COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
AT YALE UNIVERSITY

Box 2125, Yale Station
New Haven, Connecticut 06520

COWLES FOUNDATION DISCUSSION PAPER NO. 1265

MODIFIED LOCAL WHITTLE ESTIMATION OF
THE MEMORY PARAMETER IN THE NONSTATIONARY CASE

Katsumi Shimotsu and Peter C. B. Phillips

July 2000

Modified Local Whittle Estimation of the Memory Parameter in the Nonstationary Case[*]

Katsumi Shimotsu
Department of Economics
Yale University

and

Peter C. B. Phillips
Cowles Foundation for Research in Economics
Yale University

22 May 2000

Abstract

Semiparametric estimation of the memory parameter is studied in models of fractional integration in the nonstationary case, and some new representation theory for the discrete Fourier transform of a fractional process is used to assist in the analysis. A limit theory is developed for an estimator of the memory parameter that covers a range of values of d commonly encountered in applied work with economic data. The new estimator is called the modified local Whittle estimator and employs a version of the Whittle likelihood based on frequencies adjacent to the origin and modified to take into account the form of the data generating mechanism in the frequency domain. The modified local Whittle estimator is shown to be consistent for $0 < d < 2$ and is asymptotically normally distributed with variance $\frac{1}{4}$ for $\frac{1}{2} < d < \frac{7}{4}$. The approach allows for likelihood-based inference about d in a context that includes nonstationary data, is agnostic about short memory components and permits the construction of valid confidence regions for d that extend into the nonstationary region.

JEL Classification: C22
AMS 1991 subject classification: 62M10

1

1 Introduction

Fractional integration and the study of the so-called $I(d)$ processes has recently attracted a good deal of attention amongst theorists and empirical researchers. In applied econometric work, $I(d)$ processes with fractional $d > 0$ have been found to provide good empirical models for certain financial time series and volatility measures, as well as certain macroeconomic time series like inflation and interest rates. Fractional processes accommodate temporal dependence in a time series that is intermediate in form between short-memory series (the so-called $I(0)$ processes) and unit root time series ($I(1)$ processes). Fractional models encompass both stationary and nonstationary processes depending on the value of the memory parameter, and include both $I(0)$ and $I(1)$ processes as limiting cases when the memory parameter takes on the values zero and unity. For these reasons, fractional integration is attractive to empirical researchers, providing some liberation from the classical dichotomy of $I(0)$ and $I(1)$ processes. Growing evidence in applied work indicates that fractionally integrated processes can describe certain long range characteristics of economic data rather well, including the volatility of financial asset returns, forward exchange market premia, interest rate differentials, and inflation rates.

The memory parameter, d, plays a central role in the definition of fractional integration and is often the focus of empirical interest. When $-\frac{1}{2} < d < \frac{1}{2}$, the process has a stationary representation. For this range of d, two commonly used semiparametric estimators (log periodogram regression, local Whittle estimator) are shown to be consistent and asymptotically normally distributed by Robinson (1995a, 1995b). When $d \geq \frac{1}{2}$, the process is nonstationary and there are several ways of defining the observed series in terms of weakly dependent inputs. One model, which has been used in the existing literature, defines an $I(d)$ process with $\frac{1}{2} < d < \frac{3}{2}$ as a partial sum of $I(d-1)$ processes. According to this model, we can estimate d by taking first differences of the data, estimating $d-1$, and adding one to the estimate. Indeed, recent works by Velasco (1999a, 1999b)[1] extend Robinson's result to show that those two semiparametric estimators are consistent for $-\frac{1}{2} < d < 1$ and asymptotically normally distributed for $-\frac{1}{2} < d < \frac{3}{4}$. Hence, if we apply this 'differencing + adding-back' approach, the estimator is consistent for $\frac{1}{2} < d < 2$ and asymptotically normally distributed for $\frac{1}{2} < d < \frac{7}{4}$.

However, this model and approach does have some shortcomings. First, it employs different data generating mechanisms depending on the value of d, e.g. whether $d \lessgtr \frac{1}{2}$, a parameter which is not known *a priori* and which we have to estimate from the data. Second, the model for $d > \frac{1}{2}$ obscures the relationship between the observed data and the component innovations. For example, an $I(0.7)$ process is defined as a cumulative

---

[1] Velasco (1999a, 1999b) also show that the use of data tapering makes the estimators consistent and asymptotically normally distributed for $-\frac{1}{2} < d < \frac{3}{2}$, albeit at the cost of an increase in variance.

sum of an $I(-0.3)$ process, and the motivation for this construction of economic data is unclear. Appendix C in Section 9 of the paper provides some further discussion of these issues and various alternate models of fractional integration, including the model used here and another model that works from distant past rather than infinite past or fixed point initializations.

The approach taken in the present paper is to define fractionally integrated processes as weighted sums of short-memory input variables, which are treated nonparametrically[2]. This model gives a valid representation for all values of d and enables us to treat the $I(d)$ processes uniformly without any discontinuity in the data generating mechanism. It also relates the observed series directly to its component innovations, so that the impulse responses are just the weights on the short memory inputs. The paper uses a new representation and approximation theory for the discrete Fourier transform of a fractionally integrated time series (based on Phillips, 1999) which provides us with a representation that is valid in both nonstationary and (asymptotically) stationary cases. It is particularly helpful in analyzing the asymptotic behavior of the discrete Fourier transform and, hence, the periodogram of nonstationary fractionally integrated time series. So, it provides the key element in developing our theory and motivating the estimator we will use.

With this representation theory in hand, we develop a limit theory for a new estimator of the memory parameter of a fractional process allowing for nonstationary values of d. The new estimator is called the modified local Whittle estimator and employs a version of the Whittle likelihood based on frequencies adjacent to the origin and modified to take into account the form of the data generating mechanism in the frequency domain. The approach was suggested in Phillips (1999) without any formal development of its properties or asymptotic behavior. The present paper takes up this study and demonstrates that the modified local Whittle estimator is consistent for $d \in (0, 2)$ and asymptotically normally distributed with variance $\frac{1}{4}$ for $d \in \left(\frac{1}{2}, \frac{3}{2}\right)$ and $d \in \left(\frac{3}{2}, \frac{7}{4}\right)$. For $d \in \left[\frac{7}{4}, 2\right)$, the limit distribution is nonnormal and the rate of convergence decreases. Thus, the approach allows for likelihood-based inference about d in a context that allows for nonstationarity, using a limit theory that is equivalent to that which applies in the stationary region for the unmodified Whittle estimator (Robinson, 1995b). In this respect, our theory complements recent work by Velasco (1999a), extending further the domain of d where valid inference is possible. Phillips (1999) proposes another semiparametric estimator of d (an exact local Whittle estimator) that requires no prior information on the value of d. While the derivation of an asymptotic theory for the exact local Whittle estimator is very difficult, that of the modified Whittle estimator is much more feasible. Part of the motivation for the modified Whittle estimator is that it is constructed to minimize an objective function that approximates that of the exact local Whittle estimator. The analysis of this estimator therefore serves as a stepping stone towards a more general theory of estimation of d. Additionally, the modified estimator is related to an alternate 'differencing + adding-back' estimator and can be motivated in terms

---

[2]Tanaka (1999) uses a fully parametric version of this model and shows that the MLE of d is consistent and asymptotically normally distributed for any values of d.

of this approach as well.

The remainder of the paper is organized as follows. The new representation and approximation theory that we need are developed in Section 2. Section 3 defines the modified local Whittle estimator and proves its consistency. Section 4 demonstrates asymptotic normality. Section 5 reports some simulation results and gives an empirical illustration. Section 6 concludes the paper. Some technical results are given in Appendix A in Section 7. Proofs are collected together in Appendix B in Section 8. Some alternative nonstationary representations are discussed in Appendix C in Section 9.

2 Preliminary Representation Theory and Asymptotics

2.1 A Model of Nonstationary Fractional Integration

We consider the fractional process X_t generated by the model

$$(1 - L)^d (X_t - X_0) = u_t, \quad t = 0, 1, 2, ... \tag{1}$$

where X_0 is a random variable with a certain fixed distribution. Our interest is primarily in the case where X_t is nonstationary and $\frac{1}{2} < d \leq 2$, so in (1) we work from a given initial date $t = 0$, set $u_t = 0$ for all $t \leq 0$, and assume that u_t $(t \geq 1)$ is stationary with zero mean and continuous spectrum $f_u(\lambda) > 0$. Expanding the binomial in (1) gives the form

$$\sum_{k=0}^{t} \frac{(-d)_k}{k!} (X_{t-k} - X_0) = u_t, \tag{2}$$

where

$$(d)_k = \frac{\Gamma(d + k)}{\Gamma(d)} = (d)(d+1)...(d+k-1),$$

is Pochhammer's symbol for the forward factorial function and $\Gamma(\cdot)$ is the gamma function. When d is a positive integer, the series in (2) terminates, giving the usual formulae for the model (1) in terms of the differences and higher order differences of X_t. An alternate form for X_t is obtained by inversion of (1), giving

$$X_t = (1 - L)^{-d} u_t + X_0 = \sum_{k=0}^{t-1} \frac{(d)_k}{k!} u_{t-k} + X_0. \tag{3}$$

This model gives a valid representation for all values of d. When $d > \frac{1}{2}$, X_t is nonstationary, while X_t is asymptotically stationary when $0 < d < \frac{1}{2}$. The impulse responses of X_t to unit changes in u_{t-k} are given directly in (3) and we may similarly obtain impulse responses to unit changes in innovations in u_t using (3) in conjunction with (4) below. Further, the above formulation is convenient for the construction of the likelihood and an estimator of d using a likelihood-based approach is developed in Section 3.

4

Throughout this paper it will be convenient to assume that the stationary component u_t in (1) is a linear process of the form

$$u_t = C(L)\varepsilon_t = \sum_{j=0}^{\infty} c_j \varepsilon_{t-j}, \quad \sum_{j=0}^{\infty} j\,|c_j| < \infty, \quad C(1) \neq 0, \qquad (4)$$

for all t and with $\varepsilon_t = iid\left(0, \sigma^2\right)$ and $E\varepsilon_t^4 = \mu_4 < \infty$. The summability condition in (4) is satisfied by a wide class of parametric and nonparametric models for u_t and enables the use of the techniques in Phillips and Solo (1992). Under (4), the spectral density of u_t is $f_u(\lambda) = \frac{\sigma^2}{2\pi}\left|\sum_{j=0}^{\infty} c_j e^{ij\lambda}\right|^2$. In spite of its generality, the specificity of the linear process form (4) is much more restrictive than the local assumptions about $f_u(\lambda)$ at $\lambda = 0$ that are used in other work, notably Robinson (1995b), and which reflect the local nature of the semiparametric problem of estimation of the memory parameter d.

Define the discrete Fourier transform (dft) of a time series a_t evaluated at the fundamental frequencies as

$$w_a(\lambda_s) = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^{n} a_t e^{it\lambda_s}, \quad \lambda_s = \frac{2\pi s}{n}, s = 1, ..., n. \qquad (5)$$

Our approach is to algebraically manipulate (2) so that it can be rewritten in a convenient form to accommodate dft's. The following Lemma by Phillips (1999) provides an exact representation of $w_u(\lambda)$ in terms of functions of the data X_t.

2.2 Lemma

If X_t follows (1), then

$$w_u(\lambda) = w_x(\lambda) D_n\left(e^{i\lambda}; d\right) - \frac{1}{\sqrt{2\pi n}}\left(\widetilde{X}_{\lambda 0}(d) - e^{in\lambda}\widetilde{X}_{\lambda n}(d)\right) + \frac{X_0}{\sqrt{2\pi n}}\sum_{t=1}^{n} e^{it\lambda}, \quad (6)$$

where $D_n\left(e^{i\lambda}; d\right) = \sum_{k=0}^{n} \frac{(-d)_k}{k!} e^{ik\lambda}$ and

$$\widetilde{X}_{\lambda n}(d) = \widetilde{D}_{n\lambda}\left(e^{-i\lambda}L; d\right) X_n = \sum_{p=0}^{n-1} \widetilde{d}_{\lambda p} e^{-ip\lambda} X_{n-p}, \quad \widetilde{d}_{\lambda p} = \sum_{k=p+1}^{n} \frac{(-d)_k}{k!} e^{ik\lambda}.$$

The expression (6) may be interpreted as a frequency domain version of the original model (1). We can introduce a new transform

$$v_x(\lambda_s; d) = w_x(\lambda_s) - D_n\left(e^{i\lambda_s}; d\right)^{-1} \frac{1}{\sqrt{2\pi n}}\left(\widetilde{X}_{\lambda_s 0}(d) - \widetilde{X}_{\lambda_s n}(d)\right), \qquad (7)$$

for which

$$v_x(\lambda_s; d) = D_n\left(e^{i\lambda_s}; d\right)^{-1} w_u(\lambda_s),$$

holds exactly. While this representation gives an exact relationship, the terms $\widetilde{X}_{\lambda 0}\left(d\right)$ and $\widetilde{X}_{\lambda n}\left(d\right)$ in the right hand side of (7) contain involved functions of X_t and d. This makes asymptotic analysis very difficult, and hence it is useful to find approximations of them both for developing asymptotics and for suggesting simplified procedures. The following lemma gives another representation that forms the basis of the approximation in frequency domain form.

2.3 Lemma

(a) If X_t follows (1), then

$$w_x\left(\lambda\right)\left(1-e^{i\lambda}\right)=D_n\left(e^{i\lambda};f\right)w_u\left(\lambda\right)-\frac{e^{in\lambda}}{\sqrt{2\pi n}}\widetilde{U}_{\lambda n}\left(f\right)-\frac{e^{i\lambda}}{\sqrt{2\pi n}}\left(e^{in\lambda}X_n-X_0\right),\quad(8)$$

where $D_n\left(e^{i\lambda};f\right)=\sum_{k=0}^{n}\frac{(-f)_k}{k!}e^{ik\lambda}$, $f=1-d$, and

$$\widetilde{U}_{\lambda n}\left(f\right)=\widetilde{D}_{n\lambda}\left(e^{-i\lambda}L;f\right)u_n=\sum_{p=0}^{n-1}\widetilde{f}_{\lambda p}e^{-ip\lambda}u_{n-p},\quad\widetilde{f}_{\lambda p}=\sum_{k=p+1}^{n}\frac{(-f)_k}{k!}e^{ik\lambda}.\quad(9)$$

(b) If X_t follows (1) with $d=1$, then

$$w_x\left(\lambda\right)\left(1-e^{i\lambda}\right)=w_u\left(\lambda\right)-\frac{e^{i\lambda}}{\sqrt{2\pi n}}\left(e^{in\lambda}X_n-X_0\right).\quad(10)$$

The representation (8) results from algebraic manipulation and hence is valid for all values of d. However, the value of d affects the order of magnitude of the terms $\widetilde{U}_{\lambda n}\left(f\right)$ and X_n and, consequently, it affects the extent to which the (normalized) dft of the observed data, $w_x\left(\lambda\right)$, can provide an approximation of the dft of the component innovations, $w_u\left(\lambda\right)$.

2.4 The Modified Discrete Fourier Transform

The representation (8) suggests the use of the quantity

$$\begin{aligned}v_x\left(\lambda_s\right)&=w_x\left(\lambda_s\right)+\frac{e^{i\lambda_s}}{1-e^{i\lambda_s}}\frac{X_n-X_0}{\sqrt{2\pi n}}\\&=\frac{1}{\left(1-e^{i\lambda}\right)}\left[D_n\left(e^{i\lambda};f\right)w_u\left(\lambda\right)-\frac{e^{in\lambda}}{\sqrt{2\pi n}}\widetilde{U}_{\lambda n}\left(f\right)\right],\quad(11)\end{aligned}$$

and $I_v\left(\lambda_j\right)=v_x\left(\lambda_s\right)v_x\left(\lambda_s\right)^*$ to approximate $v_x\left(\lambda_s;d\right)$ and $I_v\left(\lambda_s;d\right)=v_x\left(\lambda_s;d\right)v_x\left(\lambda_s;d\right)^*$. This is done by approximating $\widetilde{X}_{\lambda_s 0}\left(d\right)$ and $\widetilde{X}_{\lambda_s n}\left(d\right)$ by $D_n\left(e^{i\lambda_s};d\right)\left(1-e^{i\lambda_s}\right)^{-1}e^{i\lambda_s}X_0$ and $D_n\left(e^{i\lambda_s};d\right)\left(1-e^{i\lambda_s}\right)^{-1}e^{i\lambda_s}X_n$. We call $v_x\left(\lambda_s\right)$ and $I_v\left(\lambda_s\right)$ the *modified discrete Fourier transform* and *modified periodogram*, respectively. In the following, we confine our attention to the case $d\in\left(0,2\right)$, which is the range of values of d commonly encountered in applied economic work. Indeed, for this range of values of d and for

the frequencies in the vicinity of the origin, the second term in (11) becomes negligible compared with the first term, and the (normalized) modified periodogram is well approximated by the periodogram of u_t and hence ε_t. The following lemmas establish this relationship and they are used in the following sections to examine the asymptotic behavior of the modified local Whittle estimator.

2.5 Lemma

Let $\widetilde{\varepsilon}_{\lambda n}(f) = \sum_{p=0}^{n-1} \widetilde{f}_{\lambda p} e^{-ip\lambda} \varepsilon_{n-p}$.

(a) For $d \in \left(\frac{1}{2}, \frac{3}{2}\right) \setminus \{1\}$,

$$\lambda_s^d v_x(\lambda_s) = e^{\frac{\pi}{2}di} C(1) w_\varepsilon(\lambda_s) - \frac{\lambda_s^d C(1)}{1 - e^{i\lambda_s}} \frac{\widetilde{\varepsilon}_{\lambda_s n}(f)}{\sqrt{2\pi n}} + r_{s,n}^a + r_{s,n}^b(d) + r_{s,n}^c(d),$$

where $E\left|r_{s,n}^a\right|^2 = O\left(\lambda_s^2\right)$, $E\left|r_{s,n}^b(d)\right|^2 = O\left(s^{2d-4}\right)$, and

$$E\left|r_{s,n}^c(d)\right|^2 = \left\{ \begin{array}{ll} O\left(s^{2d-2}n^{1-2d}\right), & \text{for } d \in \left(\frac{1}{2}, 1\right), \\ O\left(s^{2d-2}n^{-1}\right), & \text{for } d \in \left(1, \frac{3}{2}\right), \end{array} \right.$$

uniformly in s.

(b) For $d \in \left(\frac{1}{2}, \frac{3}{2}\right) \setminus \{1\}$,

$$\lambda_s^d v_x(\lambda_s) = e^{\frac{\pi}{2}di} C(1) w_\varepsilon(\lambda_s) + r_{s,n}^a + r_{s,n}^b(d),$$

where $E\left|r_{s,n}^a\right|^2 = O\left(\lambda_s^2\right)$ and

$$E\left|r_{s,n}^b(d)\right|^2 = \left\{ \begin{array}{ll} O\left(s^{2d-2}\right), & \text{for } d \in \left(\frac{1}{2}, 1\right), \\ O\left(s^{2d-3}\right), & \text{for } d \in \left(1, \frac{3}{2}\right), \end{array} \right.$$

uniformly in s.

(c) For $d = 1$,

$$\lambda_s v_x(\lambda_s) = iC(1) w_\varepsilon(\lambda_s) + r_{s,n}^a,$$

where $E\left|r_{s,n}^a\right|^2 = O\left(\lambda_s^2\right)$ uniformly in s.

(d) For $d \in \left(0, \frac{1}{2}\right]$,

$$\lambda_s^d v_x(\lambda_s) = e^{\frac{\pi}{2}di} C(1) w_\varepsilon(\lambda_s) + r_{s,n}^a + r_{s,n}^b(d),$$

where $E\left|r_{s,n}^a\right|^2 = O\left(\lambda_s^2\right)$ and $E\left|r_{s,n}^b(d)\right|^2 = O\left(s^{2d-2}n^{1-2d}\log n\right)$ uniformly in s.

(e) For $d \in \left[\frac{3}{2}, 2\right)$,

$$\lambda_s^d v_x(\lambda_s) = e^{\frac{\pi}{2}di} C(1) w_\varepsilon(\lambda_s) + r_{s,n}^a + r_{s,n}^b(d) + r_{s,n}^c(d),$$

where $E\left|r_{s,n}^a\right|^2 = O\left(\lambda_s^2\right)$, $E\left|r_{s,n}^b(d)\right|^2 = O\left(s^{2d-4}\log n\right)$, and $E\left|r_{s,n}^c(d)\right|^2 = O\left(s^{2d-2}n^{-1}\right)$ uniformly in s.

