Ann. Inst. Statist. Math.
Vol. 55, No. 4, 705–736 (2003)
©2003 The Institute of Statistical Mathematics

SEMIPARAMETRIC ESTIMATION OF THE LONG-RANGE PARAMETER*

J. HIDALGO[1] AND Y. YAJIMA[2]

[1] *Department of Economics, London School of Economics, Houghton Street, London WC2A 2AE, U.K.*
[2] *Department of Economics, University of Tokyo, Hongo 7-3-1, Bunkyo-Ku, Tokyo 113-0033, Japan*

(Received April 9, 2002; revised October 25, 2002)

Abstract. We study two estimators of the long-range parameter of a covariance stationary linear process. We show that one of the estimators achieve the optimal semiparametric rate of convergence, whereas the other has a rate of convergence as close as desired to the optimal rate. Moreover, we show that the estimators are asymptotically normal with a variance, which does not depend on any unknown parameter, smaller than others suggested in the literature. Finally, a small Monte Carlo study is included to illustrate the finite sample relative performance of our estimators compared to other suggested semiparametric estimators. More specifically, the Monte-Carlo experiment shows the superiority of the proposed estimators in terms of the Mean Squared Error.

Key words and phrases: Long-range dependence, spectral estimation.

1. Introduction

In recent years there has been a growing interest in the study of covariance stationary scalar processes x_t, $t = 0, \pm 1, \pm 2, \dots$, which are observed at times $t = 1, \dots, n$, and whose spectral density function is neither bounded nor greater than zero at some frequency $\lambda_0 \in [0, \pi]$. Denoting by $\gamma_q = E((x_0 - Ex_0)(x_q - Ex_0))$ the lag-q autocovariance of x_t, the spectral density function of x_t, $f(\lambda)$, is defined from the relation $\gamma_q = \int_{-\pi}^{\pi} f(\lambda) \cos(q\lambda) d\lambda$, $q = 0, \pm 1, \dots$.

Most of the research has focused, without loss of generality, on the particular case

$$f(\lambda) \sim C\lambda^{-2d} \quad \text{as} \quad \lambda \to 0+ \tag{1.1}$$

where $C \in (0, \infty)$, $d \in (-\frac{1}{2}, \frac{1}{2})$ and "$\sim$" means that the ratio of left- and right-hand sides tends to one. When $d = 0$, $f(\lambda) \in (0, \infty)$ and corresponds to the so-called weakly dependent stationary process. When $d > 0$, $f(\lambda)$ diverges to infinity and we say that the stationary process x_t exhibits long-range dependence and when $d < 0$, $f(\lambda)$ converges to zero as $\lambda \to 0+$ and we say that x_t exhibits the phenomenon called antipersistence or negative dependence. We will refer d as the *long-range* parameter, although this is normally used when $d > 0$ only.

When full parameterization of $f(\lambda)$ is given, that is, it depends on an unknown finite set of parameters, say θ, parametric estimators of θ for the model $f(\lambda; \theta)$ have been exhaustively studied. See Yajima (1985), Fox and Taqqu (1986), Dahlhaus (1989),

*The first author research was funded by the Economic and Social Research Council (ESRC) reference number: R000238212. The second author research was funded by the Ministry of Education, Culture, Sports and Technology of Japan, reference number: 09CE2002 and B(2)10202202.

Giraitis and Surgailis (1990) and Hosoya (1997) among others, and Cheung and Diebold (1994) for a study of their small sample performance. However, one possible criticism is that their statistical properties are very sensitive to a correct specification of the model. In particular, misspecification can lead to inconsistent estimators of the long-range parameter d in (1.1).

To avoid such criticisms and when the main interest is in the estimation of d, (semiparametric) estimators of d have been examined which only take into account the behaviour of $f(\lambda)$ at frequencies around the origin. Among such estimators, we can mention the pioneering work by Geweke and Porter-Hudak (1983) and its modification given by Robinson (1995*a*) and Robinson (1994, 1995*b*). Although the estimators of d provided in Robinson (1994, 1995*a*) have a closed (explicit) form, they have the possible drawback that the asymptotic distribution may not be Gaussian, see Lobato and Robinson (1996), or of not being as efficient as the estimator GSE given in Robinson (1995*b*). However, the latter one, compared with Robinson's (1995*a*) estimator, does not achieve the optimal semiparametric rate of convergence obtained by Giraitis *et al.* (1997), although by a logarithmic factor.

Following an earlier suggestion by Parzen (1986), our aim in this paper is to provide statistical justification of two estimators of d, denoted $\widehat{d}$ and $\widehat{d}^*$. The latter estimator achieves the optimal semiparametric rate of convergence, as that of Robinson (1995*a*), whereas the former achieves a rate of convergence as close as desired to the optimal rate, so that their rate of convergence are faster than the GSE by a logarithm factor. That is, if we denote by n^δ, say, the optimal semiparametric rate of convergence, $\widehat{d}$ has a rate of convergence (even faster than) $n^\delta[\log\log n]^{-\alpha}$ for any arbitrarily small $\alpha > 0$. However, contrary to Robinson (1995*a*) we do not need to assume that the data is Gaussian. Moreover, if we are concerned with the *Mean Square Error* (*M.S.E.*), which in finite samples can be a more appropriate measure of performance of an estimator, we observe from the Monte-Carlo simulation in Section 3, that both $\widehat{d}$ and $\widehat{d}^*$ clearly outperforms the estimator in Robinson (1995*a*). Finally, the asymptotic variances of $\widehat{d}$ and $\widehat{d}^*$ are smaller than those in Robinson (1995*a*, *b*).

The remainder of the paper is organized as follows. In the next section, we describe the estimators of the parameter d and some results regarding the statistical properties of the spectral density estimator at frequencies converging to the origin. In Section 3, we provide a Monte Carlo experiment to assert the finite-sample performance of our estimators and we compare them with those described in Robinson (1994, 1995*a*, 1995*b*). In Section 4, we provide the proofs of the results given in Section 2, whereas in Section 5, we present some technical lemmas needed for the proofs of Section 4.

2. Estimation of the long-range parameter

To introduce and give an insight to our estimators of d (see Parzen (1986)), suppose that the spectral density $f(\lambda)$ of the process x_t satisfies

$$f(\lambda) = C\lambda^{-2d} \quad \text{for} \quad \lambda \in (0, \bar{\lambda}).$$

When this is the case, it follows that

$$d = h_w \left(\bar{\lambda}^{-1} \int_0^{\bar{\lambda}} w(\bar{\lambda}^{-1}\lambda) \log f(\lambda) d\lambda - \left(\bar{\lambda}^{-1} \int_0^{\bar{\lambda}} w(\bar{\lambda}^{-1}\lambda) d\lambda \right) \log f(\bar{\lambda}) \right)$$

after straightforward calculation and where $w(u)$, with $u \in (0,1)$, is a positive weight function and $h_w = (-2\int_0^1 w(u)\log(u)du)^{-1}$. We can thus expect that the Riemann discrete approximation of the right side of the last displayed equation

$$h_w \left(\frac{1}{k} \sum_{p=1}^{k} w_p \log f_p - \left(\frac{1}{k} \sum_{p=1}^{k} w_p \right) \log f_{k+1} \right) \tag{2.1}$$

is close to d, where we abbreviate $f(\lambda_p)$ by f_p, with $\lambda_p = \frac{2\pi p}{n}$, $p = 1, \ldots, n-1$, $\bar{\lambda} = \lambda_k$ and $w_p = w(p/k)$. However (2.1) depends on the unknown spectral density function of x_t, so that to make (2.1) feasible, we need to estimate f_p.

To that end, define the discrete Fourier transform and periodogram of x_t by

$$a(\lambda) = (2\pi n)^{-1/2} \sum_{t=1}^{n} x_t e^{it\lambda} \quad \text{and} \quad I(\lambda) = |a(\lambda)|^2 \tag{2.2}$$

respectively, where the correction for an unknown mean of x_t will be unnecessary since the statistics in (2.2) are to be computed at the Fourier frequencies λ_p.

A common estimator of $f(\lambda)$ is the average periodogram

$$\ddot{f}(\lambda) = \frac{1}{(m+1)} \sum_{j=-m/2;\lambda+\lambda_j \neq 0}^{m/2} I(\lambda + \lambda_j) \tag{2.3}$$

where $m = m(n)$ is an even number such that $m^{-1} + mn^{-1} \to 0$.

When the spectral density $f(\lambda)$ of the process x_t satisfies (1.1), some statistical properties of (2.3) are known. In particular, Robinson (1994) showed that $\ddot{f}(0)/C\lambda_{m/2}^{-2d}$ converges in probability to 1, whereas Hidalgo (1996) proved that, under suitable conditions, the continuous version of $\ddot{f}(\lambda)$ is $m^{1/2}$-consistent and asymptotically normal when λ lies in any open set outside the origin. Although some statistical properties of the periodogram at frequencies $\lambda = \lambda_p \to 0$ are known, see Robinson (1995a) or Hurvich and Beltrao (1993), the statistical properties of the estimator given in (2.3) for those $\lambda_p \to 0+$ have not been studied yet. Due to the aforementioned results of Robinson (1994) and that the rate of convergence of the second moment of $\ddot{f}(\lambda)$ can be very slow for $\lambda = \lambda_p$, with $p < m/2$, see Proposition 2.4, $\ddot{f}(\lambda)$ would have some adverse consequences for the results of Theorem 2.2, in particular for the behaviour of expressions (4.17) and (4.18). So, we modify the estimator given in (2.3) by

$$\begin{aligned} \widehat{f}_p = \widehat{f}(\lambda_p) = & \frac{1}{m+1} \sum_{j=-m/2}^{m/2} I_{j+|p|} \mathcal{I}\left(\frac{m}{2} < |p|\right) \\ & + \frac{2}{m} \sum_{j=1}^{m/2} I_{j+|p|} \mathcal{I}\left(0 < |p| \leq \frac{m}{2}\right) \end{aligned} \tag{2.4}$$

where $I_j = I(\lambda_j)$ and $\mathcal{I}(\cdot)$ denotes the indicator function.

Thus, we could estimate the parameter d by

$$\widetilde{d} = h_w \left(\frac{1}{k} \sum_{p=1}^{k} w_p \log \widehat{f}_p - \left(\frac{1}{k} \sum_{p=1}^{k} w_p \right) \log \widehat{f}_{k+1} \right), \tag{2.5}$$

that is, (2.1) with f_p replaced by its estimate $\widehat{f}_p$ given in (2.4).

Although $\widetilde{d}$ is consistent, as can be easily shown from Propositions 2.1–2.3 below, it possesses the undesirable property of having a slower rate of convergence than that obtained by Giraitis *et al.* (1997). Heuristically, this is because

$$\widetilde{f}_p^{-1} f_p^* - 1 = O(1) \quad \text{for} \quad p = O(m),$$

where $f_p^* = C\lambda_p^{-2d}$ and $\widetilde{f}_p$ is as in (2.4) but with I_j being replaced by f_j, induces a "bias" term in the estimator $\widetilde{d}$ which does not converge to zero fast enough. Thus, we modify $\widetilde{d}$ as follows. Let

$$b(d) = h_w \left(\frac{1}{k} \sum_{p=1}^{k} w_p \log \left(\frac{\overline{f}_p}{f_p^*} \right) - \left(\frac{1}{k} \sum_{p=1}^{k} w_p \right) \log \left(\frac{\overline{f}_{k+1}}{f_{k+1}^*} \right) \right),$$

where

$$\overline{f}_p = \overline{f}(\lambda_p) = \frac{C}{m+1} \sum_{j=-m/2}^{m/2} \lambda_{j+|p|}^{-2d} \mathcal{I}\left(\frac{m}{2} < |p|\right) + \frac{2C}{m} \sum_{j=1}^{m/2} \lambda_{j+|p|}^{-2d} \mathcal{I}\left(0 < |p| \le \frac{m}{2}\right)$$

and define the estimator (another estimator is given in (2.7)) of d by

$$\widehat{d} = \widetilde{d} - b(\overline{d}) \tag{2.6}$$

where $\overline{d}$ is a preliminary estimator of d, say that in (2.5), but with a bandwidth number $m = k^{1/2}$ in the definition of $\widehat{f}_p$ given in (2.4).

Before we analyze the properties of (2.6) and/or (2.7), it is convenient to examine the properties of $\widehat{f}_p$. For that purpose, introduce the following regularity assumptions:

A.1. There exist $C \in (0, \infty)$, $d \in (-\frac{1}{2}, \frac{1}{2})$ and $\beta \in (0, 2]$ such that

$$f(\lambda) = C\lambda^{-2d}(1 + O(\lambda^\beta)) \quad \text{as} \quad \lambda \to 0+.$$

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

$$x_t = \sum_{j=0}^{\infty} \alpha_j e_{t-j}, \qquad \sum_{j=0}^{\infty} \alpha_j^2 < \infty, \qquad \alpha_0 = 1,$$

where $E[e_t \mid \mathcal{F}_{t-1}] = 0$; $E[e_t^2 \mid \mathcal{F}_{t-1}] = 1$; $E[|e_t|^\ell \mid \mathcal{F}_{t-1}] = \mu_\ell$, $\ell = 3, \ldots, 2r$ and $r \ge 2$, almost surely $t = 0, \pm 1, \ldots$ and where $\mathcal{F}_t$ is the σ-algebra generated by $\{e_s; s \le t\}$ and with joint fourth cumulant of $e_{t_1}, e_{t_2}, e_{t_3}$ and e_{t_4} satisfying $\mathrm{cum}(e_{t_1}, e_{t_2}, e_{t_3}, e_{t_4}) = \kappa_e \mathcal{I}(t_1 = t_2 = t_3 = t_4)$.

A.3. $\frac{\partial^\zeta}{\partial \lambda^\zeta} |\alpha(\lambda)| = O(\lambda^{-\zeta} |\alpha(\lambda)|)$ as $\lambda \to 0+$, for $\zeta = 1$ and 2 and where $\alpha(\lambda) = \sum_{j=0}^{\infty} \alpha_j e^{ij\lambda}$.

A.4. $\frac{m}{n} + \frac{1}{m} \to 0$ as $n \to \infty$.

Assumptions A.1 and A.3 are not elaborated on since they are the same as those employed by Robinson (1995*b*). Assumption A.2 is similar to that in Robinson (1995*b*) except that we allow e_t to have more than four finite moments. Its motivation comes because the rate of convergence in Proposition 2.3 depends on the number of finite moments of e_t (compare with Brillinger's (1981), Theorem 7.7.4). Assumption A.4 indicates that the bandwidth parameter m increases slowly with n.

Write $g_p = g(\lambda_p) = \widetilde{f}_p^{-1}\widehat{f}_p - 1$.

PROPOSITION 2.1. *Assuming* A.1–A.4 *with* $r = 2$ *in* A.2, *as* $n \to \infty$,

(a) *For* $m \le p \le k$ *such that* $k/n + m/k \to 0$, $E(g_p) = O(m^{-1}\log m)$.

