O\left(n^{-1}\right)$, i.e. (3.2), and hence

$$\left(2\pi k^{1/2}\right)^{-1} n\left(\widehat{\lambda}^0 - \lambda^0\right) = \widehat{\upsilon}_n + O\left(k^{-1/2}\right) \overset{d}{\to} \upsilon^* = \Psi^{1/2}X.$$

(b) As in (a), the limit process is $\xi(\upsilon)$, where from the definition of $\widehat{\lambda}^0$, $\upsilon \geq 0$. Thus, if X takes a positive value, the restriction $\upsilon \geq 0$ is not binding and the maximum of $\xi(\upsilon)$ is achieved at υ^*. However, when X takes a negative value, the restriction is binding and thus the maximum is at $\upsilon = 0$ due to the parabolic structure of $\xi(\upsilon)$.

(c) The proof is identical to part (b) once the wording positive (negative) is replaced by negative (positive). $\qquad\square$

5.3. **Proof of Theorem 3.1.** Similarly to the proof of Theorem 3.2, it suffices to show the theorem with $\overline{h}_\psi$ replaced by h_ψ. With that replacement and recalling that $\widehat{g}_p = \widetilde{f}_p^{-1}\widehat{f}_p$, $\widehat{\alpha}(\lambda_q)$ becomes

$$\frac{h_\psi^{-1}}{2k}\sum_{p=1}^{k}\psi_p\left(\log\widehat{g}_{p+q} + \log\widehat{g}_{q-p}\right) \tag{5.28}$$

$$+\frac{h_\psi^{-1}}{2k}\sum_{p=1}^{k}\psi_p\log\left(\lambda_{|p+q-s|_+}^{\alpha}\widetilde{f}_{p+q}\right) + \frac{h_\psi^{-1}}{2k}\sum_{p=1}^{k}\psi_p\log\left(\lambda_{|q-p-s|_+}^{\alpha}\widetilde{f}_{q-p}\right)$$

$$-\alpha\frac{h_\psi^{-1}}{2k}\sum_{p=1}^{k}\psi_p\left(\log\left(2\pi|p+q-s|_+/n\right) + \log\left(2\pi|q-p-s|_+/n\right)\right).$$

Because $(5.7)-(5.9)$ imply that $\sup_{\ell=0,\ldots,[n/2]}|\log\widehat{g}_\ell| = o_p(1)$, we then have that, uniformly in q, the first term of (5.28) converges to zero in probability. Next, the second term of (5.28). (Recall that $0 \leq q \leq [n/2]$.) If $q < \max\{0, s - [k_1\log k_1] - k\}$ or $s + [k_1\log k_1] < q$, this is

$$\frac{h_\psi^{-1}}{2k}\sum_{p=1}^{k}\psi_p\left\{\log\left(\frac{\widetilde{f}_{p+q}}{f_{p+q}}\right) - \log\left(\frac{g(\lambda_q)}{g(\lambda_{p+q})}\right) + \log\left(\frac{\lambda_{p+q-s}^{\alpha}}{|\lambda_{p+q} - \lambda^0|^{\alpha}}\right)\right\} + o(1) \tag{5.29}$$

because Lemma 6.10 and C.5 imply that $k^{-1}\sum_{p=1}^{k}\psi_p = O\left(k^{-1}\right)$ and by C.1, $f_p = |\lambda_p - \lambda^0|^{-\alpha}g(\lambda_p)$. But by Lemma 6.1 part (a) and Taylor expansion of $\log(z)$ around $z = 1$, (5.29) is bounded in absolute value by

$$\frac{D}{k}\sum_{p=1}^{k}|\psi_p|\left(\left|\log\left(\frac{g(\lambda_q)}{g(\lambda_{p+q})}\right)\right| + \frac{k_1^2}{|q+p-s|^2} + \alpha\log\left(\frac{\lambda_{|p+q-s|}}{|\lambda_{p+q} - \lambda^0|}\right)\right) + o(1).$$

The contribution due to the second term inside the brackets is easily shown to be $o(1)$, as is the contribution due to the third term by Taylor expansion and that $n\left|\lambda_s - \lambda^0\right| < \pi$. Finally, by the mean value theorem, the first term inside the brackets is bounded by

$$\frac{D}{k}\sum_{p=1}^{k}\frac{p}{n}\left|\psi_p\right| = O\left(\frac{k}{n}\right),$$

because by C.1, $g(\lambda)$ is continuously differentiable. Next, when $s - [k_1 \log k_1] < q \le s + [k_1 \log k_1]$, the second term of (5.28) is also $o(1)$, since there are at most $O([k_1 \log k_1])$ terms such that $|p + q - s| < [k_1 \log k_1]$ and hence by Lemma 6.1, $D^{-1}\log k_1 \le \log\left(\widetilde{f}_{p+q}\left\{\lambda_{|p+q-s|_+}^{\alpha} + f_{p+q}^{-1}\right\}\right) \le D$ whereas for the remaining ones $|p + q - s| > [k_1 \log k_1]$, so that proceeding as before, it will be $o(1)$ by C.4. Thus, we conclude that the second term of (5.28) in this region is $O\left(k^{-1}k_1 \log k_1\right) + o(1) = o(1)$ by C.4. Similarly when $\max\{0, s - [k_1 \log k_1] - k\} \le q \le s - [k_1 \log k_1]$ we obtain that the second term of (5.28) is also $o(1)$. Proceeding as with the proof of the second term of (5.28), it follows that the third term of (5.28) is $o(1)$ uniformly in q.

Using $\int_0^1 \psi(x)\,dx = 0$ and Lemma 6.10, we conclude that

$$\sup_{0 \le q \le [n/2]}\left|\widehat{\alpha}(\lambda_q) - \frac{1}{2}\delta_{+,n}(\lambda_q) - \frac{1}{2}\delta_{-,n}(\lambda_q)\right| = o_p(1)$$

where $\delta_{+,n}(\lambda_q) = -\alpha h_\psi^{-1}k^{-1}\sum_{p=1}^{k}\psi_p \log\left(\left|p+q-s\right|_+ / k\right)$ and $\delta_{-,n}(\lambda_q) = -\alpha h_\psi^{-1}k^{-1}\sum_{p=1}^{k}\psi_p \log\left(\left|q-p-s\right|_+ / k\right)$.

We now examine the properties of $\delta_{+,n}(\lambda_q)$, those of $\delta_{-,n}(\lambda_q)$ are handled similarly. First by Lemma 6.10,

$$\delta_{+,n}(\lambda_s) - \alpha = -\alpha\frac{h_\psi^{-1}}{k}\sum_{p=1}^{k}\left(\psi_p \log\left(\frac{p}{k}\right) + h_\psi\right) = O\left(k^{-1}\right).$$

Now, for arbitrarily small $\rho > 0$, $\sup_{\rho^{-1}k \le |q-s|}\delta_{+,n}(\lambda_q) < D\rho$ since by Taylor expansion,

$$\sup_{\rho^{-1}k \le |q-s|}\left|\log\left(\left|q-s\right|_+\right) - \log\left(\left|\pm p + q - s\right|_+\right)\right| \le D\rho.$$

Next, by Proposition 5.1, $\sup_{|q-s| \le \rho k}\delta_{+,n}(\lambda_q) - \alpha < -D\rho^2$, whereas since $\delta_{+,n}(\lambda_q)$ is a nonincreasing function in $|q - s|$, $\sup_{\rho k \le |q-s| \le \rho^{-1}k}\delta_{+,n}(\lambda_q) - \alpha < -D\rho^2$.

Therefore, writing $\widehat{\Lambda}_n(t) = \widehat{\alpha}(\lambda_s + (2\pi t)/n)$, we conclude that

$$\Pr\left(\sup_{|t| \ge \rho k}\left(\widehat{\Lambda}_n(t) - \widehat{\Lambda}_n(0)\right) > 0\right) \to 0,$$

i.e. $\widehat{\lambda}^0$ is a consistent estimator of λ^0. Thus, to complete the proof of the theorem, it suffices to show that for any $\varepsilon > 0$, there exists $\overline{L} > 0$ such that

$$\Pr\left(\sup_{\rho k > |t| > k^{1/2}\overline{L}}\left(\widehat{\Lambda}_n(t) - \widehat{\Lambda}_n(0)\right) > 0\right) < \varepsilon. \qquad (5.30)$$

By Theorem 3.2, cf. Propositions 5.1 to 5.3,

$$
\begin{aligned}
\frac{k}{t}\left(\widehat{\Lambda}_n\left(t\right)-\widehat{\Lambda}_n\left(0\right)\right) &= -\overline{\psi}''\frac{t\alpha h_\psi^{-1}}{k}\left(1+O\left(\left(\frac{t}{k}\right)^{1/2}\right)\right)+\frac{1}{t}\left(\widetilde{b}_t+\widetilde{\widetilde{b}}_t\right) \\
&+\frac{1}{t}\left(b_{3t}-\widetilde{b}_t\right)+\frac{1}{t}\left(\widehat{\xi}_n^{(5)}\left(t\right)-b_{3t}\right) \\
&+\frac{1}{t}\left(b_{4t}-\widetilde{\widetilde{b}}_t\right)+\frac{1}{t}\left(\widehat{\xi}_n^{(6)}\left(t\right)-b_{4t}\right).
\end{aligned}
$$

By Proposition 5.3, cf. (5.19) and (5.20), and Lemma 6.9,

$$
\sup_{|t|<\rho k}\left|\frac{1}{t}\left(\widetilde{b}_t+\widetilde{\widetilde{b}}_t\right)\right|=\frac{1}{k^{1/2}}\sup_{|t|<\rho k}\left|\frac{1}{k^{1/2}}\sum_{p=|t|}^{k-|t|}\psi_p'\left(2\pi I_{\varepsilon,p+s}-1\right)+o_p\left(1\right)\right|=O_p\left(\frac{1}{k^{1/2}}\right)
$$

since by C.5, $\psi'\left(u\right)$ is continuous, so that $\sup_{|t|<\rho k}\left|k^{-1/2}\sum_{p=|t|}^{k-|t|}\psi_p'\left(2\pi I_{\varepsilon,p+s}-1\right)\right|=O_p\left(1\right)$ by Lemma 6.7, and thence by Lemma 6.5,

$$
\sup_{|t|<\rho k}\left|\frac{1}{t}\left(\widetilde{b}_t+\widetilde{\widetilde{b}}_t\right)+\frac{1}{t}\left(b_{3t}-\widetilde{b}_t\right)+\frac{1}{t}\left(b_{4t}-\widetilde{\widetilde{b}}_t\right)\right|=O_p\left(k^{-1/2}\right). \tag{5.31}
$$