2.6   Corollary

(a) For $d \in \left(\frac{1}{2}, \frac{3}{2}\right) \setminus \{1\}$,

$$\lambda_s^{2d} I_v(\lambda_s) = \left| e^{\frac{\pi}{2} di} C(1) w_\varepsilon(\lambda_s) - \frac{\lambda_s^d C(1)}{1 - e^{i\lambda_s}} \frac{\widetilde{\varepsilon}_{\lambda_s n}(f)}{\sqrt{2\pi n}} \right|^2 + R_{s,n}^a + R_{s,n}^b(d) + R_{s,n}^c(d),$$

where $E\left|R_{s,n}^a\right| = O(\lambda_s)$, $E\left|R_{s,n}^b(d)\right| = O\left(s^{d-2}\right)$, and

$$E\left|R_{s,n}^c(d)\right| = \begin{cases} O\left(s^{d-1} n^{\frac{1}{2}-d}\right), & \text{for } d \in \left(\frac{1}{2}, 1\right), \\ O\left(s^{d-1} n^{-\frac{1}{2}}\right), & \text{for } d \in \left(1, \frac{3}{2}\right), \end{cases}$$

uniformly in s.

(b) For $d \in \left(\frac{1}{2}, \frac{3}{2}\right) \setminus \{1\}$,

$$\lambda_s^{2d} I_v(\lambda_s) = |C(1)|^2 I_\varepsilon(\lambda_s) + R_{s,n}^a + R_{s,n}^b(d),$$

where $I_a(\lambda_s) = w_a(\lambda_s) w_a(\lambda_s)^*$, $E\left|R_{s,n}^a\right| = O(\lambda_s)$, and

$$E\left|R_{s,n}^b(d)\right| = \begin{cases} O\left(s^{d-1}\right), & \text{for } d \in \left(\frac{1}{2}, 1\right), \\ O\left(s^{d-\frac{3}{2}}\right), & \text{for } d \in \left(1, \frac{3}{2}\right), \end{cases}$$

uniformly in s.

(c) For $d = 1$,

$$\lambda_s^2 I_v(\lambda_s) = |C(1)|^2 I_\varepsilon(\lambda_s) + R_{s,n}^a,$$

where $E\left|R_{s,n}^a\right| = O(\lambda_s)$ uniformly in s.

(d) For $d \in \left(0, \frac{1}{2}\right]$,

$$\lambda_s^{2d} I_v(\lambda_s) = |C(1)|^2 I_\varepsilon(\lambda_s) + R_{s,n}^a + R_{s,n}^b(d) + R_{s,n}^c(d),$$

where $E\left|R_{s,n}^a\right| = O(\lambda_s)$, $E\left|R_{s,n}^b(d)\right| = O\left(s^{d-1} n^{\frac{1}{2}-d} (\log n)^{\frac{1}{2}}\right)$, and $E\left|R_{s,n}^c(d)\right| = O\left(s^{2d-2} n^{1-2d} \log n\right)$ uniformly in s.

(e) For $d \in \left[\frac{3}{2}, 2\right)$,

$$\lambda_s^{2d} I_v(\lambda_s) = |C(1)|^2 I_\varepsilon(\lambda_s) + R_{s,n}^a + R_{s,n}^b(d) + R_{s,n}^c(d) + R_{s,n}^e(d) + R_{s,n}^g(d),$$

where

$$E\left|R_{s,n}^a\right| = O(\lambda_s), \qquad E\left|R_{s,n}^b\right| = O\left(s^{d-2} (\log n)^{\frac{1}{2}}\right), \quad E\left|R_{s,n}^c\right| = O\left(s^{d-1} n^{-\frac{1}{2}}\right),$$
$$E\left|R_{s,n}^e\right| = O\left(s^{2d-4} \log n\right), \quad E\left|R_{s,n}^g\right| = O\left(s^{2d-2} n^{-1}\right),$$

uniformly in s.

3   Modified Local Gaussian Estimation: Consistency

We propose a new estimator of d which is based on maximization of the likelihood function of u_t. Our concern is with the case where little is known about the short memory component process u_t and its spectrum $f_u(\lambda)$ is treated nonparametrically. This is accomplished by working with a set of m frequencies $\left\{\lambda_s = \frac{2\pi s}{n} : s = 1, ..., m\right\}$ that shrink slowly to origin as the sample size $n \to \infty$, and this makes the resulting estimator free from misspecification of dynamics of the component process u_t.

The (negative) Whittle likelihood based on frequencies up to λ_m and up to scale multiplication is

$$\sum_{j=1}^{m} \log f_u(\lambda_j) + \sum_{j=1}^{m} \frac{I_u(\lambda_j)}{f_u(\lambda_j)}, \tag{12}$$

where m is a number such that $\frac{1}{m} + \frac{m}{n} \to 0$ as $n \to \infty$. Using the relationship (7), we can transform (12) to be data dependent, in conjunction with the local approximation $f_u(\lambda_j) \sim f_u(0) = G$. This yields the objective function (Phillips, 1999)

$$K_m(G,d) = \frac{1}{m} \sum_{j=1}^{m} \left[\log\left(\left|D_n\left(e^{i\lambda_j};d\right)\right|^{-2} G\right) + \frac{\left|D_n\left(e^{i\lambda_j};d\right)\right|^2}{G} I_v(\lambda_j;d)\right].$$

The estimator of d that minimizes $K_m(G,d)$ does not rely on approximations of the discrete Fourier transform and may be expected to provide good semiparametric estimates for all values of d. The examination of the asymptotics is very difficult, however, so we apply the approximate relationship

$$I_v(\lambda_j) \sim I_v(\lambda_j;d), \quad D_n\left(e^{i\lambda_j};d\right) \sim \lambda_j^d,$$

to obtain the objective function

$$Q_m(G,d) = \frac{1}{m} \sum_{j=1}^{m} \left[\log\left(G\lambda_j^{-2d}\right) + \frac{\lambda_j^{2d}}{G} I_v(\lambda_j)\right]. \tag{13}$$

We call this expression the *modified local Whittle likelihood function*, because it is obtained by replacing the periodogram ordinates, $I_x(\lambda_j)$ in the local Whittle likelihood function (Künsch (1987), Robinson (1995b))

$$Q_m^*(G,d) = \frac{1}{m} \sum_{j=1}^{m} \left[\log\left(G\lambda_j^{-2d}\right) + \frac{\lambda_j^{2d}}{G} I_x(\lambda_j)\right], \tag{14}$$

by the modified periodogram ordinates, $I_v(\lambda_j)$.

We propose to estimate G and d by minimising $Q_m(G,d)$, so that

$$\left(\widehat{G},\widehat{d}\right) = \arg\min_{0<G<\infty, \, d\in\Theta} Q_m(G,d),$$

where $\Theta = [\Delta_1, \Delta_2]$ and Δ_1 and Δ_2 are numbers such that $0 < \Delta_1 < \Delta_2 < \infty$. The number Δ_1 can be chosen as close to zero as may be desired. It will be convenient

in what follows to distinguish the true values of the parameters by the notation $G_0 = f_u(0)$ and d_0.

Concentrating (13) with respect to G, we find that the estimate $\widehat{d}$ satisfies

$$\widehat{d} = \arg\min_d R(d),$$

where

$$R(d) = \log \widehat{G}(d) - 2d\frac{1}{m}\sum_{j=1}^{m}\log \lambda_j, \quad \widehat{G}(d) = \frac{1}{m}\sum_{j=1}^{m}\lambda_j^{2d}I_v(\lambda_j).$$

The following results show that $\widehat{d}$ is consistent in both the (asymptotically) stationary and nonstationary cases. When $d_0 \in \left(\frac{1}{2}, \frac{3}{2}\right)$, no condition is required on the rate of expansion of m. When $d_0 \in \left[\frac{3}{2}, 2\right)$, an additional condition $\frac{n^\alpha}{m} + \frac{m^{2d_0-2}}{n} \to 0$ is necessary in order to achieve consistency. This condition is fairly weak, though, because $m = o\left(n^{0.5}\right)$ is sufficient even when the condition is strongest, i.e. when $d_0 = 2$. When $d_0 \in \left[\Delta_1, \frac{1}{2}\right]$, however, the rate condition on m becomes stringent. Then, the condition $\frac{n^{1-2\Delta_1}\log n \log m}{m} \to 0$ implies that m has to grow fast for $\widehat{d}$ to be consistent, and it will be difficult to satisfy when Δ_1 is small.

3.1 Theorem

If $\frac{1}{m} + \frac{m}{n} \to 0$ as $n \to \infty$, then , for $d_0 \in \left(\frac{1}{2}, \frac{3}{2}\right)$, $\widehat{d} \to_p d_0$ as $n \to \infty$.

3.2 Theorem

If $d_0 \in \left[\frac{3}{2}, 2\right)$ and $\frac{1}{m} + \frac{m}{n} + \frac{n^\alpha}{m} + \frac{m^{2d_0-2}}{n} \to 0$ as $n \to \infty$ for some $\alpha > 0$, then, $\widehat{d} \to_p d_0$.

3.3 Theorem

If $d_0 \in \left[\Delta_1, \frac{1}{2}\right]$ and $\frac{1}{m} + \frac{m}{n} + \frac{n^{1-2\Delta_1}\log n \log m}{m} \to 0$ as $n \to \infty$, then, $\widehat{d} \to_p d_0$.

3.4 Theorem

If $\widehat{d} \to_p d_0$ as $n \to \infty$, then, $\widehat{G}\left(\widehat{d}\right) \to_p G_0$.

3.5 Remarks

(a) Using the result from Corbae, Ouliaris and Phillips (1999), it is straightforward to show, for $X_t = t$, that

$$v_t(\lambda_s) = \frac{1}{\sqrt{2\pi n}}\sum_{t=1}^{n}te^{i\lambda_s t} + \frac{e^{i\lambda_s}}{1-e^{i\lambda_s}}\frac{n}{\sqrt{2\pi n}} = 0.$$

Hence, the modified discrete Fourier transform and the modified local Whittle estimator are invariant to a linear trend.

10

(b) Interestingly, the modified estimator, which is derived above from the frequency domain data generating mechanism, is closely related to the ordinary Whittle estimator with the first differenced data. Indeed, it can be shown that, when $d - d_0 > -\frac{1}{2}$, $Q_m(G, d) = Q^*_{m\Delta X}(G, \delta) + o_p(1)$ holds where $\delta = d - 1$ and

$$Q^*_{m\Delta X}(G, \delta) = \frac{1}{m} \sum_{j=1}^{m} \left[\log\left(G\lambda_j^{-2\delta} \right) + \frac{\lambda_j^{2\delta}}{G} I_{\Delta x}(\lambda_j) \right],$$

which is the objective function of the local Whittle estimator with first differenced data.

4 Modified Local Gaussian Estimation: Asymptotic Normality

The following theorems establish asymptotic normality of the modified local Whittle estimator for $d_0 \in \left(\frac{1}{2}, \frac{7}{4} \right) \setminus \left\{ \frac{3}{2} \right\}$ under somewhat stronger conditions on the expansion rate of m.

4.1 Theorem

If $\frac{1}{m} + \frac{m^{\frac{3}{2}} \log m}{n} \to 0$ as $n \to \infty$, then, for $d_0 \in \left(\frac{1}{2}, 1 \right]$, we have

$$m^{\frac{1}{2}} \left(\widehat{d} - d_0 \right) \Rightarrow N\left(0, \frac{1}{4} \right).$$

If $\frac{1}{m} + \frac{m^{\frac{3}{2}} \log m}{n} + \frac{m^{2d_0 - 1}(\log m)^2}{n} \to 0$ as $n \to \infty$, then, for $d_0 \in \left(1, \frac{3}{2} \right)$, we have

$$m^{\frac{1}{2}} \left(\widehat{d} - d_0 \right) \Rightarrow N\left(0, \frac{1}{4} \right).$$

4.2 Theorem

If $\frac{1}{m} + \frac{n^\alpha}{m} + \frac{m^{\frac{3}{2}} \log m}{n} \to 0$ as $n \to \infty$ for some $\alpha > 0$, then, for $d_0 \in \left(\frac{3}{2}, \frac{7}{4} \right)$, we have

$$m^{\frac{1}{2}} \left(\widehat{d} - d_0 \right) \Rightarrow N\left(0, \frac{1}{4} \right).$$

4.3 Remarks

(a) The variance of the limiting distribution is the same as in the stationary case (see Robinson (1995b)).

(b) The rate condition $\frac{1}{m} + \frac{m^{\frac{3}{2}} \log m}{n} \to 0$ corresponds to assumption A4$'$ of Robinson (1995b) with $\beta = 1$. Indeed, since $C\left(e^{i\lambda} \right)$ is differentiable with a bounded derivative, if we define $f_x(\lambda) = \left| 1 - e^{i\lambda} \right|^{-2d} \left| C\left(e^{i\lambda} \right) \right|^2$, then it follows that $f_x(\lambda) = |C(1)|^2 \lambda^{-2d} (1 + O(\lambda))$.

(c) An additional condition on the rate of m, $\frac{m^{2d_0-1}(\log m)^2}{n} \to 0$, becomes necessary when $d_0 \in \left(1, \frac{3}{2}\right)$. When $d_0 < \frac{5}{4}$, this condition is redundant because it is dominated by $\frac{m^{\frac{3}{2}}\log m}{n} \to 0$. Indeed, this is a fairly weak condition, because $m\log m = o\left(n^{0.5}\right)$ is sufficient for it to hold even when it is strongest, i.e. when $d_0 = \frac{3}{2}$.

When $d_0 \in \left[\frac{7}{4}, 2\right)$, $\widehat{d}$ has a nonnormal distribution and the rate of convergence decreases.

4.4   Theorem

If $\frac{1}{m} + \frac{n^{\alpha}}{m} + \frac{m^{\frac{3}{2}}\log m}{n} + \frac{m^{2d_0-2}(\log m)^{12}}{n} \to 0$ as $n \to \infty$ for some $\alpha > 0$, then
(a) For $d_0 = \frac{7}{4}$, if $E\left|\varepsilon_t\right|^p < \infty$ for $p > 4$,

$$\sqrt{m}\left(\widehat{d} - d_0\right) = \xi_1 + \xi_2,$$

where

$$\xi_1 \Rightarrow N\left(0, \frac{1}{4}\right), \quad \xi_2 \Rightarrow (2\pi)^{-\frac{1}{2}} B_{-\frac{1}{4}}(1)^2.$$

(b) For $d_0 \in \left(\frac{7}{4}, 2\right)$,

$$m^{4-2d_0}\left(\widehat{d} - d_0\right) \Rightarrow \frac{(2-d_0)(2\pi)^{2d_0-4}}{(2d_0-3)^2} B_{d_0-2}(1)^2.$$

5   Simulations and An Empirical Illustration

This section reports some simulations that were conducted to examine the finite sample performance of the modified local Whittle estimator (hereafter, modified estimator) and the unmodified local Whittle estimator (hereafter, unmodified estimator), though no theoretical results are available yet for the unmodified estimator. We generate $I(d)$ processes according to (3) with $X_0 = 0$ and $u_t \sim iidN(0,1)$. The bias, standard deviation, and mean squared error (MSE) were computed using 1,000 replications. Sample size and m were chosen to be $n = 500$ and $m = n^{\alpha}$ with $\alpha = 0.55$, 0.65, and 0.75, respectively.

Tables 1 and 2 show the simulation results. For values of d smaller than 0.5, the modified estimator has positive bias, and the bias decreases as m increases. This confirms the theoretical result that a large value of m is required to achieve consistency when $d < 0.5$. For all values of d, its standard deviation is larger than the theoretical value, and becomes very large when $d = 0.2$. The unmodified estimator has little bias when $d \le 1.0$, but has a large negative bias and larger variance when $d \ge 1.2$ (see also Velasco (1999a)). For the value $0.6 \le d \le 1.0$, the variance of the two estimators are almost equal. In sum, the modified estimator gives better estimates of d unless there is a strong prior belief that the value of d is smaller than 0.5.

Table 1. Modified local Whittle estimator: $n = 500$, $m = n^\alpha$

	$\alpha = 0.55$ $(m = 30)$ Theoretical s.d.= 0.0913			$\alpha = 0.65$ $(m = 56)$ Theoretical s.d.= 0.0668			$\alpha = 0.75$ $(m = 105)$ Theoretical s.d.= 0.0488		
	bias	s.d.	MSE	bias	s.d.	MSE	bias	s.d.	MSE
$d = 0.2$	0.1325	0.1608	0.0434	0.0939	0.1157	0.0222	0.0634	0.0837	0.0110
$d = 0.4$	0.0445	0.1278	0.0183	0.0269	0.0877	0.0084	0.0111	0.0621	0.0040
$d = 0.6$	0.0018	0.1163	0.0135	−0.0016	0.0784	0.0062	−0.0092	0.0530	0.0029
$d = 0.8$	−0.0146	0.1111	0.0126	−0.0124	0.0774	0.0061	−0.0186	0.0542	0.0033
$d = 1.0$	−0.0133	0.1131	0.0130	−0.0116	0.0762	0.0059	−0.0212	0.0518	0.0031
$d = 1.2$	−0.0122	0.1125	0.0128	−0.0139	0.0752	0.0058	−0.0262	0.0512	0.0033
$d = 1.4$	−0.0143	0.1201	0.0146	−0.0143	0.0788	0.0064	−0.0279	0.0555	0.0039
$d = 1.6$	0.0015	0.1200	0.0144	−0.0045	0.0806	0.0065	−0.0246	0.0551	0.0036
$d = 1.8$	0.0203	0.1219	0.0153	0.0112	0.0809	0.0067	−0.0145	0.0586	0.0036

Note: The theoretical s.d. is valid for $0.6 \le d \le 1.6$.

Table 2. Local Whittle estimator: $n = 500$, $m = n^\alpha$

	$\alpha = 0.55$ $(m = 30)$ Theoretical s.d.= 0.0913			$\alpha = 0.65$ $(m = 56)$ Theoretical s.d.= 0.0668			$\alpha = 0.75$ $(m = 105)$ Theoretical s.d.= 0.0488		
	bias	s.d.	MSE	bias	s.d.	MSE	bias	s.d.	MSE
$d = 0.2$	−0.0147	0.1151	0.0135	−0.0091	0.0773	0.0061	−0.0080	0.0545	0.0030
$d = 0.4$	−0.0015	0.1146	0.0131	−0.0043	0.0770	0.0059	−0.0101	0.0525	0.0029
$d = 0.6$	0.0042	0.1161	0.0135	0.0018	0.0789	0.0062	−0.0054	0.0544	0.0030
$d = 0.8$	0.0138	0.1143	0.0132	0.0127	0.0805	0.0066	0.0024	0.0588	0.0035
$d = 1.0$	−0.0103	0.1048	0.0111	−0.0098	0.0695	0.0049	−0.0204	0.0469	0.0026
$d = 1.2$	−0.1127	0.1079	0.0244	−0.1211	0.0825	0.0215	−0.1400	0.0712	0.0247
$d = 1.4$	−0.2933	0.1265	0.1020	−0.3128	0.1094	0.1098	−0.3399	0.0994	0.1254
$d = 1.6$	−0.4953	0.1482	0.2673	−0.5191	0.1330	0.2872	−0.5494	0.1176	0.3157
$d = 1.8$	−0.7124	0.1533	0.5310	−0.7370	0.1314	0.5605	−0.7666	0.1104	0.5999

Note: The theoretical s.d. is the one for the modified Whittle estimator.

Figure 1 plots the empirical probability distribution function of the modified and unmodified estimator for the values of $d = 0.3, 0.7, 1.3, 1.9$. The sample size and m were chosen as $n = 500$, $m = n^{0.65} = 56$, and 10,000 replications are used. When $d = 0.3$, the distribution of the modified estimator is positively biased, whereas both estimators have an approximately unbiased normal pdf when $d = 0.7$. When d is larger than unity, the modified estimator still works well, whereas the unmodified estimator appears to converge to 1. The convergence to the squared fractional Brownian motion, which theoretically will occur when $d = 1.9$, does not show up with this sample size.