(b) *For* $m/2 < p < m$,

$$E(g_p) = O(m^{2d-1}(2p-m)^{-2d}\log m\mathcal{I}(d>0) + m^{-1}\log^2 m\mathcal{I}(d \le 0)).$$

(c) *For* $1 \le p \le m/2$,

$$E(g_p) = O\left(\frac{1}{p^{2d}}\frac{\log m}{m^{1-2d}}\mathcal{I}(d>0) + \frac{\log^2 m}{m}\mathcal{I}(d \le 0)\right).$$

PROPOSITION 2.2. *Assuming* A.1–A.4 *with* $r = 2$ *in* A.2, *as* $n \to \infty$,

(a) *For* $m \le p \le q \le k$ *such that* $k/n + m/k \to 0$,

$$m\,\mathrm{Cov}(g_p, g_q) = \begin{cases} O(1) & \text{if} \quad |q-p| \le m \\ O(n^{-1}m + mp^{-1}q^{-1/2} + mp^{-2}\log^2 q) & \text{if} \quad m < |q-p|. \end{cases}$$

(b) *For* $m/2 < p \le q < m$,

$$\begin{aligned} m\,\mathrm{Cov}(g_p, g_q) &= O\left(\mathcal{I}\left(d < \frac{1}{4}\right) + \log\left(\frac{2p+m}{2p-m}\right)\mathcal{I}\left(d = \frac{1}{4}\right)\right) \\ &\quad + O\left(\left(\frac{m}{2p-m}\right)^{4d-1}\mathcal{I}\left(d > \frac{1}{4}\right)\right) + O\left(\frac{m^{2d}\log^2(m)}{(2p-m)^{2d+1}}\mathcal{I}(d \ge 0)\right). \end{aligned}$$

(c) *For* $1 \le p \le q \le m/2$,

$$\begin{aligned} m\,\mathrm{Cov}(g_p, g_q) &= O\left(\mathcal{I}\left(d < \frac{1}{4}\right) + \log\left(\frac{2p+m}{2p}\right)\mathcal{I}\left(d = \frac{1}{4}\right)\right) \\ &\quad + O\left(\left(\frac{m}{p}\right)^{4d-1}\mathcal{I}\left(d > \frac{1}{4}\right)\right) + O\left(\frac{m^{2d}\log^2(m)}{p^{2d+1}}\mathcal{I}(d \ge 0)\right). \end{aligned}$$

PROPOSITION 2.3. *Let* $k = k(n)$ *be such that* $k/n \to 0$. *Assuming* A.1–A.4, *as* $n \to \infty$, $\sup_{p=1,\ldots,k}|g_p| = O_p(m^{-1/2-1/r^2}k^{1/r}) + o_p(1)$.

THEOREM 2.1. *Let* $p = p(n)$ *be such that* $p^{-1} + p^{-1}m + n^{-1}p \to 0$. *Assuming* A.1–A.4, *as* $n \to \infty$, $m^{1/2}g_p \xrightarrow{d} N(0,1)$.

To study the properties of (2.6), introduce the additional assumptions:

A.5. $\frac{1}{m} + \frac{m}{k} + \frac{k^{1+2\beta}}{n^{2\beta}} + \frac{k}{m^{(r^2+2)/(2r)}} \to 0$, where β is as in A.1 and r as in A.2.

Let $\mathcal{L}(\zeta)$ denote the set of continuous Lipchitz functions of order ζ.

A.6. $w(u)$ is a positive weight function which belongs to $\mathcal{L}(1)$ and such that $w(0) = 1$.

THEOREM 2.2. *Assuming* A.1–A.3, A.5 *and* A.6, *as* $n \to \infty$,

$$m^{1/2}(\widehat{d} - d) \xrightarrow{d} N(0, \Phi_w^2); \qquad \Phi_w = h_w \int_0^1 w(u)du.$$

From Theorem 2.2 we observe that the asymptotic variance of $\widehat{d}$ is smaller than those in Robinson (1995*a*, *b*). Indeed, suppose that $w(u) = 1 - u^c$. Then, when $c = \infty$, that is $w(u) = 1$, $h_w^2 = 1/4$, our estimator in (2.6) is as efficient as that of Robinson (1995*b*). However for $1 \leq c < \infty$, the asymptotic variance of $\widehat{d}$ is smaller than $1/4$. For instance, in the lower end of admissible values of c, that is $c = 1$, $h_w = 2/3$, so that the asymptotic standard deviation is equal to $1/3$, which is 33% smaller than $1/2$ obtained by Robinson (1995*b*). Moreover, we observe that the asymptotic distribution of $\widehat{d}$ does not depend on any unknown quantity or parameter. Finally we observe that A.5 implies that the rate of convergence of $\widehat{d}$ is as close as desired to the optimal semiparametric rate of convergence obtained by Giraitis *et al.* (1997) and achieved by the estimator in Robinson (1995*a*). In particular choosing, say, $m = k[\log\log k]^{-\alpha}$ for any arbitrarily small $\alpha > 0$, the rate of convergence of Robinson's (1995*a*) estimator is faster than $\widehat{d}$ by the factor $[\log\log k]^{\alpha/2}$, which in finite samples is negligible. For instance, for $k = 10^{3.2}$, that is $n = 10^4$ taken $\beta = 2$ in A.1 and $\alpha = .01$, $[\log\log k]^{\alpha/2} = 1.003$. This, in terms of the *M.S.E.*, which in finite samples gives a better and more accurate measure of the performance of an estimator, implies that the *M.S.E.* of $\widehat{d}$ will still be smaller than that of Robinson (1995*a*). This is confirmed in the Monte-Carlo experiment, where $\widehat{d}$ tends to be much better than the estimators of Robinson (1995*a*, *b*) in terms of *M.S.E.* On the other hand, it should be mentioned that the rate of convergence of $\widehat{d}$ is faster than that of Robinson (1995*b*) by a logarithm factor.

One, possibly, undesirable feature of the estimator of d given in (2.6) is that it depends on a preliminary estimator $\overline{d}$. However, it is worth observing that the choice of $k^{1/2}$ as the original bandwidth in $\overline{d}$ appears not to be very crucial. Indeed, this is the case as we can always iterate the estimator, that is,

$$\widehat{d}^{(i)} = \widetilde{d} - b(\widehat{d}^{(i-1)}),$$

where $\widehat{d}^{(i)}$ is the estimator of d at the i-th iteration. More specifically, from the proof of Theorem 2.2, in particular part (*b*), we observe that as long as the original choice of m in $\widehat{f}_p$, say m^*, satisfies that $m/(m^*)^\alpha \to 0$ for some $\alpha \in (0, 1)$, after a finite number of iterations $\widehat{d}^{(i)}$ will have the same asymptotic properties of $\widehat{d}$.

A second undesirable feature, from a theoretical point of view, is that $\widehat{d}$ does not achieve the semiparametric optimal rate of convergence. So, it might also be convenient to provide an estimator of d which avoids the need for a preliminary estimator, as well as achieving the optimal semiparametric rate of convergence.

To that end, we consider the following estimator

$$\widehat{d}^* = \overline{v}^{-1} \frac{1}{m} \sum_{p=1}^{m} v_p \widetilde{d}_p \tag{2.7}$$

where $v_p = v(p/m)$, $\overline{v} = m^{-1}\sum_{p=1}^{m} v_p$ and

$$\widetilde{d}_p = h_w \left(\frac{1}{p} \sum_{\ell=1}^{p} w_\ell \log \ddot{f}_\ell - \overline{w}_p \log \ddot{f}_{p+1} \right),$$

with $h_w = (-2\int_0^1 w(u)(\log u)du)^{-1}$, $w_\ell = w(\ell/p)$, $\overline{w}_p = p^{-1}\sum_{\ell=1}^{p} w_\ell$ and $\ddot{f}_p = \ddot{f}(\lambda_p)$ as defined in (2.3) with $m = m_1$ there.

Remark 1. It is worth mentioning that the results of Theorem 2.3 below follow if instead of $\ddot{f}(\lambda)$ we use $\widehat{f}(\lambda)$. However, the Monte-Carlo experiment indicates that the finite sample properties of $\widehat{d}^*$ are better when $\ddot{f}(\lambda)$ is used instead of $\widehat{f}(\lambda)$. This is the main motivation to use $\ddot{f}(\lambda)$ instead of $\widehat{f}(\lambda)$ in the definition of $\widetilde{d}_p$.

Before examining the properties of $\widehat{d}^*$, we need to modify slightly the results of Propositions 2.1 to 2.3 for those frequencies λ_p such that $2p < m_1$. Define

$$\widetilde{f}_p = \widetilde{f}(\lambda_p) = \frac{1}{m_1+1} \sum_{j=-m_1/2; j\neq -p}^{m_1/2} f_{j+p}.$$

and write $\ddot{g}_p = \ddot{g}(\lambda_p) = \widetilde{f}_p^{-1}\ddot{f}_p - 1$.

PROPOSITION 2.4. *Assume* A.1–A.4 *with* $r = 2$ *in* A.2. *For* $1 \leq p \leq \frac{m_1}{2}$, *as* $n \to \infty$,

$$E(\ddot{g}_p) = O\left(\frac{\log(m_1)}{m_1^{1-2d}} \mathcal{I}(d > 0) + \frac{\log^2(m_1)}{m_1} \mathcal{I}(d \leq 0) \right).$$

$$m_1 \operatorname{Var}(\ddot{g}_p) = O\left(\mathcal{I}\left(d < \frac{1}{4}\right) + \log(m_1)\mathcal{I}\left(d = \frac{1}{4}\right) + m_1^{4d-1}\mathcal{I}\left(d > \frac{1}{4}\right) \right).$$

PROOF. The proof of this proposition is omitted since it follows by an easy modification of Propositions A.1 and A.2 of Hidalgo and Robinson (2002). □

PROPOSITION 2.5. *Assuming* A.1–A.4, *as* $n \to \infty$,

$$\sup_{p=1,\ldots,m} |\ddot{g}_p| = o_p(1).$$

Let us introduce the following assumptions:

A.7. The weight functions $w(u)$ and $v(u)$ belong to $\mathcal{L}(1/4)$ and $\mathcal{L}(1)$ respectively and satisfy that $w(u) \sim cu^\zeta$, for some $1/4 \leq \zeta < 1$, and $v(u) \sim cu$ as $u \to 0+$.

A.8. $\frac{1}{m} + \frac{m}{m_1^2} + \frac{m^{1+2\beta}}{n^{2\beta}} + \frac{m_1^5}{m^3} + \frac{m}{m_1^{(r^2+2)/(2r)}} \to 0$, where β is as in A.1 and $r \geq 3$ in A.2.

Two comments about A.7 are in place. First, the reason to require that $w(u) \sim cu^\zeta$ as $u \to 0+$ with $\zeta \geq 1/4$ is due to a bias problem of our estimate of the long range parameter d, cf. the second term on the right of (4.19), that otherwise it would exist in the limiting distribution of $m^{1/2}(\widehat{d}^* - d)$. Second, it is worth noting that we could generalize the weight $v(u)$ to $v(u) \sim cu^{\mu}$ with the requirement that $\zeta < \mu$ to guarantee that the function $x^{-1}w(1/x)v(x)$ is integrable, but at the expense of strengthening the rates of m and m_1 in A.8. However, for simplicity we keep A.7 as it stands.

THEOREM 2.3. *Assuming* A.1–A.3, A.7 *and* A.8, *as* $n \to \infty$,

$$m^{1/2}(\widehat{d}^* - d) \xrightarrow{d} N(0, h_w^2 \Phi^2)$$

where Φ *is*

$$\left(\int_0^1 v(x)dx\right)^{-1} \int_0^1 \left(\left(v(u)\int_0^1 w(x)dx\right) - \int_u^1 \frac{w(u/x)v(x)}{x}dx\right) du.$$

Theorem 2.3 indicates that, in contrast with the estimator given in (2.6), $\widehat{d}^*$ does not require a preliminary estimator of d, nor to iterate (2.6) starting from $\overline{d}$, and attains the optimal semiparametric rate of convergence. Moreover, its limit distribution does not depend on any unknown parameter and it is more efficient than those in Robinson (1995a, b). However, the asymptotic variance of $\widehat{d}^*$ is greater than the estimator given in (2.6), although the difference of the asymptotic variance of $\widehat{d}^*$ and $\widehat{d}$ is small. As an example, choose $v(u) = u$ and $w(u) = u^{1/3}(1 - u^{1/6})$, which implies that $|h_w|\Phi \simeq 7/17$ instead of $\Phi_w = 1/3$.

We finish this section indicating that the results of the asymptotic distribution of both $\widehat{d}$ and $\widehat{d}^*$ does not depend on the location of the singularity of the spectral density $f(\lambda)$. To that end, suppose that model (1.1) is modified to

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

where $C \in (0, \infty)$ and $d \in (-1/2, 1/2)$.

Write

$$\widehat{f}_p = \widehat{f}(\lambda_p) = \frac{1}{m+1} \sum_{j=-m/2}^{m/2} I_{j+p+s}\mathcal{I}\left(\frac{m}{2} < p\right) + \frac{2}{m}\sum_{j=1}^{m/2} I_{j+p+s}\mathcal{I}\left(0 < p \le \frac{m}{2}\right)$$

and

$$\ddot{f}_p = \ddot{f}(\lambda_p) = \frac{1}{m_1+1} \sum_{j=-m_1/2; j+p\neq 0}^{m_1/2} I_{j+p+s}$$

with λ_s the closest Fourier frequency to λ^0 and define

$$\widehat{d}(\lambda_s) = \frac{h_w}{2}\left(\frac{1}{k}\sum_{p=1}^{k} w_p \log(\widehat{f}_p\widehat{f}_{-p}) - \overline{w}\log(\widehat{f}_{k+1}\widehat{f}_{-k-1})\right) - b(\overline{d})$$

$$\widehat{d}^*(\lambda_s) = \overline{v}^{-1}\frac{1}{2m}\sum_{p=1}^{m} v_p \widetilde{d}_p$$

with

$$\widetilde{d}_p = \frac{h_w}{2}\left(\frac{1}{p}\sum_{\ell=1}^{p} w_\ell \log(\ddot{f}_\ell \ddot{f}_{-\ell}) - \overline{w}_p \log(\ddot{f}_{p+1}\ddot{f}_{-p-1})\right),$$

and $\widehat{f}_p$ and $\ddot{f}_p$ defined as above. Let us introduce

A.1′. There exist $C \in (0, \infty)$, $d \in (-\frac{1}{2}, \frac{1}{2})$ and $\beta \in (0, 2]$ such that as $\lambda \to \lambda^0$

$$f(\lambda) = \begin{cases} C|\lambda - \lambda^0|^{-2d}(1 + O(|\lambda - \lambda^0|)) & \text{if } \beta \leq 1 \\ C|\lambda - \lambda^0|^{-2d}(1 + C|\lambda - \lambda^0| + O(|\lambda - \lambda^0|^{\beta})) & \text{if } 1 < \beta \leq 2. \end{cases}$$

A.3′. $\frac{\partial^{\varsigma}}{\partial \lambda^{\varsigma}} |\alpha(\lambda)| = O(|\lambda - \lambda^0|^{-\varsigma} |\alpha(\lambda)|)$ as $\lambda \to \lambda^0$, for $\varsigma = 1, 2$, and where $\alpha(\lambda) = \sum_{j=0}^{\infty} \alpha_j e^{ij\lambda}$.

Remark 2. Assumption A.1′ indicates that as the spectral density function does not need to be symmetric around $\lambda^0 \neq 0$ as is the case when $\lambda^0 = 0$, so that the approximation of $f(\lambda)$ by $C|\lambda - \lambda^0|^{-2d}$ cannot be better than $O(|\lambda - \lambda^0|)$, see also Hidalgo (2002).