By C.4, there exists a finite positive integer r such that $k_1^{(r-1)\beta}<k^{1/2}<k_1^{(r+1)\beta}$. Consider first the case $2k_1^{1+r\beta}<k$. Then, the left side of (5.30) is bounded by

$$
\sum_{\ell=1}^{r}\Pr\left(\sup_{\frac{L_{\ell-1}k}{k_1^{(\ell-1)\beta}}\geq|t|>\frac{L_\ell k}{k_1^{\ell\beta}}}\frac{k}{t}\left(\widehat{\Lambda}_n\left(t\right)-\widehat{\Lambda}_n\left(0\right)\right)>0\right) \tag{5.32}
$$

$$
+\Pr\left(\sup_{\frac{L_r k}{k_1^{r\beta}}\geq|t|>2k_1\widetilde{L}}\frac{k}{t}\left(\widehat{\Lambda}_n\left(t\right)-\widehat{\Lambda}_n\left(0\right)\right)>0\right)+\Pr\left(\sup_{2k_1\widetilde{L}\geq|t|>k^{1/2}\overline{L}}\frac{k}{t}\left(\widehat{\Lambda}_n\left(t\right)-\widehat{\Lambda}_n\left(0\right)\right)>0\right),
$$

where $L_0=\rho$, $L_\ell>0$ for $\ell>1$ and $\widetilde{L}>0$. Since $h_\psi^{-1}\alpha\overline{\psi}''>0$ and $\widehat{\Lambda}_n\left(t\right)-\widehat{\Lambda}_n\left(0\right)>0$, the third term of (5.32) is bounded by

$$
\Pr\left\{\left|O_p\left(\frac{k_1^2}{k^{3/2}\log k_1}+\frac{k_1}{k_1^\beta k^{1/2}}+1\right)\right|>\alpha\left|\overline{\psi}''\right|\inf_{2k_1\widetilde{L}\geq|t|>k^{1/2}\overline{L}}\left|\frac{t}{k^{1/2}}\right|\right\}<\varepsilon
$$

for $\overline{L}$ large enough, by Lemma 6.8 part (b) and (5.31), and because C.5 implies that $k_1^2k^{-3/2}=o\left(1\right)$, and $k^{-1/2}k_1^{1-\beta}=k^{1/2}k_1^{1+r\beta}/\left(kk_1^{(1+r)\beta}\right)=O\left(1\right)$. Next, the second term of (5.32) is bounded by

$$
\Pr\left\{\left|O_p\left(\frac{k}{k_1^{(r+1)\beta+1}}+\frac{1}{k_1^{r\beta}\log k_1}+\frac{k^{1/2}}{k_1}\right)\right|=o_p\left(1\right)>\alpha\left|\overline{\psi}''\right|\inf_{\frac{L_r k}{k_1^{r\beta}}>|t|>2\widetilde{L}k_1}\left|\frac{t}{k_1}\right|\right\}<\varepsilon
$$

for $\widetilde{L}$ large enough, by Lemma 6.8 part (a), (5.31) and that $k^{1/2}=o\left(k_1\right)$, $k^{1/2}<k_1^{(r+1)\beta}$ and C.4.

Finally, the first term of (5.32), whose typical element is, proceeding as before, bounded by

$$
\Pr\left\{\left|O_p\left(\frac{k_1^{\ell\beta}}{k_1^{\ell\beta}}+\frac{k_1^{1+\beta}}{k\log k_1}+\frac{k_1^{\ell\beta}}{k^{1/2}}\right)\right|=O_p\left(1\right)>\alpha\left|\overline{\psi}''\right|\inf_{\frac{L_{\ell-1}k}{k_1^{(\ell-1)\beta}}>|t|>\frac{L_\ell k}{k_1^{\ell\beta}}}\left|\frac{tk_1^{\ell\beta}}{k}\right|\right\}<\varepsilon
$$

because $\frac{k_1^{\ell\beta}}{k^{1/2}} \leq \frac{k_1^{\ell\beta+1}}{k^{1/2}k_1} = o\left(k^{-1}k_1^{\ell\beta+1}\right) = o\left(1\right)$ by C.4 and that $\ell \leq r$.

Next, consider the case $k < 2k_1^{1+r\beta}$. In this case, let $\widetilde{r}$ be the biggest integer such that $2k_1^{1+\widetilde{r}\beta} < k < k_1^{1+r\beta}$ but $k < k_1^{1+(\widetilde{r}+1)\beta}$. The proof now proceed identically as in the previous case, but now the sum of the first term of (5.32) runs from $\ell = 1$ to $\widetilde{r}$. Note that if $k < 2k_1^{1+\beta}$, so that $r = 1$, then the left side of (5.30) is bounded by

$$\Pr\left(\sup_{\rho k > |t| > 2\widetilde{L}k_1} \frac{k}{t}\left(\widehat{\Lambda}_n\left(t\right) - \widehat{\Lambda}_n\left(0\right)\right) > 0\right) + \Pr\left(\sup_{2\widetilde{L}k_1 \geq |t| > k^{1/2}\overline{L}} \frac{k}{t}\left(\widehat{\Lambda}_n\left(t\right) - \widehat{\Lambda}_n\left(0\right)\right) > 0\right),$$

and then proceed as in the proof of the second and third terms of (5.32), since C.4 implies that $k_1^3 = o\left(k^2\right)$, and recalling that now $k < 2k_1^{1+\beta}$, we have that $k_1^{1-\beta}k^{-1/2} = o\left(1\right)$ by C.4. $\qquad\square$

5.4. **Proof of Theorem 3.4.** We begin with part (**a**). Observing that C.7 implies that $\left|\overline{h}_w - h_w\right| = O\left(m^{-1}\right)$ and $m^{-1}\sum_{p=1}^m w_p = O\left(m^{-1}\right)$ by Lemma 6.10, the behaviour of $(2m)^{1/2}\left(\breve{\alpha}\left(\lambda_s\right) - \alpha\right)$ is governed by that of

$$\frac{h_w^{-1}}{(2m)^{1/2}}\sum_{p=1}^m w_p \log\left(\widehat{g}_{p+s}\widehat{g}_{-p}\right) + \frac{h_w^{-1}}{(2m)^{1/2}}\sum_{p=1}^m w_p \log\left(\frac{\widetilde{f}_{p+s}}{\lambda_p^{-\alpha}}\frac{\widetilde{f}_{s-p}}{\lambda_p^{-\alpha}}\right)$$

$$-\frac{(2m)^{1/2}\alpha}{h_w}\left(\frac{1}{m}\sum_{p=1}^m w_p \log\left(\frac{p}{m}\right) + h_w\right). \tag{5.33}$$

Recall that $\widetilde{f}_p^{-1}\widehat{f}_p = \widehat{g}_p$. By Lemma 6.10, the last term of (5.33) is $o\left(1\right)$. Denoting $m_1^* = [m_1 \log m_1]$, the second term of (5.33) is

$$\frac{h_w^{-1}}{(2m)^{1/2}}\left\{\sum_{p=1}^{m_1^*} w_p \log\left(\frac{\widetilde{f}_{p+s}}{\lambda_p^{-\alpha}}\frac{\widetilde{f}_{s-p}}{\lambda_p^{-\alpha}}\right) + \sum_{p=m_1^*+1}^m w_p \log\left(\frac{\widetilde{f}_{p+s}}{\lambda_p^{-\alpha}}\frac{\widetilde{f}_{s-p}}{\lambda_p^{-\alpha}}\right)\right\}. \tag{5.34}$$

Next, because by Lemma 6.1, $D^{-1} < \lambda_{m_1}^\alpha \widetilde{f}_{s\pm p} < D$ for $|p| \leq 2m_1^*$, we have that the first term of (5.34) is $O\left(m^{-1/2}\log\left(m_1\right)\sum_{p=1}^{m_1^*}|w_p|\right) = O\left(m_1^{1+\zeta}m^{-(2\zeta+1)/2}\log^{1+\zeta}\left(m_1\right)\right) = o\left(1\right)$ by C.6 and C.7 since $\zeta \geq 1/3$. Denoting $g\left(\lambda_p\right) = g_p$, the second term of (5.34) is

$$\frac{h_w^{-1}}{(2m)^{1/2}}\sum_{p=m_1^*+1}^m w_p \log\left(\frac{\widetilde{f}_{p+s}}{f_{p+s}}\frac{\widetilde{f}_{s-p}}{f_{s-p}}\right) + \frac{h_w^{-1}}{(2m)^{1/2}}\sum_{p=m_1^*+1}^m w_p \log\left(g_{p+s}g_{-p}\right)$$

$$-\frac{\alpha h_w^{-1}}{(2m)^{1/2}}\sum_{p=m_1^*+1}^m w_p\left\{\log\left(\lambda_p^{-1}\left|\lambda_p + \lambda_s - \lambda^0\right|\right) + \log\left(\lambda_p^{-1}\left|\lambda_s - \lambda^0 - \lambda_p\right|\right)\right\}.$$

Because, for $|p| > 2m_1^*$, Lemma 6.1 part (a) implies that $D^{-1} \leq m_1^{-2}p^2\left|f_{p+s}^{-1}\widetilde{f}_{p+s} - 1\right| \leq D$, we obtain by the mean value theorem that $\log f_{p+s}^{-1}\widetilde{f}_{p+s} = O\left(m_1^2 p^{-2}\right)$ and so the first term of the last displayed expression is bounded in absolute value by

$$\frac{Dm_1^2}{(2m)^{1/2}}\sum_{p=m_1^*+1}^m |w_p|p^{-2} = o\left(m_1^{1+\zeta}m^{-(2\zeta+1)/2}\right),$$

whereas the second term is $4\pi^2 Bc^{5/2}/\left(2^{1/2}h_w\right) + O\left(m^{-1/2}\right)$ because

$$\frac{1}{(2m)^{1/2}}\sum_{p=m_1^*+1}^m w_p \log\left(g_{p+s}g_{s-p}\right) = \frac{2}{(2m)^{1/2}}\sum_{p=1}^{m_1^*} w_p \log\left(g_s\right) + \frac{4\pi^2 Bc^{5/2}}{2^{1/2}h_w} + O\left(\frac{1}{m^{1/2}}\right)$$

by C.6 and Taylor expansion of $\log\left(g_{p+s}\right)$ and $\log\left(g_{s-p}\right)$ around $\log\left(g_s\right)$ and that by Lemma 6.10 and C.7, $\sum_{p=1}^{m} w_p = O\left(1\right)$. So, the second term of (5.33) is $4\pi^2 B c^{5/2}/\left(2^{1/2}h_w\right) + o\left(1\right)$. Finally, proceeding as with the second term of (5.4), the third term is easily shown to be bounded by $Dm^{-1/2}\sum_p |w_p|\, p^{-1} = o\left(1\right)$.