As an empirical illustration, the modified local Whittle estimator was applied to the logarithm of the monthly UK wholesale price index. The series constituted 797 observations over the period 1885:1-1951:5. The first panel of Figure 2 graphs the series. The second panel of Figure 2 plots $\widehat{d}$ for different values of m (specifically, $m = n^{0.5}, ..., n^{0.65}$ were used). As m increases, $\widehat{d}$ initially increases and then stays
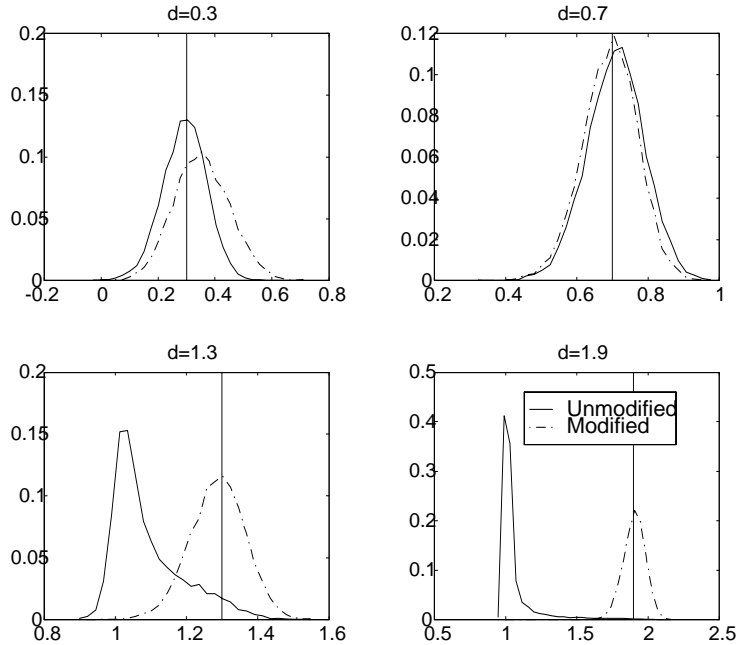
Figure 1: Modified and unmodified local Whittle estimates

around the same level. The estimates of the memory parameter over the stable area are in the region $(1.3, 1.4)$, indicating the series is $I(d)$ with $d > 1$. The third panel shows the residual fractionally differenced series $\widehat{u}_t = (1 - L)^{\widehat{d}}(X_t - X_0)$, where $\widehat{d}$ is the estimate with $m = n^{0.6}$. The spectral density estimates of $\widehat{u}_t$ are displayed in the fourth panel.

6   Conclusion

This paper explores the properties of a new semiparametric estimator, the modified local Whittle estimator, of the memory parameter in models of fractional integration. An alternate model of fractionally integrated processes that has some advantages as a generating mechanism is employed, and some new representation theory for the discrete Fourier transform of a fractional process is used to assist in the analysis. The new estimator is simple and convenient to use and involves only a minor adjustment over the well known local Whittle estimator. The limit theory for the modified estimator covers a range of values of d that is commonly encountered in applied work with economic data and is the same as that which is known to apply to the local Whittle estimator in the stationary range. It is therefore more efficient than the modified log periodogram regression estimator analysed in Kim and Phillips (1999), which is also suitable for use over a similar range of nonstationary values of d.

As suggested in Phillips (1999), a further possibility is to use the exact form of the discrete Fourier transform (8) in constructing the local Whittle likelihood. Such

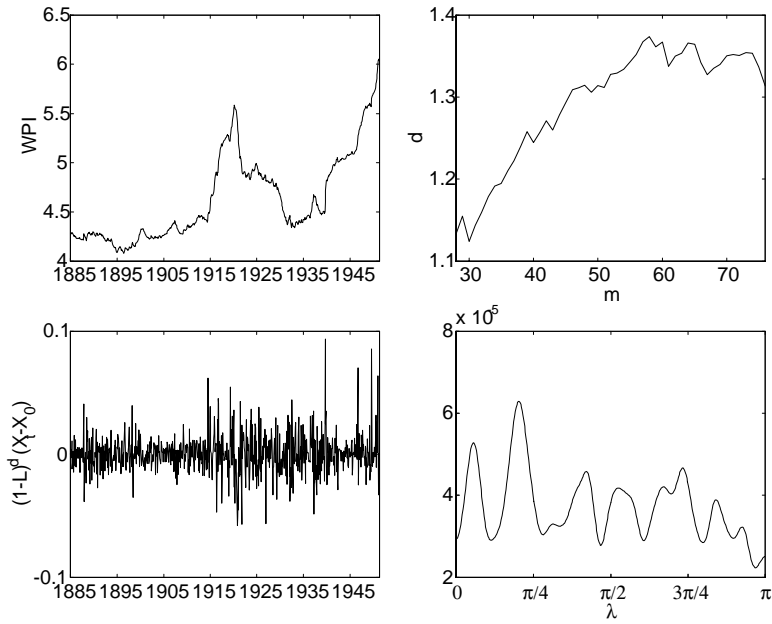Figure 2: Wholesale price index data and estimates of d

a likelihood does not rely on approximations of the discrete Fourier transform and may therefore be expected to provide good semiparametric estimates for all values of d. However, this method involves much more demanding computations than the modified Whittle estimator discussed here and an asymptotic theory is yet to be worked out.

15

7 Appendix A: Technical Lemmas

This section provides technical lemmas that are useful in the evaluation of the modified discrete Fourier transform on frequencies $\lambda_s = \frac{2\pi s}{n} \to 0$. The lemmas are divided into two groups. The first gives approximate representations of the sinusoidal polynomials $D_n\left(e^{i\lambda_s};d\right)$ and $\widehat{f}_{\lambda p}$ in (8). The other gives asymptotic approximations to the term $\widetilde{U}_{\lambda n}\left(f\right)$ and X_n in (8).

7.1 Component Approximations (deterministic part)

7.2 Lemma

For $f > -1$ and $\lambda_s = \frac{2\pi s}{n} \to 0$,

$$D_n\left(e^{i\lambda_s};f\right) = \left(1 - e^{i\lambda_s}\right)^f + O\left(n^{-f}s^{-1}\right), \tag{15}$$

uniformly in s.

7.3 Proof

$$
\begin{aligned}
D_n\left(e^{i\lambda_s};f\right) &= \sum_{k=0}^{n} \frac{(-f)_k}{k!} e^{ik\lambda_s} \\
&= \sum_{k=0}^{\infty} \frac{(-f)_k}{k!} e^{ik\lambda_s} - \sum_{k=n+1}^{\infty} \frac{(-f)_k}{k!} e^{ik\lambda_s} \\
&= {}_2F_1\left(-f,1;1;e^{i\lambda_s}\right) - \frac{1}{\Gamma(-f)} \sum_{k=n+1}^{\infty} k^{-f-1} e^{ik\lambda_s} + O\left(\sum_{k=n+1}^{\infty} k^{-f-2}\right) \tag{16}
\end{aligned}
$$

where the third line follows from the fact that (Erdélyi, 1953, p.47)

$$\frac{(-f)_k}{k!} = \frac{\Gamma(-f+k)}{\Gamma(-f)\,\Gamma(k+1)} = \frac{1}{\Gamma(-f)} k^{-f-1}\left(1 + O\left(k^{-1}\right)\right). \tag{17}$$

Because $f > -1$ and $s \neq 0$, the first term in (16) converges and equals to $\left(1 - e^{i\lambda_s}\right)^f$ (Erdélyi, 1953, p.57). For the second term in (16), by Theorem 2.2 of Zygmund (1959) we have

$$\left|\sum_{k=n+1}^{\infty} k^{-f-1} e^{2\pi i s k / n}\right| \leq (n+1)^{-f-1} \max_N \left|\sum_{k=n+1}^{n+N} e^{2\pi i s k / n}\right|,$$

and the ordinary summation formula gives

$$\left|\sum_{k=n+1}^{n+N} e^{2\pi i s k / n}\right| = \left|\sum_{k=1}^{N} e^{2\pi i s k / n}\right| = O\left(\frac{n}{s}\right).$$

uniformly in N. The third term in (16) is $O\left(n^{-f}s^{-1}\right)$ because $\sum_{k=n+1}^{\infty} k^{-f-2} = O\left(n^{-f-1}\right)$ and $s/n \to 0$. ∎

16

7.4 Lemma

For $\lambda \downarrow 0$, uniformly in λ,

$$\lambda^{-f}\left(1 - e^{i\lambda}\right)^f \;=\; e^{-\frac{\pi}{2}fi} + O\left(\lambda\right), \qquad (18)$$

$$\lambda^{-f}\left(1 - e^{-i\lambda}\right)^f \;=\; e^{\frac{\pi}{2}fi} + O\left(\lambda\right).$$

7.5 Corollary

For $f > -1$ and $\lambda_s = \frac{2\pi s}{n} \to 0$,

$$\lambda_s^{-f} D_n\left(e^{i\lambda_s}; f\right) \;=\; \lambda_s^{-f}\left(1 - e^{i\lambda_s}\right)^f + \lambda_s^{-f} O\left(n^{-f} s^{-1}\right)$$

$$=\; e^{-\frac{\pi}{2}fi} + O\left(\lambda_s\right) + O\left(s^{-1-f}\right), \qquad (19)$$

uniformly in s.

7.6 Proof

Note that $\left|1 - e^{\pm i\lambda}\right| = \left|2\sin\left(\frac{\lambda}{2}\right)\right|$. An elementary geometric argument (see the attached figure) implies that, for $0 \le \lambda < \pi$,

$$\arg\left(1 - e^{i\lambda}\right) = \frac{\lambda - \pi}{2} \quad \text{and} \quad \arg\left(1 - e^{-i\lambda}\right) = \frac{\pi - \lambda}{2}.$$

Hence we can write $\left(1 - e^{i\lambda}\right)^f$ in polar form as

$$\left(1 - e^{i\lambda}\right)^f \;=\; \left\{\left|2\sin\left(\frac{\lambda}{2}\right)\right| e^{i\left(\frac{\lambda}{2} - \frac{\pi}{2}\right)}\right\}^f$$

$$=\; \left|2\sin\left(\frac{\lambda}{2}\right)\right|^f e^{if\left(\frac{\lambda}{2} - \frac{\pi}{2}\right)}$$

$$=\; \left|2\sin\left(\frac{\lambda}{2}\right)\right|^f \left[\cos\left(\frac{\lambda f}{2} - \frac{\pi f}{2}\right) + i\sin\left(\frac{\lambda f}{2} - \frac{\pi f}{2}\right)\right].$$

Taylor expansion yields

$$2\sin\left(\frac{\lambda}{2}\right) \;=\; 2\cos\left(0\right)\cdot\frac{\lambda}{2} - \frac{1}{3}\cos\left(\tilde{\lambda}\right)\cdot\left(\frac{\lambda}{2}\right)^3 = \lambda + O\left(\lambda^3\right),$$

$$\cos\left(\frac{\lambda f}{2} - \frac{\pi f}{2}\right) \;=\; \cos\left(-\frac{\pi f}{2}\right) - \sin\left(\tilde{\lambda}\right)\cdot\left(\frac{\lambda f}{2}\right) = \cos\left(-\frac{\pi f}{2}\right) + O\left(\lambda\right),$$

$$\sin\left(\frac{\lambda f}{2} - \frac{\pi f}{2}\right) \;=\; \sin\left(-\frac{\pi f}{2}\right) + \cos\left(\tilde{\lambda}\right)\cdot\left(\frac{\lambda f}{2}\right) = \sin\left(-\frac{\pi f}{2}\right) + O\left(\lambda\right),$$

17

and all the reminder terms are uniform in λ. Therefore, uniformly in λ,

$$
\begin{aligned}
\lambda^{-f}\left(1-e^{i\lambda}\right)^{f} &= \lambda^{-f}\left(\lambda+O\left(\lambda^{3}\right)\right)^{f}\left[\cos\left(-\frac{\pi f}{2}\right)+O\left(\lambda\right)+i\sin\left(-\frac{\pi f}{2}\right)+iO\left(\lambda\right)\right] \\
&= \left(1+O\left(\lambda^{2}\right)\right)^{f}\left[e^{-\frac{\pi}{2}fi}+O\left(\lambda\right)\right] \\
&= \left(1+O\left(\lambda^{2}\right)\right)\left[e^{-\frac{\pi}{2}fi}+O\left(\lambda\right)\right] \\
&= e^{-\frac{\pi}{2}fi}+O\left(\lambda\right).
\end{aligned}
$$

The approximation of $\lambda^{-f}\left(1-e^{-i\lambda}\right)^{f}$ follows the same line of argument. $\blacksquare$

7.7 Lemma

Uniformly in p and s,

$$
(a) \quad \widetilde{f}_{\lambda_{s}p} = \begin{cases} O\left(p^{-f}\right), & \text{for } f>0, \\ O\left(n^{-f}\right), & \text{for } f\in(-1,0), \end{cases} \tag{20}
$$

$$
(b) \quad \widetilde{f}_{\lambda_{s}p} = O\left(\frac{n}{p^{f+1}s}\right), \quad \text{for } f>-1. \tag{21}
$$

7.8 Proof

The approximation (17) yields

$$
\begin{aligned}
\widetilde{f}_{\lambda_{s}p} &= \frac{1}{\Gamma\left(-f\right)}\sum_{k=p+1}^{n}k^{-f-1}\left(1+O\left(k^{-1}\right)\right)e^{2\pi isk/n} \\
&= \frac{1}{\Gamma\left(-f\right)}\sum_{k=p+1}^{n}k^{-f-1}e^{2\pi isk/n}+O\left(\sum_{k=p+1}^{n}k^{-f-2}\right).
\end{aligned}
$$

Using the results derived in the proof of Lemma 7.2, we obtain

$$
\sum_{k=p+1}^{n}k^{-f-1}\left(1+O\left(k^{-1}\right)\right)e^{2\pi isk/n}=O\left(\sum_{k=p+1}^{n}k^{-f-1}\right)=\begin{cases} O\left(p^{-f}\right), & \text{for } f>0, \\ O\left(n^{-f}\right), & \text{for } f\in(-1,0), \end{cases}
$$

giving part (a). Part (b) follows from

$$
\left|\sum_{k=p+1}^{n}k^{-f-1}e^{2\pi isk/n}\right| \leq (p+1)^{-f-1}\max_{N}\left|\sum_{k=p+1}^{p+N}e^{2\pi isk/n}\right|,
$$

$$
\left|\sum_{k=p+1}^{p+N}e^{2\pi isk/n}\right| = O\left(\frac{n}{s}\right),
$$

$$
\sum_{k=p+1}^{n}k^{-f-2} = O\left(p^{-f-1}\right)=O\left(\frac{n}{p^{f+1}s}\right),
$$

18

since $s \leq n$. ∎

7.9 Component Approximations (stochastic part)

The following lemmas give asymptotic approximations to the term $\widetilde{U}_{\lambda n}(f)$ and X_n in (8) when $\lambda_s = \frac{2\pi s}{n} \to 0$. Note that the stochastic order of magnitude of $\widetilde{U}_{\lambda n}(f)$ changes depending on the value of f.

7.10 Lemma

For $\lambda_s = \frac{2\pi s}{n} \to 0$,

$$\widetilde{U}_{\lambda_s n}(f) = C(1)\widetilde{\varepsilon}_{\lambda_s n}(f) + r_{s,n}(f),$$

where

$$\widetilde{\varepsilon}_{\lambda n}(f) = \sum_{p=0}^{n-1} \widetilde{f}_{\lambda p} e^{-ip\lambda} \varepsilon_{n-p},$$

and

$$E\left|r_{s,n}(f)\right|^2 = \begin{cases} O(1), & \text{for } f > 0, \\ O\left(n^{-2f}\right) = O\left(n^{2d-2}\right), & \text{for } f \in (-1,0), \end{cases}$$

uniformly in s.

7.11 Proof

Applying the BN decomposition

$$u_t = C(L)\varepsilon_t = C(1)\varepsilon_t - (1-L)\widetilde{\varepsilon}_t, \quad \widetilde{\varepsilon}_t = \sum_{j=0}^{\infty} \widetilde{c}_j \varepsilon_{t-j}, \quad \widetilde{c}_j = \sum_{s=j+1}^{\infty} c_s, \qquad (22)$$

to $\widetilde{U}_{\lambda_s n}(f)$ yields

$$
\begin{aligned}
\widetilde{U}_{\lambda_s n}(f) &= \sum_{p=0}^{n-1} \widetilde{f}_{\lambda_s p} e^{-ip\lambda_s} u_{n-p} \\
&= \sum_{p=0}^{n-1} \widetilde{f}_{\lambda_s p} e^{-ip\lambda_s} \left[C(1)\varepsilon_{n-p} - (1-L)\widetilde{\varepsilon}_{n-p}\right] \\
&= C(1)\widetilde{\varepsilon}_{\lambda_s n}(f) - \sum_{p=0}^{n-1} \widetilde{f}_{\lambda_s p} e^{-ip\lambda_s} (1-L)\widetilde{\varepsilon}_{n-p}.
\end{aligned}
$$

Note that the assumption $\sum_{j=0}^{\infty} j\,|c_j| < \infty$ implies that $\sum_{j=0}^{\infty} \widetilde{c}_j^2 < \infty$ hence $E\left[\widetilde{\varepsilon}_t\right]^2 < \infty$. Rewrite the second term as follows:

$$\sum_{p=0}^{n-1} \widetilde{f}_{\lambda_s p} e^{-ip\lambda_s} (1-L)\widetilde{\varepsilon}_{n-p}$$

19

$$= \quad \widetilde{f}_{\lambda_s 0} \widetilde{\varepsilon}_n + \sum_{p=1}^{n-1} \widetilde{f}_{\lambda_s p} e^{-ip\lambda_s} \widetilde{\varepsilon}_{n-p} - \sum_{p=1}^{n-1} \widetilde{f}_{\lambda_s(p-1)} e^{-i(p-1)\lambda_s} \widetilde{\varepsilon}_{n-p} - \widetilde{f}_{\lambda_s(n-1)} e^{-i(n-1)\lambda_s} \widetilde{\varepsilon}_0$$

$$= \quad \sum_{p=1}^{n-1} \left[\widetilde{f}_{\lambda_s p} e^{-ip\lambda_s} - \widetilde{f}_{\lambda_s(p-1)} e^{-i(p-1)\lambda_s} \right] \widetilde{\varepsilon}_{n-p} + \widetilde{f}_{\lambda_s 0} \widetilde{\varepsilon}_n - \widetilde{f}_{\lambda_s(n-1)} e^{-i(n-1)\lambda_s} \widetilde{\varepsilon}_0.$$

In view of the results in the proof of Lemma 7.7,

$$\begin{cases} \widetilde{f}_{\lambda_s 0} = O\left(1\right), \quad \widetilde{f}_{\lambda_s(n-1)} = O\left(n^{-f-1}\right), & \text{for } f > 0, \\ \widetilde{f}_{\lambda_s 0} = O\left(n^{-f}\right), \quad \widetilde{f}_{\lambda_s(n-1)} = O\left(n^{-f-1}\right), & \text{for } f \in (-1, 0). \end{cases}$$

Hence, using the fact that $E\left[\widetilde{\varepsilon}_t\right]^2 < \infty$, we have

$$\begin{cases} E\left|\widetilde{f}_{\lambda_s 0} \widetilde{\varepsilon}_n\right|^2 = O\left(1\right), \quad E\left|\widetilde{f}_{\lambda_s(n-1)} e^{-i(n-1)\lambda_s} \widetilde{\varepsilon}_0\right|^2 = O\left(n^{-2f-2}\right), & \text{for } f > 0, \\ E\left|\widetilde{f}_{\lambda_s 0} \widetilde{\varepsilon}_n\right|^2 = O\left(n^{-2f}\right), \quad E\left|\widetilde{f}_{\lambda_s(n-1)} e^{-i(n-1)\lambda_s} \widetilde{\varepsilon}_0\right|^2 = O\left(n^{-2f-2}\right), & \text{for } f \in (-1, 0). \end{cases}$$

>From Lemma 7.7,

$$\begin{aligned} \widetilde{f}_{\lambda_s p} e^{-ip\lambda_s} - \widetilde{f}_{\lambda_s(p-1)} e^{-i(p-1)\lambda_s} &= \quad \widetilde{f}_{\lambda_s p} \left[e^{-ip\lambda_s} - e^{-i(p-1)\lambda_s} \right] + e^{-i(p-1)\lambda_s} \left[\widetilde{f}_{\lambda_s p} - \widetilde{f}_{\lambda_s(p-1)} \right] \\ &= \quad \widetilde{f}_{\lambda_s p} e^{-ip\lambda_s} \left(1 - e^{i\lambda_s} \right) - e^{-i(p-1)\lambda_s} \frac{(-f)_p}{p!} e^{ip\lambda_s} \\ &= \quad O\left(\frac{n}{p^{f+1}s}\right) O\left(\frac{s}{n}\right) + O\left(p^{-f-1}\right) = O\left(p^{-f-1}\right). \end{aligned}$$