COROLLARY 2.1. *Under the same conditions of Theorem* 2.2 *with* A.1′ *and* A.3′ *replacing* A.1 *and* A.3 *respectively*

$$\text{(a)} \quad m^{1/2}(\widehat{d}(\lambda_s) - d) \xrightarrow{d} N(0, \Phi_w^2)$$

and under the same conditions of Theorem 2.3 *with* A.1′ *and* A.3′ *replacing* A.1 *and* A.3

$$\text{(b)} \quad m^{1/2}(\widehat{d}^*(\lambda_s) - d) \xrightarrow{d} N(0, h_w^2 \Phi^2).$$

PROOF. The proof follows by identical arguments to those of Theorems 2.2 and 2.3, and thus it is omitted. □

3. Monte Carlo simulation

In this section we shall perform a Monte Carlo experiment to shed some light on the finite sample behaviour of our estimators (2.6) and (2.7) of the long range parameter d introduced in the previous section. Also, we will examine their relative performance compared with some previous semiparametric estimators of d.

In particular, the estimators $\widehat{d}$ and $\widehat{d}^*$ are compared with three other semiparametric estimators. Consider the estimator of the spectral distribution $F(\lambda)$,

$$\widehat{F}(\lambda) = \frac{2\pi}{n} \sum_{j=1}^{[n\lambda/2\pi]} I_j.$$

Thus, we can obtain an estimator of d by

$$\widehat{d}_{AVE} = \frac{1}{2} \left\{ 1 - \frac{\log\{\widehat{F}(q\lambda_m)/\widehat{F}(\lambda_m)\}}{\log q} \right\}$$

where $q \in (0, 1)$. See Robinson (1994) and Lobato and Robinson (1996) for the asymptotic properties of $\widehat{d}_{AVE}$. The second semiparametric estimator of d to be considered is the simple closed-form estimator

$$\widehat{d}_{LOG} = -\frac{1}{2}\frac{\sum_{j=1}^{m}\log I_j\left(\log j - \frac{1}{m}\sum_{\ell=1}^{m}\log \ell\right)}{\sum_{j=1}^{m}\log j\left(\log j - \frac{1}{m}\sum_{\ell=1}^{m}\log \ell\right)}.$$

This estimator is a slight modification of that of Geweke and Porter-Hudak (1983). Although, to provide asymptotic theory for $\widehat{d}_{LOG}$, as suggested by Künsch (1986) and Robinson (1995*a*), requires that an increasing number of frequencies $\lambda_1, \ldots, \lambda_\ell$ should be deleted, we have decided to keep them, on the grounds of easier comparability of $\widehat{d}_{LOG}$ with our estimators and the estimator proposed by Künsch (1987) and studied by Robinson (1995*b*), which we now introduce.

Table 1. Bias of the estimators.

Sample size		$n = 64$			$n = 128$			$n = 256$		
Bandwidth		$m = 4$	$m = 8$	$m = 16$	$m = 8$	$m = 16$	$m = 32$	$m = 16$	$m = 32$	$m = 64$
$d = -0.4$	AVE	−.015	−.120	−.170	−.023	−.110	−.190	−.040	−.100	−.200
	LOG	−.011	−.080	−.130	−.008	−.070	−.128	−.010	−.060	−.120
	GAU	.132	.018	−.048	−.048	−.022	−.071	−.006	−.044	−.081
	$\hat{d}_1$	.164	.091	.053	.121	.061	.035	.097	.047	.026
	$\hat{d}_2$	.196	.122	.080	.150	.089	.059	.123	.072	.048
	$\hat{d}_3$	.003	.019	−.003	.053	.047	.011	.108	.056	.004
	$\hat{d}_4$	.002	.014	−.010	.048	.039	.001	.100	.045	−.009
$d = -0.2$	AVE	−.042	−.032	.047	−.011	−.025	−.043	−.003	−.021	−.049
	LOG	.002	−.018	−.036	.003	−.013	−.032	.000	−.009	−.030
	GAU	.041	−.011	.050	−.006	−.028	−.044	−.016	−.026	−.037
	$\hat{d}_1$	.065	.041	.022	.053	.035	.018	.048	.034	.016
	$\hat{d}_2$	.083	.058	.037	.068	.049	.032	.061	.046	.028
	$\hat{d}_3$	−.075	−.010	−.008	.004	.025	.014	.052	.037	.015
	$\hat{d}_4$	−.076	−.014	−.013	.001	.021	.008	.048	.031	.009
$d = 0.0$	AVE	−.102	−.052	−.023	−.046	−.026	.010	−.020	−.012	−.005
	LOG	.004	−.004	−.001	.002	−.001	.000	.001	.002	.000
	GAU	−.031	−.026	−.020	−.022	−.020	−.011	−.022	−.013	−.005
	$\hat{d}_1$	−.047	−.031	−.018	−.030	−.018	−.009	−.017	−.009	−.005
	$\hat{d}_2$	−.042	−.028	−.017	−.028	−.017	−.009	−.016	−.008	−.004
	$\hat{d}_3$	−.133	−.055	−.030	−.054	−.026	−.013	−.022	−.013	−.007
	$\hat{d}_4$	−.134	−.056	−.030	−.054	−.027	−.013	−.023	−.014	−.007
$d = 0.2$	AVE	−.180	−.110	−.050	−.100	−.059	−.023	−.060	−.036	−.010
	LOG	.010	.006	.017	.008	.008	.017	.005	.009	.015
	GAU	−.099	−.044	−.004	−.041	−.017	.004	−.020	−.001	.009
	$\hat{d}_1$	−.156	−.099	−.049	−.104	−.061	−.022	−.066	−.035	−.006
	$\hat{d}_2$	−.164	−.112	−.063	−.116	−.074	−.036	−.079	−.048	−.020
	$\hat{d}_3$	−.176	−.099	−.051	−.103	−.072	−.034	−.086	−.051	−.020
	$\hat{d}_4$	−.175	−.097	−.047	−.101	−.068	−.028	−.082	−.045	−.012
$d = 0.4$	AVE	−.270	−.177	−.120	−.180	−.125	−.080	−.130	−.100	−.060
	LOG	.021	.019	.034	.021	.019	.032	.014	.017	.027
	GAU	−.181	−.084	−.019	−.088	−.030	.005	−.034	−.007	.016
	$\hat{d}_1$	−.258	−.157	−.066	−.163	−.084	−.010	−.090	−.029	.029
	$\hat{d}_2$	−.281	−.185	−.095	−.191	−.112	−.039	−.118	−.057	.002
	$\hat{d}_3$	−.222	−.143	−.069	−.148	−.104	−.038	−.132	−.064	−.006
	$\hat{d}_4$	−.221	−.138	−.061	−.144	−.095	−.026	−.122	−.051	.008

Consider the objective function

$$Q(C,d) = \frac{1}{m}\sum_{j=1}^{m}\{\log C\lambda_j^{-2d} + C^{-1}\lambda_j^{2d}I_j\}.$$

Then, Robinson's (1995*b*) estimator is defined as

$$\widehat{d}_{GAU} = \underset{d\in(-1/2,1/2)}{\arg\min}\left(\log\left\{\frac{1}{m}\sum_{j=1}^{m}\lambda_j^{2d}I_j\right\} - 2d\frac{1}{m}\sum_{j=1}^{m}\log\lambda_j\right).$$

Using an algorithm of Davies and Harte (1987) and the random generator *G05DDF* from the NAG library, Gaussian time series were generated with mean zero, variance unity and lag-j autocovariance

$$\gamma_j = \frac{1}{2}(|j+1|^{2d+1} - 2|j|^{2d+1} + |j-1|^{2d+1}).$$

Table 2. Standard deviation of the estimators.

Sample size		$n = 64$			$n = 128$			$n = 256$		
Bandwidth		$m = 4$	$m = 16$	$m = 32$	$m = 8$	$m = 16$	$m = 32$	$m = 16$	$m = 32$	$m = 64$
$d = -0.4$	AVE	.560	.407	.303	.382	.297	.216	.277	.209	.155
	LOG	.627	.350	.216	.350	.216	.139	.217	.140	.089
	GAU	.307	.173	.084	.190	.104	.051	.121	.071	.034
	$\hat{d}_1$	.141	.113	.086	.115	.086	.062	.084	.060	.043
	$\hat{d}_2$	.125	.104	.081	.104	.080	.058	.077	.056	.040
	$\hat{d}_3$	.149	.115	.086	.124	.091	.066	.084	.064	.046
	$\hat{d}_4$	.148	.114	.085	.123	.091	.066	.085	.065	.047
$d = -0.2$	AVE	.493	.340	.239	.325	.239	.168	.230	.164	.119
	LOG	.628	.341	.207	.350	.212	.135	.212	.134	.089
	GAU	.345	.239	.159	.244	.161	.109	.166	.109	.073
	$\hat{d}_1$	.139	.114	.088	.113	.085	.060	.081	.058	.041
	$\hat{d}_2$	.119	.101	.080	.100	.077	.056	.074	.053	.038
	$\hat{d}_3$	.209	.142	.106	.140	.094	.068	.080	.062	.046
	$\hat{d}_4$	.209	.143	.108	.141	.095	.070	.081	.063	.047
$d = 0.0$	AVE	.436	.284	.188	.276	.193	.128	.187	.130	.091
	LOG	.628	.340	.208	.348	.214	.135	.210	.134	.089
	GAU	.366	.270	.175	.268	.175	.111	.173	.110	.071
	$\hat{d}_1$	.141	.119	.094	.117	.088	.063	.084	.060	.043
	$\hat{d}_2$	.120	.105	.084	.103	.080	.058	.077	.056	.040
	$\hat{d}_3$	.247	.150	.111	.146	.096	.071	.081	.064	.048
	$\hat{d}_4$	.248	.152	.113	.148	.098	.073	.083	.066	.049
$d = 0.2$	AVE	.383	.232	.144	.228	.150	.094	.146	.099	.066
	LOG	.631	.342	.209	.348	.213	.135	.209	.134	.089
	GAU	.362	.266	.173	.264	.175	.111	.172	.109	.073
	$\hat{d}_1$	.151	.131	.104	.129	.100	.073	.086	.070	.051
	$\hat{d}_2$	.128	.116	.094	.114	.090	.067	.095	.065	.047
	$\hat{d}_3$	.268	.160	.121	.155	.105	.079	.089	.072	.054
	$\hat{d}_4$	.270	.163	.124	.158	.108	.081	.091	.075	.056
$d = 0.4$	AVE	.333	.183	.105	.178	.109	.065	.109	.065	.050
	LOG	.625	.343	.213	.350	.216	.135	.213	.137	.090
	GAU	.333	.222	.134	.221	.138	.091	.142	.094	.065
	$\hat{d}_1$	.162	.144	.114	.143	.112	.083	.110	.084	.059
	$\hat{d}_2$	.140	.131	.108	.129	.105	.081	.103	.081	.061
	$\hat{d}_3$	.262	.164	.124	.162	.117	.089	.105	.088	.066
	$\hat{d}_4$	.263	.165	.125	.164	.119	.090	.108	.089	.066

The corresponding spectral density function satisfies A.1 with $\beta = 2$. Five different values of d were employed, $d = -0.4$, -0.2, 0, 0.2 and 0.4. The sample sizes chosen were $n = 64$, 128 and 256. When the estimator $\widehat{d}$ given in (2.6) was employed, for each sample size, three different values of k were chosen: $n/16$, $n/8$ and $n/4$, while m was chosen as $k/\max\{1; [\log\log k]\}$ where $[z]$ means the integer part of z. Meanwhile for the estimator $\widehat{d}^*$ in (2.7), for each sample size three different values of m were chosen: $n/16$, $n/8$ and $n/4$, while $m_1 = m^{0.60}[\log\log m]$. Observe that this choice of m_1 satisfies A.8 with $r = 3$. For each (d, n, k) and (d, n, m) combination, 5000 replications were generated. In Tables 1–3, we have reported the *bias*, *standard deviation* and *M.S.E.* of the estimators used in the Monte Carlo experiment.

We should point out that although no asymptotic theory is available for $\widehat{d}_{AVE}$ when $d = -0.4$ or -0.2, or that the asymptotic distribution of $\widehat{d}_{AVE}$ when $d = 0.4$ is not Gaussian, see Lobato and Robinson (1996), (it follows a Rosenblatt distribution, see Taqqu (1975) for a definition), we have included it to gain some insight about its finite

Table 3. MSE of the estimators.

Sample size		$n = 64$			$n = 128$			$n = 256$		
Bandwidth		$m = 4$	$m = 8$	$m = 16$	$m = 8$	$m = 16$	$m = 32$	$m = 16$	$m = 32$	$m = 64$
$d = -0.4$	AVE	.314	.180	.121	.147	.100	.083	.078	.054	.064
	LOG	.393	.130	.064	.122	.051	.036	.047	.023	.022
	GAU	.117	.030	.009	.040	.011	.008	.015	.007	.007
	$\hat{d}_1$	.047	.021	.010	.028	.011	.005	.016	.006	.002
	$\hat{d}_2$	.054	.026	.013	.033	.014	.007	.021	.008	.004
	$\hat{d}_3$	.022	.013	.007	.018	.010	.004	.019	.007	.002
	$\hat{d}_4$	.022	.013	.007	.018	.010	.004	.017	.006	.002
$d = -0.2$	AVE	.245	.117	.059	.106	.058	.030	.053	.027	.017
	LOG	.394	.117	.044	.123	.045	.019	.045	.018	.008
	GAU	.121	.070	.028	.060	.027	.014	.028	.013	.007
	$\hat{d}_1$	.023	.015	.008	.016	.008	.004	.009	.004	.002
	$\hat{d}_2$	.021	.014	.008	.015	.008	.004	.009	.005	.002
	$\hat{d}_3$	.049	.020	.011	.020	.009	.005	.009	.005	.002
	$\hat{d}_4$	.049	.021	.012	.020	.009	.005	.009	.005	.002
$d = 0.0$	AVE	.200	.083	.036	.080	.038	.016	.035	.017	.008
	LOG	.394	.116	.043	.121	.046	.018	.044	.018	.008
	GAU	.135	.074	.031	.072	.031	.012	.030	.012	.005
	$\hat{d}_1$	.022	.015	.009	.015	.008	.004	.007	.004	.002
	$\hat{d}_2$	.016	.012	.007	.011	.007	.003	.006	.003	.002
	$\hat{d}_3$	.079	.026	.013	.024	.010	.005	.007	.004	.002
	$\hat{d}_4$	.080	.026	.014	.025	.010	.005	.007	.005	.002
$d = 0.2$	AVE	.180	.065	.023	.062	.026	.010	.025	.011	.004
	LOG	.398	.117	.044	.121	.045	.018	.044	.018	.008
	GAU	.141	.073	.030	.071	.031	.012	.030	.012	.005
	$\hat{d}_1$	.047	.027	.013	.027	.014	.006	.013	.006	.003
	$\hat{d}_2$	.043	.026	.013	.026	.014	.006	.014	.007	.003
	$\hat{d}_3$	.103	.035	.017	.035	.016	.007	.015	.008	.003
	$\hat{d}_4$	.103	.036	.018	.035	.016	.007	.015	.008	.003
$d = 0.4$	AVE	.183	.065	.025	.064	.028	.010	.029	.014	.006
	LOG	.391	.118	.046	.123	.047	.020	.050	.019	.008
	GAU	.176	.056	.018	.057	.020	.008	.021	.009	.004
	$\hat{d}_1$	.093	.045	.017	.047	.020	.007	.020	.008	.004
	$\hat{d}_2$	.099	.051	.021	.053	.024	.008	.024	.010	.004
	$\hat{d}_3$	.118	.047	.020	.048	.024	.009	.028	.012	.004
	$\hat{d}_4$	.118	.046	.019	.047	.023	.009	.027	.011	.004

sample statistical properties. For our estimator $\widehat{d}$, two different weights w_p have been used. Namely, $w_p = 1 - p/k$ and $w_p = 1 - (p/k)^{1/2}$. These would be denoted as $\widehat{d}_1$ and $\widehat{d}_2$ respectively in Tables 1–3, while for $\widehat{d}^*$, we have chosen the weights $w_\ell = (\ell/p)^{1/3} - (\ell/p)^{1/4}$ and $w_\ell = (\ell/p)^{1/3} - (\ell/p)^{1/2}$ and $v_p = p/m$. These estimators are denoted by $\widehat{d}_3$ and $\widehat{d}_4$ in the aforementioned tables. A word of caution is needed at this stage. Since the purpose of this Monte Carlo experiment is to gain some insight with regard to the finite sample performance of our estimators, for $\widehat{d}_1$ and $\widehat{d}_2$ we have used the "unfeasible" estimator $\widetilde{d} - b(d)$, that is, (2.6) but with $b(\cdot)$ evaluated at d instead of $\overline{d}$. Finally, in Tables 1–3, AVE, LOG and GAU will refer to the estimators $\widehat{d}_{AVE}$, $\widehat{d}_{LOG}$ and $\widehat{d}_{GAU}$, respectively.