Denoting $\vartheta_b = \widehat{g}_b - 1$, and proceeding as with the proof of Lemma 6.5, to complete the proof of the theorem, it suffices to show that

$$\frac{h_w^{-1}}{(2m)^{1/2}}\sum_{p=1}^{m} w_p\left(\vartheta_{p+s} + \vartheta_{s-p}\right) \xrightarrow{d} \mathcal{N}\left(0, h_w^{-2}\Phi^2\right) \qquad (5.35)$$

$$\frac{1}{(2m)^{1/2}}\left\{\sum_{p=1}^{2m_1} + \sum_{p=2m_1+1}^{m}\right\} w_p\left(\vartheta_{p+s}^2 + \vartheta_{s-p}^2\right) \xrightarrow{P} 0. \qquad (5.36)$$

We begin with (5.36). By Lemma 6.2 part (b), the first moment of the first sum inside the braces on the left of (5.36) is $o\left(m^{-1/2}\sum_{p=1}^{2m_1}|w_p|\right) = o\left(m_1^{1+\zeta}m^{-(2\zeta+1)/2}\right) = o\left(1\right)$ by C.7 and C.6 since $\zeta \geq 1/3$, whereas the contribution due to the second sum inside the braces on the left of (5.36) is $O_p\left(m_1^{-1}m^{1/2}\right) = o_p\left(1\right)$ by Lemma 6.2 part (a), C.6 and Markov's inequality. So, it remains to show (5.35), whose left side is

$$\frac{h_w^{-1}}{(2m)^{1/2}}\left(\sum_{p=1}^{2m_1} w_p\left(\vartheta_{p+s} + \vartheta_{s-p}\right) + \sum_{p=2m_1+1}^{m} w_p\left(\vartheta_{p+s} + \vartheta_{s-p}\right)\right). \qquad (5.37)$$

Because Lemma 6.2 part (b) implies that $E\left|\vartheta_b\right| = o\left(1\right)$ for $|b - s| < 2m_1$, the first term of (5.37) is $o_p\left(m_1^{1+\zeta}m^{-(2\zeta+1)/2}\right) = o_p\left(1\right)$ by C.7, C.6 and Markov's inequality. But proceeding as in the proof of Proposition 5.3, cf. (5.21), the second term of (5.37) converges in distribution to $\mathcal{N}\left(0, h_w^{-2}\Phi^2\right)$, which completes the proof of part (a).

Part (b). Dropping the constant h_w and $(2m)^{-1/2}$, it suffices to show that

$$\sum_{p=1}^{m} w_p\left\{\log\left(\widehat{g}_{p+s-t}/\widehat{g}_{p+s}\right) + \log\left(\widehat{g}_{s-p-t}/\widehat{g}_{s-p}\right)\right\} = o_p\left(m^{1/2}\right) \qquad (5.38)$$

holds true uniformly in $|t| \leq \left[k^{1/2}M\right] = o\left(m^{1/2}\right)$. We only examine the contribution due to the first term on the left of (5.38), the contribution due to the second term follows by identical steps. The first term on the left of (5.38) is

$$\sum_{p=1}^{2t} w_p\log\left(\widehat{g}_{p+s-t}/\widehat{g}_{p+s}\right) + \sum_{p=2t+1}^{m} w_p\log\left(\widehat{g}_{p+s-t}/\widehat{g}_{p+s}\right). \qquad (5.39)$$

Using $(5.7)-(5.9)$ and Markov's inequality, the first term of (5.39) is, uniformly in t, $o_p\left(\sum_{p=1}^{\left[2k^{1/2}M\right]}|w_p|\right) = o_p\left(m^{1/2}\right)$ by C.6. Next, the second term of (5.39) is

$$\sum_{p=t+1}^{2t} w_{p+t}\log\widehat{g}_{p+s} - \sum_{p=m-t+1}^{m} w_p\log\widehat{g}_{p+s} + \sum_{p=2t+1}^{m-t}\left(w_{p+t} - w_p\right)\log\widehat{g}_{p+s}.$$

By $(5.7)-(5.9)$, the first two terms of the last displayed expression, uniformly in t, are $o_p\left(k^{1/2}\right) = o_p\left(m^{1/2}\right)$ by C.6. Finally, the third term on the last displayed expression.

Let $\vartheta_p = \widehat{g}_p - 1$. Since by Lemma 6.3 and Markov's inequality, $\sup_{p=1,\ldots,[n/2]}|\vartheta_p| = o_p\left(1\right)$, except in a set Ω_n such that $\lim_{n\to\infty}\Pr\left\{\Omega_n\right\} = 0$, it implies that $\log\widehat{g}_{p+s} =$

$\vartheta_{p+s} - 2^{-1}\vartheta_{p+s}^2 \left(1 + o_p\left(1\right)\right)$ by Taylor expansion. So, the third term of the last displayed expression is

$$\left\{\sum_{p=2m_1+1}^{m-t} + \sum_{p=2t+1}^{2m_1}\right\} \left(w_{p+t} - w_p\right)\vartheta_{p+s} + D\sum_{p=2t+1}^{m-t}|w_{p+t} - w_p|\vartheta_{p+s}^2 \left(1 + o_p\left(1\right)\right).$$
(5.40)

Since C.7 implies that $|w_{p+t} - w_p| \leq D\left(t/m\right)^\zeta$, we have that

$$\sup_{t\leq[k^{1/2}M]}\left|\sum_{p=2t+1}^{m-t}|w_{p+t} - w_p|\vartheta_{p+s}^2\right| \leq D \sup_{t\leq[k^{1/2}M]}\left(\frac{t}{m}\right)^\zeta \left(\sum_{p=2m_1+1}^{m}\vartheta_{p+s}^2\right)$$
$$+ \sup_{t\leq[k^{1/2}M]}\left|\sum_{p=2t+1}^{2m_1} w_p\vartheta_{p+s}^2\right|.$$

But by Lemma 6.2, $E\left|\vartheta_{p+s}\right|^2 = O\left(m_1^{-1}\right)$ if $|p| > 2m_1$, whereas Lemma 6.3 implies that $\sup_{p=1,\ldots,[n/2]} \vartheta_{p+s}^2 = o_p\left(1\right)$. Hence, by Markov's inequality, C.6 and C.7, the third term of (5.40) is $o_p\left(k^{\zeta/2}m^{1-\zeta/2}m_1^{-1} + m^{-\zeta}m_1^{1+\zeta}\right) = o_p\left(m^{1/2}\right)$. Proceeding similarly and in view of C.7, the second term of (5.40) is $o_p\left(m_1^{1/2}\left(k^{1/2}/m\right)^\zeta\right) = o_p\left(m^{1/2}\right)$. So, denoting $b_t = \sum_{p=2m_1+1}^{m-t}\left(w_{p+t} - w_p\right)\vartheta_{p+s}$, that is the first term of (5.40), to complete the proof we need to show that $\sup_{t\leq[k^{1/2}M]}|b_t| = \sup_{q=1,\ldots,[Mk^{1/4}]}\sup_{(q-1)k^{1/4}\leq t\leq qk^{1/4}}|b_t|$ is $o_p\left(m^{1/2}\right)$. Now, by the triangle inequality, $\sup_{t\leq[k^{1/2}M]}|b_t|$ is bounded by

$$\sup_{q=1,\ldots,[Mk^{1/4}]}\sup_{(q-1)k^{1/4}\leq t\leq qk^{1/4}}\left|\sum_{p=2m_1+1}^{m-t}\left(w_{p+t} - w_{p+qk^{1/4}}\right)\vartheta_{p+s}\right|$$
(5.41)
$$+ \sup_{q=1,\ldots,[Mk^{1/4}]}\sup_{(q-1)k^{1/4}\leq t\leq qk^{1/4}}\left|\left\{\sum_{p=2m_1+1}^{m} - \sum_{p=m-t+1}^{m}\right\}\left(w_{p+qk^{1/4}} - w_p\right)\vartheta_{p+s}\right|.$$

Because $\left(\sup_j |c_j|\right)^\mu = \sup_j |c_j|^\mu \leq \sum_j |c_j|^\mu$ for $\mu > 0$, the second moment of the second term of (5.41) is bounded by

$$\sum_{q=1}^{[Mk^{1/4}]}\left\{E\left|\sum_{p=2m_1+1}^{m}\left(w_{p+qk^{1/4}} - w_p\right)\vartheta_{p+s}\right|^2 + \sum_{\ell=1}^{[Mk^{1/2}]}E\left|\sum_{p=m-\ell+1}^{m}\left(w_{p+qk^{1/4}} - w_p\right)\vartheta_{p+s}\right|^2\right\}$$

$$\leq D\sum_{q=1}^{[Mk^{1/4}]}\left\{m + \sum_{\ell=1}^{[Mk^{1/2}]}\ell\right\}\left(\frac{qk^{1/4}}{m}\right)^{2\zeta} = D\frac{k^{\zeta+1/4}}{m^{2\zeta-1}} = o\left(m\right),$$

proceeding as with the proof of (5.16) and noting that C.7 implies that $\left|\left(m/p\right)^\zeta w_p\right| \leq D$, and $\left|\left(m/\left(qk^{1/4}\right)\right)^\zeta \left(w_{p+qk^{1/4}} - w_p\right)\right| \leq D$ with $\zeta \geq 1/3$. The second moment of the first term of (5.41) is bounded by

$$\sum_{q=1}^{[Mk^{1/4}]}\sum_{t=(q-1)k^{1/4}+1}^{qk^{1/4}}E\left|\sum_{p=2m_1+1}^{m-t}\left(w_{p+t} - w_{p+qk^{1/4}}\right)\vartheta_{p+s}\right|^2$$

$$= \frac{k^{\zeta/2}}{m^{2\zeta}}\sum_{q=1}^{[Mk^{1/4}]}\sum_{t=(q-1)k^{1/4}+1}^{qk^{1/4}}E\left|\sum_{p=2m_1+1}^{m-t}w_{p,t}^*\vartheta_{p+s}\right|^2$$

where $\left|w_{p,t}^{*}\right| = \left|\left(m/k^{1/4}\right)^{\zeta}\left(w_{p+t} - w_{p+qk^{1/4}}\right)\right| \leq D$ by C.7. Now proceeding as with the proof of (5.16), the right side of the last displayed equation is $O\left(m^{1-2\zeta}k^{(1+\zeta)/2}\right) = o\left(m\right)$ since $\zeta \geq 1/3$ and $k = o\left(m\right)$ by C.6. Using Markov's inequality we conclude that (5.41) is $o_{p}\left(m^{1/2}\right)$ and the proof. $\qquad\square$

6. TECHNICAL LEMMAS

From now on $\sum_j$ denotes $\sum_{j=-k_1}^{k_1}$ and $k_1 n^{-1} \to 0$.