It follows that

$$E\left|\sum_{p=1}^{n-1} \left[\widetilde{f}_{\lambda_s p} e^{-ip\lambda_s} - \widetilde{f}_{\lambda_s(p-1)} e^{-i(p-1)\lambda_s} \right] \widetilde{\varepsilon}_{n-p}\right|^2 \leq \sum_{p=1}^{n-1} \sum_{q=1}^{n-1} O\left(p^{-f-1}\right) O\left(q^{-f-1}\right) E\left|\widetilde{\varepsilon}_{n-p} \widetilde{\varepsilon}_{n-q}\right|$$

$$= \begin{cases} O\left(1\right), & \text{for } f > 0, \\ O\left(n^{-2f}\right), & \text{for } f \in (-1, 0), \end{cases}$$

and the result follows from Loève's c_r inequality (Davidson, 1994, p.140). ∎

7.12   Lemma

For $f \in \left(0, \frac{1}{2}\right)$ and any number L such that $L \to \infty$ and $\frac{L}{n} \to 0$, the following hold uniformly in s:

$$\begin{aligned} (a) \quad E\left|\widetilde{\varepsilon}_{\lambda_s n}\left(f\right)\right|^2 &= \quad O\left(n^{1-2f}\right) = O\left(n^{2d-1}\right), \\ (b) \quad E\left|\widetilde{\varepsilon}_{\lambda_s n}\left(f\right)\right|^2 &= \quad O\left(L^{1-2f} + \frac{n}{s}L^{-2f}\right) = O\left(L^{2d-1} + \frac{n}{s}L^{2d-2}\right), \\ (c) \quad E\left|\widetilde{\varepsilon}_{\lambda_s n}\left(f\right)\right| &= \quad O\left(L^{\frac{1}{2}-f} + \left(\frac{n}{s}\right)^{\frac{1}{2}} L^{-f}\right) = O\left(L^{d-\frac{1}{2}} + \left(\frac{n}{s}\right)^{\frac{1}{2}} L^{d-1}\right). \end{aligned}$$

20

7.13 Proof

For part (a), it follows from Lemma 7.7 that

$$E\left|\widetilde{\varepsilon}_{\lambda_s n}(f)\right|^2 = \sum_{p=0}^{n-1}\left|\widetilde{f}_{\lambda_s p}e^{-ip\lambda_s}\right|^2 = O\left(\sum_{p=1}^{n-1}p^{-2f}\right) = O\left(n^{1-2f}\right).$$

For part (b), we write $\widetilde{\varepsilon}_{\lambda_s n}(f)$ as the sum of two components, the first involving $L+1$ components. Specifically,

$$\begin{aligned}
E\left|\widetilde{\varepsilon}_{\lambda_s n}(f)\right|^2 &= E\left|\sum_{p=0}^{L}\widetilde{f}_{\lambda_s p}e^{-ip\lambda_s}\varepsilon_{n-p} + \sum_{p=L+1}^{n-1}\widetilde{f}_{\lambda_s p}e^{-ip\lambda_s}\varepsilon_{n-p}\right|^2 \\
&\leq 2E\left|\sum_{p=0}^{L}\widetilde{f}_{\lambda_s p}e^{-ip\lambda_s}\varepsilon_{n-p}\right|^2 + 2E\left|\sum_{p=L+1}^{n-1}\widetilde{f}_{\lambda_s p}e^{-ip\lambda_s}\varepsilon_{n-p}\right|^2, \quad (23)
\end{aligned}$$

where the second line follows from Loève's c_r inequality. By Lemma 7.7, each of the terms in (23) is bounded by

$$\sum_{p=0}^{L}\left(\widetilde{f}_{\lambda_s p}\right)^2 = O\left(\sum_{p=1}^{L}p^{-2f}\right) = O\left(L^{1-2f}\right),$$

$$\sum_{p=L+1}^{n-1}\left(\widetilde{f}_{\lambda_s p}\right)^2 = O\left(\sum_{p=L+1}^{n-1}\frac{1}{p^f}\frac{n}{p^{f+1}s}\right) = O\left(\frac{n}{s}\sum_{p=L+1}^{n-1}\frac{1}{p^{2f+1}}\right) = O\left(\frac{n}{s}L^{-2f}\right).$$

For part (c), Minkowski's inequality yields

$$\begin{aligned}
E\left|\widetilde{\varepsilon}_{\lambda_s n}(f)\right| &\leq \left(E\left|\sum_{p=0}^{L}\widetilde{f}_{\lambda_s p}e^{-ip\lambda_s}\varepsilon_{n-p}\right|^2\right)^{1/2} + \left(E\left|\sum_{p=L+1}^{n-1}\widetilde{f}_{\lambda_s p}e^{-ip\lambda_s}\varepsilon_{n-p}\right|^2\right)^{1/2} \\
&= O\left(L^{\frac{1}{2}-f}\right) + O\left(\left(\frac{n}{s}\right)^{\frac{1}{2}}L^{-f}\right),
\end{aligned}$$

giving the required result. ∎

7.14 Lemma

(a) For $f \in (-1,0)$, the following holds uniformly in s:

$$E\left|\widetilde{\varepsilon}_{\lambda_s n}(f)\right|^2 = O\left(n^{1-2f}s^{-1}\right) = O\left(n^{2d-1}s^{-1}\right).$$

(b) For $f \in \left(-\frac{1}{2},0\right)$ and any number L such that $L \to \infty$ and $\frac{L}{n} \to 0$, the following holds uniformly in s:

$$E\left|\widetilde{\varepsilon}_{\lambda_s n}(f)\right|^2 = O\left(\frac{n^{1-f}}{s}L^{-f} + \frac{n^2}{s^2}L^{-2f-1}\right) = O\left(\frac{n^d}{s}L^{d-1} + \frac{n^2}{s^2}L^{2d-3}\right).$$

7.15 Proof

For part (a), using Lemma 7.7 we get

$$E\left|\widetilde{\varepsilon}_{\lambda_s n}\left(f\right)\right|^2 = \sum_{p=0}^{n-1}\left|\widetilde{f}_{\lambda_s p}e^{-ip\lambda_s}\right|^2 = O\left(\sum_{p=1}^{n-1}n^{-f}\frac{n}{p^{f+1}s}\right) = O\left(\frac{n^{1-f}}{s}\sum_{p=1}^{n-1}p^{-f-1}\right) = O\left(n^{1-2f}s^{-1}\right).$$

Part (b) is proved by the same argument as used in Lemma 7.12. Specifically, we have

$$E\left|\widetilde{\varepsilon}_{\lambda_s n}\left(f\right)\right|^2 \leq 2E\left|\sum_{p=0}^{L}\widetilde{f}_{\lambda_s p}e^{-ip\lambda_s}\varepsilon_{n-p}\right|^2 + 2E\left|\sum_{p=L+1}^{n-1}\widetilde{f}_{\lambda_s p}e^{-ip\lambda_s}\varepsilon_{n-p}\right|^2, \qquad (24)$$

and

$$\sum_{p=0}^{L}\left(\widetilde{f}_{\lambda_s p}\right)^2 = O\left(\sum_{p=1}^{L}n^{-f}\frac{n}{p^{f+1}s}\right) = O\left(\frac{n^{1-f}}{s}\sum_{p=1}^{L}p^{-f-1}\right) = O\left(\frac{n^{1-f}}{s}L^{-f}\right),$$

$$\sum_{p=L+1}^{n-1}\left(\widetilde{f}_{\lambda_s p}\right)^2 = O\left(\sum_{p=L+1}^{n-1}\frac{n^2}{p^{2f+2}s^2}\right) = O\left(\frac{n^2}{s^2}\sum_{p=L+1}^{n-1}p^{-2f-2}\right) = O\left(\frac{n^2}{s^2}L^{-2f-1}\right),$$

giving the required result. ∎

7.16 Lemma

(a) For $f \in \left[\frac{1}{2},1\right)$, the following holds:

$$E\left|\widetilde{\varepsilon}_{\lambda_s n}\left(f\right)\right|^2 = \begin{cases} O\left(\log n\right), & \text{for } f = \frac{1}{2}, \\ O\left(1\right), & \text{for } f \in \left(\frac{1}{2},1\right). \end{cases}$$

(b) For $f \in \left(-1,-\frac{1}{2}\right]$, the following holds uniformly in s:

$$E\left|\widetilde{\varepsilon}_{\lambda_s n}\left(f\right)\right|^2 = \begin{cases} O\left(n^{1-2f}s^{-2}\right) = O\left(n^{2d-1}s^{-2}\right), & \text{for } f \in \left(-1,-\frac{1}{2}\right), \\ O\left(s^{-2}n^2\log n\right), & \text{for } f = -\frac{1}{2}. \end{cases}$$

7.17 Proof

Using Lemma 7.7, we have, for part (a),

$$E\left|\widetilde{\varepsilon}_{\lambda_s n}\left(f\right)\right|^2 = \sum_{p=0}^{n-1}\left|\widetilde{f}_{\lambda_s p}e^{-ip\lambda_s}\right|^2 = O\left(\sum_{p=1}^{n-1}p^{-2f}\right) = \begin{cases} O\left(\log n\right), & \text{for } f = \frac{1}{2}, \\ O\left(1\right), & \text{for } f = \left(\frac{1}{2},1\right), \end{cases}$$

and for part (b),

$$\begin{aligned} E\left|\widetilde{\varepsilon}_{\lambda_s n}\left(f\right)\right|^2 &= \sum_{p=0}^{n-1}\left|\widetilde{f}_{\lambda_s p}e^{-ip\lambda_s}\right|^2 = O\left(\sum_{p=1}^{n-1}\frac{n^2}{p^{2f+2}s^2}\right) = O\left(\frac{n^2}{s^2}\sum_{p=1}^{n-1}p^{-2f-2}\right) \\ &= \begin{cases} O\left(n^{1-2f}s^{-2}\right), & \text{for } f = \left(-1,-\frac{1}{2}\right), \\ O\left(s^{-2}n^2\log n\right), & \text{for } f = -\frac{1}{2}, \end{cases} \end{aligned}$$

22

giving the required result. ∎

7.18 Lemma

For $d \in \left(\frac{1}{2}, 2\right)$ and $1 \leq t \leq n,$ we have

(a) $X_t - X_0 = C(1) X_t^\varepsilon + r_t,$ where $X_t^\varepsilon = \sum_{k=0}^{t-1} \frac{(d)_k}{k!} \varepsilon_{t-k}$ and

$$E|r_t|^2 = \begin{cases} O(1), & \text{for } d \in \left(\frac{1}{2}, 1\right], \\ O\left(t^{2d-2}\right), & \text{for } d \in (1, 2), \end{cases}$$

uniformly in t,

(b) $E|X_t^\varepsilon|^2 = O\left(n^{2d-1}\right),$

(c) $E|X_t|^2 = O\left(n^{2d-1}\right).$

7.19 Proof

When $d = 1$, see Phillips and Solo (1992) page 976. For $d \neq 1$, applying the BN decomposition to u_t and substituting into (3) yields

$$X_t - X_0 = \sum_{k=0}^{t-1} \frac{(d)_k}{k!} u_{t-k} = \sum_{k=0}^{t-1} \frac{(d)_k}{k!} \left[C(1)\varepsilon_{t-k} - (1-L)\widetilde{\varepsilon}_{t-k}\right] = C(1) X_t^\varepsilon - \sum_{k=0}^{t-1} \frac{(d)_k}{k!}(1-L)\widetilde{\varepsilon}_{t-k}.$$

Rearrange the second term as follows:

$$
\begin{aligned}
\sum_{k=0}^{t-1} \frac{(d)_k}{k!}(1-L)\widetilde{\varepsilon}_{t-k} &= \sum_{k=0}^{t-1} \frac{(d)_k}{k!}\widetilde{\varepsilon}_{t-k} - \sum_{k=0}^{t-1} \frac{(d)_k}{k!}\widetilde{\varepsilon}_{t-k-1} \\
&= \frac{(d)_0}{0!}\widetilde{\varepsilon}_t + \sum_{k=1}^{t-1} \frac{(d)_k}{k!}\widetilde{\varepsilon}_{t-k} - \sum_{k=1}^{t-1} \frac{(d)_{k-1}}{(k-1)!}\widetilde{\varepsilon}_{t-k} - \frac{(d)_{t-1}}{(t-1)!}\widetilde{\varepsilon}_0 \\
&= \sum_{k=1}^{t-1} \frac{(d-1)_k}{k!}\widetilde{\varepsilon}_{t-k} + \widetilde{\varepsilon}_t - \frac{(d)_{t-1}}{(t-1)!}\widetilde{\varepsilon}_0, \quad (25)
\end{aligned}
$$

where the fourth line follows from the fact that

$$
\begin{aligned}
\frac{(d)_k}{k!} - \frac{(d)_{k-1}}{(k-1)!} &= \frac{1}{\Gamma(d)}\left[\frac{\Gamma(d+k)}{\Gamma(k+1)} - \frac{\Gamma(d+k-1)}{\Gamma(k)}\right] \\
&= \frac{\Gamma(d+k-1)}{\Gamma(d)\Gamma(k+1)}\left[(d+k-1) - k\right] \\
&= \frac{\Gamma(d+k-1)}{\Gamma(d-1)\Gamma(k+1)} = \frac{(d-1)_k}{k!}. \quad (26)
\end{aligned}
$$

The mean square of the first term in (25) is

$$E\left[\sum_{k=1}^{t-1} k^{d-2}\widetilde{\varepsilon}_{t-k}\right]^2 = E\left[\sum_{k=1}^{t-1} k^{d-2}\widetilde{\varepsilon}_{t-k}\right]\left[\sum_{l=1}^{t-1} l^{d-2}\widetilde{\varepsilon}_{t-l}\right] = \sum_{k=1}^{t-1} k^{d-2}\sum_{l=1}^{t-1} l^{d-2} E\left(\widetilde{\varepsilon}_{t-k}\widetilde{\varepsilon}_{t-l}\right),$$

$$(27)$$

23

and the result follows from the fact that $E\tilde{\varepsilon}^2_{t-k} < \infty$ and Cauchy-Schwartz inequality. Trivially $E\tilde{\varepsilon}^2_t = O(1)$ and $E\left|\frac{(d)_{t-1}}{(t-1)!}\tilde{\varepsilon}_0\right|^2 = O(1)$, and part (a) follows from Loève's c_r inequality.

For part (b) and (c), $E[X^\varepsilon_t]^2$ is bounded by $\sigma^2 \sum_{k=1}^{t-1} k^{2(d-1)} = O\left(t^{2d-1}\right) = O\left(n^{2d-1}\right)$, giving the required result. $\blacksquare$

8   Appendix B: Proofs

8.1   Proof of Lemma 2.2 and Lemma 2.3

See Theorem 2.2 and 2.7 of Phillips (1999). $\blacksquare$

8.2   Proof of Lemma 2.5

Multiplying both sides of (8) by $\lambda_s^d \left(1 - e^{i\lambda_s}\right)^{-1}$ yields

$$\lambda_s^d w_x\left(\lambda_s\right) + \frac{\lambda_s^d}{1 - e^{i\lambda_s}} \frac{e^{i\lambda_s}\left(X_n - X_0\right)}{\sqrt{2\pi n}} = \frac{\lambda_s^d D_n\left(e^{i\lambda_s}; f\right)}{1 - e^{i\lambda_s}} w_u\left(\lambda_s\right) - \frac{\lambda_s^d}{1 - e^{i\lambda_s}} \frac{\widetilde{U}_{\lambda_s n}\left(f\right)}{\sqrt{2\pi n}}. \tag{28}$$

Using Lemma 7.4 and Corollary 7.5, we have

$$\frac{\lambda_s^d D_n\left(e^{i\lambda_s}; f\right)}{1 - e^{i\lambda_s}} = \frac{\lambda_s^{-f} D_n\left(e^{i\lambda_s}; f\right)}{\lambda_s^{-1}\left(1 - e^{i\lambda_s}\right)} = \frac{e^{-\frac{\pi}{2}fi} + O\left(\lambda_s\right) + O\left(s^{-1-f}\right)}{e^{-\frac{\pi}{2}i} + O\left(\lambda_s\right)}$$

$$= e^{\frac{\pi}{2}di} + O\left(\lambda_s\right) + O\left(s^{-1-f}\right), \tag{29}$$

$$\frac{\lambda_s^d}{1 - e^{i\lambda_s}} = O\left(\lambda_s^{d-1}\right). \tag{30}$$

Since $E\varepsilon^4 < \infty$ and $\sum_{j=0}^{\infty} j\left|c_j\right| < \infty$, $w_u\left(\lambda_s\right)$ can be approximated as follows (Hannan, 1970, p.248):

$$w_u\left(\lambda_s\right) = C\left(e^{i\lambda_s}\right) w_\varepsilon\left(\lambda_s\right) + r_n\left(\lambda_s\right),$$

where

$$E\left|w_\varepsilon\left(\lambda_s\right)\right|^2 = \frac{\sigma^2}{2\pi}, \quad E\left|r_n\left(\lambda_s\right)\right|^2 = O\left(n^{-2}\right),$$

uniformly in s. $C\left(e^{i\lambda}\right)$ is differentiable with a bounded derivative because $\sum_{j=0}^{\infty} j\left|c_j\right| < \infty$. Therefore, Taylor expansion gives $C\left(e^{i\lambda_s}\right) = C\left(1\right) + O\left(\lambda_s\right)$ uniformly in s. It follows that

$$w_u\left(\lambda_s\right) = C\left(1\right) w_\varepsilon\left(\lambda_s\right) + O\left(\lambda_s\right) w_\varepsilon\left(\lambda_s\right) + r_n\left(\lambda_s\right) = C\left(1\right) w_\varepsilon\left(\lambda_s\right) + r_n^1\left(\lambda_s\right), \tag{31}$$

where $E\left|r_n^1\left(\lambda_s\right)\right|^2 = O\left(\lambda_s^2\right)$. Combining (??) and (??), we obtain the approximation of the first term in (??), viz.

$$\frac{\lambda_s^d D_n\left(e^{i\lambda_s}; f\right)}{1 - e^{i\lambda_s}} w_u\left(\lambda_s\right)$$

24

$$
\begin{aligned}
&= e^{\frac{\pi}{2}di}C\left(1\right)w_\varepsilon\left(\lambda_s\right)+e^{\frac{\pi}{2}di}r_n^1\left(\lambda_s\right)+\left[O\left(\lambda_s\right)+O\left(s^{-1-f}\right)\right]\left[C\left(1\right)w_\varepsilon\left(\lambda_s\right)+r_n^1\left(\lambda_s\right)\right]\\
&= e^{\frac{\pi}{2}di}C\left(1\right)w_\varepsilon\left(\lambda_s\right)+r_n^a\left(\lambda_s\right)+r_n^2\left(\lambda_s\right),
\end{aligned}
$$

where $E\left|r_n^a\left(\lambda_s\right)\right|^2=O\left(\lambda_s^2\right)$ and $E\left|r_n^2\left(\lambda_s\right)\right|^2=O\left(s^{-2-2f}\right)=O\left(s^{2d-4}\right)$.

Now we derive the bound of the second term in (??). It follows from Lemma 7.10 and (??) that

$$
\frac{\lambda_s^d}{1-e^{i\lambda_s}}\frac{\widetilde{U}_{\lambda_s n}\left(f\right)}{\sqrt{2\pi n}}=\frac{\lambda_s^d}{1-e^{i\lambda_s}}\frac{C\left(1\right)\widetilde{\varepsilon}_{\lambda_s n}\left(f\right)}{\sqrt{2\pi n}}+\frac{\lambda_s^d}{1-e^{i\lambda_s}}\frac{r_{s,n}\left(f\right)}{\sqrt{2\pi n}},
$$

where

$$
E\left|\frac{\lambda_s^d}{1-e^{i\lambda_s}}\frac{r_{s,n}\left(f\right)}{\sqrt{2\pi n}}\right|^2=\left\{\begin{array}{ll}O\left(\lambda_s^{2d-2}n^{-1}\right)=O\left(s^{2d-2}n^{1-2d}\right), & \text{for } d\in\left(\frac{1}{2},1\right),\\ O\left(\lambda_s^{2d-2}n^{2d-3}\right)=O\left(s^{2d-2}n^{-1}\right), & \text{for } d\in\left(1,\frac{3}{2}\right),\end{array}\right.
$$

uniformly in s, giving part (a).