From Table 1, we observe that for $\widehat{d}_i (i = 1, \ldots, 4)$, there is a negative bias if $d \geq 0$ while that bias turns positive for $d < 0$. It seems that $\widehat{d}_1$ and $\widehat{d}_2$ have a bias bigger than that for $\widehat{d}_3$ and $\widehat{d}_4$ and $\widehat{d}_{AVE}$, $\widehat{d}_{LOG}$ and $\widehat{d}_{GAU}$. The standard deviations in Table 2 decrease as both n and m or k increase, being the performance of $\widehat{d}_1$ and $\widehat{d}_2$ better than that of $\widehat{d}_3$ and $\widehat{d}_4$ as the results of Theorems 2.2 and 2.3 suggest. In almost all cases, $\widehat{d}_i (i = 1, \ldots, 4)$ are better than $\widehat{d}_{AVE}$, $\widehat{d}_{LOG}$ and $\widehat{d}_{GAU}$. The $M.S.E.$ in Table 3 shows similar results, being the overall picture from the Tables, that our estimators qualitatively tend to outperform $\widehat{d}_{AVE}$, $\widehat{d}_{LOG}$ and $\widehat{d}_{GAU}$, not only asymptotically but in small samples too. In many cases, the reduction of the $M.S.E.$ of $\widehat{d}_i (i = 1, \ldots, 4)$ is very substantial compared to the $M.S.E.$ of $\widehat{d}_{AVE}$, $\widehat{d}_{LOG}$ and $\widehat{d}_{GAU}$.

4. Proofs

4.1 *Proof of Proposition* 2.1

(a) By Robinson's (1995a) Theorem 2(a), and noting that we have f_j instead of its approximation $C\lambda_j^{-2d}$,

$$E(g_p) = \frac{K\widetilde{f}_p^{-1}}{(m+1)} \sum_{j=-m/2}^{m/2} \frac{\log(j+p)}{j+p} f_{j+p} = O(m^{-1}\log m),$$

since $(j+p)^{-1}\log(j+p) \leq Km^{-1}\log m$ for $m \leq p$ where henceforth K denotes a generic positive finite constant.

(b) By Robinson's (1995a) Theorem 2(a), $E(g_p)$ is bounded by

$$\frac{K}{m+1} \sum_{j=-m/2}^{m/2} \frac{\log(j+p+1)}{j+p} \frac{f_{j+p}}{\widetilde{f}_p} = O\left(\frac{\log^2 m}{m}\mathcal{I}(d \leq 0) + \frac{m^{2d}\log m}{(2p-m)^{2d}m}\mathcal{I}(d > 0)\right),$$

by Lemma 5.2 with $\psi = 1$ and $a = 1$ there and since by Lemma 5.4 $K^{-1} \leq \widetilde{f}_p\lambda_m^{2d} \leq K$.

(c) Let ℓ be a finite number $0 < \ell \leq m/2$. Because $p \leq m/2$, $E(g_p)$ is

$$\left(\sum_{j=1}^{m/2} f_{j+p}\right)^{-1} \left\{\sum_{j=1}^{\ell}(EI_{j+p} - f_{j+p}) + \sum_{j=\ell+1}^{m/2}(EI_{j+p} - f_{j+p})\right\}. \tag{4.1}$$

By Theorems 1 and 2 of Robinson (1995a), the first term of (4.1) is bounded by

$$K\left(\sum_{j=1}^{m/2} f_{j+p}\right)^{-1} \sum_{j=1}^{\ell} f_{j+p} = O\left(m^{2d-1}p^{-2d}\ell\mathcal{I}(d > 0) + \frac{\ell}{m}\mathcal{I}(d \leq 0)\right)$$

because $f_{j+p} \leq Kn^{2d}(p^{-2d}\mathcal{I}(d>0)+m^{-2d}\mathcal{I}(d \leq 0))$ and by Lemma 5.5, $K^{-1} \leq \widetilde{f}_p \lambda_m^{2d} \leq K$. The second term of (4.1) is, by Theorem 2(a) of Robinson (1995a),

$$K\left(\sum_{j=1}^{m/2} f_{j+p}\right)^{-1} \sum_{j=\ell+1}^{m/2} \frac{\log m}{j+p} f_{j+p} = O\left(\frac{\log m}{p^{2d}m^{1-2d}}\mathcal{I}(d>0) + \frac{\log^2 m}{m}\mathcal{I}(d \leq 0)\right),$$

using Lemma 5.3 with $\psi = 1$ and $a = 1$ for the second factor on the left of the last displayed equation and that Lemma 5.5 implies $K^{-1} \leq \widetilde{f}_p \lambda_m^{2d} \leq K$. Now conclude since ℓ is finite. □

4.2 *Proof of Proposition* 2.2

We begin estimating the covariance of the spectral density estimator. Writing $\phi_{j,p} = (m+1)^{-1}\mathcal{I}(m/2 < p) + 2/m\mathcal{I}(p \leq m/2)\mathcal{I}(j > 0)$,

$$\begin{aligned}(4.2) \qquad \operatorname{Cov}&\left(\sum_{j=-m/2}^{m/2} \phi_{j,p} I_{j+p}, \sum_{j=-m/2}^{m/2} \phi_{j,q} I_{j+q}\right) \\ &= \sum_{j,k=-m/2}^{m/2} \phi_{j,p}\phi_{k,q} \frac{1}{4\pi^2 n^2} \\ &\quad \times \sum_{t_1,t_2,s_1,s_2=1}^{n} \{\gamma(t_1-t_2)\gamma(s_1-s_2) + \gamma(t_1-s_2)\gamma(s_1-t_2) \\ &\qquad\qquad + \operatorname{cum}(x_{t_1}, x_{t_2}, x_{s_1}, x_{s_2})\} \\ &\quad \times \exp(-i(t_1-s_1)\lambda_{j+p} + i(t_2-s_2)\lambda_{k+q}).\end{aligned}$$

Because $\gamma(t) = \int_{-\pi}^{\pi} e^{itw} f(w)dw$, by an obvious change of variables, the first term on the right of (4.2) is

$$\begin{aligned}(4.3) \quad \frac{1}{4\pi^2 n^2} \sum_{j,k=-m/2}^{m/2} \phi_{j,p}\phi_{k,q} &\left\{\left(\int H(\theta_1)H(\lambda_{k+q-j-p} - \theta_1) f(\theta_1 + \lambda_{j+p}) d\theta_1\right)\right. \\ &\left.\times \left(\int H(-\theta_2)H(\theta_2 + \lambda_{j+p-k-q}) f(\theta_2 - \lambda_{k+q}) d\theta_2\right)\right\},\end{aligned}$$

where $H(\theta) = \sum_{\ell=1}^{n} e^{i\ell\theta}$ is the Dirichlet's kernel. We examine the first factor inside the braces of (4.3), the second being identical. Adding and subtracting $f_{j+p}\int H(\theta_1)H(\lambda_{k+q-j-p} - \theta_1)d\theta_1$, that factor is

$$\begin{aligned}(4.4) \qquad &\int H(\theta_1)H(\lambda_{k+q-j-p} - \theta_1)(f(\theta_1 + \lambda_{j+p}) - f_{j+p})d\theta_1 \\ &\quad + f_{j+p}\int H(\theta_1)H(\lambda_{k+q-j-p} - \theta_1)d\theta_1.\end{aligned}$$

By Theorem 2 part (c) of Robinson (1995a), the absolute value of the first term of (4.4) is bounded by

$$Knf_{j+p}\frac{\max(\log(k+q), \log(j+p))}{\min((j+p),(k+q))},$$

whereas the second term of (4.4) is zero unless $(k+q)=(j+p)$, in which case is $2\pi n f_{j+p}$. Hence, as $n\to\infty$, (4.3) and thus the first term on the right of (4.2) is bounded in absolute value by

$$(4.5)\qquad K\sum_{j,k=-m/2}^{m/2}\phi_{j,p}\phi_{k,q}f_{j+p}f_{k+q}\left(\frac{\max(\log(k+q),\log(j+p))}{\min((j+p),(k+q))}\right)^2 + K\sum_{j+p=k+q}\phi_{j,p}\phi_{k,q}f_{j+p}f_{k+q}\left(\frac{\log(j+p)}{j+p}+1\right)\mathcal{I}(q-p\le m),$$

where $\sum_{j+p=k+q}$ means the terms in $\sum_{j,k=-m/2}^{m/2}$ such that $j+p=k+q$.

By the same arguments, the second term on the right of (4.2) is in absolute value bounded by

$$(4.6)\qquad K\sum_{j,k=-m/2}^{m/2}\phi_{j,p}\phi_{k,q}f_{j+p}f_{k+q}\left(\frac{\max(\log(k+q),\log(j+p))}{\min((j+p),(k+q))}\right)^2 + K\sum_{j+p=-(k+q)}\phi_{j,p}\phi_{k,q}f_{j+p}f_{k+q}\left(\frac{\log(j+p)}{j+p}+1\right)\mathcal{I}(q+p\le m),$$

where $\sum_{j+p=-(k+q)}$ denotes the terms in the double sum such that $j+p=-(k+q)$.

Next, we examine the contribution from the third term on the right of (4.2). Applying formulae of Brillinger (1981, (2.6.3) page 26, and (2.10.3) page 39), the contribution is bounded in absolute value by

$$(4.7)\qquad \frac{K}{n^2}\sum_{j,k=-m/2}^{m/2}\phi_{j,p}\phi_{k,q}\left|\int_{[-\pi,\pi]^3}\alpha(-\lambda)\alpha(-\mu)\alpha(-\zeta)\alpha(\lambda+\mu+\zeta)\times H(\lambda+\lambda_{j+p})H(\mu-\lambda_{j+p})H(\zeta-\lambda_{k+q})\times H(\lambda_{k+q}-(\lambda+\mu+\zeta))d\lambda d\mu d\zeta\right|$$

$$=\frac{K}{n^2}\sum_{j,k=-m/2}^{m/2}\phi_{j,p}\phi_{k,q}|\alpha(\lambda_{j+p})\alpha(-\lambda_{j+p})\alpha(\lambda_{k+q})\alpha(-\lambda_{k+q})| \times\left|\int_{[-\pi,\pi]^3}\left(\frac{\alpha(-\lambda)\alpha(-\mu)\alpha(-\zeta)\alpha(\lambda+\mu+\zeta)}{\alpha(\lambda_{j+p})\alpha(-\lambda_{j+p})\alpha(\lambda_{k+q})\alpha(-\lambda_{k+q})}\times H(\lambda+\lambda_{j+p})H(\mu-\lambda_{j+p})H(\zeta-\lambda_{k+q})\times H(\lambda_{k+q}-(\lambda+\mu+\zeta))\right)d\lambda d\mu d\zeta\right|,$$

where we note that since $f(\lambda)>0$ for $\lambda>0$, it implies that $|\alpha(\lambda_{j+p})|$ and $|\alpha(\lambda_{k+q})|>0$ since $(j+p)$ and $(k+q)>0$. Because

$$\int_{[-\pi,\pi]^3}H(\lambda+\lambda_{j+p})H(\mu-\lambda_{j+p})H(\zeta-\lambda_{k+q})H(\lambda_{k+q}-(\lambda+\mu+\zeta))d\lambda d\mu d\zeta$$

is equal to $(2\pi)^3 n$, using the identity

$$c_1c_2c_3c_4 = (c_1c_2 - 1)(c_3c_4 - 1) + \sum_{\ell=1}^{2}(c_{2\ell-1} - 1)(c_{2\ell} - 1) + \sum_{i=1}^{4}(c_i - 1) + 1,$$

and proceeding as in the proof of (4.31) in Robinson (1995b), the right side of (4.7) is bounded by

$$K \sum_{j,k=-m/2}^{m/2} \phi_{j,p}\phi_{k,q}|\alpha(\lambda_{j+p})\alpha(-\lambda_{j+p})\alpha(\lambda_{k+q})\alpha(-\lambda_{k+q})| \times \left(\frac{1}{n} + \frac{1}{|j+p||k+q|} + |j+p|^{-1/2}|k+q|^{-1} + |j+p|^{-1}|k+q|^{-1/2} + n^{-1/2}|j+p|^{-1/2}|k+q|^{-1/2}\right). \tag{4.8}$$

Thus,

$$(4.2) = (4.8) + (4.5) + (4.6). \tag{4.9}$$

With these preliminaries, let us examine the covariance of the spectral density estimator.

We begin with part (a). The first term on the right of (4.9) is $O((n^{-1} + p^{-1}q^{-1/2})f_pf_q)$ by Lemma 5.1 with $\psi = 0$, $a = 1$ and $b = 0$ there and observing that $(j+p) > p/2$ for $p \geq m$. Next the second term on the right of (4.9), that is (4.5). Proceeding as with (4.8), the first term of (4.5) is $O(f_pf_qp^{-2}\log^2 q)$, whereas the second term of (4.5) is zero unless $q - p \leq m$, in which case is bounded by

$$\frac{K}{(m+1)^2}\sum_{j=-m/2}^{m/2} f_{j+p}^2\mathcal{I}(q-p\leq m) = O(m^{-1}f_p^2\mathcal{I}(q-p\leq m))$$

since $(j+p)^{-1}\log(j+p) \leq K$ and by Lemma 5.1 with $\psi = 0$, $a = b = 1$ and $p = q$ there. Finally, proceeding as with the second term of (4.9), the third term on the right of (4.9) is $O(f_pf_qp^{-2}\log^2 q)$ after observing that because $p + q \geq m$, the contribution of the second term of (4.6) is zero.