Lemma 6.1. *Let $\widetilde{f}_p$ be as defined in the proof of Theorem 3.2. Then,*

(a) $\begin{cases} D^{-1} < f_p^{-1} \widetilde{f}_p < D \\ \left(f_p^{-1} \widetilde{f}_p - 1\right) = O\left(k_1^2 / |p-s|^2\right) \end{cases}$ $|p-s| \geq 2k_1.$

(b) $D^{-1} \leq \lambda_{k_1}^\alpha \widetilde{f}_p \leq D$ *if $|p-s| < 2k_1$.*

Proof. First observe that by definition of $\widetilde{f}_p$, $f_p^{-1}\widetilde{f}_p = (2k_1+1)^{-1}\sum_j f_p^{-1} f_{j+p}$. We begin with (a). We first show that $D^{-1} < (2k_1+1)^{-1}\sum_j f_p^{-1} f_{j+p} < D$. Because $\left|\lambda^0 - \lambda_s\right| \leq \frac{\pi}{n}$, $D^{-1} < \left|1 + \frac{n|\lambda^0 - \lambda_s|}{2\pi(j+p-s)}\right| < D$, for $|j| \leq k_1$, so that C.1 implies that

$$\frac{D^{-1}}{k_1} \sum_{j=[k_1/4]}^{[k_1/2]} \left|\frac{p-s}{j+p-s}\right|^\alpha \leq \frac{1}{2k_1+1}\sum_j \frac{f_{j+p}}{f_p} \leq \frac{D}{k_1}\sum_j \left|\frac{p-s}{j+p-s}\right|^\alpha.$$

But $|p-s| \geq 2k_1$ and $|j| \leq k_1$ imply that $2/3 < |(p-s)/(j+p-s)| < 2$. From here the conclusion is standard since $\alpha > 0$, we conclude the first part of (a). Next, we show the second part of (a). By Taylor expansion of f_{j+p}, the left side is

$$\frac{1}{2k_1+1}\sum_j \left\{\frac{(2\pi)j}{n}\frac{f_p'}{f_p} + \frac{(2\pi)^2 j^2}{2n^2}\frac{f''\left(\overline{\lambda}\right)}{f_p}\right\} \leq \frac{D}{k_1}\sum_j \left(\frac{j}{p-s+\delta j}\right)^2 (1+o(1))$$

where $\overline{\lambda} = \overline{\lambda}(p+\delta j)$ is an intermediate point between λ_p and λ_{p+j} and $\delta = \delta(j) \in (0,1)$, by C.1 and that $f_p^{-1} f\left(\overline{\lambda}\right)$ is bounded. The conclusion follows since $|p-s+\delta j| \geq |p-s| - |\delta j| > |p-s|/2$.

(b) It is immediate since by C.1, $f_{(j+p)\mathcal{I}(j+p\neq s)} = D\lambda_{|j+p-s|_+}^{-\alpha}(1+o(1))$ and $\lambda_1^{-\alpha}\lambda_{k_1}^\alpha = o(k_1)$. $\qquad\square$

Lemma 6.2. *Denote $\varphi(k_1) = O\left(k_1^{-1/2}\right) + O\left(k_1^{\alpha-1}\right)\mathcal{I}(\alpha \geq 1/2)\mathcal{I}(|p-s| < 2k_1)$. Then,*

$$(a)\ E\left|\widetilde{f}_p^{-1}\ddot{f}_p - 1\right| = \varphi(k_1) \quad and\ (b)\ E\left|\widetilde{f}_p^{-1}\widehat{f}_p - 1\right| = \varphi(k_1).$$

Proof. We begin with (a). $\widetilde{f}_p^{-1}\ddot{f}_p - 1$ is

$$\widetilde{f}_p^{-1}(2k_1+1)^{-1}\sum_{j+p\neq s} f_{j+p}\left(\frac{I_{j+p}}{f_{j+p}} - 1\right) + \widetilde{f}_p^{-1}(2k_1+1)^{-1}(I_s - f_{s+1})\mathcal{I}(|p-s| \leq k_1).$$
$$(6.1)$$

In view of Propositions A.1 and A.2 of Hidalgo and Robinson (2002) and Lemma 6.1, the first term of (6.1) is $\varphi(k_1)$, whereas the second term of (6.1) is also $\varphi(k_1)$ by Lemma 6.1 part $(\mathbf{b})$ and that $E(n^{-\alpha}I_s) < D$.

To show part (b), it suffices to examine $\widetilde{f}_p^{-1}\left(\widehat{f}_p - \ddot{f}_p\right)$, which is by definition

$$\widetilde{f}_p^{-1}\left(n^{-1} - \ddot{f}_p\right)\mathcal{I}\left(\ddot{f}_p < n^{-1}\right) = \left(\left(\widetilde{f}_p^{-1}n^{-1} - 1\right) - \left(\widetilde{f}_p^{-1}\ddot{f}_p - 1\right)\right)\mathcal{I}\left(\ddot{f}_p < n^{-1}\right).$$
$$(6.2)$$

By Cauchy-Schwarz inequality, the second moment of the right side of (6.2) is bounded by

$$2E\left(\widetilde{f}_p^{-1}\ddot{f}_p - 1\right)^2 + 2\left(\widetilde{f}_p^{-1}n^{-1} - 1\right)^2 E\left(\mathcal{I}\left(\ddot{f}_p < n^{-1}\right)\right) \leq DE\left(\widetilde{f}_p^{-1}\ddot{f}_p - 1\right)^2,$$

using that $E\left(\mathcal{I}\left(\ddot{f}_p < n^{-1}\right)\right)$ is

$$\Pr\left\{\widetilde{f}_p^{-1}\ddot{f}_p - 1 < \widetilde{f}_p^{-1}n^{-1} - 1\right\} \leq \Pr\left\{\left|\widetilde{f}_p^{-1}\ddot{f}_p - 1\right| > \left|1 - \widetilde{f}_p^{-1}n^{-1}\right|\right\}$$

because by Lemma 6.1, $\widetilde{f}_p^{-1}n^{-1} - 1 < -D$ for n large enough. Now use part (a) and Markov's inequality to conclude. $\qquad\square$

Lemma 6.3. *Let* $2k_1 < v < u \leq [n/2]$ *and* $p = 0, 1, ..., [n/2]$. *Denoting* $\psi(v, u) = O\left(\frac{\max(u-v, k_1)^{1/\tau}}{k_1^{(2+\tau^2)/2\tau^2}}\right)$ *and* $\varphi(k) = O\left(\log^{-\mu-1} k\right)$,

$$(a) \quad \begin{cases} E\left(\sup_{p:|\ p-s|\leq 2k_1}\left|\widetilde{f}_p^{-1}\left(\ddot{f}_p - \widetilde{f}_p\right)\right|^\mu\right) = \varphi(k_1) \\ E\left(\sup_{p:2k_1<|p-s|=v+1,...,u}\left|\widetilde{f}_p^{-1}\left(\ddot{f}_p - \widetilde{f}_p\right)\right|\right) = \psi(v, u) \end{cases}$$

$$(b) \quad \begin{cases} E\sup_{p:|p-s|\leq 2k_1}\left|\widetilde{f}_p^{-1}\left(\ddot{f}_p - \widehat{f}_p\right)\right|^\mu = \varphi(k_1) \\ E\sup_{p:2k_1<|p-s|=v+1,...,u}\left|\widetilde{f}_p^{-1}\left(\ddot{f}_p - \widehat{f}_p\right)\right| = \psi(v, u). \end{cases}$$

Proof. For notational simplicity, we shall take $s = 0$. We begin with part $(\mathbf{a})$. From Hidalgo and Robinson's (2002) Proposition A.1, it suffices to examine the behaviour of $\widetilde{f}_p^{-1}\left(\ddot{f}_p - E\ddot{f}_p\right)$. On the other hand, Hidalgo and Robinson's (2002) Proposition A.3 parts (a,b) imply that it suffices to examine the behaviour of $\widetilde{f}_p^{-1}\left(\ddot{f}_{\varepsilon,p} - E\ddot{f}_{\varepsilon,p}\right)$, where

$$\ddot{f}_{\varepsilon,p} = \frac{1}{2k_1 + 1}\sum_j f_{j+p}I_{\varepsilon,j+p}$$

and $I_{\varepsilon,p} = I_\varepsilon(\lambda_p)$ is the periodogram of $\{\varepsilon_t\}_{t=1}^n$. We examine $\sup_{p=v+1,...,u}\left|\widetilde{f}_p^{-1}\left(\ddot{f}_{\varepsilon,p} - E\ddot{f}_{\varepsilon,p}\right)\right|$ only, that of $\sup_{p=1,...,2k_1}\left|\widetilde{f}_p^{-1}\left(\ddot{f}_{\varepsilon,p} - E\ddot{f}_{\varepsilon,p}\right)\right|$ is similarly handled. Because $\sup_j|a_j| = \left(\sup_j|a_j|^\tau\right)^{1/\tau}$, the τth power of $\sup_{p=v+1,...,u}\left|\widetilde{f}_p^{-1}\left(\ddot{f}_{\varepsilon,p} - E\ddot{f}_{\varepsilon,p}\right)\right|$ is, except constants,

$$\sup_{p=v+1,...,u}\left|\frac{1}{2k_1 + 1}\sum_j \phi_{j+p,p}\left((2\pi)I_{\varepsilon,j+p} - 1\right)\right|^\tau,$$

where $\phi_{j,p} = \widetilde{f}_p^{-1}f_j$. The last displayed expression is bounded by

$$2^{\tau-1}\sup_q\sup_p\left|\frac{1}{2k_1 + 1}\sum_j\left(\phi_{j+p,p}\left((2\pi)I_{\varepsilon,j+p} - 1\right) - \phi_{j+b,p}\left((2\pi)I_{\varepsilon,j+b} - 1\right)\right)\right|^\tau$$

$$+ 2^{\tau-1}\sup_q\sup_p\left|\frac{1}{2k_1 + 1}\sum_j\phi_{j+b,p}\left((2\pi)I_{\varepsilon,j+b} - 1\right)\right|^\tau \qquad (6.3)$$

where $\sup_q$ and $\sup_p$ denote $\sup_{q=1+v/k_1^{1/\tau},...,u/k_1^{1/\tau}}$ and $\sup_{p=1+b-k_1^{1/\tau},...,b}$ respectively, and $b = qk_1^{1/\tau}$.