For part (b), using Lemma 7.12 (a) and 7.14 (a), we get

$$
E\left|\frac{\lambda_s^d C\left(1\right)\widetilde{\varepsilon}_{\lambda_s n}\left(f\right)}{1-e^{i\lambda_s}}\frac{\widetilde{\varepsilon}_{\lambda_s n}\left(f\right)}{\sqrt{2\pi n}}\right|^2=\left\{\begin{array}{ll}O\left(\lambda_s^{2d-2}n^{2d-2}\right)=O\left(s^{2d-2}\right), & \text{for } d\in\left(\frac{1}{2},1\right),\\ O\left(\lambda_s^{2d-2}n^{2d-2}s^{-1}\right)=O\left(s^{2d-3}\right), & \text{for } d\in\left(1,\frac{3}{2}\right).\end{array}\right.
$$

It follows from Minkowski's inequality that

$$
E\left|r_n^2\left(\lambda_s\right)+\frac{\lambda_s^d}{1-e^{i\lambda_s}}\frac{\widetilde{U}_{\lambda_s n}\left(f\right)}{\sqrt{2\pi n}}\right|^2=\left\{\begin{array}{ll}O\left(s^{2d-2}\right), & \text{for } d\in\left(\frac{1}{2},1\right),\\ O\left(s^{2d-3}\right), & \text{for } d\in\left(1,\frac{3}{2}\right),\end{array}\right.
$$

uniformly in s, giving the required result.

For part (c), a straightforward application of Lemma 2.3 (b) yields

$$
\begin{aligned}
\lambda_s w_x\left(\lambda_s\right)+\frac{\lambda_s}{1-e^{i\lambda_s}}\frac{e^{i\lambda_s}\left(X_n-X_0\right)}{\sqrt{2\pi n}} &= \frac{\lambda_s}{1-e^{i\lambda_s}}w_u\left(\lambda_s\right) \quad\quad\quad (32)\\
&= \left(e^{\frac{\pi}{2}i}+O\left(\lambda_s\right)\right)\left[C\left(1\right)w_\varepsilon\left(\lambda_s\right)+r_n^1\left(\lambda_s\right)\right]\\
&= iC\left(1\right)w_\varepsilon\left(\lambda_s\right)+O\left(\lambda_s\right)\left[C\left(1\right)w_\varepsilon\left(\lambda_s\right)+r_n^1\left(\lambda_s\right)\right]+e^{\frac{\pi}{2}i}r_n^1\left(\lambda_s\right)\\
&= iC\left(1\right)w_\varepsilon\left(\lambda_s\right)+r_{s,n}^c.
\end{aligned}
$$

For part (d), using Lemmas 7.10 and 7.16, we have

$$
E\left|\frac{\lambda_s^d}{1-e^{i\lambda_s}}\frac{r_{s,n}\left(f\right)}{\sqrt{2\pi n}}\right|^2=O\left(\lambda_s^{2d-2}n^{-1}\right)=O\left(s^{2d-2}n^{1-2d}\right),
$$

$$
E\left|\frac{\lambda_s^d C\left(1\right)\widetilde{\varepsilon}_{\lambda_s n}\left(f\right)}{1-e^{i\lambda_s}}\frac{\widetilde{\varepsilon}_{\lambda_s n}\left(f\right)}{\sqrt{2\pi n}}\right|^2=\left\{\begin{array}{ll}O\left(\lambda_s^{2d-2}n^{-1}\right)=O\left(s^{2d-2}n^{1-2d}\right), & \text{for } d\in\left(0,\frac{1}{2}\right),\\ O\left(\lambda_s^{2d-2}n^{-1}\log n\right)=O\left(s^{2d-2}n^{1-2d}\log n\right), & \text{for } d=\frac{1}{2}.\end{array}\right.
$$

It follows that

$$
E\left|r_n^2(\lambda_s) + \frac{\lambda_s^d}{1-e^{i\lambda_s}}\frac{\widetilde{U}_{\lambda_s n}(f)}{\sqrt{2\pi n}}\right|^2 = O\left(s^{2d-2}n^{1-2d}\log n\right),
$$

giving the required result.

For part (e), a similar calculation yields

$$
E\left|\frac{\lambda_s^d}{1-e^{i\lambda_s}}\frac{r_{s,n}(f)}{\sqrt{2\pi n}}\right|^2 = O\left(\lambda_s^{2d-2}n^{2d-3}\right) = O\left(s^{2d-2}n^{-1}\right),
$$

$$
E\left|\frac{\lambda_s^d C(1)}{1-e^{i\lambda_s}}\frac{\widetilde{\varepsilon}_{\lambda_s n}(f)}{\sqrt{2\pi n}}\right|^2
$$
$$
= \begin{cases} O\left(\lambda_s^{2d-2}n^{2d-2}s^{-2}\right) = O\left(s^{2d-4}\right), & \text{for } d \in \left(\frac{3}{2},2\right), \\ O\left(\lambda_s^{2d-2}s^{-2}n\log n\right) = O\left(s^{2d-4}n^{3-2d}\log n\right) = O\left(s^{2d-4}\log n\right), & \text{for } d = \frac{3}{2}. \end{cases}
$$

Thus

$$
E\left|r_n^2(\lambda_s) + \frac{\lambda_s^d C(1)}{1-e^{i\lambda_s}}\frac{\widetilde{\varepsilon}_{\lambda_s n}(f)}{\sqrt{2\pi n}}\right|^2 = O\left(s^{2d-4}\log n\right),
$$

and the stated result follows. ∎

8.3 Proof of Theorem 3.1

We follow the general approach developed by Robinson (1995b) for the stationary case. Define $G(d) = G_0\frac{1}{m}\sum_{j=1}^m \lambda_j^{2(d-d_0)}$ and $S(d) = R(d) - R(d_0)$. Rewrite $S(d)$ as follows:

$$
\begin{aligned}
S(d) &= R(d) - R(d_0) \\
&= \log\widehat{G}(d) - \log\widehat{G}(d_0) - (2d-2d_0)\frac{1}{m}\sum_{j=1}^m \log\lambda_j \\
&= \log\frac{\widehat{G}(d)}{G_0\frac{1}{m}\sum_{j=1}^m \lambda_j^{2(d-d_0)}} - \log\frac{\widehat{G}(d_0)}{G_0} + \log\left(\frac{1}{m}\sum_{j=1}^m \lambda_j^{2(d-d_0)}\right) \\
&\quad -(2d-2d_0)\frac{1}{m}\sum_{j=1}^m \log\lambda_j \\
&= \log\frac{\widehat{G}(d)}{G(d)} - \log\frac{\widehat{G}(d_0)}{G_0} + \log\left(\frac{1}{m}\sum_{j=1}^m j^{2d-2d_0}\Big/\frac{m^{2(d-d_0)}}{2(d-d_0)+1}\right) \\
&\quad -(2d-2d_0)\frac{1}{m}\sum_{j=1}^m \log j + \log\left(\frac{m^{2(d-d_0)}}{2(d-d_0)+1}\right)
\end{aligned}
$$

26

$$= \log \frac{\widehat{G}(d)}{G(d)} - \log \frac{\widehat{G}(d_0)}{G_0} + \log \left(\frac{1}{m} \sum_{j=1}^{m} j^{2d-2d_0} \Big/ \frac{m^{2(d-d_0)}}{2(d-d_0)+1} \right)$$

$$- (2d - 2d_0) \left[\frac{1}{m} \sum_{j=1}^{m} \log j - (\log m - 1) \right]$$

$$+ (2d - 2d_0) - \log \left(2(d - d_0) + 1 \right).$$

For arbitrary small $\Delta > 0$, define $\Theta_1 = \left\{ d : d_0 - \frac{1}{2} + \Delta < d \leq \Delta_2 \right\}$ and $\Theta_2 = \left\{ d : \Delta_1 \leq d \leq d_0 - \frac{1}{2} + \Delta \right\}$. Without loss of generality, we assume $\Delta < \frac{1}{4}$ hereafter. Since the function $x - \log x$ achieves a unique relative and absolute minimum on $(-1, \infty)$ at $x = 0$, and $x - \log x \geq 0.5\delta^2$ if $|x| > \delta$, $\widehat{d} \rightarrow_p d_0$ if

$$\sup_{\Theta_1} |T(d)| \rightarrow_p 0,$$

and

$$\Pr \left(\inf_{\Theta_2} S(d) \leq 0 \right) \rightarrow 0,$$

as $n \rightarrow \infty$, where

$$T(d) = \log \frac{\widehat{G}(d_0)}{G_0} - \log \frac{\widehat{G}(d)}{G(d)} - \log \left(\frac{1}{m} \sum_{j=1}^{m} j^{2d-2d_0} \Big/ \frac{m^{2(d-d_0)}}{2(d-d_0)+1} \right)$$

$$+ (2d - 2d_0) \left[\frac{1}{m} \sum_{j=1}^{m} \log j - (\log m - 1) \right].$$

>From Lemma 1 and Lemma 2 of Robinson (1995b), for $d \in \Theta_1$, we have

$$\frac{1}{m} \sum_{j=1}^{m} \log j - (\log m - 1) = O \left(\frac{\log m}{m} \right),$$

$$\frac{2(d - d_0) + 1}{m} \sum_{j=1}^{m} \left(\frac{j}{m} \right)^{2d-2d_0} - 1 = O \left(\frac{1}{m^{2\Delta_1}} \right). \tag{33}$$

Note that

$$\frac{\widehat{G}(d) - G(d)}{G(d)}$$

$$= \frac{\frac{1}{m} \sum_{j=1}^{m} \lambda_j^{2d} I_v(\lambda_j) - G_0 \frac{1}{m} \sum_{j=1}^{m} \lambda_j^{2(d-d_0)}}{G_0 \frac{1}{m} \sum_{j=1}^{m} \lambda_j^{2(d-d_0)}}$$

$$= \frac{\frac{1}{m} \sum_{j=1}^{m} \lambda_j^{2d_0} \left(\frac{j}{m} \right)^{2(d-d_0)} I_v(\lambda_j) - G_0 \frac{1}{m} \sum_{j=1}^{m} \left(\frac{j}{m} \right)^{2(d-d_0)}}{G_0 \frac{1}{m} \sum_{j=1}^{m} \left(\frac{j}{m} \right)^{2(d-d_0)}}$$

27

$$= \frac{[2(d-d_0)+1]\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0}\left[\lambda_j^{2d_0}I_v(\lambda_j)-G_0\right]}{[2(d-d_0)+1]G_0\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2(d-d_0)}}$$

$$= \frac{A(d)}{B(d)}. \tag{34}$$

Therefore, by the fact that $\Pr(|\log Y| \geq \varepsilon) \leq \Pr(|Y-1| \geq \varepsilon/2)$ for any nonnegative random variable Y and $\varepsilon \leq 1$, $\sup_{\Theta_1}|T(d)| \to_p 0$ if

$$\sup_{\Theta_1}\left|\frac{A(d)}{B(d)}\right| \to_p 0.$$

>From Corollary 2.6 (b) and (c), we have

$$\lambda_j^{2d_0}I_v(\lambda_j) = |C(1)|^2 I_\varepsilon(\lambda_j) + R_{j,n}^a + R_{j,n}^b(d_0),$$

where $E\left|R_{j,n}^a\right| = O(\lambda_j)$ and

$$E\left|R_{j,n}^b(d_0)\right| = \begin{cases} O\left(j^{d_0-1}\right), & \text{for } d_0 \in \left(\frac{1}{2},1\right), \\ 0, & \text{for } d_0 = 1, \\ O\left(j^{d_0-\frac{3}{2}}\right), & \text{for } d_0 \in \left(1,\frac{3}{2}\right), \end{cases}$$

uniformly in j. Thus, in view of the fact that $G_0 = f_u(0) = \frac{\sigma^2}{2\pi}|C(1)|^2$, $A(d)$ can be written as

$$\begin{aligned} A(d) &= [2(d-d_0)+1]\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0}\left[\lambda_j^{2d_0}I_v(\lambda_j)-G_0\right] \\ &= A_1(d) + A_2(d) + A_3(d), \end{aligned}$$

where

$$\begin{aligned} A_1(d) &= g\frac{|C(1)|^2}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0}\left[I_\varepsilon(\lambda_j)-\frac{\sigma^2}{2\pi}\right], \\ A_2(d) &= g\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0}R_{j,n}^a, \quad A_3(d) = g\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0}R_{j,n}^b(d_0), \end{aligned}$$

and $g = 2(d-d_0)+1$. We proceed to consider the successive terms $A_i(d)$ $i = 1,...,3$. For the first term $A_1(d)$, since $E\varepsilon^4 < \infty$, we have, uniformly in j and k, (Priestley, 1981, p.405)

$$\begin{aligned} EI_\varepsilon(\lambda_j) &= \frac{\sigma^2}{2\pi}, \tag{35} \\ Var(I_\varepsilon(\lambda_j)) &= O(1), \tag{36} \\ Cov(I_\varepsilon(\lambda_j), I_\varepsilon(\lambda_k)) &= O\left(n^{-1}\right), \quad j \neq k. \tag{37} \end{aligned}$$

28

>From (??), (??), and (??) and the fact that $\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0} = O(1)$ for $d \in \Theta_1$ (see (??)), it follows that

$$E\left[\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0}\left[I_\varepsilon(\lambda_j) - \frac{\sigma^2}{2\pi}\right]\right]^2$$

$$= \frac{1}{m^2}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{4d-4d_0} Var\left[I_\varepsilon(\lambda_j)\right]$$

$$+ \frac{1}{m^2}\sum_{j \neq k}^{m}\left(\frac{j}{m}\right)^{2d-2d_0}\left(\frac{k}{m}\right)^{2d-2d_0} Cov\left[I_\varepsilon(\lambda_j), I_\varepsilon(\lambda_k)\right]$$

$$= O\left(\frac{1}{m^2}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{4d-4d_0}\right) + O\left(\frac{1}{m^2}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0}\sum_{k=1}^{m}\left(\frac{k}{m}\right)^{2d-2d_0}n^{-1}\right)$$

$$= \left\{\begin{array}{ll} O(\log m/m) & \text{for } 4d - 4d_0 \geq -1 \\ O\left(m^{-2-4d+4d_0}\right) & \text{for } -2 + 4\Delta < 4d - 4d_0 < -1 \end{array}\right\} + O\left(n^{-1}\right)$$

$$= O\left(m^{-4\Delta} + m^{-1}\log m + n^{-1}\right) = O\left(m^{-4\Delta} + n^{-1}\right). \tag{38}$$

Therefore, for all $d \in \Theta_1$ we have

$$A_1(d) = O_p\left(m^{-2\Delta} + n^{-\frac{1}{2}}\right). \tag{39}$$

Next consider $A_2(d)$ and $A_3(d)$. $E|A_2(d)|$ is bounded by

$$\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0}\frac{j}{n} = O\left(\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0}\frac{m}{n}\right) = O\left(\frac{m}{n}\right).$$

$E|A_3(d)| = 0$ for $d_0 = 1$, and for $d_0 \in \left(\frac{1}{2}, 1\right)$, we get

$$E|A_3(d)| = O\left(\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0}j^{d_0-1}\right)$$

$$= O\left(m^{2d_0-2d-1}\sum_{j=1}^{m}j^{2d-d_0-1}\right)$$

$$= \left\{\begin{array}{ll} O\left(m^{d_0-1}\right) & \text{for } 2d - d_0 > 0 \\ O\left(m^{2d_0-2d-1}\log m\right) & \text{for } 2d - d_0 \leq 0 \end{array}\right.$$

$$= O\left(m^{d_0-1} + m^{-2\Delta}\log m\right).$$

For $d_0 \in \left(1, \frac{3}{2}\right)$, we obtain

$$E|A_3(d)| = O\left(\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0}j^{d_0-\frac{3}{2}}\right)$$

29

$$= O\left(m^{2d_0 - 2d - 1} \sum_{j=1}^{m} j^{2d - d_0 - \frac{3}{2}}\right)$$

$$= \begin{cases} O\left(m^{d_0 - \frac{3}{2}}\right) & \text{for } 2d - d_0 - \frac{3}{2} > -1 \\ O\left(m^{2d_0 - 2d - 1} \log m\right) & \text{for } 2d - d_0 - \frac{3}{2} \le -1 \end{cases}$$

$$= O\left(m^{d_0 - \frac{3}{2}} + m^{-2\Delta} \log m\right).$$

Thus, $A_2(d) = O_p\left(n^{-1}m\right)$ and

$$A_3(d) = \begin{cases} O_p\left(m^{d_0 - 1} + m^{-2\Delta} \log m\right), & \text{for } d_0 \in \left(\frac{1}{2}, 1\right), \\ 0, & \text{for } d_0 = 1, \\ O_p\left(m^{d_0 - \frac{3}{2}} + m^{-2\Delta} \log m\right), & \text{for } d_0 \in \left(1, \frac{3}{2}\right), \end{cases}$$

for all $d \in \Theta_1$.

In sum, $A(d)$ is bounded uniformly for all $d \in \Theta_1$ as follows:

$$A(d) = O_p\left(m^{-2\Delta} \log m + n^{-\frac{1}{2}} + n^{-1}m\right) + \begin{cases} O_p\left(m^{d_0 - 1}\right), & \text{for } d_0 \in \left(\frac{1}{2}, 1\right), \\ 0, & \text{for } d_0 = 1, \\ O_p\left(m^{d_0 - \frac{3}{2}}\right), & \text{for } d_0 \in \left(1, \frac{3}{2}\right). \end{cases} \tag{40}$$

Finally, observe that

$$B(d) = [2(d - d_0) + 1] G_0 \frac{1}{m} \sum_{j=1}^{m} \left(\frac{j}{m}\right)^{2(d - d_0)} = G_0 + O\left(m^{-2\Delta}\right), \tag{41}$$

uniformly for all $d \in \Theta_1$, hence $\Pr\left(\inf_{\Theta_1} B(d) \le G_0/2\right) \to 0$.