Thus, because for $q, p \geq m$, $K^{-1} \leq f_p^{-1}\widetilde{f}_p - 1 \leq K$ by Lemma 5.1, when $q - p > m$,

$$\mathrm{Cov}(g_p, g_q) = O(p^{-2}\log^2 q + n^{-1} + p^{-1}q^{-1/2}) = O(n^{-1} + p^{-1}q^{-1/2} + p^{-2}\log^2 q),$$

whereas if $q \pm p \leq m$

$$\mathrm{Cov}(g_p, g_q) = O(p^{-2}\log^2 q + n^{-1} + p^{-1}q^{-1/2} + m^{-1}) = O(m^{-1}).$$

This concludes the proof of part (a).

Next (b). The first term on the right of (4.9) has three typical components, namely,

$$\sum_{j=-m/2}^{m/2}\phi_{j,p}f_{j+p}, \qquad \sum_{j=-m/2}^{m/2}\phi_{j,p}f_{j+p}|j+p|^{-1/2} \quad \text{and} \quad \sum_{j=-m/2}^{m/2}\phi_{j,p}f_{j+p}|j+p|^{-1}.$$

So by Lemma 5.2 with $a = 1$, for $d > 1/4$, the first term on the right of (4.9) is λ_m^{-4d} times

$$O(n^{-1} + n^{-1/2}m^{4d-2}(2p-m)^{1-4d} + m^{4d-2}(2p-m)^{-4d} + m^{4d-2}(2p-m)^{1/2-4d}).$$

When $d = 1/4$, it is λ_m^{-4d} times

$$O\left(\frac{1}{n} + \frac{1}{n^{1/2}m}\log^2\left(\frac{2p+m}{2p-m}\right) + \frac{1}{m}\left(\log\left(\frac{2p+m}{2p-m}\right)\frac{1}{(2p-m)^{1/2}} + \frac{1}{(2p-m)}\right)\right)$$

whereas for $d < 1/4$, it is $o(m^{-1}\lambda_m^{-4d})$. So, the first term on the right of (4.9) is

$$\begin{aligned}(4.10)\qquad & O\left(\frac{\lambda_m^{-4d}}{n} + \frac{\lambda_m^{-4d}}{m^{2-4d}}(2p-m)^{1-4d}\right)\mathcal{I}(d > 1/4) + o\left(\frac{\lambda_m^{-4d}}{m}\right)\mathcal{I}(d < 1/4) \\ & + O\left(\frac{\lambda_m^{-4d}}{n} + \frac{\lambda_m^{-4d}}{m}\log\left(\frac{2p+m}{2p-m}\right)(2p-m)^{-1/2}\right)\mathcal{I}(d = 1/4).\end{aligned}$$

Next, the second term on the right of (4.9), that is (4.5), whose first term, by Lemma 5.2 with $\psi = 2$ and $a = 1$ there, is λ_m^{-4d} times

$$\begin{aligned}& O\left(\frac{\log^2 m}{m^2}\mathcal{I}(d < 0) + \frac{\log^4 m}{m^2}\mathcal{I}(d = 0)\right) + O(m^{4d-2}(2p-m)^{-4d}\mathcal{I}(d > 0)) \\ & \quad + O(m^{2d-1}(2p-m)^{-2d-1}\log^2 m\mathcal{I}(d > 0)) \\ & = O(m^{4d-2}(2p-m)^{-4d}\mathcal{I}(d > 1/4)) + o(m^{-1}) \\ & \quad + O(m^{2d-1}(2p-m)^{-2d-1}\log^2 m\mathcal{I}(d \geq 0))\end{aligned}$$

using for the second term on the left that $x^{-b}\log x < K$ for $x > 1$ and $b > 0$. The second term of (4.5) is bounded by $K\sum_{j=-m/2}^{m/2}\phi_{j,p}^2 f_{j+p}^2$ which is λ_m^{-4d} times

$$(4.11)\qquad O\left(\frac{\mathcal{I}(d < 1/4)}{m} + \log\left(\frac{2p+m}{2p-m}\right)\frac{\mathcal{I}(d = 1/4)}{m} + (2p-m)^{1-4d}\frac{\mathcal{I}(d > 1/4)}{m^{2-4d}}\right).$$

Finally, the third term on the right of (4.9), proceeding as with the second term, is λ_m^{-4d} times

$$(4.12)\qquad O(m^{2d-1}(2p-m)^{-2d-1}\log^2 m\mathcal{I}(d > 0)) + (4.11)$$

Then, gathering (4.10)–(4.12) and that by Lemma 5.4, $K^{-1} \leq \lambda_m^{2d}\widetilde{f}_p \leq K$,

$$\begin{aligned} m\operatorname{Cov}(g_p, g_q) = {} & O\left(\mathcal{I}\left(d < \frac{1}{4}\right) + \log\left(\frac{2p+m}{2p-m}\right)\mathcal{I}\left(d = \frac{1}{4}\right)\right) \\ & + O\left(\left(\frac{m}{2p-m}\right)^{4d-1}\mathcal{I}\left(d > \frac{1}{4}\right)\right) + O\left(\frac{m^{2d}\log^2 m}{(2p-m)^{1+2d}}\mathcal{I}(d \geq 0)\right),\end{aligned}$$

which concludes the proof of part (b).

Finally, part (c) follows by identical arguments to those of part (b) but using Lemmas 5.3 and 5.5 instead of Lemmas 5.2 and 5.4 respectively. □

4.3 *Proof of Proposition* 2.3

From Hidalgo and Robinson's (2002) Proposition A.1 and Proposition 2.1, it suffices to examine $\widetilde{f}_p^{-1}(\widehat{f}_p - E\widehat{f}_p)$. On the other hand, Hidalgo and Robinson's (2002) Proposition A.3 part (a,b) implies that it suffices to examine the behaviour of $\widetilde{f}_p^{-1}(\widehat{f}_{e,p} - E\widehat{f}_{e,p})$, where

$$\widehat{f}_{e,p} = \widehat{f}_e(\lambda_p) = \frac{1}{m+1}\sum_{j=-m/2}^{m/2} f_{j+p} I_{e,j+p}\mathcal{I}\left(\frac{m}{2} < p\right) + \frac{2}{m}\sum_{j=1}^{m/2} f_{j+p} I_{e,j+p}\mathcal{I}\left(0 < p \le \frac{m}{2}\right)$$

and $I_{e,p} = I_e(\lambda_p)$ denotes the periodogram of e_t. We only examine $\sup_{p=1+m,\ldots,k} \times |\widetilde{f}_p^{-1}(\widehat{f}_{e,p} - E\widehat{f}_{e,p})|^2$, being $\sup_{p=1,\ldots,m} |\widetilde{f}_p^{-1}(\widehat{f}_{e,p} - E\widehat{f}_{e,p})|^2$ similarly handled.

Because $\sup_j |a_j|^2 = (\sup_j |a_j|^r)^{2/r}$, $(\sup_{p=1+m,\ldots,k} |\widetilde{f}_p^{-1}(\widehat{f}_{e,p} - E\widehat{f}_{e,p})|^2)^{r/2}$ is

$$\sup_{p=1+m,\ldots,k}\left|\frac{1}{m+1}\sum_{j=-m/2}^{m/2}\phi_{j+p,p}((2\pi)I_{e,j+p} - 1)\right|^r,$$

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

$$\begin{aligned}(4.13)\qquad & 2^{r-1}\sup_q\sup_p\left|\left(\frac{1}{m+1}\sum_{j=-m/2}^{m/2}(\phi_{j+p,p}((2\pi)I_{e,j+p} - 1)\right.\right. \\ &\qquad\left.\left. - \phi_{j+s,p}((2\pi)I_{e,j+s} - 1))\right)\right|^r \\ &+ 2^{r-1}\sup_q\sup_p\left|\left(\frac{1}{m+1}\sum_{j=-m/2}^{m/2}\phi_{j+s,p}((2\pi)I_{e,j+s} - 1)\right)\right|^r,\end{aligned}$$

where $\sup_q$ and $\sup_p$ mean $\sup_{q=1+m^{(r-1)/r},\ldots,k/m^{1/r}}$ and $\sup_{p=1+s-m^{1/r},\ldots,s}$ respectively, and $s = qm^{1/r}$.

The second term of (4.13) is bounded by

$$\begin{aligned}(4.14)\qquad & K\sup_q\sup_p\left|\left(\frac{1}{m+1}\sum_{j=0}^{m}\phi_{j+s-m/2,p}((2\pi)I_{e,j+s-m/2} - 1)\right)\right|^r \\ &\le K\sup_q\sup_p\left|\frac{1}{m+1}\sum_{j=0}^{m-1}(\phi_{j+s-m/2,p} - \phi_{j+s+1-m/2,p})\right. \\ &\qquad\left.\times\left(\sum_{\ell=0}^{j}((2\pi)I_{e,\ell+s-m/2} - 1)\right)\right|^r \\ &+ K\sup_q\sup_p|\phi_{j+s+m/2,p}|^r\left|\frac{1}{m+1}\sum_{j=0}^{m}((2\pi)I_{e,j+s-m/2} - 1)\right|^r,\end{aligned}$$

by Abel summation by parts. Now by A.1 and A.3 and that Lemma 5.1 implies that $K^{-1} < |f_p^{-1}\widetilde{f}_p| < K$, we have that

$$|\phi_{j+s-m/2,p} - \phi_{j+s+1-m/2,p}| \le K\widetilde{f}_p^{-1}(j+s-m/2)^{-1-2d}n^{2d}$$

$$\leq K\left(\frac{p}{j+s-m/2}\right)^{2d}(j+s-m/2)^{-1}.$$

So, using that $\sup_\ell a_\ell^2 \leq \sum_\ell a_\ell^2$ the right side of (4.14) is bounded by

$$\begin{aligned}
&\frac{K}{m+1}\sum_{q=1+m^{(r-1)/r}}^{k/m^{1/r}}\sum_{j=0}^{m}\sup_p\left(\frac{p}{j+s-m/2}\right)^{2rd}(j+s-m/2)^{-r}\\
&\qquad\times\left|\sum_{\ell=0}^{j}((2\pi)I_{e,\ell+s-m/2}-1)\right|^r\\
&\qquad+K\sum_{q=1+m^{(r-1)/r}}^{k/m^{1/r}}\left|\frac{1}{m+1}\sum_{j=0}^{m}((2\pi)I_{e,j+s-m/2}-1)\right|^r
\end{aligned}$$

whose expectation, proceeding as in the proof of Brillinger's (1981) Theorem 7.4.4, is bounded by

$$\begin{aligned}
&K\sum_{q=1+m^{(r-1)/r}}^{k/m^{1/r}}\left(\left(\frac{s}{s-m/2}\right)^{2rd}\frac{1}{m}\sum_{j=1}^{m}\frac{j^{r/2}}{(j+s-m/2)^r}+m^{-r/2}\right)\\
&\qquad\leq K\sum_{q=1+m^{(r-1)/r}}^{k/m^{1/r}}\left(\left(\frac{1}{s-m/2}\right)^{r/2}+m^{-r/2}\right)=O\left(\frac{k}{m^{r/2+1/r}}\right),
\end{aligned}$$

because $d<1/2$ and $s\leq 2(s-m/2)$ since $s\leq 2(s-m/2)$ and $q\geq 1+m^{(r-1)/r}$. Thus, we conclude that the second term of (4.13) is $O(k/m^{r/2+1/r})$.

Next, we examine the first term of (4.13). Because

$$a_{p,m}=\sum_{j=0}^{m}(\phi_{j+p-m/2,p}((2\pi)I_{e,j+p-m/2}-1)-\phi_{j+s-m/2,p}((2\pi)I_{e,j+s-m/2}-1))$$

has at most $m^{1/r}/2$ terms, each of which is $O_p(1)$ uniformly in j by Chen and Hannan (1980), that term is bounded by

$$\frac{K}{m}\sum_{q=1+m^{(r-1)/r}}^{k/m^{1/r}}\sup_p\left|\frac{a_{p,m}}{(m+1)^{(r-1)/r}}\right|^r=O_p\left(\frac{1}{m^{r-1}}\sum_{q=1+m^{(r-1)/r}}^{k/m^{1/r}}1\right)=O_p\left(\frac{k}{m^{r/2+1/r}}\right).$$

Thus, to conclude the proof, we are left with $\sup_{p=1,\ldots,m}|\widetilde{f}_p^{-1}(\widehat{f}_{e,p}-E\widehat{f}_{e,p})|^2$. Proceeding as with $\sup_{p=1+m,\ldots,k}|\widetilde{f}_p^{-1}(\widehat{f}_{e,p}-E\widehat{f}_{e,p})|^2$ and observing that $\sup_q$ runs, in this case, for $q=1,\ldots,m^{(r-1)/r}$, the only term which is slightly different is the one corresponding to (4.14), whose expectation is bounded by

$$\frac{K}{m+1}\sum_{q=1}^{m^{(r-1)/r}}\sum_{j=1}^{m}\sup_p\left(\frac{1}{j+p+1-m/2}\right)^r E\left|\sum_{\ell=1}^{j}((2\pi)I_{e,\ell+s-m/2}-1)\right|^r\mathcal{I}\left(p>\frac{m}{2}\right)$$

$$+\frac{K}{m+1}\sum_{q=1}^{m^{(r-1)/r}}\sum_{j=1}^{m/2}\sup_p\left(\frac{1}{j+p+1}\right)^r E\left|\sum_{\ell=1}^{j}((2\pi)I_{e,\ell+s}-1)\right|^r \mathcal{I}\left(p\le\frac{m}{2}\right)$$

$$\le\frac{K}{m+1}\sum_{q=1}^{m^{(r-1)/r}}\sum_{j=1}^{m}j^{-r/2}=O(m^{-1/r}\log m),$$

using Lemmas 5.2 and 5.3 instead of Lemma 5.1 and $r\ge 2$. (Observe that Lemmas 5.1 and 5.2 coincide for $p=m$ there.) Thus

$$\sup_{p=1,\ldots,k}|\widetilde{f}_p^{-1}(\widehat{f}_p-\widetilde{f}_p)|=O_p\left(\frac{k^{1/r}}{m^{1/2+1/r^2}}\right)+o_p(1). \qquad \square$$

4.4 *Proof of Theorem* 2.1

Write

$$g_p=\left(\frac{1}{m+1}\sum_{j=-m/2}^{m/2}\left(\frac{f_{p+j}}{f_p}\right)\right)^{-1}\bar{g}_p;\qquad \bar{g}_p=\frac{1}{m+1}\sum_{j=-m/2}^{m/2}\left(\frac{f_{p+j}}{f_p}\right)\left\{\frac{I_{p+j}}{f_{p+j}}-1\right\}.$$

Because uniformly in $j=-m/2,\ldots,m/2$, $f_p^{-1}f_{p+j}\to 1$ by A.1, we have that when $m/p\to 0$, by a direct use of Toeplitz's Lemma (see Stout (1974)) and Proposition 2.2 part (a), the limiting distribution of $m^{1/2}g_p$ is that of $m^{1/2}\bar{g}_p$ where

$$m^{1/2}\bar{g}_p=\frac{m^{1/2}}{m+1}\sum_{j=-m/2}^{m/2}\left\{\frac{I_{p+j}}{f_{p+j}}-1\right\}.$$

But, by an obvious extension of Robinson's (1995*b*) Theorem 2, $m^{1/2}\bar{g}_p\xrightarrow{d}N(0,1)$. $\square$

4.5 *Proof of Theorem* 2.2

The theorem is proved if

$$\text{(a)}\quad m^{1/2}(\widetilde{d}-b(d)-d)\xrightarrow{d}N(0,\Phi_w^2)\qquad\text{and}\qquad\text{(b)}\quad m^{1/2}(b(d)-b(\bar{d}))\xrightarrow{p}0.$$

We begin with (a). By definition $\widetilde{d}-b(d)-d=h_w(a_1+a_2+a_3)$ where

$$a_1=\frac{1}{k}\sum_{p=1}^{k}w_p(\log\widehat{f}_p/\widetilde{f}_p)-\overline{w}(\log\widehat{f}_{k+1}/\widetilde{f}_{k+1})$$

$$a_2=\frac{1}{k}\sum_{p=1}^{k}w_p(\log\widetilde{f}_p/\overline{f}_p)-\overline{w}(\log\widetilde{f}_{k+1}/\overline{f}_{k+1})$$

$$a_3=\frac{1}{k}\sum_{p=1}^{k}w_p\log(f_p^*/f_{k+1}^*)-h_w^{-1}d.$$

So, it suffices to show that $m^{1/2}a_1\xrightarrow{d}N(0,\Phi_w^2)$, $a_2=o(m^{-1/2})$ and $a_3=o(m^{-1/2})$.