After the change of indices $j = j' - k_1$, the second term of (6.3) is bounded by

$$D \sup_q \sup_p \left| \frac{1}{2k_1 + 1} \sum_{j=0}^{2k_1 - 1} \left(\phi_{j+b-k_1,p} - \phi_{j+b+1-k_1,p} \right) \sum_{a=0}^{j} \left((2\pi) I_{\varepsilon, a+b-k_1} - 1 \right) \right|^{\tau}$$

$$+ D \sup_q \sup_p \left| \phi_{b+k_1,p} \right|^{\tau} \left| \frac{1}{2k_1 + 1} \sum_{j=0}^{2k_1} \left((2\pi) I_{\varepsilon, j+b-k_1} - 1 \right) \right|^{\tau}, \tag{6.4}$$

by Abel summation by parts. On the other hand, by C.1 and C.3,

$$\begin{aligned} \left| \phi_{j+b-k_1,p} - \phi_{j+b+1-k_1,p} \right| &\leq D \widetilde{f}_p^{-1} (j+b-k_1)^{-1-\alpha} n^{\alpha} \\ &\leq D \left(\frac{p}{j+b-k_1} \right)^{\alpha} (j+b-k_1)^{-1}, \end{aligned}$$

since by Lemma 6.1 part (a) $D^{-1} < \left| f_p^{-1} \widetilde{f}_p \right| < D$. So, using that $\sup_j |a_j|^{\tau} \leq \sum_j |a_j|^{\tau}$, by Hölder's inequality and that $D^{-1} < \left| \phi_{b+k_1,p} \right| < D$ by Lemma 6.1 part (a), we obtain that the first moment of (6.4) is bounded by

$$\frac{D}{2k_1 + 1} \sum_{q=1+v/k_1^{1/\tau}}^{u/k_1^{1/\tau}} \sum_{j=0}^{2k_1} \sup_p \left(\frac{p}{j+b-k_1} \right)^{\tau\alpha} (j+b-k_1)^{-\tau} E \left| \sum_{a=0}^{j} \left((2\pi) I_{\varepsilon, a+b-k_1} - 1 \right) \right|^{\tau}$$

$$+ D \sum_{q=1+v/k_1^{1/\tau}}^{u/k_1^{1/\tau}} E \left| \frac{1}{2k_1 + 1} \sum_{j=0}^{2k_1} \left((2\pi) I_{\varepsilon, j+b-k_1} - 1 \right) \right|^{\tau}$$

which, because $(2\pi) E I_{\varepsilon, j+p} = 1$ and proceeding as in the proof of Brillinger's (1981) Theorem 7.4.4, is bounded by

$$D \sum_{q=1+v/k_1^{1/\tau}}^{u/k_1^{1/\tau}} \left(\left(\frac{b}{b-k_1} \right)^{\tau\alpha} \frac{1}{2k_1} \sum_{j=0}^{2k_1} \frac{(j+1)^{\tau/2}}{(j+b-k_1)^{\tau}} + k_1^{-\tau/2} \right)$$

$$\leq D \sum_{q=1+v/k_1^{1/\tau}}^{u/k_1^{1/\tau}} \left(\left(\frac{1}{b} \right)^{\tau/2} + \left(\frac{1}{k_1} \right)^{\tau/2} \right) = O \left(\frac{\max(u-v, k_1)}{k_1^{\tau/2+1/\tau}} \right),$$

because $\alpha < 1$, $b \leq 2(b-k_1)$, $q \geq 1 + v/k_1^{1/\tau}$ and $b = qk_1^{1/\tau}$. Thus, we conclude that the second term of (6.3) is $O \left(\max(u-v, k_1)/k_1^{\tau/2+1/\tau} \right)$.

Next, we examine the first term of (6.3). Because

$$a_{p,k_1} = \sum_{j=0}^{2k_1} \left(\phi_{j+p-k_1,p} \left((2\pi) I_{\varepsilon, j+p-k_1} - 1 \right) - \phi_{j+b-k_1,p} \left((2\pi) I_{\varepsilon, j+b-k_1} - 1 \right) \right)$$

has at most $k_1^{1/\tau}$ terms, and because $(2\pi) E I_{\varepsilon, j+p} = 1$ and proceeding as in the proof of Brillinger's (1981) Theorem 7.4.4, its τth moment is bounded by $k_1^{1/2}$, so that the expectation of the first term of (6.3) is bounded by

$$D \sum_{q=1+v/k_1^{1/\tau}}^{u/k_1^{1/\tau}} \sum_{p=1+b-k_1^{1/\tau}}^{b} k_1^{1/2-\tau} = o_p \left(\frac{u-v}{k_1^{\tau/2+1/\tau}} \right),$$

because $\tau > 2$. This completes the proof of part (a).

To show part (**b**), denoting $a_p = \tilde{f}_p^{-1} n^{-1} - 1$ and using (6.2), it suffices to examine

$$\sup_p |a_p| \, \mathcal{I} \left(\ddot{f}_p < n^{-1} \right) \leq D \sup_p \mathcal{I} \left(\tilde{f}_p^{-1} \ddot{f}_p - 1 < a_p \right).$$

But the expectation of the right side is bounded by

$$
\begin{aligned}
D E \mathcal{I} \left(\sup_p \left| \tilde{f}_p^{-1} \ddot{f}_p - 1 \right| > \min_p |a_p| \right) &= D \Pr \left\{ \sup_p \left| \tilde{f}_p^{-1} \ddot{f}_p - 1 \right| > \min_p |a_p| \right\} \\
&\leq D \left(\min_p |a_p| \right)^{-\tau} E \left(\sup_p \left| \tilde{f}_p^{-1} \ddot{f}_p - 1 \right| \right)^{\tau} \\
&\leq D \left(\varphi(k_1) \mathcal{I}(p \leq 2k_1) + \psi(v, u) \mathcal{I}(p > 2k_1) \right)
\end{aligned}
$$

by Markov's inequality and because $\left(\sup_p |c_p| \right)^{\tau} = \sup_p |c_p|^{\tau}$. □

Lemma 6.4. *Let $h(u)$ be a twice continuously differentiable function in $(0,1)$ such that $h(0) = h'(0) = h(1) = 0$ where $h'(u) = \frac{d}{du} h(u)$. Consider a sequence $\{\nu_j\}$ such that $\left| j^2 \nu_j \right| \leq D$ for all j. Then, for $a > 2t$, and denoting $q = p - t + 1$,*

$$\sum_{j=a+1}^{p} h(j/p)(\nu_j - \nu_{j-t}) = O\left(\frac{t}{p^2} \log\left(\frac{p}{t} \right) + \frac{t^2}{pq^2} + \frac{t}{p^2} \right). \tag{6.5}$$

Proof. The left side of (6.5) is

$$\sum_{j=p-t+1}^{p} h(j/p)\nu_j - \sum_{j=a-t+1}^{a} h((j+t)/p)\nu_j + \sum_{j=a+1}^{p-t} (h(j/p) - h((j+t)/p))\nu_j. \tag{6.6}$$

Since the first derivative of $h(u)$ is continuous and $h(1) = 0$, from the mean value theorem it follows that the absolute value of the first term of (6.6) is bounded by

$$D \sum_{j=p-t+1}^{p} |(p-j)/p| \, |\nu_j| = O\left(t^2 / \left(p(p-t+1)^2 \right) \right) = O\left(p^{-1} q^{-2} t^2 \right),$$

using $\left| j^2 \nu_j \right| \leq D$. The absolute value of the second term of (6.6) is bounded by

$$D \sum_{j=a-t+1}^{a} \left(\frac{j+t}{p} \right)^2 |\nu_j| = O\left(\frac{t}{p^2} \right),$$

since $h(0) = h'(0) = 0$ and $\left| j^2 \nu_j \right| \leq D$, whereas the absolute value of the third term of (6.6) is bounded by

$$D \frac{t}{p} \sum_{j=a+1}^{p-t} \left| h'\left(\frac{j}{p} \right) + \frac{t}{p} h''\left(\frac{j}{p} + \xi \frac{t}{p} \right) \right| |\nu_j| = O\left(\frac{t}{p^2} \log\left(\frac{p}{t} \right) \right)$$

by Taylor expansion of $h'(x)$ and using that $h'(0) = 0$, where $\xi = \xi(j) \in (0,1)$. □

Lemma 6.5. *Let $\tilde{b}_t$, $\tilde{\tilde{b}}_t$, b_{3t} and b_{4t} be given in (5.16), (5.10) and (5.15) respectively. Then, for $\rho < 1/3$,*

$$(a) \ \sup_{t \leq \rho k} t^{-1} \left| b_{3t} - \tilde{b}_t \right| = o_p\left(k^{-1/2} \right), \quad (b) \ \sup_{t \leq \rho k} t^{-1} \left| b_{4t} - \tilde{\tilde{b}}_t \right| = o_p\left(k^{-1/2} \right).$$

Proof. We only examine part (a), part (b) is identical. Because by Lemma 6.3, $\sup_{\ell=1,\ldots,[n/2]} |\hat{g}_\ell - 1| = O_p\left(\log^{-2} k_1 \right)$, then except in a set Ω_n such that $\lim_n \Pr\{\Omega_n\} =$

0, $\log \widehat{g}_{\cdot+s} = \vartheta_{\cdot+s} - 2^{-1}\vartheta_{\cdot+s}^2 (1 + o_p(1))$ by Taylor expansion, which implies that for n sufficiently large, by definition of b_{3t},

$$\sup_{t \leq \rho k} t^{-1} \left| b_{3t} - \widetilde{b}_t \right| \leq D \sup_{t \leq \rho k} t^{-1} \sum_{p=2t+1}^{k-t} |\psi_p - \psi_{p+t}| \vartheta_{p+s}^2.$$