>From (??)-(??) we deduce that, uniformly over $d \in \Theta_1$,

$$\sup_{\Theta_1} \left|\frac{A(d)}{B(d)}\right| = o_p(1). \tag{42}$$

Also we have established

$$\frac{\widehat{G}(d)}{G(d)} = 1 + \frac{\widehat{G}(d) - G(d)}{G(d)} \to_p 1.$$

Now we consider $\Theta_2 = \left\{d : \Delta_1 \le d \le d_0 - \frac{1}{2} + \Delta\right\}$. Using the same notation and technique as Robinson (1995b), we proceed to prove that,

$$\Pr\left(\inf_{\Theta_2} S(d) \le 0\right) \to 0.$$

Let $p = \exp\left(m^{-1} \sum_{1}^{m} \log j\right)$ and $S(d) = \log\left\{\widehat{D}(d) / \widehat{D}(d_0)\right\}$, where

$$\widehat{D}(d) = \frac{1}{m} \sum_{j=1}^{m} \left(\frac{j}{p}\right)^{2(d - d_0)} j^{2d_0} I_v(\lambda_j).$$

30

It follows that

$$\inf_{\Theta_2} \widehat{D}(d) \geq \frac{1}{m} \sum_{j=1}^{m} a_j j^{2d_0} I_v(\lambda_j),$$

where

$$a_j = \begin{cases} \left(\frac{j}{p}\right)^{2\Delta - 1}, & \text{for } 1 \leq j \leq p, \\ \left(\frac{j}{p}\right)^{-2d_0 - 1}, & \text{for } p < j \leq m. \end{cases}$$

Then,

$$\Pr\left(\inf_{\Theta_2} S(d) \leq 0\right) \leq \Pr\left(\frac{1}{m} \sum_{j=1}^{m} (a_j - 1) j^{2d_0} I_v(\lambda_j) \leq 0\right) = \Pr\left(\frac{1}{m} \sum_{j=1}^{m} (a_j - 1) \frac{I_v(\lambda_j)}{\lambda_j^{-2d_0}} \leq 0\right).$$

(43)

>From Corollary 2.6 (b) and (c), (??) is equal to

$$\Pr(B_1 + B_2 + B_3 \leq 0),$$

where

$$B_1 = \frac{|C(1)|^2}{m} \sum_{j=1}^{m} (a_j - 1) I_\varepsilon(\lambda_j),$$

$$B_2 = \frac{1}{m} \sum_{j=1}^{m} (a_j - 1) R_{j,n}^a, \quad B_3 = \frac{1}{m} \sum_{j=1}^{m} (a_j - 1) R_{j,n}^b(d_0).$$

We proceed to consider the successive terms as above. For B_1,

$$B_1 = \frac{|C(1)|^2}{m} \sum_{j=1}^{m} (a_j - 1) \left[I_\varepsilon(\lambda_j) - \frac{\sigma^2}{2\pi}\right] + G_0 \frac{1}{m} \sum_{j=1}^{m} (a_j - 1).$$

As $m \to \infty$, $p \sim m/e$ and $\sum_{1 \leq j \leq p} a_j \sim \frac{m}{2\Delta e}$. In view of the magnitudes of the moments of $I_\varepsilon(\lambda_j)$ discussed above and the fact that (note that $\Delta < 1/4$)

$$\sum_{j=1}^{m} a_j = \sum_{1 \leq j \leq p} a_j + \sum_{p+1 \leq j \leq m} a_j = O(m) + O\left(p^{2d_0+1} \int_p^m x^{-2d_0-1} dx\right) = O(m),$$

$$\sum_{j=1}^{m} a_j^2 = p^{2-4\Delta} \sum_{j=1}^{p} j^{4\Delta-2} + p^{4d_0+2} \sum_{j=p+1}^{m} j^{-4d_0-2} = O(m^{2-4\Delta} + m),$$

we have

$$E\left[\frac{1}{m} \sum_{j=1}^{m} (a_j - 1) \left[I_\varepsilon(\lambda_j) - \frac{\sigma^2}{2\pi}\right]\right]^2$$

$$= O\left(\frac{1}{m^2} \sum_{j=1}^{m} (a_j - 1)^2\right) + O\left(\frac{1}{m^2} \sum_{j=1}^{m} (a_j - 1) \sum_{k=1}^{m} (a_k - 1) \frac{1}{n}\right)$$

$$= O\left(m^{-4\Delta} + m^{-1}\right) + O\left(n^{-1}\right).$$

(44)

31

Thus
$$B_1 \to_p G_0 \frac{1}{m} \sum_{j=1}^{m} (a_j - 1).$$

$E\,|B_2|$ is bounded by
$$\frac{1}{m} \sum_{j=1}^{m} (a_j - 1) \frac{j}{n} = O\left(\frac{1}{m} \sum_{j=1}^{m} (a_j - 1) \frac{m}{n}\right) = O\left(\frac{m}{n}\right).$$

For B_3, we have
$$E\,|B_3| = \begin{cases} O\left(\frac{1}{m} \sum_{j=1}^{m} (a_j - 1) j^{d_0 - 1}\right), & \text{for } d_0 \in \left(\frac{1}{2}, 1\right), \\ 0, & \text{for } d_0 = 1, \\ O\left(\frac{1}{m} \sum_{j=1}^{m} (a_j - 1) j^{d_0 - \frac{3}{2}}\right), & \text{for } d_0 \in \left(1, \frac{3}{2}\right). \end{cases}$$

This is $o\,(1)$ because
$$\frac{1}{m} \sum_{j=1}^{m} a_j j^{d_0 - 1} = \frac{p^{1-2\Delta}}{m} \sum_{j=1}^{p} j^{2\Delta + d_0 - 2} + \frac{p^{1+2d_0}}{m} \sum_{j=p+1}^{m} j^{-2-d_0} = O\left(m^{-2\Delta} \log m + m^{d_0 - 1}\right),$$
$$\frac{1}{m} \sum_{j=1}^{m} a_j j^{d_0 - \frac{3}{2}} = \frac{p^{1-2\Delta}}{m} \sum_{j=1}^{p} j^{2\Delta + d_0 - \frac{5}{2}} + \frac{p^{1+2d_0}}{m} \sum_{j=p+1}^{m} j^{-\frac{5}{2} - d_0} = O\left(m^{-2\Delta} \log m + m^{d_0 - \frac{3}{2}}\right),$$

and
$$\frac{1}{m} \sum_{j=1}^{m} j^{d_0 - 1} = O\left(m^{d_0 - 1}\right), \quad \frac{1}{m} \sum_{j=1}^{m} j^{d_0 - \frac{3}{2}} = O\left(m^{d_0 - \frac{3}{2}}\right).$$

Choose $\Delta < 1/\,(2e) < 1/4$ with no loss of generality, then for sufficiently large m,
$$\frac{1}{m} \sum_{j=1}^{m} (a_j - 1) \ge \frac{1}{m} \sum_{1 \le j \le p} a_j - 1 \sim \frac{1}{2\Delta e} - 1 > \delta > 0.$$

Hence,
$$B_1 + B_2 + B_3 \to_p G_0 \frac{1}{m} \sum_{j=1}^{m} (a_j - 1) \ge G_0 \delta > 0.$$

It follows that
$$\Pr\left(B_1 + B_2 + B_3 \le 0\right) \to 0 \quad \text{as } m \to \infty. \tag{45}$$

Therefore, $\widehat{d} \to_p d_0$, giving the stated result. $\blacksquare$

8.4 Proof of Theorem 3.2

The proof has the same structure as that of Theorem 3.1 and we therefore provide only the relevant parts. First, it follows from Corollary 2.6 (e) that

$$
\begin{aligned}
A(d) &= [2(d-d_0)+1]\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0}\left[\lambda_j^{2d_0} I_v(\lambda_j) - G_0\right]\\
&= A_1(d) + A_2(d) + A_3(d) + A_4(d) + A_5(d) + A_6(d),
\end{aligned}
$$

where

$$
A_1(d) = g\frac{|C(1)|^2}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0}\left[I_\varepsilon(\lambda_j) - \frac{\sigma^2}{2\pi}\right],
$$

$$
A_2(d) = g\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0} R_{j,n}^a, \quad A_3(d) = g\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0} R_{j,n}^b(d_0),
$$

$$
A_4(d) = g\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0} R_{j,n}^c(d_0), \quad A_5(d) = g\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0} R_{j,n}^e(d_0),
$$

$$
A_6(d) = g\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0} R_{j,n}^g(d_0).
$$

It has already shown that $A_1(d) \to_p 0$ and $A_2(d) \to_p 0$. For $A_i(d_0)$ $i = 3, ..., 6$, we obtain

$$
\begin{aligned}
E\left|A_3(d)\right| &= O\left(\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0} j^{d_0-2}(\log n)^{\frac{1}{2}}\right) = O\left((\log n)^{\frac{1}{2}} m^{2d_0-2d-1}\sum_{j=1}^{m} j^{2d-d_0-2}\right)\\
&= O\left((\log n)^{\frac{1}{2}} m^{d_0-2} + (\log n)^{\frac{1}{2}} m^{-2\Delta}\log m\right),
\end{aligned}
$$

$$
\begin{aligned}
E\left|A_4(d)\right| &= O\left(\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0} j^{d_0-1} n^{-\frac{1}{2}}\right) = O\left(n^{-\frac{1}{2}} m^{2d_0-2d-1}\sum_{j=1}^{m} j^{2d-d_0-1}\right)\\
&= O\left(n^{-\frac{1}{2}} m^{d_0-1} + n^{-\frac{1}{2}} m^{-2\Delta}\log m\right),
\end{aligned}
$$

$$
\begin{aligned}
E\left|A_5(d)\right| &= O\left(\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0} j^{2d_0-4}\log n\right) = O\left((\log n) m^{2d_0-2d-1}\sum_{j=1}^{m} j^{2d-4}\right)\\
&= O\left((\log n) m^{2d_0-4} + (\log n) m^{-2\Delta}\log m\right),
\end{aligned}
$$

$$
\begin{aligned}
E\left|A_6(d)\right| &= O\left(\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0} j^{2d_0-2} n^{-1}\right) = O\left(n^{-1} m^{2d_0-2d-1}\sum_{j=1}^{m} j^{2d-2}\right)\\
&= O\left(n^{-1} m^{2d_0-2} + n^{-1} m^{-2\Delta}\log m\right).
\end{aligned}
$$

33

It follows that $\sum_{i=1}^{6} A_i (d) \to_p 0$ if $\frac{n^\alpha}{m} + \frac{m^{2d_0-2}}{n} \to 0$ as $n \to \infty$.

Similarly, we have

$$
\begin{aligned}
B_1 &= \frac{|C(1)|^2}{m} \sum_{j=1}^{m} (a_j - 1) I_\varepsilon (\lambda_j) \to_p G_0 \frac{1}{m} \sum_{j=1}^{m} (a_j - 1), \\
B_2 &= \frac{1}{m} \sum_{j=1}^{m} (a_j - 1) R_{j,n}^a \to_p 0, \\
B_3 &= \frac{1}{m} \sum_{j=1}^{m} (a_j - 1) R_{j,n}^b (d_0), \quad B_4 = \frac{1}{m} \sum_{j=1}^{m} (a_j - 1) R_{j,n}^c (d_0), \\
B_5 &= \frac{1}{m} \sum_{j=1}^{m} (a_j - 1) R_{j,n}^e (d_0), \quad B_6 = \frac{1}{m} \sum_{j=1}^{m} (a_j - 1) R_{j,n}^g (d_0).
\end{aligned}
$$

$B_3, ..., B_6$ converge to zero in probability, because

$$
E|B_3| = O\left(\frac{1}{m} \sum_{j=1}^{m} (a_j - 1) j^{d_0-2} (\log n)^{\frac{1}{2}} \right) = O\left((\log n)^{\frac{1}{2}} m^{-2\Delta} \log m + (\log n)^{\frac{1}{2}} m^{d_0-2} \right),
$$

$$
E|B_4| = O\left(\frac{1}{m} \sum_{j=1}^{m} (a_j - 1) j^{d_0-1} n^{-\frac{1}{2}} \right) = O\left(n^{-\frac{1}{2}} m^{-2\Delta} \log m + n^{-\frac{1}{2}} m^{d_0-1} \right),
$$

$$
E|B_5| = O\left(\frac{1}{m} \sum_{j=1}^{m} (a_j - 1) j^{2d_0-4} \log n \right) = O\left((\log n) m^{-2\Delta} \log m + (\log n) m^{2d_0-4} \right),
$$

$$
E|B_6| = O\left(\frac{1}{m} \sum_{j=1}^{m} (a_j - 1) j^{2d_0-2} n^{-1} \right) = O\left(n^{-1} m^{-2\Delta} \log m + n^{-1} m^{2d_0-2} \right).
$$

Hence $\sum_{i=1}^{6} B_i \to_p G_0 \frac{1}{m} \sum_{j=1}^{m} (a_j - 1) > 0$ if $\frac{n^\alpha}{m} + \frac{m^{2d_0-2}}{n} \to 0$ as $n \to \infty$ and $\widehat{d} \to_p d_0$ follows. ∎

8.5 Proof of Theorem 3.3

As above, we deal only with the relevant parts. It follows from Corollary 2.6 (d) that

$$
A(d) = A_1 (d) + A_2 (d) + A_3 (d) + A_4 (d),
$$

where $A_1 (d) \to_p 0$, $A_2 (d) \to_p 0$, and

$$
A_3 (d) = g \frac{1}{m} \sum_{j=1}^{m} \left(\frac{j}{m} \right)^{2d-2d_0} R_{j,n}^b (d_0), \quad A_4 (d) = g \frac{1}{m} \sum_{j=1}^{m} \left(\frac{j}{m} \right)^{2d-2d_0} R_{j,n}^c (d_0).
$$

For $A_3 (d)$ and $A_4 (d)$, we obtain

$$
E|A_3 (d)| = O\left(\frac{1}{m} \sum_{j=1}^{m} \left(\frac{j}{m} \right)^{2d-2d_0} j^{d_0-1} n^{\frac{1}{2}-d_0} (\log n)^{\frac{1}{2}} \right)
$$

34

$$= O\left(n^{\frac{1}{2}-d_0}(\log n)^{\frac{1}{2}}m^{2d_0-2d-1}\sum_{j=1}^{m}j^{2d-d_0-1}\right)$$

$$= \begin{cases} O\left(n^{\frac{1}{2}-d_0}(\log n)^{\frac{1}{2}}m^{d_0-1}\right), & \text{for } 2d-d_0 > 0, \\ O\left(n^{\frac{1}{2}-d_0}(\log n)^{\frac{1}{2}}m^{2d_0-2d-1}\log m\right), & \text{for } 2d-d_0 \leq 0, \end{cases}$$

$$E\left|A_4(d)\right| = O\left(\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0}j^{2d_0-2}n^{1-2d_0}\log n\right)$$

$$= O\left(n^{1-2d_0}(\log n)m^{2d_0-2d-1}\sum_{j=1}^{m}j^{2d-2}\right)$$

$$= \begin{cases} O\left(n^{1-2d_0}(\log n)m^{2d_0-2}\right), & \text{for } d > 1/2, \\ O\left(n^{1-2d_0}(\log n)m^{2d_0-2d-1}\log m\right), & \text{for } d \leq 1/2. \end{cases}$$

Note that, for $d_0 \leq \frac{1}{2}$ we have

$$n^{\frac{1}{2}-d_0}(\log n)^{\frac{1}{2}}m^{2d_0-2d-1}\log m = n^{d_0-\frac{1}{2}}(\log n)^{-\frac{1}{2}}\left(n^{1-2d_0}(\log n)m^{2d_0-2d-1}\log m\right)$$

$$= O\left(n^{1-2d_0}(\log n)m^{2d_0-2d-1}\log m\right).$$

Hence, $E\left|A_3(d)\right| \to 0$ and $E\left|A_4(d)\right| \to 0$ if (note that $d \geq \Delta_1$)

$$\frac{n^{1-2d_0}\log n}{m^{2-2d_0}} \to 0 \quad \text{and} \quad \frac{n^{1-2d_0}\log n\log m}{m^{1-2d_0+2\Delta_1}} \to 0.$$

Since $\Delta_1 \leq d_0 \leq \frac{1}{2}$, this is satisfied if

$$\frac{n^{1-2\Delta_1}\log n\log m}{m} \to 0.$$

For the parameter space Θ_2, we change the definition of a_j as follows:

$$a_j = \begin{cases} \left(\frac{j}{p}\right)^{2(d-d_0)}, & \text{for } 1 \leq j \leq p, \\ \left(\frac{j}{p}\right)^{-2d_0-1}, & \text{for } p < j \leq m. \end{cases}$$

It still holds that

$$\inf_{\Theta_2}\widehat{D}(d) \geq \frac{1}{m}\sum_{j=1}^{m}a_j j^{2d_0}I_v(\lambda_j),$$

and hence

$$\Pr\left(\inf_{\Theta_2}S(d) \leq 0\right) \leq \Pr\left(\frac{1}{m}\sum_{j=1}^{m}(a_j-1)\frac{I_v(\lambda_j)}{\lambda_j^{-2d_0}} \leq 0\right).$$

35

Since $d \geq \Delta_1 > 0$ and $d_0 \leq \frac{1}{2}$, we have $2d - 2d_0 > -1$. Thus

$$
\begin{aligned}
\sum_{1 \leq j \leq p} a_j &= p^{2(d_0-d)} \sum_{1 \leq j \leq p} j^{2(d-d_0)} \sim \frac{p}{2(d-d_0)+1} \sim \frac{m}{e} \frac{1}{2(d-d_0)+1} \geq \frac{m}{2\Delta e}, \\
\sum_{1 \leq j \leq p} a_j{}^2 &= p^{4(d_0-d)} \sum_{1 \leq j \leq p} j^{4(d-d_0)} = \begin{cases} O(m) & \text{for } 4(d-d_0) > -1 \\ O\left(m^{4(d_0-d)} \log m\right) & \text{for } 4(d-d_0) \leq -1 \end{cases} \\
&= O\left(m^2 \left(m^{-1} + m^{4(d_0-d)-2} \log m\right)\right) = o\left(m^2\right),
\end{aligned}
$$

and it follows that $\sum_{j=1}^m a_j = O(m)$ and $\sum_{j=1}^m a_j^2 = o\left(m^2\right)$. Therefore, using the same argument as above, we have $B_1 \to_p G_0 \frac{1}{m} \sum_{j=1}^m (a_j - 1)$, $B_2 \to_p 0$, and

$$
\begin{aligned}
E\left|B_3(d)\right| &= O\left(\frac{1}{m} \sum_{j=1}^m (a_j - 1) j^{d_0-1} n^{\frac{1}{2}-d_0} (\log n)^{\frac{1}{2}}\right) \\
&= O\left(\frac{1}{m} \sum_{j=1}^p \left(\frac{j}{p}\right)^{2(d-d_0)} j^{d_0-1} n^{\frac{1}{2}-d_0} (\log n)^{\frac{1}{2}} + \frac{1}{m} \sum_{j=p+1}^m \left(\frac{j}{p}\right)^{-2d_0-1} j^{d_0-1} n^{\frac{1}{2}-d_0} (\log n)^{\frac{1}{2}}\right) \\
&\quad + O\left(\frac{1}{m} \sum_{j=1}^m j^{d_0-1} n^{\frac{1}{2}-d_0} (\log n)^{\frac{1}{2}}\right) \\
&= O\left(n^{\frac{1}{2}-d_0} (\log n)^{\frac{1}{2}} \left(m^{2(d_0-d)-1} \sum_{j=1}^p j^{2d-d_0-1} + m^{2d_0} \sum_{j=p+1}^m j^{-d_0-2} + \frac{1}{m} \sum_{j=1}^m j^{d_0-1}\right)\right) \\
&= O\left(n^{\frac{1}{2}-d_0} (\log n)^{\frac{1}{2}} \left(m^{d_0-1} + m^{2d_0-2d-1} \log m\right)\right) \\
&= O\left(n^{1-2d_0} \log n \left(m^{d_0-1} + m^{2d_0-2d-1} \log m\right)\right) = o(1), \\
E\left|B_4(d)\right| &= O\left(\frac{1}{m} \sum_{j=1}^m (a_j - 1) j^{2d_0-2} n^{1-2d_0} \log n\right) \\
&= O\left(\frac{1}{m} \sum_{j=1}^p \left(\frac{j}{p}\right)^{2(d-d_0)} j^{2d_0-2} n^{1-2d_0} \log n + \frac{1}{m} \sum_{j=p+1}^m \left(\frac{j}{p}\right)^{-2d_0-1} j^{2d_0-2} n^{1-2d_0} \log n\right) \\
&\quad + O\left(\frac{1}{m} \sum_{j=1}^m j^{2d_0-2} n^{1-2d_0} \log n\right) \\
&= O\left(n^{1-2d_0} \log n \left(m^{2d_0-2d-1} \sum_{j=1}^p j^{2d-2} + m^{2d_0} \sum_{j=p+1}^m j^{-3} + \frac{1}{m} \sum_{j=1}^m j^{2d_0-2}\right)\right) \\
&= O\left(n^{1-2d_0} \log n \left(m^{2d_0-2} + m^{2d_0-2d-1} \log m + m^{-1}\right)\right) \\
&= O\left(n^{1-2d_0} \log n \left(m^{-1} + m^{2d_0-2d-1} \log m\right)\right) = o(1),
\end{aligned}
$$

36

because $\frac{n^{1-2\Delta_1} \log m \log n}{m} \to 0$. Therefore, $\sum_{i=1}^4 B_i \to_p G_0 \frac{1}{m} \sum_{j=1}^m (a_j - 1) > 0$ and $\widehat{d} \to_p d_0$ follows. ■

8.6   Proof of Theorem 3.4

Since $\widehat{d} \to_p d_0$ and $\widehat{G}(d)$ is a continuous function of d, we may analyse $\widehat{G}(d_0)$. We have
$$\frac{\widehat{G}(d_0) - G(d_0)}{G(d_0)} = \frac{\frac{1}{m} \sum_{j=1}^m \lambda_j^{2d_0} I_v(\lambda_j) - G_0}{G_0} = \frac{A(d_0)}{B(d_0)} \to_p 0.$$

So
$$\widehat{G}(d_0) \to_p G_0,$$

which gives the required result. ■

8.7   Proof of Theorem 4.1

We work from the first order conditions for $\widehat{d}$, viz.
$$0 = R'\left(\widehat{d}\right) = R'(d_0) + R''(d^*)\left(\widehat{d} - d_0\right), \tag{46}$$

where $|d^* - d_0| \le \left|\widehat{d} - d_0\right|$. As in the proof of Theorem 2 of Robinson (1995b) we get the following expression for $R''(d)$

$$R''(d) = \frac{4\left[\widehat{F}_2(d) \widehat{F}_0(d) - \widehat{F}_1(d)^2\right]}{\widehat{F}_0(d)^2},$$

where
$$\widehat{F}_k(d) = \frac{1}{m} \sum_{j=1}^m (\log j)^k \lambda_j^{2d} I_v(\lambda_j).$$

>From (??) and (??), we have
$$\sup_{\Theta_1} \left|\frac{A(d)}{B(d)}\right| = o_p\left((\log m)^{-6}\right),$$

so, by the argument on pages 1642-43 of Robinson (1995b), $R''(d^*) = R''(d_0) + o_p(1)$. Now, from Corollary 2.6 (b), we find
$$\widehat{F}_k(d_0) = \frac{1}{m} \sum_{j=1}^m (\log j)^k \lambda_j^{2d_0} I_v(\lambda_j) = C_1 + C_2 + C_3,$$

where
$$C_1 = \frac{|C(1)|^2}{m} \sum_{j=1}^m (\log j)^k I_\varepsilon(\lambda_j),$$

$$C_2 = \frac{1}{m} \sum_{j=1}^m (\log j)^k R_{j,n}^a, \quad C_3 = \frac{1}{m} \sum_{j=1}^m (\log j)^k R_{j,n}^b(d_0).$$

37

We proceed to consider the successive terms as above for $k = 0, 1, 2$. For C_1,

$$C_1 = \frac{|C(1)|^2}{m} \sum_{j=1}^{m} (\log j)^k \left[I_\varepsilon (\lambda_j) - \frac{\sigma^2}{2\pi} \right] + G_0 \frac{1}{m} \sum_{j=1}^{m} (\log j)^k .$$

A similar argument to that above and the fact that

$$\frac{1}{m} \sum_{j=1}^{m} \log j \sim \log m, \quad \frac{1}{m} \sum_{j=1}^{m} (\log j)^2 \sim (\log m)^2 ,$$

yield

$$E \left[\frac{|C(1)|^2}{m} \sum_{j=1}^{m} (\log j)^k \left[I_\varepsilon (\lambda_j) - \frac{\sigma^2}{2\pi} \right] \right]^2$$

$$= O \left(\frac{1}{m^2} \sum_{j=1}^{m} (\log j)^{2k} \right) + O \left(\frac{1}{m^2} \sum_{j=1}^{m} (\log j)^k \sum_{l=1}^{m} (\log l)^k \frac{1}{n} \right)$$

$$= O \left(\frac{(\log m)^{2k}}{m} \right) + O \left(\frac{(\log m)^{2k}}{n} \right) = O \left(\frac{(\log m)^{2k}}{m} \right) ,$$

giving $C_1 = \frac{1}{m} \sum_{j=1}^{m} (\log j)^k (G_0 + o_p(1))$.