We begin showing that $a_3 = o(m^{-1/2})$. By the definition of f_p^* and h_w,

$$a_3 = -2d\left(\frac{1}{k}\sum_{p=1}^{k} w_p \log\left(\frac{p}{k+1}\right) - \int_0^1 w(u)(\log u)du\right).$$

By Lemma 2 of Robinson (1995b), $k^{-1}\sum_{p=1}^{k}\log(p/k)+1 = O(k^{-1}\log k)$, whereas since A.6 implies that $W(u) = (w(u)-1)\log u$ has an integrable derivative, $k^{-1}\sum_{p=1}^{k}(w_p - 1)\log(p/k) - \int_0^1 W(u)du = O(k^{-1})$ by Brillinger ((1981), p. 15). Thus, by A.5, $a_3 = O(k^{-1}\log k) = o(m^{-1/2})$.

Next, we examine $a_2 = a_{21} + a_{22}$, where

$$a_{21} = \frac{1}{k}\sum_{p=1}^{m/2} w_p(\log \widetilde{f}_p/\overline{f}_p)$$

$$a_{22} = \frac{1}{k}\sum_{p=m/2+1}^{k} w_p(\log \widetilde{f}_p/\overline{f}_p) - \left(\frac{1}{k}\sum_{p=1}^{k} w_p\right)(\log \widetilde{f}_{k+1}/\overline{f}_{k+1}).$$

By A.1, if $m/2 < p \le (k+1) = o(n)$

$$\overline{f}_p^{-1}\widetilde{f}_p - 1 = \overline{f}_p^{-1}\left(\frac{C}{m+1}\sum_{j=-m/2}^{m/2}\lambda_{j+p}^{-2d}\right) - 1 + K\overline{f}_p^{-1}\frac{1}{m+1}\sum_{j=-m/2}^{m/2}\lambda_{j+p}^{-2d+\beta}$$

$$= K\overline{f}_p^{-1}\frac{1}{m+1}\sum_{j=-m/2}^{m/2}\lambda_{j+p}^{-2d+\beta} \le K\lambda_{p+m/2}^{\beta} = O(\lambda_{2p}^{\beta}) = O(\lambda_{2k}^{\beta}).$$

Thus, by the mean value theorem,

$$\left|\log\left(\frac{\widetilde{f}_p}{\overline{f}_p} - 1 + 1\right)\right| \le \left|\frac{\widetilde{f}_p}{\overline{f}_p} - 1\right|\left|\frac{1}{1+\xi(\overline{f}_p^{-1}\widetilde{f}_p - 1)}\right|$$

for some $\xi = \xi(p)$ with $|\xi| < 1$, which implies that $m^{1/2}a_{22} = O(m^{1/2}\lambda_{2k}^{\beta}) = o(1)$ by A.5.

For $1 \le p \le m/2$,

$$\overline{f}_p^{-1}\widetilde{f}_p - 1 = \overline{f}_p^{-1}\frac{1}{m}\sum_{j=1}^{m/2}(f_{j+p} - C\lambda_{j+p}^{-2d}) \le K\overline{f}_p^{-1}\frac{1}{m}\sum_{j=1}^{m/2}\lambda_{j+p}^{-2d+\beta} = O(\lambda_m^{\beta}).$$

Then, proceeding as with a_{22}, we conclude that $m^{1/2}a_{21} = O(m^{3/2}\lambda_m^{\beta}/k) = o(m^{1/2}\lambda_m^{\beta})$ by A.5, which implies that $m^{1/2}a_2 = o(1)$.

To complete the proof of part (a), we need to show that $m^{1/2}a_1 \xrightarrow{d} N(0,\Phi_w^2)$. To this end, it suffices to show that

$$b_k = m^{1/2}\left(\frac{1}{k}\sum_{p=1}^{k} w_p g_p - \overline{w}g_{k+1}\right) \xrightarrow{d} N(0, h_w^{-2}\Phi_w^2) \tag{4.15}$$

and

$$m^{1/2}a_1 - b_k = o_p(1). \tag{4.16}$$

We show (4.15) first. Write $c_k = (m^{1/2}/k)\sum_{p=1}^{k} w_p g_p$. Then,

$$E(c_k) = \frac{m^{1/2}}{k}\left\{\sum_{p=1}^{m/2} w_p E(g_p) + \sum_{p=1+m/2}^{m-1} w_p E(g_p) + \sum_{p=m}^{k} w_p E(g_p)\right\}.$$

By Proposition 2.1 part (c), the first term on the right of the last displayed equation is

$$O\left(\frac{1}{k}\sum_{p=1}^{m/2}\left(m^{2d-1/2}\frac{\log m}{p^{2d}}\mathcal{I}(d>0) + \frac{\log^2 m}{m^{1/2}}\mathcal{I}(d\le 0)\right)\right) = O\left(\frac{m^{1/2}\log^2 m}{k}\right)$$

which is $o(1)$ by A.5, whereas Proposition 2.1 part (b) implies that the second term is bounded by

$$K\frac{m^{1/2}}{k}\sum_{p=1+m/2}^{m-1}\left(\frac{\log^2 m}{m}\mathcal{I}(d\le 0) + \frac{\log m\mathcal{I}(d>0)}{m^{1-2d}(2p-m)^{2d}}\right) = O\left(\frac{m^{1/2}\log^2 m}{k}\right).$$

Finally, the third term is $O(m^{-1/2}\log m) = o(1)$ by Proposition 2.1 part (a) and A.5.

Next, the variance of c_k is bounded by

$$mK\left(\operatorname{Var}\left(\frac{1}{k}\sum_{p=1}^{m/2} w_p g_p\right) + \operatorname{Var}\left(\frac{1}{k}\sum_{p=1+m/2}^{m-1} w_p g_p\right) + \operatorname{Var}\left(\frac{1}{k}\sum_{p=m}^{k} w_p g_p\right)\right). \tag{4.17}$$

The third term of (4.17) is

$$m\left\{\frac{K}{k^2}\sum_{p=m}^{k} w_p^2 \operatorname{Var}(g_p) + \frac{2K}{k^2}\sum_{m\le p<q}^{k} w_p w_q \operatorname{Cov}(g_p, g_q)\right\}$$
$$= O\left(\frac{1}{k}\right) + \frac{Km}{k^2}\sum_{m\le p<q}^{k}\operatorname{Cov}(g_p, g_q)$$

because Proposition 2.2 part (a) implies that $m\operatorname{Var}(g_p) = O(1)$ for $p \ge m$. The second term is bounded in absolute value by

$$\frac{Km}{k^2}\left(\sum_{m\le p<q,|q-p|\le m}^{k}\operatorname{Var}^{1/2}(g_p)\operatorname{Var}^{1/2}(g_q) + \sum_{m\le p<q,|q-p|>m}^{k}|\operatorname{Cov}(g_p, g_q)|\right) = o(1)$$

as we now show. By Proposition 2.2 part (a), the second term on the left is $O(k^{-1}m\sum_{p=m}^{k}(n^{-1}+m^{-3/2})) = o(1)$ because $p, q > m$ implies $p^{-1}q^{-1/2} < m^{-3/2}$, whereas the first term is also $o(1)$ because the sum $\sum_{m\le p<q,|q-p|\le m}^{k}$ has at most km

terms and by Proposition 2.2 part (a) and by A.5, $\mathrm{Var}(g_p) = O(m^{-1})$ and $m/k \to 0$ respectively. So, we conclude that the third term of (4.17) is $o(1)$.

Next, the second term of (4.17) is bounded by

$$K\frac{m^2}{k^2}\sum_{p=1+m/2}^{m-1}\mathrm{Var}(g_p) = o(1),$$

since by Proposition 2.2 part (b), it is $O(k^{-2}(m^2+m^{1+2d}\log^2 m)) = o(1)$ if $d \in (-\frac{1}{2}, \frac{1}{4})$, for $d = 1/4$ using that $|m^{-1}\sum_{p=1}^{3m/2}\log((p+m)/p)| \to |\int_0^{3/2}\log(v^{-1}(v+1))dv| \le K$, and for $d \in (\frac{1}{4}, \frac{1}{2})$, since

$$O\left(\frac{m^2}{k^2}\sum_{p=1}^{3m/2}\frac{m^{4d-2}}{p^{4d-1}} + \frac{m^{1+2d}\log^2 m}{k^2}\right) = O\left(\frac{m^2+m^{1+2d}\log^2 m}{k^2}\right) = o(1),$$

using that $m^{-1}\sum_{p=1}^{m-1}(p/m)^{1-4d} \le K$. Finally, the first term of (4.17) is also $o(1)$ using Proposition 2.2 part (c) instead of part (b). Thus, we obtain that

$$(4.15) = -\left(\frac{1}{k}\sum_{p=1}^{k} w_p\right) m^{1/2} g_{k+1} + o_p(1) \xrightarrow{d} N(0, h_w^{-2}\Phi_w^2)$$

since by Theorem 2.1 the first term on the right converges in distribution to a $N(0, h_w^{-2}\Phi_w^2)$.

Next we show (4.16). First, for any arbitrary $\varepsilon, \eta > 0$,

$$\begin{aligned}\Pr\{|m^{1/2}a_1 - b_k| > \varepsilon\} &= \Pr\left\{|m^{1/2}a_1 - b_k| > \varepsilon; \sup_p |g_p| \ge \eta\right\} \\ &\quad + \Pr\left\{|m^{1/2}a_1 - b_k| > \varepsilon; \sup_p |g_p| < \eta\right\}.\end{aligned}$$

The first term on the right converges to zero since by Proposition 2.3, with $r = 2$ there, $\sup_p |g_p = \widehat{f}_p/\widetilde{f}_p - 1| = o_p(1)$. To show that the first term on the right converges to zero, since $|\log x - (x-1)| \le 2^{-1}(x-1)^2$ for $x \simeq 1$, by Markov inequality it suffices to show that

$$\begin{aligned}(4.18)\qquad &\frac{m^{1/2}}{k}\sum_{p=1}^{k}(Eg_p^2 + Eg_{k+1}^2) \\ &= \frac{m^{1/2}}{k}\left(\sum_{p=1}^{k}(\mathrm{Var}(g_p) + \mathrm{Var}(g_{k+1})) + \sum_{p=1}^{k}(E^2(g_p) + E^2(g_{k+1}))\right)\end{aligned}$$

is $o(1)$, which is the case by Propositions 2.2 and 2.1 respectively. This concludes the proof of (a).

Part (b). By mean value theorem $m^{1/2}(b(d) - b(\overline{d})) = O_p(m^{1/2}(d-\overline{d})b'(d)) + o_p(1)$. But $b'(d)$, except the constant h_w, is

$$\frac{2}{k}\left(-\sum_{p=1}^{k^{1/2}/2} w_p \frac{\sum_{j=1}^{k^{1/2}/2}\varphi_{j,p}}{\sum_{j=1}^{k^{1/2}/2}(j+p)^{-2d}} - \sum_{p=k^{1/2}/2+1}^{k} w_p \frac{\sum_{j=-k^{1/2}/2}^{k^{1/2}/2}\varphi_{j,p}}{\sum_{j=-k^{1/2}/2}^{k^{1/2}/2}(j+p)^{-2d}}\right)$$

$$
\begin{aligned}
&+ \left(\frac{1}{k}\sum_{p=1}^{k} w_p\right) \frac{2\sum_{j=-k^{1/2}/2}^{k^{1/2}/2} \varphi_{j,k}}{\sum_{j=-k^{1/2}/2}^{k^{1/2}/2} (j+k)^{-2d}} \\
&= O(k^{-1/2}\log k),
\end{aligned}
$$

where $\varphi_{j,\ell} = (j+\ell)^{-2d}(\log((j+\ell)/\ell) - 1)$ uniformly in $d \in (-1/2, 1/2)$. So,

$$m^{1/2}(b(d) - b(\overline{d})) = O_p(k^{-3/4} m^{1/2} \log k) = o_p(1)$$

by A.5, because $(\overline{d} - d) = (\overline{d} - d - b(d)) + b(d) = O_p(k^{-1/4} + k^{-1/2}\log k)$ from the proof of part (a) with $m = k^{1/2}$ there and that by straightforward algebra $b(d) = O_p(k^{-1/2}\log k)$. This concludes the proof of part (b) and the theorem. □

4.6 *Proof of Proposition* 2.4

Let

$$\ddot{f}_{e,p} = \ddot{f}_e(\lambda_p) = (m_1+1)^{-1} \sum_{j=-m_1/2; j\neq -p}^{m_1/2} f_{|j+p|} I_{e,|j+p|}.$$

Proceeding as in the proof of Proposition 2.3, it suffices to examine $\sup_{p=1,\dots,m_1} |\widetilde{f}_p^{-1}(\ddot{f}_{e,p} - E\ddot{f}_{e,p})|$. Proceeding as with $\sup_{p=1+m,\dots,k} |\widetilde{f}_p^{-1}(\widehat{f}_{e,p} - E\widehat{f}_{e,p})|^2$ in Proposition 2.3 and observing that $\sup_q$ runs, in this case, for $q = 1, \dots, m_1^{(r-1)/r}$ with $s = q m_1^{1/r}$, the only term which is slightly different is the one corresponding to (4.14), whose expectation is bounded by

$$
\begin{aligned}
&\frac{K}{m_1+1} \sum_{q=1}^{m_1^{(r-1)/r}} \sum_{j=1}^{m_1/2} \sup_p \left(\frac{1}{j+p+1}\right)^r E\left|\sum_{\ell=1}^{j}((2\pi)I_{e,\ell+s} - 1)\right|^r \mathcal{I}\left(p \le \frac{m_1}{2}\right) \\
&\le \frac{K}{m_1+1} \sum_{q=1}^{m_1^{(r-1)/r}} \sum_{j=1}^{m_1} j^{-r/2} = O(m_1^{-1/r}\log m),
\end{aligned}
$$

using Lemma 5.6 part (b) instead of Lemmas 5.3 and 5.5 and that $r \ge 2$. Thus,

$$\sup_{p=1,\dots,m} |\widetilde{f}_p^{-1}(\ddot{f}_p - \widetilde{f}_p)| = O_p\left(\frac{m^{1/r}}{m_1^{1/2+1/r^2}}\right) + o_p(1). \qquad \square$$

4.7 *Proof of Theorem* 2.3

Denoting $\overline{w}_p = p^{-1}\sum_{\ell=1}^{p} w_\ell$, by definition,

$$
\begin{aligned}
(4.19) \qquad m^{1/2}(\widehat{d}^* - d) &= \frac{h_w}{\overline{v} m^{1/2}} \sum_{p=1}^{m} v_p \left\{\frac{1}{p}\sum_{\ell=1}^{p} w_\ell \log\frac{\ddot{f}_\ell}{\widetilde{f}_\ell} - \overline{w}_p \log\frac{\ddot{f}_{p+1}}{\widetilde{f}_{p+1}}\right\} \\
&+ \frac{h_w}{\overline{v} m^{1/2}} \sum_{p=1}^{m} v_p \left\{\frac{1}{p}\sum_{\ell=1}^{p} w_\ell \log\frac{\widetilde{f}_\ell}{f_\ell^*} - \overline{w}_p \log\frac{\widetilde{f}_{p+1}}{f_{p+1}^*}\right\} \\
&+ \frac{h_w}{\overline{v} m^{1/2}} \sum_{p=1}^{m} v_p \left\{\frac{1}{p}\sum_{\ell=1}^{p} \left(w_\ell \log\frac{f_\ell^*}{f_{p+1}^*} - h_w^{-1} d\right)\right\}.
\end{aligned}
$$

The third term on the right of (4.19) is

$$\frac{h_w}{\overline{v}m^{1/2}}\left\{\sum_{p=1}^{m^{1/4}}+\sum_{p=1+m^{1/4}}^{m}\right\}\left(v_p\frac{1}{p}\sum_{\ell=1}^{p}\left(w_\ell\log\frac{f_\ell^*}{f_{p+1}^*}-h_w^{-1}d\right)\right),$$

whose first term is $o(1)$ since $|w(u)\log u|\leq K$ and $\sum_{p=1}^{m^{1/4}}|v_p|=o(1)$ by A.7, whereas the second term is also $o(1)$, because by Brillinger ((1981), p. 15) the term in parenthesis is $O(p^{-1}|v_p|)$ and A.7 implies that $\sum_{p=1+m^{1/4}}^{m}p^{-1}|v_p|=O(1)$.