Because by Lemma 6.2, for $|p| > 2k_1$, $\vartheta_{p+s}^2 = O_p\left(k_1^{-1}\right)$, for $|p| < 2k_1$, $\vartheta_{p+s}^2 = O_p\left(k_1^{2(\alpha-1)}\mathcal{I}(\alpha \geq 1/2) + k_1^{-1}\right)$ and by C.5, $|\psi_p - \psi_{p+t}| \leq D |\psi_\xi'| t/k$, where $\psi_p' = \psi'(p/k)$ and $p \leq \xi \leq p+t$, the last displayed expression is

$$O_p\left(\frac{\mathcal{I}(\alpha \geq 1/2)}{k_1^{2(1-\alpha)}} \sum_{p=1}^{2k_1} \frac{\sup_{t < 2k_1} |\psi_\xi'|}{k} + \frac{1}{k_1}\sum_{p=1}^{k} \frac{\sup_{t \leq \rho k} |\psi_\xi'|}{k}\right) = o_p\left(k^{-1/2}\right)$$

by C.5 and C.4. $\qquad\square$

Lemma 6.6. *Let* $\phi_p = \phi(p/k)$ *where* $\phi(u)$ *is a continuous function in* $(0,1)$. *Define*

$$c_r(\mu; \vartheta) = \frac{2}{nk^{1/2}} \sum_{p=[k\mu]+1}^{[k\vartheta]} \phi_p \cos(r\lambda_p)$$

where $0 \leq \mu < \vartheta \leq 1$. *For any* $\mu < \vartheta_1 < \vartheta_2 \leq 1$, *if* $k/n \to 0$, *then*

$$\sum_{r_1=1}^{n-1}\sum_{r_2=1}^{n-r_1} c_{r_2}(\mu; \vartheta_1) c_{r_2}(\mu; \vartheta_2) = \int_\mu^{\vartheta_1} \phi^2(u)\, du\, (1 + o(1)). \qquad (6.7)$$

Proof. The left side of (6.7) is

$$\frac{4}{n^2 k} \sum_{p_1=[k\mu]+1}^{[k\vartheta_1]} \phi_{p_1} \sum_{p_2=[k\mu]+1}^{[k\vartheta_2]} \phi_{p_2} \sum_{r_1=1}^{n-1}\sum_{r_2=1}^{n-r_1} \cos(r_2\lambda_{p_1})\cos(r_2\lambda_{p_2})$$

$$= \frac{4}{n^2 k} \sum_{p=[k\mu]+1}^{[k\vartheta_1]} \phi_p^2 \sum_{r_1=1}^{n-1}\sum_{r_2=1}^{n-r_1} \cos^2(r_2\lambda_p) \qquad (6.8)$$

$$+ \frac{2}{n^2 k} \sum_{p_1=[k\mu]+1}^{[k\vartheta_1]} \phi_{p_1} \sum_{p_2=[k\mu]+1, p_2 \neq p_1}^{[k\vartheta_2]} \phi_{p_2} \sum_{r_1=1}^{n-1}\sum_{r_2=1}^{n-r_1} \{\cos(r_2\lambda_{p_1+p_2}) + \cos(r_2\lambda_{p_1-p_2})\}.$$

Because, see Robinson (1995b), $\sum_{r_1=1}^{n-1}\sum_{r_2=1}^{n-r_1}\cos^2(r_2\lambda_p) = (n-1)^2/4$, and

$$\sum_{r_1=1}^{n-1}\sum_{r_2=1}^{n-r_1} \{\cos(r_2\lambda_{p_1+p_2}) + \cos(r_2\lambda_{p_1-p_2})\} = -n, \text{ for } p_1 \neq p_2,$$

the right side of (6.8) is

$$\frac{(n-1)^2}{n^2} \left(\frac{1}{k}\sum_{p=[k\mu]+1}^{[k\vartheta_1]} \phi_p^2\right) - \frac{2}{nk} \sum_{p_1=[k\mu]+1}^{[k\vartheta_1]} \phi_{p_1} \sum_{p_2=[k\mu]+1, p_2 \neq p_1}^{[k\vartheta_2]} \phi_{p_2} = \int_\mu^{\vartheta_1} \phi^2(u)\, du\, (1 + o(1)),$$

because $\phi(u)$ is continuous in u and $k/n \to 0$. $\qquad\square$

Lemma 6.7. *Denote* $\eta_p = (2\pi)I_{\varepsilon,p} - 1$ *and* $\phi(u)$ *as in Lemma 6.6. The process*

$$R_n(\vartheta) = \frac{1}{k^{1/2}} \sum_{p=[k\vartheta]+1}^{k-[k\vartheta]} \phi_p \eta_p; \qquad 0 \leq \vartheta \leq 1/2$$

is tight.

Proof. Since by Proposition 5.3, the finite limit distributions of $R_n(\vartheta)$ converge to those of a Gaussian process with continuous paths, then by Billingsley's (1968) Theorem 15.6, it suffices to check the moment condition

$$E\left(\left|R_n(\vartheta_2) - R_n(\vartheta)\right|^\tau \left|R_n(\vartheta) - R_n(\vartheta_1)\right|^\tau\right) \le D(\vartheta_2 - \vartheta_1)^\psi \qquad (6.9)$$

for some $\tau > 0, \psi > 1$, where $0 \le \vartheta_1 < \vartheta < \vartheta_2 \le 1/2$. Because

$$R_n(\vartheta) - R_n(\vartheta_2) = \frac{1}{k^{1/2}} \sum_{p=[k\vartheta]+1}^{[k\vartheta_2]} \phi_p \eta_p + \frac{1}{k^{1/2}} \sum_{p=k-[k\vartheta_2]+1}^{k-[k\vartheta]} \phi_p \eta_p$$

a sufficient condition for (6.9) to hold true is

$$E\left(\left|\frac{1}{k^{1/2}} \sum_{p=[k\vartheta]+1}^{[k\vartheta_2]} \phi_p \eta_p\right|^\tau \left|\frac{1}{k^{1/2}} \sum_{p=[k\vartheta_1]+1}^{[k\vartheta]} \phi_p \eta_p\right|^\tau\right) \le D(\vartheta_2 - \vartheta_1)^\psi \quad (6.10)$$

$$E\left(\left|\frac{1}{k^{1/2}} \sum_{p=k-[k\vartheta_2]+1}^{k-[k\vartheta]} \phi_p \eta_p\right|^\tau \left|\frac{1}{k^{1/2}} \sum_{p=k-[k\vartheta]+1}^{k-[k\vartheta_1]} \phi_p \eta_p\right|^\tau\right) \le D(\vartheta_2 - \vartheta_1)^\psi.$$

We will examine (6.10) only, being the last displayed inequality similarly handled.

By definition of η_p,

$$\frac{1}{k^{1/2}} \sum_{p=[k\vartheta]+1}^{[k\vartheta_2]} \phi_p \eta_p = \left(\frac{1}{k} \sum_{p=[k\vartheta]+1}^{[k\vartheta_2]} \phi_p\right)\left(\frac{k^{1/2}}{n} \sum_{r=1}^{n} (\varepsilon_r^2 - 1)\right) + \sum_{r=2}^{n} \varepsilon_r \sum_{a=1}^{r-1} \varepsilon_a c_{r-a}(\vartheta, \vartheta_2)$$

$$: = \mathcal{E}_{1,n}(\vartheta, \vartheta_2) + \mathcal{E}_{2,n}(\vartheta, \vartheta_2),$$

where $c_r(\vartheta, \vartheta_2)$ was defined in Lemma 6.6. Because $\left|\sum_{p=[k\vartheta]+1}^{[k\vartheta_2]} \phi_p\right| \le Dk|\vartheta_2 - \vartheta|$ by continuity of $\phi(x)$ and $E\left(\sum_{r=1}^{n} (\varepsilon_r^2 - 1)\right)^2 < Dn$ by C.2, $E\left(\left|\mathcal{E}_{1,n}(\vartheta, \vartheta_2)\right|\left|\mathcal{E}_{1,n}(\vartheta_1, \vartheta)\right|\right) \le (\vartheta_2 - \vartheta_1)^2$ by the Cauchy-Schwarz inequality and that $|\vartheta_2 - \vartheta||\vartheta - \vartheta_1| < |\vartheta_2 - \vartheta_1|^2$. That is, $\mathcal{E}_{1,n}(\vartheta, \vartheta_2)$ satisfies the inequality in (6.10) with $\tau = 1$ and $\psi = 2$. So, to complete the proof, it suffices to examine that the inequality in (6.10) holds true for $\mathcal{E}_{2,n}(\vartheta, \vartheta_2)$. The fourth moment of $\mathcal{E}_{2,n}(\vartheta, \vartheta_2)$ is

$$E\left[\sum_{2=r_1 \le r_2 \le r_3 \le r_4}^{n} \prod_{j=1}^{4} \varepsilon_{r_j} \left(\sum_{a_j=1}^{r_j-1} \varepsilon_{a_j} c_{r_j-a_j}(\vartheta, \vartheta_2)\right)\right] \le D \prod_{j=1}^{4} \left(\sum_{1 \le a_j \le r_j \le n} c_{r_j-a_j}^2(\vartheta, \vartheta_2)\right)^{1/2}$$

$$= D\left(\sum_{1 \le a \le r \le n} c_{r-a}^2(\vartheta, \vartheta_2)\right)^2,$$

proceeding as in the proof of Lemma 5.4 of Giraitis et al. (2001). But proceeding as in Lemma 6.6, the right side of the last displayed equation is bounded by

$$D\left(\int_{\vartheta}^{\vartheta_2} \phi^2(u) \, du\right)^2 \le D(\vartheta_2 - \vartheta_1)^2$$

since $\phi(u)$ is continuous. This concludes the proof choosing $\tau = \psi = 2$. $\qquad \square$

Lemma 6.8. *Let $2k_1 \le t_0 < \rho k$ for some arbitrarily small $\rho > 0$. Then,*

$$(a) \quad \begin{cases} \sup_{2k_1 < t \le t_0} \left| t^{-1} \left(\widehat{\xi}_n^{(5)}(t) - b_{3t} \right) \right| = O_p \left(\frac{t_0}{kk_1^\beta} + \frac{k_1 t_0}{k^2 \log k_1} \right) \\ \sup_{2k_1 < t \le t_0} \left| t^{-1} \left(\widehat{\xi}_n^{(6)}(t) - b_{4t} \right) \right| = O_p \left(\frac{t_0}{kk_1^\beta} + \frac{k_1 t_0}{k^2 \log k_1} \right) \end{cases} \qquad (6.11)$$

$$(b) \quad \begin{cases} \sup_{Lk^{1/2} < t \le 2k_1} \left| t^{-1} \left(\widehat{\xi}_n^{(5)}(t) - b_{3t} \right) \right| = O_p \left(\frac{k_1^2}{k^2 \log k_1} + \frac{k_1}{kk_1^\beta} \right) \\ \sup_{Lk^{1/2} < t \le 2k_1} \left| t^{-1} \left(\widehat{\xi}_n^{(6)}(t) - b_{4t} \right) \right| = O_p \left(\frac{k_1^2}{k^2 \log k_1} + \frac{k_1}{kk_1^\beta} \right) \end{cases} \qquad (6.12)$$

where b_{3t} and b_{4t} are given by (5.10) and (5.15) respectively, and $L > 0$ and $\beta > 0$.