For C_2, $E |C_2|$ is bounded by

$$\frac{1}{m} \sum_{j=1}^{m} (\log j)^k \frac{j}{n} = O \left(\frac{1}{m} \sum_{j=1}^{m} (\log j)^k \frac{m}{n} \right) = o \left((\log m)^k \right) .$$

$E |C_3| = 0$ for $d_0 = 1$, and for $d_0 \neq 1$, we obtain

$$E |C_3| = \begin{cases} O \left(\frac{1}{m} \sum_{j=1}^{m} (\log j)^k j^{d_0-1} \right) = O \left(\frac{(\log m)^k}{m} \sum_{j=1}^{m} j^{d_0-1} \right), & \text{for } d_0 \in \left(\frac{1}{2}, 1 \right), \\ O \left(\frac{1}{m} \sum_{j=1}^{m} (\log j)^k j^{d_0-\frac{3}{2}} \right) = O \left(\frac{(\log m)^k}{m} \sum_{j=1}^{m} j^{d_0-\frac{3}{2}} \right), & \text{for } d_0 \in \left(1, \frac{3}{2} \right). \end{cases}$$

Hence

$$\widehat{F}_k (d_0) = G_0 \left[\frac{1}{m} \sum_{j=1}^{m} (\log j)^k \right] [1 + o_p(1)] .$$

Then

$$R'' (d_0) = \frac{4 G_0^2 \left[\frac{1}{m} \sum_{j=1}^{m} (\log j)^2 - \left(\frac{1}{m} \sum_{j=1}^{m} (\log j) \right)^2 \right]}{G_0^2} [1 + o_p(1)]$$

$$= \left\{ 4 \left[\frac{1}{m} \left\{ \left(m + \frac{1}{2} \right) (\log m)^2 - 2m \log m + 2m + O(1) \right\} \right] \right.$$

$$\left. - 4 \left[\frac{1}{m^2} \left\{ \left(m + \frac{1}{2} \right) (\log m) - m + O(1) \right\}^2 \right] \right\} [1 + o_p(1)]$$

$$= 4 + o_p(1) . \tag{47}$$

38

Next we consider the first term on the right side of (??). We have

$$R'(d_0) = 2\frac{\widehat{G}_1(d_0)}{\widehat{G}(d_0)} - \frac{2}{m}\sum_{j=1}^{m}\log\lambda_j,$$

where

$$\widehat{G}_1(d_0) = \frac{1}{m}\sum_{j=1}^{m}(\log\lambda_j)\,\lambda_j^{2d_0}I_v(\lambda_j)\,, \quad \widehat{G}(d_0) = \frac{1}{m}\sum_{j=1}^{m}\lambda_j^{2d_0}I_v(\lambda_j)\,.$$

Then

$$
\begin{aligned}
m^{\frac{1}{2}}R'(d_0) &= m^{\frac{1}{2}}\left[2\frac{\widehat{G}_1(d_0)}{\widehat{G}(d_0)} - \frac{2}{m}\sum_{j=1}^{m}\log\lambda_j\right]\\
&= \frac{2}{\sqrt{m}}\frac{\sum_{j=1}^{m}(\log\lambda_j)\,\lambda_j^{2d_0}I_v(\lambda_j) - \left(\sum_{j=1}^{m}\log\lambda_j\right)\widehat{G}(d_0)}{\widehat{G}(d_0)}\\
&= \frac{2}{\sqrt{m}}\frac{\sum_{j=1}^{m}(\log\lambda_j)\,\lambda_j^{2d_0}I_v(\lambda_j) - \left(\frac{1}{m}\sum_{j=1}^{m}\log\lambda_j\right)\sum_{j=1}^{m}\lambda_j^{2d_0}I_v(\lambda_j)}{\widehat{G}(d_0)}\\
&= \frac{2}{\sqrt{m}}\frac{\sum_{j=1}^{m}\left(\log\lambda_j - \frac{1}{m}\sum_{j=1}^{m}\log\lambda_j\right)\lambda_j^{2d_0}I_v(\lambda_j)}{\widehat{G}(d_0)}\\
&= \frac{\frac{2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\left[\lambda_j^{2d_0}I_v(\lambda_j) - G_0\right]}{\widehat{G}(d_0)},
\end{aligned}
$$

where

$$\nu_j = \log\lambda_j - \frac{1}{m}\sum_{j=1}^{m}\log\lambda_j = \log j - \frac{1}{m}\sum_{j=1}^{m}\log j,$$

and $\sum_{j=1}^{m}\nu_j = 0$. For the denominator, from Theorem 3.4 we have

$$\widehat{G}(d_0) \to_p G_0. \tag{48}$$

By Corollary 2.6 (a) and (c), the numerator can be decomposed as follows:

$$\frac{2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\left[\lambda_j^{2d_0}I_v(\lambda_j) - G_0\right] = \begin{cases} D_1 + D_2 + D_3 + D_4 + D_5 + D_6, & \text{for } d_0 \in \left(\frac{1}{2}, \frac{3}{2}\right)\setminus\{1\}, \\ D_1 + D_4, & \text{for } d_0 = 1, \end{cases}$$

where

$$
\begin{aligned}
D_1 &= \frac{2\,|C(1)|^2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\left[I_\varepsilon(\lambda_j) - \frac{\sigma^2}{2\pi}\right],\\
D_2 &= \frac{2\,|C(1)|^2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\frac{\lambda_j^{2d_0}}{\left|1 - e^{i\lambda_j}\right|^2}\frac{\left|\widetilde{\varepsilon}_{\lambda_j n}(f_0)\right|^2}{2\pi n},
\end{aligned}
$$

39

$$D_3 = -\frac{2\left|C\left(1\right)\right|^2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\left[e^{\frac{\pi}{2}d_0 i}w_\varepsilon\left(\lambda_j\right)\frac{\lambda_j^{d_0}}{1-e^{-i\lambda_j}}\frac{\widetilde{\varepsilon}_{\lambda_j n}\left(f_0\right)^*}{\sqrt{2\pi n}}+\frac{\lambda_j^{d_0}}{1-e^{i\lambda_j}}\frac{\widetilde{\varepsilon}_{\lambda_j n}\left(f_0\right)}{\sqrt{2\pi n}}e^{-\frac{\pi}{2}d_0 i}w_\varepsilon\left(\lambda_j\right)^*\right],$$

$$D_4 = \frac{2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j R_{j,n}^a,\quad D_5 = \frac{2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j R_{j,n}^b\left(d_0\right),\quad D_6 = \frac{2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j R_{j,n}^c\left(d_0\right).$$

For D_1, $\{\varepsilon_t\}$ satisfies the assumptions

$$E\left(\varepsilon_t|F_{t-1}\right) = 0,\quad E\left(\varepsilon_t^2|F_{t-1}\right) = \sigma^2,\quad E\left(\varepsilon_t^3|F_{t-1}\right) = \mu_3,\quad a.s.,\quad E\left(\varepsilon_t^4\right) = \mu_4,$$

for $t = 1, 2, ...$, thus we can apply the result in Robinson (1995b) pp.1644-1647 to conclude

$$D_1 = \frac{2G_0}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\left[\frac{2\pi}{\sigma^2}I_\varepsilon\left(\lambda_j\right) - 1\right] \to_d N\left(0, 4G_0^2\right). \tag{49}$$

>From Lemma 7.12 (b), for $d_0 \in \left(\frac{1}{2}, 1\right)$, $E\left|D_2\right|$ is bounded by

$$\frac{1}{\sqrt{m}}\sum_{j=1}^{m}\left|\nu_j\right|\lambda_j^{2d_0-2}\frac{E\left|\widetilde{\varepsilon}_{\lambda_j n}(f_0)\right|^2}{2\pi n}$$

$$= O\left(\frac{\log m}{\sqrt{m}}\sum_{j=1}^{m}\left(\frac{j}{n}\right)^{2d_0-2}\frac{1}{n}\left(L^{2d_0-1}+\frac{n}{j}L^{2d_0-2}\right)\right)$$

$$= O\left(\frac{\log m}{\sqrt{m}}\left(m^{2d_0-1}n^{1-2d_0}L^{2d_0-1}+n^{2-2d_0}L^{2d_0-2}\right)\right)$$

$$= O\left(\log m\left(m^{-\frac{1}{2}}\left(\frac{mL}{n}\right)^{2d_0-1}+m^{\frac{3}{2}-2d_0}\left(\frac{mL}{n}\right)^{2d_0-2}\right)\right).$$

When $d_0 > \frac{3}{4}$, this is $o\left(1\right)$ by choosing $L = \frac{n}{m}$. When $d_0 \le \frac{3}{4}$, choose $L = n\left(\log m\right)^{\frac{-2}{2d_0-1}}$, then we have

$$\left(\log m\right)m^{-\frac{1}{2}}\left(\frac{mL}{n}\right)^{2d_0-1} = m^{2d_0-\frac{3}{2}}\left(\log m\right)^{-1} \to 0,$$

$$\left(\log m\right)m^{\frac{3}{2}-2d_0}\left(\frac{mL}{n}\right)^{2d_0-2} = m^{-\frac{1}{2}}\left(\log m\right)^{\frac{4-4d_0}{2d_0-1}+1} \to 0.$$

Therefore, $D_2 = o_p\left(1\right)$ for $d_0 \in \left(\frac{1}{2}, 1\right)$. For $d_0 \in \left(1, \frac{5}{4}\right)$, from Lemma 7.14 (a), $E\left|D_2\right|$ is bounded by

$$\frac{1}{\sqrt{m}}\sum_{j=1}^{m}\left|\nu_j\right|\lambda_j^{2d_0-2}\frac{E\left|\widetilde{\varepsilon}_{\lambda_j n}\left(f_0\right)\right|^2}{2\pi n} = O\left(\frac{\log m}{\sqrt{m}}\sum_{j=1}^{m}j^{2d_0-3}\right) = O\left(\log m\left(m^{2d_0-\frac{5}{2}}\right)\right) = o\left(1\right),$$

and for $d_0 \in \left[\frac{5}{4}, \frac{3}{2}\right)$, from Lemma 7.14 (b), we have

$$\frac{1}{\sqrt{m}}\sum_{j=1}^{m}\left|\nu_j\right|\frac{\lambda_j^{2d_0}}{\left|1-e^{i\lambda_j}\right|^2}\frac{E\left|\widetilde{\varepsilon}_{\lambda_j n}\left(f_0\right)\right|^2}{2\pi n}$$

40

$$
= O\left(\frac{\log m}{\sqrt{m}}\sum_{j=1}^{m}\left(\left(\frac{j}{n}\right)^{2d_0-2}\frac{L^{d_0-1}n^{d_0-1}}{j}+\left(\frac{j}{n}\right)^{2d_0-2}\frac{L^{2d_0-3}n}{j^2}\right)\right)
$$

$$
= O\left(\frac{\log m}{\sqrt{m}}\sum_{j=1}^{m}\left(j^{2d_0-3}L^{d_0-1}n^{1-d_0}+j^{2d_0-4}L^{2d_0-3}n^{3-2d_0}\right)\right)
$$

$$
= O\left(\log m\left(\left(\frac{L}{n}\right)^{d_0-1}m^{2d_0-\frac{5}{2}}+\left(\frac{L}{n}\right)^{2d_0-3}m^{-\frac{1}{2}}\right)\right)
$$

$$
= O\left(\log m\left(\left(\frac{Lm}{n}\right)^{d_0-1}m^{d_0-\frac{3}{2}}+\left(\frac{Lm}{n}\right)^{2d_0-3}m^{\frac{5}{2}-2d_0}\right)\right)=o\left(1\right),
$$

by letting $L=\frac{n}{m}\left(\log m\right)^{\frac{-2}{2d_0-3}}$, giving $D_2=o_p\left(1\right)$.

For D_3, in view of the fact that

$$
w_\varepsilon\left(\lambda_j\right)^*=\frac{1}{\sqrt{2\pi n}}\sum_{p=1}^{n}e^{-ip\lambda_j}\varepsilon_p=\frac{1}{\sqrt{2\pi n}}\sum_{n-p=0}^{n-1}e^{i(n-p)\lambda_j}\varepsilon_{n-(n-p)}=\frac{1}{\sqrt{2\pi n}}\sum_{q=0}^{n-1}e^{iq\lambda_j}\varepsilon_{n-q},
$$

we obtain the decomposition

$$
\frac{2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\frac{\lambda_j^{d_0}}{1-e^{i\lambda_j}}\frac{\widetilde{\varepsilon}_{\lambda_j n}(f_0)}{\sqrt{2\pi n}}w_\varepsilon\left(\lambda_j\right)^*
$$

$$
= \frac{2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\frac{\lambda_j^{d_0}}{1-e^{i\lambda_j}}\frac{1}{2\pi n}\left(\sum_{p=0}^{n-1}\widetilde{f}_{\lambda_j p}e^{-ip\lambda_j}\varepsilon_{n-p}\right)\left(\sum_{q=0}^{n-1}e^{iq\lambda_j}\varepsilon_{n-q}\right).
$$

It follows that

$$
E\left|\frac{2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\frac{\lambda_j^{d_0}}{1-e^{i\lambda_j}}\frac{\widetilde{\varepsilon}_{\lambda_j n}(f_0)}{\sqrt{2\pi n}}w_\varepsilon\left(\lambda_j\right)^*\right|^2
$$

$$
= \frac{1}{\pi^2 mn^2}E\left[\sum_{j=1}^{m}\nu_j\frac{\lambda_j^{d_0}}{1-e^{i\lambda_j}}\left(\sum_{p=0}^{n-1}\widetilde{f}_{\lambda_j p}e^{-ip\lambda_j}\varepsilon_{n-p}\right)\left(\sum_{q=0}^{n-1}e^{iq\lambda_j}\varepsilon_{n-q}\right)\right]
$$

$$
\times\left[\sum_{k=1}^{m}\nu_k\frac{\lambda_k^{d_0}}{1-e^{-i\lambda_k}}\left(\sum_{r=0}^{n-1}\widetilde{f}_{-\lambda_k r}e^{ir\lambda_k}\varepsilon_{n-r}\right)\left(\sum_{s=0}^{n-1}e^{-is\lambda_k}\varepsilon_{n-s}\right)\right]. \qquad (50)
$$

Because $\{\varepsilon_t\}$ are independent,

$$
E\left(\varepsilon_p\varepsilon_q\varepsilon_r\varepsilon_s\right)=\begin{cases}\mu_4, & \text{if } p=q=r=s,\\ \sigma^4, & \text{if } p=q\neq r=s,\ p=s\neq q=r,\ p=r\neq q=s,\\ 0, & \text{otherwise.}\end{cases}
$$

41

Therefore, (??) is bounded by

$$\frac{\mu_4}{mn^2}\sum_{j=1}^{m}\sum_{k=1}^{m}|\nu_j|\,|\nu_k|\,\lambda_j^{d_0-1}\lambda_k^{d_0-1}\left(\sum_{p=0}^{n-1}\left|\widetilde{f}_{\lambda_j p}\right|\left|\widetilde{f}_{-\lambda_k p}\right|\right) \tag{51}$$

$$+\frac{2\sigma^4}{mn^2}\sum_{j=1}^{m}|\nu_j|\,\lambda_j^{d_0-1}\left(\sum_{p=0}^{n-1}\left|\widetilde{f}_{\lambda_j p}\right|\right)\sum_{k=1}^{m}|\nu_k|\,\lambda_k^{d_0-1}\left(\sum_{q=0}^{n-1}\left|\widetilde{f}_{-\lambda_k q}\right|\right) \tag{52}$$

$$+\frac{\sigma^4}{mn^2}\sum_{j=1}^{m}\sum_{k=1}^{m}|\nu_j|\,|\nu_k|\,\lambda_j^{d_0-1}\lambda_k^{d_0-1}\left(\sum_{p=0}^{n-1}\left|\widetilde{f}_{\lambda_j p}\right|\left|\widetilde{f}_{-\lambda_k p}\right|\right)\left(\sum_{q=0}^{n-1}e^{iq(\lambda_j-\lambda_k)}\right). \tag{53}$$

(??) is bounded by, for $d_0\in\left(\frac{1}{2},1\right)$,

$$\frac{(\log m)^2}{m}\sum_{j=1}^{m}\sum_{k=1}^{m}\left(\frac{j}{n}\right)^{d_0-1}\left(\frac{k}{n}\right)^{d_0-1}\frac{1}{n^2}\sum_{p=1}^{n-1}p^{-2f_0}$$

$$= O\left(\frac{(\log m)^2}{m}\sum_{j=1}^{m}\sum_{k=1}^{m}j^{d_0-1}k^{d_0-1}n^{-2d_0}n^{1-2f_0}\right)$$

$$= O\left(\frac{(\log m)^2}{m}m^{2d_0}n^{-1}\right)$$

$$= O\left(\frac{m^{2d_0-1}(\log m)^2}{n}\right)$$

$$= O\left(\left(\frac{m}{n}\right)^{2d_0-1}\frac{(\log m)^2}{n^{2-2d_0}}\right) = o(1),$$

and for $d_0\in\left(1,\frac{3}{2}\right)$,

$$\frac{(\log m)^2}{m}\sum_{j=1}^{m}\sum_{k=1}^{m}\left(\frac{j}{n}\right)^{d_0-1}\left(\frac{k}{n}\right)^{d_0-1}\frac{1}{n^2}\sum_{p=1}^{n-1}n^{-f_0}\frac{n}{p^{f_0+1}j}$$

$$= O\left(\frac{(\log m)^2}{m}\sum_{j=1}^{m}\sum_{k=1}^{m}j^{d_0-2}k^{d_0-1}n^{-2d_0}n^{1-2f_0}\right)$$

$$= O\left(\frac{(\log m)^2}{m}m^{d_0-1}m^{d_0}n^{-1}\right)$$

$$= O\left(\frac{m^{2d_0-2}(\log m)^2}{n}\right)$$

$$= O\left(\left(\frac{m}{n}\right)^{2d_0-2}\frac{(\log m)^2}{n^{3-2d_0}}\right) = o(1).$$