Denoting $m_1^*=[m_1\log m_1]$, the second term on the right of (4.19) is

$$\frac{h_w}{\overline{v}m^{1/2}}\left\{\sum_{p=1}^{m_1^*}+\sum_{p=m_1^*+1}^{m}\right\}\frac{v_p}{p}\sum_{\ell=1}^{p}\left(w_\ell\log\left(\frac{\widetilde{f}_\ell}{f_\ell}\right)-\overline{w}_p\log\left(\frac{\widetilde{f}_{p+1}}{f_{p+1}}\right)\right) \tag{4.20}$$

$$+\frac{h_w}{\overline{v}m^{1/2}}\sum_{p=1}^{m}\frac{v_p}{p}\sum_{\ell=1}^{p}\left(w_\ell\log\left(\frac{f_\ell}{f_\ell^*}\right)-\overline{w}_p\log\left(\frac{f_{p+1}}{f_{p+1}^*}\right)\right). \tag{4.21}$$

First, (4.21) $=o(1)$ proceeding as with the proof of a_2 in Theorem 2.2. Next, since by Lemma 5.6, $K^{-1}<f_{m_1/2}^{*-1}\widetilde{f}_j<K$ for $|j|\leq m_1^*$ and $\log|f_j^{-1}f_{m_1/2}^*|=O(\log m_1)$, the first term of (4.20) is $O(m^{-1/2}\log(m_1)\sum_{p=1}^{m_1^*}|v_p|)=O(m_1^2m^{-3/2}\log(m_1))=o(1)$ by A.7 and A.8. Finally, the second term of (4.20) is

$$\frac{h_w}{\overline{v}m^{1/2}}\sum_{p=m_1^*+1}^{m}v_p\left(\frac{1}{p}\sum_{\ell=1}^{m_1^*}w_\ell\log\left(\frac{\widetilde{f}_\ell}{f_\ell}\right)+\frac{1}{p}\sum_{\ell=m_1^*+1}^{p}w_\ell\log\left(\frac{\widetilde{f}_\ell}{f_\ell}\right)-\overline{w}_p\log\left(\frac{\widetilde{f}_{p+1}}{f_{p+1}}\right)\right)$$

which is $o(1)$ as we now show. The last term is bounded in absolute value by $m_1^2/m^{1/2}\sum_{p=m_1^*+1}^{m}|v_p|p^{-2}=O(m^{-3/2}m_1^2\log(m/m_1))=o(1)$ by A.8 and because, for $|\ell|>m_1^*$, Lemma 5.6 part (a) implies that $K^{-1}\leq m_1^{-2}p^2|f_p^{-1}\widetilde{f}_p-1|\leq K$, so that by the mean value theorem, $\log f_p^{-1}\widetilde{f}_p=O(m_1^2p^{-2})$. Similarly, because A.7 implies that $|w_\ell|\leq K(\ell/p)^\zeta$, the second term is $O(m^{-\zeta-1/2}m_1^{1+\zeta})=o(1)$ by A.8 since $\zeta\geq 1/3$. Finally, the first term of the last displayed expression is bounded in absolute value by

$$Dm^{-1/2}m_1^{1+\zeta}\sum_{p=m_1^*+1}^{m}v_pp^{-1-\zeta}\log m=O(m^{-\zeta-1/2}m_1^{1+\zeta}\log m)=o(1)$$

proceeding as with the first term of (4.20) by A.7 and A.8 in view that $\zeta\geq 1/3$ by A.7. So, the second term on the right of (4.19) is $o(1)$.

To complete the proof we need to show that the first term on the right of (4.19) converges in distribution to $\mathcal{N}(0,h_w^2\Phi^2)$. Proceeding as with the proof of (4.16), but using Proposition 2.5 instead of Proposition 2.3 when needed, it suffices to show that

$$\text{(a)}\quad \frac{1}{\overline{v}m^{1/2}}\sum_{p=1}^{m}v_p\left(\frac{1}{p}\sum_{\ell=1}^{p}w_\ell\left(\frac{\ddot{f}_\ell-\widetilde{f}_\ell}{\widetilde{f}_\ell}\right)-\overline{w}_p\left(\frac{\ddot{f}_{p+1}-\widetilde{f}_{p+1}}{\widetilde{f}_{p+1}}\right)\right) \xrightarrow{d}\mathcal{N}(0,\Phi^2) \tag{4.22}$$

$$(4.23)\quad \text{(b)}\quad \frac{K}{m^{1/2}}\left\{\sum_{p=1}^{m_1}+\sum_{p=m_1+1}^{m}\right\}v_p\left(\frac{1}{p}\sum_{\ell=1}^{p}w_\ell\left(\frac{\ddot{f}_\ell-\tilde{f}_\ell}{\tilde{f}_\ell}\right)^2+\overline{w}_p\left(\frac{\ddot{f}_{p+1}-\tilde{f}_{p+1}}{\tilde{f}_{p+1}}\right)^2\right) \xrightarrow{P} 0.$$

We begin with part (b). By Proposition 2.4 and Proposition 2.2 part (b) the first moment of the first term on the left of (4.23) is $o(m^{-1/2}\sum_{p=1}^{m_1}v_p)=o(m^{-3/2}m_1^2)=o(1)$ by A.7 and A.8, whereas the second term of (4.23) is

$$(4.24)\qquad \frac{K}{m^{1/2}}\sum_{p=m_1+1}^{m}v_p\frac{1}{p}\left\{\sum_{\ell=1}^{m_1}+\sum_{\ell=m_1+1}^{p}\right\}|w_\ell|\left(\frac{\ddot{f}_\ell-\tilde{f}_\ell}{\tilde{f}_\ell}\right)^2+\left(\frac{\ddot{f}_{p+1}-\tilde{f}_{p+1}}{\tilde{f}_{p+1}}\right)^2.$$

By Proposition 2.2 part (a) and A.8 the second term of (4.24) is $O_p(m_1^{-1}m^{1/2})=o_p(1)$, whereas Proposition 2.2 part (b), Proposition 2.4 and A.7 imply that the first term of (4.24) is

$$\begin{aligned}O_p\left(\frac{1}{m^{1/2}\log m_1}\sum_{p=m_1+1}^{m}v_p\frac{1}{p^{1+\zeta}}m_1^{1+\zeta}\right)&=O_p\left(\frac{m_1^{1+\zeta}}{m^{3/2}\log m_1}\sum_{p=m_1+1}^{m}p^{-\zeta}\right)\\&=o_p(m^{-1/2-\zeta}m_1^{1+\zeta}\log^{-1}m)\end{aligned}$$

which is $o_p(1)$ by A.8 and since $\zeta\geq 1/4$ by A.7.

So, to complete the proof we need to show part (a). The left side of (4.22) is

$$\begin{aligned}(4.25)\qquad &\frac{1}{\overline{v}m^{1/2}}\left[\sum_{p=1}^{m_1}v_p\left(\frac{1}{p}\sum_{\ell=1}^{p}w_\ell\vartheta_\ell-\overline{w}_p\vartheta_p\right)+\sum_{p=m-m_1+1}^{m}v_p\left(\frac{1}{p}\sum_{\ell=1}^{p}w_\ell\vartheta_\ell-\overline{w}_p\vartheta_p\right)\right]\\&+\frac{1}{\overline{v}m^{1/2}}\sum_{p=m_1+1}^{m-m_1}v_p\frac{1}{p}\left(\sum_{\ell=1}^{p}w_\ell\vartheta_\ell\right)-\frac{1}{\overline{v}m^{1/2}}\sum_{p=m_1+1}^{m-m_1}v_p\overline{w}_p\vartheta_p\\&+\frac{1}{\overline{v}m^{1/2}}\sum_{p=1}^{m}v_p\left(\overline{w}_p-\int_0^1 w(u)du\right)(\vartheta_p-\vartheta_{p+1})\\&+\frac{\int_0^1 w(u)du}{\overline{v}m^{1/2}}\sum_{p=1}^{m}v_p(\vartheta_p-\vartheta_{p+1}),\end{aligned}$$

where $\vartheta_b=\tilde{f}_b^{-1}(\ddot{f}_b-\tilde{f}_b)$. The last term of (4.25) is except constants

$$\frac{1}{m^{1/2}}\sum_{p=2}^{m}(v_{p+1}-v_p)\vartheta_p+m^{-1/2}v_1\vartheta_1+m^{-1/2}v_m\vartheta_{p+1}=o_p(1)$$

by Propositions 2.2 and 2.4 and that by A.7, $|v_{p+1}-v_p|\leq K/m$. Because Proposition 2.4 and Proposition 2.2 part (b) imply that $E|\vartheta_b|=o(1)$ for $b<m_1$, by A.7 and A.8 the first term of (4.25) is $o_p(m^{-3/2}m_1^2)=o_p(1)$. Since by Brillinger ((1981), p. 15), $\overline{w}_p-\int_0^1 w(u)du=O(p^{-1})$ and $E\sup_b|\vartheta_b|=o(1)$, the fifth term is $o_p(m^{-1/2}\log m)=o_p(1)$. The second term of (4.25) is

$$\begin{aligned}&\frac{1}{\overline{v}m^{1/2}}\left(\sum_{p=m-m_1+1}^{m}v_p\frac{1}{p}\sum_{\ell=1}^{m_1}w_\ell\vartheta_\ell+\sum_{p=m-m_1+1}^{m}v_p\left[\frac{1}{p}\sum_{\ell=m_1+1}^{p}w_\ell\vartheta_\ell-\overline{w}_p\vartheta_p\right]\right)\\&=b_{1,m}+b_{2,m}.\end{aligned}$$

Proceeding as with the first term of (4.24) $b_{1,m} = o_p(1)$ and $b_{2,m} = O_p(m_1^{1/2}/m^{1/2})$ since $\sum_{p=m-m_1+1}^{m} v_p = O(m_1)$ by A.7 and because by Proposition 2.2 part (a) $E|\vartheta_p| = O(m_1^{-1/2})$ for $p > m_1$.

The third term of (4.25), after rearranging subindexes and standard calculations, is

$$\frac{1}{\overline{v}m^{1/2}}\sum_{\ell=1}^{m_1}\vartheta_\ell \sum_{p=m_1+1}^{m-m_1} p^{-1}w_\ell v_p + \frac{1}{\overline{v}m^{1/2}}\sum_{\ell=m_1+1}^{m-m_1}\vartheta_\ell \sum_{p=\ell}^{m-m_1} p^{-1}w_\ell v_p, \tag{4.26}$$

whose first term is $o_p(1)$ by A.8 and similar arguments to those for the first term of (4.25).

Thus, gathering the second term of (4.26) and the fourth term of (4.25), we conclude that

$$m^{1/2}(\widehat{d}^* - d) = -\frac{h_w}{\overline{v}m^{1/2}}\sum_{\ell=m_1+1}^{m-m_1}\vartheta_\ell h_\ell + o_p(1),$$

where $h_\ell = \overline{w}_\ell v_\ell - \sum_{p=\ell}^{m-m_1} p^{-1}w_\ell v_p$. Because $\overline{v} \to \int_0^1 v(x)dx$, the proof is completed if

$$b_m = \frac{1}{m^{1/2}}\sum_{\ell=m_1+1}^{m-m_1}\vartheta_\ell h_\ell \xrightarrow{d} \mathcal{N}(0,\Phi^2).$$

Denoting $\eta_j = f_j^{-1}I_j - 1$, b_m is

$$\begin{aligned}
&\frac{1}{m^{1/2}}\sum_{\ell=m_1+1}^{m-m_1} h_\ell \frac{1}{m_1+1}\sum_{j=-m_1/2}^{m_1/2}(\eta_{\ell+j} + (\widetilde{f}_\ell^{-1}f_{j+\ell} - 1)\eta_{\ell+j}) \\
&= \frac{1}{m^{1/2}}\sum_{\ell=m_1+1}^{m-m_1} h_\ell \frac{1}{m_1+1}\sum_{j=-m_1/2}^{m_1/2}\eta_{\ell+j} + o_p(1),
\end{aligned} \tag{4.27}$$

as we now show. Denoting $\eta_j^* = \sigma_e^{-2}(2\pi)I_{e,j} - 1$ and $\widetilde{f}_\ell^{-1}f_{j+\ell} - 1 = \zeta_{j+\ell,\ell}$,

$$\begin{aligned}
\frac{1}{m_1+1}\sum_{j=-m_1/2}^{m_1/2}\zeta_{j+\ell,\ell}\eta_{\ell+j} &= \frac{1}{m_1+1}\sum_{j=-m_1/2}^{m_1/2}\zeta_{j+\ell,\ell}(\eta_{\ell+j} - \eta_{\ell+j}^*) \\
&\quad + \frac{1}{m_1+1}\sum_{j=-m_1/2}^{m_1/2}\zeta_{j+\ell,\ell}\eta_{\ell+j}^*.
\end{aligned}$$