Proof. We begin with (**a**). We only examine the first equality in (6.11), being the proof of the second equality similarly handled. By definition, see Proposition 5.3, $\widehat{\xi}_n^{(5)}(t) = b_{1t} + b_{2t}$. First, $\sup_{2k_1 < t \le t_0} \left| t^{-1} b_{1t} \right|$ satisfies the equality in (6.11) by $C.5$ and since the sum in p has at most $2k_1$ terms, say $p^* = 1, ..., 2k_1$, for which $\sup_{p^* = 1, ..., 2k_1} \left| \log \widehat{g}_{p^* + s} \right| = O_p \left(\log^{-1} k_1 \right)$, whereas for the remaining terms $\sup_{p = 1, ..., k; p \ne p^*} \left| \log \widehat{g}_{p+s} \right| = O_p \left(k_1^{-\beta} \right)$ using (5.7) and (5.9).

Next we estimate $b_{2t} - b_{3t}$. First by (5.11), $\sup_{2k_1 < t \le t_0} t^{-1} \left| b_{2t} - b_{3t} \right|$ is

$$\sup_{2k_1 < t \le t_0} t^{-1} \left| \sum_{p=k-t+1}^{k} \psi_p \log \left(\widehat{g}_{p+s} \right) - \sum_{p=t+1}^{2t} \psi_{p+t} \log \left(\widehat{g}_{p+s} \right) \right| = O_p \left(t_0 k^{-1} k_1^{-\beta} + t_0^2 k^{-2} k_1^{-\beta} \right),$$

by $C.5$ and using (5.7) and (5.9) and Markov's inequality.

Part (**b**). As was done in part (a), we only examine the first equality in (6.12), being the second similarly handled. By C.5 and (5.7) and (5.9) together with Markov's inequality, it follows easily that $\sup_{Lk^{1/2} < t \le 2k_1} \left| t^{-1} b_{1t} \right| = O_p \left(k^{-2} k_1^2 \log^{-1} k_1 \right)$. Finally, we estimate $b_{2t} - b_{3t}$. As was done in part (a), it suffices to examine

$$\sup_{Lk^{1/2} < t \le 2k_1} t^{-1} \left\{ \left| \sum_{p=k-t+1}^{k} \psi_p \log \left(\widehat{g}_{p+s} \right) \right| + \left| \sum_{p=t+1}^{2t} \psi_{p+t} \log \left(\widehat{g}_{p+s} \right) \right| \right\}.$$

The second term of the last displayed expression is $O_p \left(k_1^2 k^{-2} \log^{-1} k_1 \right)$ by C.5 and using $(5.7) - (5.8)$, whereas the first term is $O_p \left(k_1 k^{-1} k_1^{-\beta} \right)$ using C.5 and (5.7) and (5.9), which concludes the proof. $\square$

Lemma 6.9. *Let $\phi(u)$ be as in Lemma 6.6. Then,*

$$\sup_{\omega \in [0,1]} \left| \frac{1}{k^{1/2}} \sum_{j=1}^{[k\omega]} \phi_j \left(\frac{I_{j+s}}{f_{j+s}} - 2\pi I_{\varepsilon, j+s} \right) \right| = o_p(1).$$

Proof. Writing $u_j = f_{j+s}^{-1/2} \omega_{j+s,x}$ and $v_j = (2\pi)^{1/2} \omega_{j+s,\varepsilon}$ where $\omega_{j+s,x}$ and $\omega_{j+s,\varepsilon}$ are the discrete Fourier transforms of x_r and ε_r respectively the left side of the last displayed expression is, by triangle inequality, bounded by

$$\sup_{\omega \in [0,1]} \frac{1}{k^{1/2}} \sum_{j=1}^{[k\omega]} |\phi_j| \left| u_j - v_j \right|^2 + 2 \sup_{\omega \in [0,1]} \left| \frac{1}{k^{1/2}} \sum_{j=1}^{[k\omega]} \phi_j v_j \left(\overline{u}_j - \overline{v}_j \right) \right|, \qquad (6.13)$$

where $\overline{c}$ denotes the conjugate of the complex number c.

The first term of (6.13) is $o_p(1)$ since its expectation is bounded by

$$k^{-1/2} \sum_{j=1}^{k} |\phi_j| \left\{ \left(E|u_j|^2 - 1 \right) - (E(u_j \overline{v}_j) - 1) - (E(\overline{u}_j v_j) - 1) + \left(E|v_j|^2 - 1 \right) \right\}$$

$$= O\left(k^{-1/2} \sum_{j=1}^{k} \frac{\log j}{j} \right) \tag{6.14}$$

because $E|v_j|^2 = 1$, $|\phi_j| \leq D$ and the extension of Theorems 1 and 2 of Robinson (1995a) given in Lemma 4.4 of Giraitis et al. (2001).

Next, to show that the second term of (6.13) is $o_p(1)$, it suffices to show that the finite dimensional distributions of the term inside the absolute value converge to zero and the tightness condition. First, choosing ω_1^* such that $[k\omega_1^*] = \max\left([k^\zeta], [k\omega_1] \right)$ for some $0 < \zeta < 1/4$, then for any $0 < \omega_1 < \omega_2 < 1$, $E\left| k^{-1/2} \sum_{j=1+[k\omega_1]}^{[k\omega_2]} \phi_j v_j (\overline{u}_j - \overline{v}_j) \right|^2$ is bounded by

$$2E\left| k^{-1/2} \sum_{j=1+[k\omega_1^*]}^{[k\omega_2]} \phi_j v_j (\overline{u}_j - \overline{v}_j) \right|^2 + 2k^{-1}[k\omega_1^*]([k\omega_1^*] - [k\omega_1]) \tag{6.15}$$

$$\leq Dk^{-1} \log^2 k \left(\left([k\omega_2]^{1/2} - [k\omega_1^*]^{1/2} \right) (\log(k\omega_2) - \log(k\omega_1^*)) + [k\omega_1^*]([k\omega_1^*] - [k\omega_1]) \right),$$

proceeding as with (4.8) in Robinson (1995b). So, the finite dimensional distributions of the second term of (6.13) converge to zero in probability by Markov's inequality.

To complete the proof we need to show tightness. Since the limiting process has continuous paths, by Billingsley's (1968) Theorem 15.6, it suffices to show that

$$E\left| k^{-1/2} \sum_{j=1+[k\omega_1]}^{[k\omega_2]} \phi_j v_j (\overline{u}_j - \overline{v}_j) \right|^4 \leq D(H(\omega_2) - H(\omega_1))^{1+\delta} \tag{6.16}$$

where $\delta > 0$, $0 < \omega_1 < \omega_2 < 1$ and $H(\omega)$ is a nondecreasing continuous function. The left side of (6.16) is bounded by

$$k^{-2} \left(|M_4| + 3M_2^2 \right)$$

where M_r denotes the rth cumulant of $\sum_{j=1+[k\omega_1]}^{[k\omega_2]} \phi_j v_j (\overline{u}_j - \overline{v}_j)$. Using the inequality in (6.15), $k^{-2}M_2^2 \leq D(H(\omega_2) - H(\omega_1))^{1+\delta}$, so it remains to show that $k^{-2}|M_4|$ satisfies the inequality in (6.16). Now $k^{-2}|M_4|$ is

$$\frac{1}{k^2} \sum_{j_1,j_2,j_3,j_4=1+[k\omega_1]}^{[k\omega_2]} \left(\prod_{i=1}^{4} \phi_{j_i} \right) cum\left(v_{j_1}\overline{z}_{j_1}, v_{j_2}\overline{z}_{j_2}, v_{j_3}\overline{z}_{j_3}, v_{j_4}\overline{z}_{j_4} \right) \tag{6.17}$$

where we have abbreviated $u_j - v_j$ by z_j. By Theorem 2.3.2 of Brillinger (1981) and denoting $X_{j1} = \phi_j v_j$ and $X_{j2} = \phi_j z_j$,

$$(6.17) = \frac{1}{k^2} \sum_{\vartheta} cum\left(X_{j\ell}; j\ell \in \vartheta_1 \right) ... cum\left(X_{j\ell}; j\ell \in \vartheta_p \right)$$

where the summation is over all indecomposable partitions $\vartheta = \vartheta_1 \cup ... \cup \vartheta_p$. A typical component in $cum\left(X_{j\ell}; j\ell \in \vartheta_1 \right)$ has q_1 elements v_j and q_2 elements z_j, so applying formulae of Brillinger [(1981), (2.6.3), page 26 and (2.10.3), page 39], we

deduce after straightforward calculations that $cum\left(X_{j\ell}; j\ell \in \vartheta_1\right)$ is $\prod_{j \in \upsilon_1} \phi_j$ times

$$\frac{\mu_{q_1+q_2}}{k^{(q_1+q_2)/2}} \int_{[-\pi,\pi]^{q_1+q_2-1}} \left(\frac{\beta\left(\lambda^1 + ... + \lambda^{(q_1-1)} + \nu^1 + ... + \nu^{q_2}\right)\beta\left(-\lambda^1\right)...\beta\left(-\lambda^{q_1-1}\right)}{\beta_{j_1}...\beta_{j_{q_1}}} \right.$$

$$\times \widetilde{\beta}\left(-\nu^1\right)...\widetilde{\beta}\left(-\nu^{q_2}\right) E_{j_1...j_{q_1}\ell_1...\ell_{q_2}}\left(\lambda^1,...,\lambda^{(q_1-1)},\nu^1,...,\nu^{q_2}\right) d\lambda^1...d\lambda^{(q_1-1)}d\nu^1...d\nu^{q_2},$$

where $E_{j_1...j_q\ell_1...\ell_p}\left(\lambda^1,...,\lambda^{(q-1)},\nu^1,...,\nu^p\right)$ is

$$G\left(\lambda_{j_1} - \left[\lambda^1 + ... + \lambda^{(q-1)} + \nu^1 + ... + \nu^p\right]\right) G\left(\lambda_{j_2} + \lambda^1\right)$$

$$\times ...G\left(\lambda_{j_q} + \lambda^{(q-1)}\right) G\left(\nu^1 - \lambda_{\ell_1}\right) \times ... \times G\left(\nu^p - \lambda_{\ell_p}\right),$$

with $G\left(\lambda\right) = \sum_{t=1}^n e^{it\lambda}$ and, say, $\widetilde{\beta}\left(-\nu^1\right) = \beta_{\ell_1}^{-1}\beta\left(-\nu^1\right) - 1$. But by a routine extension of Lemma 3 of Robinson (1995b) and observing that in each partitioned υ, the subindex $j_i, i = 1, ..., 4$, appears only once

$$(6.17) \leq Dk^{-2} \left(\sum_{j=1+[k\omega_1]}^{[k\omega_2]} \frac{1}{j^{1/2}} \right)^4 \leq D\left(H\left(\omega_2\right) - H\left(\omega_1\right)\right)^4,$$

where $H\left(\omega\right) = \omega^{1/2}$ which is a nondecreasing continuous function. $\qquad\square$

Remark 6.1. *An alternative proof of this lemma can be found in Lemma 4 of Delgado et al. (2004).*

Lemma 6.10. *(Brillinger, 1981 p. 15) Let $h\left(x\right)$, $0 \leq x \leq 1$, be integrable and have an integrable derivative $h^{(1)}\left(x\right)$. Then*

$$\frac{1}{n}\sum_{j=0}^n h\left(\frac{j}{n}\right) - \int_0^1 h\left(x\right)dx = \frac{1}{2n}\left(h\left(0\right) + h\left(1\right)\right) + \frac{1}{n}\int_0^1 \left(nx - [nx] - \frac{1}{2}\right)h^{(1)}\left(x\right)dx. \ \square$$

7. CONCLUSIONS

In this paper we have studied a nonparametric estimator for the pole of a long memory process under mild conditions on the spectral density $f\left(\lambda\right)$. Specifically, we have only assumed that $f\left(\lambda\right) \sim C\left|\lambda - \lambda^0\right|^{-\alpha}$ with $C > 0$, but smooth elsewhere, and where α, the memory parameter, belongs to the interval $(0, 1)$. We have shown that the estimator $\widehat{\lambda}^0$ of the pole λ^0 is consistent and we have characterized its limit distribution. More precisely, $\widehat{\lambda}^0$, centered around λ^0 and appropriately renormalized, is asymptotically normal when $\lambda^0 \in (0, \pi)$, whereas if $\lambda^0 = 0$ or π, the asymptotic distribution is a mixture of a discrete and continuous random variable. In particular, when $\lambda^0 = 0$ the asymptotic distribution takes the value 0 with probability $1/2$ and behaves as a (truncated) normal random variable for positive values. In addition, we have shown that the asymptotic statistical properties of a two step estimator of α are the same as when λ^0 is known.

REFERENCES

[1] Adenstedt, I. (1974). On large-sample estimation for the mean of a stationary random sequence. Annals of Mathematical Statistics, 2, 1095-1107.

[2] Andel, J. (1986). Long-memory time series models. Kybernetika, 22, 105-123.

[3] Arteche, J. and Robinson, P.M. (2000). Semiparametric inference in seasonal and cyclical long memory processes. Journal of Time Series Analysis, 21, 1-25.

[4] Billingsley, P. (1968). Convergence of Probability Measures. John Wiley, New York.

[5] Brillinger, D.R. (1981). Time Series, Data Analysis and Theory. McGraw-Hill.

[6] Chen, Z-G. (1988). Consistent estimate for hidden frequencies in a linear process. Adv. in Applied Probability, 20, 295-314.

[7] Dahlhaus, R. (1989). Efficient parameter estimation for self-similar processes. Ann. of Statist., 17, 1749-1766.

[8] Davies, R.B. and Harte, D.S. (1987). Tests for Hurst Effect. Biometrika, 74, 95-101.

[9] Delgado, M.A., Hidalgo, J. and Velasco, C. (2004). Distribution free goodness-of-fit tests for linear processes. Preprint.

[10] Eddy, W.F. (1980). Optimum kernel estimators of the mode. Ann. of Statist., 8, 870-882.

[11] Fox, R. and Taqqu, M.S. (1986). Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. Ann. of Statist., 14, 517-532.

[12] Giraitis, L., Hidalgo, J. and Robinson, P.M. (2001). Gaussian estimation of parametric spectral density with unknown pole. Ann. of Statist., 29, 987-1023.

[13] Giraitis, L. and Leipus, R. (1995). A generalized fractionally differencing approach in long-memory modelling. Liet. Matem. Rink, 35, 65-81.

[14] Giraitis, L., Robinson, P.M. and Samarov, A. (1997). Rate optimal semiparametric estimation of the memory parameter of the Gaussian time series with long-range dependence. Journal of Time Series Analysis, 18, 49-60.

[15] Giraitis, L. and Surgailis, D. (1990). A central limit theorem for quadratic forms in strongly dependent linear variables and its application to asymptotic normality of Whittle's estimate. Prob. Theory and Related Fields, 86, 87-104.

[16] Granger, C.W.J. and Joyeux, R. (1980). An introduction to long memory time series and fractional differencing. Journal of Time Series Analysis, 1, 15-30.

[17] Gray, H.L., Zhang, N-F. and Woodward, W.A. (1989). On generalized fractional processes. Journal of Time Series Analysis, 10, 233-257.

[18] Grenander, U. and Rosenblatt, M. (1957). Statistical analysis of stationary time series. John Wiley, New York.

[19] Hannan, E.J. (1961). Testing for a jump in the spectral function. J. Royal Statist. Soc., Series B, 23, 394-404.

[20] Hannan, E.J. (1973). The estimation of frequency. J. Appl. Probab., 10, 510-519.

[21] Hannan, E.J. (1976). The asymptotic distribution of serial covariates. Ann. of Statist., 4, 396-399.

[22] Hidalgo, J. and Robinson, P.M. (2002). Adapting to unknown error autocorrelation in regression with long memory. Econometrica, 70, 1545-1581.

[23] Hosking, J. (1981). Fractional differencing. Biometrika, 68, 165-176.

[24] Hosoya, Y. (1997). A limit theory with long-range dependence and statistical inference on related models. Ann. of Statist., 25, 105-137.

[25] Kavalieris, L. and Hannan, E.J. (1994). Determining the number of terms in a trigonometric regression. Journal of Time Series Analysis, 15, 613-625.

[26] Müller, H-G. (1992). Change-points in nonparametric regression analysis. Ann. of Statist., 20, 737-761.

[27] Müller, H-G. and K. Prewitt (1992). Weak convergence and adaptive peak estimation for spectral densities. Ann. of Statist., 20, 1329-1349.

[28] Pollard, D. (1984). Convergence of stochastic processes. Springer, New York.

[29] Quinn, B.G. (1989). Estimating the number of terms in a sinusoidal regression. Journal of Time Series Analysis, 10, 71-75.

[30] Quinn, B.G. and Hannan, E.J. (2001). The estimation and tracking of frequency. Cambridge University Press, New York.

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

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

[33] Taqqu, M. (1975). Weak convergence to fractional Brownian motion and to the Rosenblatt process. Zeit. für Wahrs. und Verwandte Gebiete, 31, 287-305.

[34] Van der Vaart, A.W and Wellner, J.A. (1996). Weak convergence and empirical processes. Springer-Verlag, New York.

[35] Whittle, P. (1952). The simultaneous estimation of a time series harmonic components and covariance structure. Trab. Estadist., 3, 43-57.

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

Table 4.1

Bias and Standard Deviation of $\widehat{q}$ and $\widetilde{q}$

			α							
			0.2		0.4		0.6		0.8	
$\lambda^0 = 0$	n	256	9.35	(8.33)	6.38	(6.96)	4.24	(5.39)	2.80	(4.04)
			9.26	(7.88)	7.32	(6.85)	5.94	(6.01)	4.85	(5.25)
		1024	15.40	(15.50)	8.43	(10.74)	4.81	(7.64)	2.62	(5.76)
			22.91	(25.31)	15.55	(20.89)	9.60	(14.02)	6.73	(9.96)
$\lambda^0 = \frac{\pi}{2}$	n	256	0.003	(7.64)	-0.084	(5.33)	-0.091	(2.96)	-0.054	(1.56)
			0.209	(9.59)	0.270	(9.21)	0.272	(8.66)	0.320	(7.28)
		1024	0.051	(11.87)	0.117	(4.77)	0.063	(1.89)	0.216	(1.13)
			0.435	(27.89)	0.144	(25.77)	-0.213	(21.30)	-0.060	(13.52)

The first row in each cell corresponds to $\widehat{q}$, whereas the second row is that of $\widetilde{q}$.

Table 4.2

Bias, Standard Deviation and MSE of the long-memory parameter estimators

λ^0		n	α											
			0.2			0.4			0.6			0.8		
			BIAS	S.D	M.S.E.	BIAS	S.D	M.S.E.	BIAS	S.D	M.S.E.	BIAS	S.D	M.S.E.
0		256	-0.020	0.064	0.004	-0.022	0.067	0.005	-0.017	0.071	0.005	-0.006	0.072	0.005
			-0.019	0.057	0.004	-0.030	0.065	0.005	-0.024	0.074	0.006	-0.006	0.075	0.006
			-0.001	0.089	0.008	-0.003	0.089	0.008	-0.003	0.089	0.008	-0.007	0.082	0.007
			-0.015	0.084	0.007	-0.043	0.090	0.010	-0.064	0.099	0.014	-0.079	0.105	0.017
		1024	-0.006	0.024	0.001	-0.003	0.025	0.001	0.007	0.031	0.001	0.026	0.031	0.002
			-0.015	0.030	0.001	-0.014	0.035	0.001	0.002	0.040	0.002	0.032	0.045	0.003
			-0.002	0.042	0.002	-0.003	0.042	0.002	-0.005	0.042	0.002	-0.005	0.042	0.002
			-0.022	0.045	0.003	-0.039	0.054	0.004	-0.046	0.059	0.006	-0.051	0.066	0.007
$\frac{\pi}{2}$		256	-0.020	0.055	0.003	-0.035	0.059	0.005	-0.041	0.064	0.006	-0.040	0.070	0.006
			-0.010	0.046	0.002	-0.020	0.053	0.003	-0.004	0.062	0.004	0.043	0.059	0.005
			0.002	0.094	0.009	0.000	0.094	0.009	0.000	0.093	0.009	0.005	0.084	0.007
			-0.050	0.098	0.012	-0.083	0.121	0.022	-0.100	0.156	0.034	-0.083	0.182	0.040
		1024	-0.012	0.022	0.001	-0.015	0.024	0.001	-0.007	0.028	0.001	0.014	0.034	0.001
			-0.014	0.018	0.001	-0.017	0.020	0.001	-0.003	0.024	0.001	0.044	0.035	0.003
			-0.002	0.038	0.001	-0.004	0.038	0.001	-0.006	0.038	0.001	-0.007	0.038	0.001
			-0.039	0.046	0.004	-0.064	0.069	0.009	-0.061	0.096	0.013	-0.023	0.097	0.010

The first row in each cell corresponds to the estimator $\widehat{\alpha}(\lambda^0)$, whereas the second, third and fourth correspond to the estimators $\overset{\vee}{\alpha}(\widehat{\lambda}^0)$, $\widehat{\alpha}_{LOG}(\lambda^0)$ and $\widehat{\alpha}_{LOG}(\widetilde{\lambda}^0)$ respectively.