Because Robinson's (1995*b*) Theorems 1 and 2 imply that $E|\eta_{\ell+j} - \eta_{\ell+j}^*| \le (\ell+j)^{-1/2}\log^{1/2}(\ell+j)$ and Lemma 5.6 part (a) implies that $|\zeta_{j+\ell,\ell}| \le K(m_1^2/\ell^2)$, the first absolute moment of the first term on the right is bounded by $Km_1^{3/2}\ell^{-2}\log m_1$ whereas using the Schwarz inequality, the first moment of the second term on the right is bounded by $K\ell^{-2}m_1^{3/2}$ because $|\operatorname{Cov}(\eta_{j_1}^*, \eta_{j_2}^*)| \le K\mathcal{I}(j_1 = j_2) + Kn^{-1}\mathcal{I}(j_1 \neq j_2)$. So,

$$\begin{aligned}
&E\left|\frac{1}{m^{1/2}}\sum_{\ell=m_1+1}^{m-m_1} h_\ell \frac{1}{m_1+1}\sum_{j=-m_1/2}^{m_1/2}\zeta_{j+\ell,\ell}\eta_{\ell+j}\right| \\
&\le \frac{Km_1^{1/2}\log^{1/2} m_1}{m^{1/2}}\sum_{\ell=m_1+1}^{m-m_1}\frac{h_\ell}{\ell} + \frac{Km_1^{1/2}}{m^{1/2}}\sum_{\ell=m_1+1}^{m-m_1}\frac{h_\ell}{\ell} = O\left(\left(\frac{m_1}{m}\right)^{1/2}\log m\right)
\end{aligned}$$

which is $o(1)$ A.7 and A.8. So, it suffices to examine the first term on the right of (4.27) which, after rearranging subindexes, is

$$\frac{1}{m^{1/2}}\sum_{p=3/2m_1+1}^{m-3/2m_1}\frac{\eta_p}{m_1+1}\sum_{j=-m_1/2}^{m_1/2}h_{p+j}+\frac{1}{m^{1/2}}\sum_{p=1+m_1/2}^{3/2m_1}\frac{\eta_p}{m_1+1}\sum_{j=1}^{p-m_1/2}h_{j+m_1}$$
$$+\frac{1}{m^{1/2}}\sum_{p=m-3/2m_1+1}^{m-m_1/2}\frac{\eta_p}{m_1+1}\sum_{j=p}^{m-m_1/2}h_{j-m_1/2}.$$

The last two terms are $O_p(m^{-1/2}m_1^{1/2})$, because they involve at most m_1 number of terms and by routine extension of Robinson's (1995*b*) Theorem 2, $m^{-1/2}\sum_{p=1}^{a}\eta_p = O_p(a^{1/2}m^{-1/2})$, for any $1 < a \leq m$. On the other hand, the first term is

$$\frac{1}{m^{1/2}}\sum_{p=3/2m_1+1}^{m-3/2m_1}h_p\eta_p+\sum_{p=3/2m_1+1}^{m-3/2m_1}h_p\eta_p\left(\frac{1}{h_p}\left(\frac{1}{m_1+1}\sum_{j=-m_1/2}^{m_1/2}h_{p+j}\right)-1\right)$$
$$=\frac{1}{m^{1/2}}\sum_{p=3/2m_1+1}^{m-3/2m_1}h_p\eta_p(1+o(1))$$

by continuity of $\eta(u)=v(u)\int_0^1 w(x)dx-\int_u^1 x^{-1}w(u/x)v(x)dx$ and Toeplitz Lemma. But, since A.7 implies the squared integrability of $\eta(u)$, by routine modification of Robinson's (1995*b*) Theorem 2 $m^{-1/2}\sum_{p=m_1+1}^{m-m_1}\overline{v}\,\overline{w}_p\eta_p \xrightarrow{d} N(0,\Phi^2)$. □

5. Technical lemmas

LEMMA 5.1. *Assuming* A.1 *and* A.3, *for* $m \leq p \leq q \leq k$, $k/n \to 0$, $\psi \in [0,1]$ *and* $a, b = 0, 1$,

$$\frac{f_p^{-a}f_q^{-b}}{m+1}\sum_{j=-m/2}^{m/2}\frac{1}{(j+p)^{\psi}}(f_{j+p}^a f_{j+q}^b - f_p^a f_q^b) = O(p^{-\psi-1}m). \tag{5.1}$$

PROOF. The proof is quite straightforward. By A.3 and mean value theorem,

$$f_{j+p} = f_p + \left(\frac{2\pi j}{n}\right) f'(\overline{\lambda}(p,j)),$$

where $\overline{\lambda}(p,j) \in [\lambda_p, \lambda_{j+p}]$. But also by A.3, $f'(\lambda) \leq K\lambda^{-1}f(\lambda)$, so

$$|f_{j+p}-f_p| \leq K\lambda_j\overline{\lambda}^{-1}(p,j)f(\overline{\lambda}(p,j)) = O(p^{-1}mf_p),$$

because $\lambda_j\overline{\lambda}^{-1}(p,j) = O(p^{-1}m)$ and $f_p^{-1}f(\overline{\lambda}(p,j)) = O(1)$ for $p \geq m$ by A.1. The conclusion is now immediate since $p^{\psi}/(p+j)^{\psi} \leq K$. □

LEMMA 5.2. *Assuming* A.1 *and* A.4, *for* $m/2 < p < m$,

$$\frac{\lambda_m^{2da}m^{1-2da}}{m+1}\sum_{j=-m/2}^{m/2}\frac{f_{j+p}^a}{(j+p)^{\psi}} = O\left(\log\left(\frac{2p+m}{2p-m}\right)\mathcal{I}(\vartheta=1)\right) \tag{5.2}$$

$$+ O((2p-m)^{1-\vartheta}\mathcal{I}(\vartheta>1)) + O((2p+m)^{1-\vartheta}\mathcal{I}(\vartheta<1)),$$

where $\vartheta = 2da + \psi$, $\psi \in [0,2]$ *and* $a = 1,2$.

PROOF. Because $\lambda_{j+p} \to 0$ as $m/2 < p < m$, the left side of (5.2) is

$$K \sum_{j=p-m/2}^{p+m/2} j^{-\psi-2da}(1+o(1))$$

by A.1 and A.4. From here the conclusion is immediate since $j \geq 1$. □

LEMMA 5.3. *Assuming* A.1 *and* A.4, *for* $1 \leq p \leq m/2$

$$\frac{\lambda_m^{2da} m^{1-2da}}{m} \sum_{j=1}^{m/2} \frac{f_{j+p}^a}{(j+p)^{\psi}} = O\left(\log\left(\frac{2p+m}{2p}\right)\mathcal{I}(\vartheta=1)\right) + O(p^{1-\vartheta}\mathcal{I}(\vartheta>1)) + O((2p+m)^{1-\vartheta}\mathcal{I}(\vartheta<1)), \tag{5.3}$$

where ϑ, ψ *and* a *are as in Lemma* 5.2.

PROOF. Because $f_{j+p} = \lambda_{j+p}^{-2d}(1+o(1))$ by A.1 and A.4, the left side of (5.3) is bounded by

$$K \sum_{j=1}^{m/2} (j+p)^{-\psi-2da}(1+o(1)) \leq K \sum_{j=p+1}^{p+m/2} j^{-\psi-2da}(1+o(1)).$$

From here, the conclusion is standard. □

LEMMA 5.4. *Assuming* A.1 *and* A.4, *for* $m/2 < p < m$

$$K^{-1} \leq \frac{\lambda_m^{2d}}{m+1} \sum_{j=-m/2}^{m/2} f_{j+p} \leq K.$$

PROOF. From Lemma 5.2, we only need to show the inequality on the left. Because $f_{j+p} = \lambda_{j+p}^{-2d}(1+o(1))$ by A.1 and A.4, the middle term is bounded from below by

$$\frac{K^{-1}}{m^{1-2d}} \sum_{j=p-m/2}^{p+m/2} j^{-2d} \geq \frac{K^{-1}}{m^{1-2d}} \left\{ \sum_{j=p-m/2}^{p} j^{-2d}\mathcal{I}(d \geq 0) + \sum_{j=p}^{p+m/2} j^{-2d}\mathcal{I}(d<0) \right\}$$

which is greater than or equal to $(m/p)^{2d} \geq K^{-1}$ since $m > p > m/2$. □

LEMMA 5.5. *Assuming* A.1 *and* A.4, *for* $1 \leq p \leq m/2$

$$K^{-1} \leq \frac{\lambda_m^{2d}}{m} \sum_{j=1}^{m/2} f_{j+p} \leq K.$$

PROOF. From Lemma 5.3, we only need to show the inequality on the left. Because $f_{j+p} = \lambda_{j+p}^{-2d}(1+o(1))$ by A.1 and A.4, the middle term is bounded from below by

$$\frac{K^{-1}}{m^{1-2d}}\sum_{j=1}^{m/2}(j+p)^{-2d} \geq \frac{K^{-1}}{m^{1-2d}}\left\{\sum_{j=1}^{m/2}\left(\left(\frac{m}{2}+p\right)^{-2d}\mathcal{I}(d\geq 0)+\frac{\mathcal{I}(d<0)}{j^{2d}}\right)\right\}$$

which is greater than or equal to K^{-1} since $1 \leq p \leq m/2$. □

LEMMA 5.6. *Under* A.1 *and* A.3, *for* $\mu = 1$ *and* 2

$$\text{(a)}\quad \text{For } m_1 \leq p,\ \begin{cases} K^{-1} < (m_1+1)^{-1}\displaystyle\sum_{j=m_1/2}^{m_1/2}\frac{f_{j+p}^{\mu}}{f_p^{\mu}} < K \\ (m_1+1)^{-1}\displaystyle\sum_{j=m_1/2;j\neq p}^{m_1/2}\left(\frac{f_{j+p}^{\mu}}{f_p^{\mu}}-1\right) = O(m_1^2p^{-2}). \end{cases}$$

$$\text{(b)}\quad \text{For } p < m_1,\ \begin{cases} K^{-1} \leq \lambda_{m_1/2}^{2d\mu}(2k_1+1)^{-1}\displaystyle\sum_{j=m_1/2;j\neq p}^{m_1/2} f_{j+p}^{\mu} \leq K & \textit{if}\ \ d\mu < \dfrac{1}{2} \\ n^{-2d\mu}m_1(m_1+1)^{-1}\displaystyle\sum_{j=m_1/2;j\neq p}^{m_1/2} f_{j+p}^{\mu} \leq K & \textit{if}\ \ d\mu \geq \dfrac{1}{2}. \end{cases}$$

PROOF. (a) We first show that $K^{-1} < (m_1+1)^{-1}\sum_{j=m_1/2}^{m_1/2} f_p^{-\mu} f_{j+p}^{\mu} < K$. Suppose first that $d \geq 0$. By A.1,

$$\frac{1}{(m_1+1)}\sum_{j=m_1/2}^{m_1/2}\frac{f_{j+p}^{\mu}}{f_p^{\mu}} \leq \frac{K}{(m_1+1)}\sum_{j=m_1/2}^{m_1/2}\left|\frac{p}{j+p}\right|^{2d\mu} \leq K$$

because for $p \geq m_1$ and $|j| \leq m_1/2$, $2^{-1} < |(j+p)/p|$. On the other hand,

$$\frac{1}{(m_1+1)}\sum_{j=m_1/2}^{m_1/2}\frac{f_{j+p}^{\mu}}{f_p^{\mu}} \geq K^{-1}\frac{1}{m_1}\sum_{j=[m_1/8]}^{[m_1/4]}\left|\frac{p}{j+p}\right|^{2d\mu} \geq K^{-1}$$

because for $p \geq m_1$ and $|j| \leq m_1/2$, $|(j+p)/p| < 3/2$. The case $d < 0$ follows by similar arguments. Next, we show the second part of (a). By Taylor expansion of f_{j+p}, the left side is

$$\frac{1}{(m_1+1)}\sum_{j=m_1/2}^{m_1/2}\left\{\frac{\mu(2\pi)j}{n}\frac{f_p'}{f_p}\right.$$
$$\left.+\frac{\mu(2\pi)^2}{2}\frac{j^2}{n^2}\left(\frac{f''(\overline{\lambda})f^{\mu-1}(\overline{\lambda})}{f_p^{\mu}}+(\mu-1)\frac{f'^2(\overline{\lambda})f^{\mu-2}(\overline{\lambda})}{f_p^{\mu}}\right)\right\}$$
$$\leq K\frac{\pi^2\mu^2}{(m_1+1)}\sum_{j=m_1/2}^{m_1/2}\left(\frac{j}{p-s+\delta j}\right)^2(1+o(1))$$

where $\overline{\lambda} = \overline{\lambda}(p + \delta j) \in (\lambda_p, \lambda_{p+j})$ and $\delta = \delta(j) \in (0, 1)$, by A.1, A.3 and that $f_p^{-1} f(\overline{\lambda})$ is bounded. The conclusion is now standard since $|p + \delta j| \geq p - |\delta j| > p/2$.

(b) It is immediate since by A.1, $f^{\mu}_{(j+p)\mathcal{I}(j+p\neq 0)} = K\lambda^{-2\mu d}_{|j+p|_+}(1 + o(1))$, where $|q|_+ = \max(q, 1)$. □

Acknowledgements

We are grateful to an Associate Editor and a referee for their very helpful comments. Of course, all remaining errors are our sole responsibility.

References

Brillinger, D. R. (1981). *Time Series: Data Analysis and Theory*, McGraw-Hill, New York.

Chen, Z.-G. and Hannan, E. J. (1980). The distribution of periodogram ordinates, *J. Time Ser. Anal.*, **1**, 73–82.

Cheung, Y.-W. and Diebold, F. X. (1994). On maximum-likelihood estimation of the differencing parameter of fractionally-integrated noise with unknown mean, *J. Econometrics*, **62**, 317–350.

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

Davies, R. B. and Harte, D. S. (1987). Tests for Hurst effect, *Biometrica*, **74**, 95–101.

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

Geweke, J. and Porter-Hudak, S. (1983). The estimation and application of long-memory time series models, *J. Time Ser. Anal.*, **4**, 221–238.

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, *Probab. Theory Related Fields*, **86**, 87–104.

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, *J. Time Ser. Anal.*, **18**, 49–60.

Hidalgo, J. (1996). Spectral analysis for bivariate time series with long memory, *Econometric Theory*, **12**, 773–792.

Hidalgo, J. (2002). Semiparametric estimation for stationary processes whose spectra have an unknown pole (preprint).

Hidalgo, J. and Robinson, P. M. (2002). Adapting to unknown disturbance autocorrelation in regression with long memory, *Econometrica*, **70**, 1545–1581.

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

Hurvich, C. M. and Beltrao, K. I. (1993). Asymptotics for the low-frequency ordinates of the periodogram of a long memory time series, *J. Time Ser. Anal.*, **14**, 455–472.

Künsch, H. R. (1986). Discrimination between monotonic trends and long-range dependence, *J. Appl. Probab.*, **23**, 1025–1030.

Künsch, H. R. (1987). Statistical aspects of self-similar processes, *Proceedings of the 1st World Congress of the Bernoulli Society*, 67–74, VNU Science Press, Utrecht.

Lobato, I. and Robinson, P. M. (1996). Average periodogram estimation of long memory, *J. Econometrics*, **73**, 303–324.

Parzen, E. (1986). Quantile spectral analysis and long-memory time series, *J. Appl. Probab.*, Special Vol. **23A**, 41–54.

Robinson, P. M. (1994). Semiparametric analysis of long-memory time series, *Ann. Statist.*, **22**, 515–539.

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

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

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

Taqqu, M. S. (1975). Weak convergence to fractional Brownian to the Rosenblatt process, *Z. Wahrsch. Verw. Gebiete*, **31**, 203–238.

Yajima, Y. (1985). On estimation of long-memory time series models, *Austral. J. Statist.*, **27**, 303–320.