For $d_0 \in \left(\frac{1}{2}, 1\right)$, (??) is bounded by

$$\left[\frac{\log m}{\sqrt{m}} \sum_{j=1}^{m} \left(\frac{j}{n}\right)^{d_0-1} \frac{1}{n} \left(\sum_{p=0}^{n-1} \left|\widetilde{f}_{\lambda_j p}\right|\right)\right]^2$$

$$= \left[\frac{\log m}{\sqrt{m}} \sum_{j=1}^{m} \left[\left(\frac{j}{n}\right)^{d_0-1} \frac{1}{n} \sum_{p=0}^{L} \left|\widetilde{f}_{\lambda_j p}\right| + \left(\frac{j}{n}\right)^{d_0-1} \frac{1}{n} \sum_{p=L+1}^{n-1} \left|\widetilde{f}_{\lambda_j p}\right|\right]\right]^2.$$

>From Lemma 7.7, the first term in the bracket is

$$O\left(\left(\frac{j}{n}\right)^{d_0-1} \frac{1}{n} \sum_{p=0}^{L} \frac{1}{p^{f_0}}\right) = O\left(\left(\frac{j}{n}\right)^{d_0-1} \frac{1}{n} L^{1-f_0}\right) = O\left(j^{d_0-1} n^{-d_0} L^{d_0}\right),$$

and the second term is

$$O\left(\left(\frac{j}{n}\right)^{d_0-1} \frac{1}{n} \sum_{p=L+1}^{n-1} \frac{n}{p^{f_0+1} j}\right) = O\left(\left(\frac{j}{n}\right)^{d_0-1} \frac{1}{j} L^{-f_0}\right) = O\left(j^{d_0-2} n^{1-d_0} L^{d_0-1}\right).$$

It follows that

$$\frac{\log m}{\sqrt{m}} \sum_{j=1}^{m} \left(\frac{j}{n}\right)^{d_0-1} \frac{1}{n} \left(\sum_{p=0}^{n-1} \left|\widetilde{f}_{\lambda_j p}\right|\right)$$

$$= O\left(\frac{\log m}{\sqrt{m}} \sum_{j=1}^{m} \left(j^{d_0-1} n^{-d_0} L^{d_0} + j^{d_0-2} n^{1-d_0} L^{d_0-1}\right)\right)$$

$$= O\left(\frac{\log m}{\sqrt{m}} m^{d_0} \left(\frac{L}{n}\right)^{d_0} + \frac{\log m}{\sqrt{m}} \left(\frac{n}{L}\right)^{1-d_0}\right).$$

Choose $L = \frac{n}{\sqrt{m}}$, and we obtain

$$m^{d_0-\frac{1}{2}} \left(\frac{L}{n}\right)^{d_0} = m^{d_0-\frac{1}{2}} m^{-\frac{d_0}{2}} = m^{\frac{2d_0-1-d_0}{2}} = m^{\frac{d_0-1}{2}} = o\left(\frac{1}{\log m}\right),$$

$$m^{-\frac{1}{2}} \left(\frac{n}{L}\right)^{1-d_0} = m^{-\frac{1}{2}} m^{\frac{1-d_0}{2}} = m^{\frac{-d_0}{2}} = o\left(\frac{1}{\log m}\right),$$

thus (??) is $o\left(1\right)$. For $d_0 \in \left(1, \frac{3}{2}\right)$, (??) is bounded by

$$\left[\frac{\log m}{\sqrt{m}} \sum_{j=1}^{m} \left(\frac{j}{n}\right)^{d_0-1} \frac{1}{n} \left(\sum_{p=0}^{n-1} \left|\widetilde{f}_{\lambda_j p}\right|\right)\right]^2.$$

43

Using Lemma 7.7, we have

$$\frac{\log m}{\sqrt{m}} \sum_{j=1}^{m} \left(\frac{j}{n}\right)^{d_0-1} \frac{1}{n} \left(\sum_{p=0}^{n-1} \left|\widetilde{f}_{\lambda_j p}\right|\right)$$

$$= O\left(\frac{\log m}{\sqrt{m}} \sum_{j=1}^{m} \left(\frac{j}{n}\right)^{d_0-1} \frac{1}{n} \sum_{p=1}^{n-1} \frac{n}{p^{f_0+1} j}\right)$$

$$= O\left(\frac{\log m}{\sqrt{m}} \sum_{j=1}^{m} \left(\frac{j}{n}\right)^{d_0-1} \frac{n^{-f_0}}{j}\right)$$

$$= O\left(\frac{\log m}{\sqrt{m}} \sum_{j=1}^{m} j^{d_0-2}\right) = O\left(m^{d_0-\frac{3}{2}} \log m\right),$$

and, hence, (??) is $o(1)$.

In view of the fact that $\sum_{q=0}^{n-1} e^{iq(\lambda_j-\lambda_k)} = n\mathbf{1}(j=k)$, (??) is bounded by, for $d_0 \in \left(\frac{1}{2}, 1\right)$,

$$\frac{(\log m)^2}{mn} \sum_{j=1}^{m} \lambda_j^{2d_0-2} \sum_{p=0}^{n-1} \left|\widetilde{f}_{\lambda_j p}\right|^2 = O\left(\frac{(\log m)^2}{mn} \sum_{j=1}^{m} \left(\frac{j}{n}\right)^{2d_0-2} \sum_{p=1}^{n} p^{-2f_0}\right)$$

$$= O\left(\frac{(\log m)^2}{mn} \sum_{j=1}^{m} \left(\frac{j}{n}\right)^{2d_0-2} n^{1-2f_0}\right)$$

$$= O\left((\log m)^2 m^{2d_0-2}\right),$$

and for $d_0 \in \left(1, \frac{3}{2}\right)$,

$$\frac{(\log m)^2}{mn} \sum_{j=1}^{m} \lambda_j^{2d_0-2} \sum_{p=0}^{n-1} \left|\widetilde{f}_{\lambda_j p}\right|^2 = O\left(\frac{(\log m)^2}{mn} \sum_{j=1}^{m} \left(\frac{j}{n}\right)^{2d_0-2} \left(\sum_{p=1}^{n-1} n^{-f_0} \frac{n}{p^{f_0+1} j}\right)\right)$$

$$= O\left(\frac{(\log m)^2}{mn} \sum_{j=1}^{m} \left(\frac{j}{n}\right)^{2d_0-2} n^{2d_0-1} j^{-1}\right)$$

$$= O\left(\frac{(\log m)^2}{m} \sum_{j=1}^{m} j^{2d_0-3}\right)$$

$$= O\left((\log m)^2 m^{2d_0-3}\right).$$

Therefore, (??) converges to zero, and thus $D_3 = o_p(1)$.

D_4, D_5 and D_6 are all $o_p(1)$ because

$$E|D_4| = O\left(\frac{1}{\sqrt{m}} \sum_{j=1}^{m} |\nu_j| \lambda_j\right) = O\left(\frac{\log m}{\sqrt{m}} \sum_{j=1}^{m} \frac{j}{n}\right) = O\left(\frac{m^{\frac{3}{2}} \log m}{n}\right),$$

$$E\left|D_5\right| \;=\; O\left(\frac{1}{\sqrt{m}}\sum_{j=1}^{m}|\nu_j|\,j^{d_0-2}\right) = \begin{cases} O\left(\frac{\log m}{\sqrt{m}}\right), & \text{for } d_0 \in \left(\tfrac{1}{2},1\right), \\[2mm] O\left(m^{d_0-\frac{3}{2}}\log m\right), & \text{for } d_0 \in \left(1,\tfrac{3}{2}\right), \end{cases}$$

$$E\left|D_6\right| \;=\; \begin{cases} O\left(\frac{2}{\sqrt{m}}\sum_{j=1}^{m}|\nu_j|\,j^{d_0-1}n^{\frac{1}{2}-d_0}\right) = O\left(n^{\frac{1}{2}-d_0}m^{d_0-\frac{1}{2}}\log m\right), & \text{for } d_0 \in \left(\tfrac{1}{2},1\right), \\[2mm] O\left(\frac{2}{\sqrt{m}}\sum_{j=1}^{m}|\nu_j|\,j^{d_0-1}n^{-\frac{1}{2}}\right) = O\left(n^{-\frac{1}{2}}m^{d_0-\frac{1}{2}}\log m\right), & \text{for } d_0 \in \left(1,\tfrac{3}{2}\right). \end{cases}$$

Furthermore,

$$\frac{m^{d_0-\frac{1}{2}}\log m}{n^{d_0-\frac{1}{2}}} = \left(\frac{m^{\frac{3}{2}}}{n}\right)^{d_0-\frac{1}{2}} m^{-\frac{1}{2}\left(d_0-\frac{1}{2}\right)}\log m \to 0,$$

if $\frac{m^{\frac{3}{2}}\log m}{n} \to 0$.

Therefore, we obtain

$$m^{\frac{1}{2}}R'(d_0) \Rightarrow \frac{1}{G_0}N\left(0,4G_0^2\right). \tag{54}$$

It follows from (??), (??) and (??) that

$$m^{\frac{1}{2}}\left(\widehat{d}-d_0\right) = -\frac{m^{\frac{1}{2}}R'(d_0)}{R''(d^*)} \Rightarrow \frac{1}{4G_0}N\left(0,4G_0^2\right) \equiv N\left(0,\frac{1}{4}\right),$$

giving the required result. $\blacksquare$

8.8 Proof of Theorem 4.2

The proof follows the same line of argument as Theorem 4.1. The condition $\frac{m^{\frac{3}{2}}\log m}{n} \to 0$ implies that

$$\frac{m^{2d_0-2}(\log m)^{12}}{n} = \frac{m^{\frac{3}{2}}\log m}{n}\frac{(\log m)^{11}}{m^{\frac{7}{2}-2d_0}} \to 0.$$

Thus, $\widehat{d}$ is consistent and also we have $\sup_{\Theta_1}\left|\frac{A(d)}{B(d)}\right| = o_p\left((\log m)^{-6}\right)$, which gives $R''(d^*) = R''(d_0) + o_p(1)$. It follows from Corollary 2.6 (e) that

$$\widehat{F}_k(d_0) = \frac{1}{m}\sum_{j=1}^{m}(\log j)^k\,\lambda_j^{2d_0}\,I_v(\lambda_j) = \sum_{j=1}^{6}C_j$$

where $C_1 = \frac{1}{m}\sum_{j=1}^{m}(\log j)^k\left(G_0 + o_p(1)\right)$, $C_2 = o_p\left((\log m)^k\right)$, and

$$E\left|C_3\right| \;=\; O\left(\frac{1}{m}\sum_{j=1}^{m}(\log j)^k\,j^{d_0-2}(\log n)^{\frac{1}{2}}\right) = O\left((\log m)^k\,m^{d_0-2}(\log n)^{\frac{1}{2}}\right),$$

$$E\left|C_4\right| \;=\; O\left(\frac{1}{m}\sum_{j=1}^{m}(\log j)^k\,j^{d_0-1}n^{-\frac{1}{2}}\right) = O\left((\log m)^k\,m^{d_0-1}n^{-\frac{1}{2}}\right),$$

$$
E\left|C_5\right| \;=\; O\left(\frac{1}{m}\sum_{j=1}^{m}(\log j)^k\, j^{2d_0-4}\log n\right) = O\left((\log m)^k\, m^{2d_0-4}\log n\right),
$$

$$
E\left|C_6\right| \;=\; O\left(\frac{1}{m}\sum_{j=1}^{m}(\log j)^k\, j^{2d_0-2}n^{-1}\right) = O\left((\log m)^k\, m^{2d_0-2}n^{-1}\right),
$$

giving $\widehat{F}_k\left(d_0\right) = G_0\left[\frac{1}{m}\sum_{j=1}^{m}(\log j)^k\right]\left[1+o_p\left(1\right)\right].$

Before we evaluate $\frac{2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\left[\lambda_j^{2d_0}I_v\left(\lambda_j\right)-G_0\right]$, we derive the approximation of $\lambda_j^{2d_0}I_v\left(\lambda_j\right)$ for $d_0\in\left(\frac{3}{2},2\right)$. First, note that (see (26))

$$
\begin{aligned}
\Delta X_t \;&=\; X_t - X_{t-1} \\
&=\; \sum_{k=0}^{t-1}\frac{(d)_k}{k!}u_{t-k} - \sum_{k=0}^{(t-1)-1}\frac{(d)_k}{k!}u_{t-k-1} \\
&=\; \sum_{k=0}^{t-1}\frac{(d)_k}{k!}u_{t-k} - \sum_{k+1=1}^{t-1}\frac{(d)_{k+1-1}}{(k+1-1)!}u_{t-(k+1)} \\
&=\; u_t + \sum_{k=1}^{t-1}\frac{(d)_k}{k!}u_{t-k} - \sum_{k=1}^{t-1}\frac{(d)_{k-1}}{(k-1)!}u_{t-k} \\
&=\; u_t + \sum_{k=1}^{t-1}\frac{(d-1)_k}{k!}u_{t-k} = \sum_{k=0}^{t-1}\frac{(d-1)_k}{k!}u_{t-k},
\end{aligned}
$$

and $\Delta X_0 = 0$. Let $\overline{d}_0 = d_0 - 1 \in \left(\frac{1}{2},1\right)$ and $\overline{f}_0 = 1 - \overline{d}_0$. From Lemma 2.5 (a) we have

$$
\lambda_s^{\overline{d}_0}w_{\Delta x}\left(\lambda_s\right) = e^{\frac{\pi}{2}\overline{d}_0 i}C\left(1\right)w_\varepsilon\left(\lambda_s\right) - \frac{\lambda_s^{\overline{d}_0}C\left(1\right)\widetilde{\varepsilon}_{\lambda_s n}\left(\overline{f}_0\right)}{1-e^{i\lambda_s}}\frac{}{\sqrt{2\pi n}} - \frac{\lambda_s^{\overline{d}_0}e^{i\lambda_s}}{1-e^{i\lambda_s}}\frac{\Delta X_n}{\sqrt{2\pi n}} + r_{s,n}^a + r_{s,n}^b\left(\overline{d}_0\right) + r_{s,n}^c\left(\overline{d}_0\right),
$$

where $E\left|r_{s,n}^a\right|^2 = O\left(\lambda_s^2\right), E\left|r_{s,n}^b\left(\overline{d}_0\right)\right|^2 = O\left(s^{2\overline{d}_0-4}\right),$ and $E\left|r_{s,n}^c\left(\overline{d}_0\right)\right|^2 = O\left(s^{2\overline{d}_0-2}n^{1-2\overline{d}_0}\right)$ uniformly in s. Lemma 7.18 (c) yields

$$
E\left|\frac{\lambda_s^{\overline{d}_0}e^{i\lambda_s}}{1-e^{i\lambda_s}}\frac{\Delta X_n}{\sqrt{2\pi n}}\right|^2 = O\left(\lambda_s^{2\overline{d}_0-2}n^{2\overline{d}_0-2}\right) = O\left(s^{2\overline{d}_0-2}\right) = o\left(1\right),
$$

hence we obtain

$$
\begin{aligned}
\lambda_s^{2\overline{d}_0}I_{\Delta x}\left(\lambda_s\right) \;=\;& \left|e^{\frac{\pi}{2}\overline{d}_0 i}C\left(1\right)w_\varepsilon\left(\lambda_s\right) - \frac{\lambda_s^{\overline{d}_0}C\left(1\right)\widetilde{\varepsilon}_{\lambda_s n}\left(\overline{f}_0\right)}{1-e^{i\lambda_s}}\frac{}{\sqrt{2\pi n}} - \frac{\lambda_s^{\overline{d}_0}e^{i\lambda_s}}{1-e^{i\lambda_s}}\frac{\Delta X_n}{\sqrt{2\pi n}}\right|^2 \\
& + R_{s,n}^a + R_{s,n}^b\left(\overline{d}_0\right) + R_{s,n}^c\left(\overline{d}_0\right),
\end{aligned}
$$

where $E\left|R_{s,n}^a\right| = O\left(\lambda_s\right), E\left|R_{s,n}^b\left(\overline{d}_0\right)\right| = O\left(s^{\overline{d}_0-2}\right),$ and $E\left|R_{s,n}^c\left(\overline{d}_0\right)\right| = O\left(s^{\overline{d}_0-1}n^{\frac{1}{2}-\overline{d}_0}\right)$ uniformly in s.

>From Lemma 2.3 (b), we have

$$v_x(\lambda_s) = \left(1 - e^{i\lambda_s}\right)^{-1} w_{\Delta x}(\lambda_s).$$

Thus

$$I_v(\lambda_s) = \left|1 - e^{i\lambda_s}\right|^{-2} I_{\Delta x}(\lambda_s),$$

and, in view of the fact that $E\left|\lambda_s^{2\overline{d}_0} I_{\Delta x}(\lambda_s)\right| = O(1)$, it follows that

$$
\begin{aligned}
\lambda_s^{2d_0} I_v(\lambda_s) &= \lambda_s^{2d_0-2}\left(1 + O(\lambda_s)\right) I_{\Delta x}(\lambda_s) \\
&= \left| e^{\frac{\pi}{2}\overline{d}_0 i} C(1) w_\varepsilon(\lambda_s) - \frac{\lambda_s^{\overline{d}_0} C(1)}{1 - e^{i\lambda_s}} \frac{\widetilde{\varepsilon}_{\lambda_s n}\left(\overline{f}_0\right)}{\sqrt{2\pi n}} - \frac{\lambda_s^{\overline{d}_0} e^{i\lambda_s}}{1 - e^{i\lambda_s}} \frac{\Delta X_n}{\sqrt{2\pi n}} \right|^2 \\
&\quad + R_{s,n}^a + R_{s,n}^b\left(\overline{d}_0\right) + R_{s,n}^c\left(\overline{d}_0\right),
\end{aligned}
$$

where the order of magnitude of the reminder terms is the same as above.

Finally, we obtain the expression

$$\frac{2}{\sqrt{m}} \sum_{j=1}^{m} \nu_j \left[\lambda_j^{2d_0} I_v(\lambda_j) - G_0\right] = \sum_{k=1}^{9} D_k,$$

where

$$D_1 = \frac{2\left|C(1)\right|^2}{\sqrt{m}} \sum_{j=1}^{m} \nu_j \left[I_\varepsilon(\lambda_j) - \frac{\sigma^2}{2\pi}\right],$$

$$D_2 = \frac{2\left|C(1)\right|^2}{\sqrt{m}} \sum_{j=1}^{m} \nu_j \frac{\lambda_j^{2\overline{d}_0}}{\left|1 - e^{i\lambda_j}\right|^2} \frac{\left|\widetilde{\varepsilon}_{\lambda_j n}(\overline{f}_0)\right|^2}{2\pi n}, \quad D_3 = \frac{2}{\sqrt{m}} \frac{\Delta X_n^2}{2\pi n} \sum_{j=1}^{m} \nu_j \frac{\lambda_j^{2\overline{d}_0}}{\left|1 - e^{i\lambda_j}\right|^2}$$

$$D_4 = -\frac{2\left|C(1)\right|^2}{\sqrt{m}} \sum_{j=1}^{m} \nu_j \left[e^{\frac{\pi}{2}\overline{d}_0 i} w_\varepsilon(\lambda_j) \frac{\lambda_j^{\overline{d}_0}}{1 - e^{-i\lambda_j}} \frac{\widetilde{\varepsilon}_{\lambda_j n}\left(\overline{f}_0\right)^*}{\sqrt{2\pi n}} + \frac{\lambda_j^{\overline{d}_0}}{1 - e^{i\lambda_j}} \frac{\widetilde{\varepsilon}_{\lambda_j n}\left(\overline{f}_0\right)}{\sqrt{2\pi n}} e^{-\frac{\pi}{2}\overline{d}_0 i} w_\varepsilon(\lambda_j)^* \right],$$

$$D_5 = -\frac{2C(1)}{\sqrt{m}} \frac{\Delta X_n}{\sqrt{2\pi n}} \sum_{j=1}^{m} \nu_j \left[e^{\frac{\pi}{2}\overline{d}_0 i} w_\varepsilon(\lambda_j) \frac{\lambda_j^{\overline{d}_0} e^{-i\lambda_j}}{1 - e^{-i\lambda_j}} + \frac{\lambda_j^{\overline{d}_0} e^{i\lambda_j}}{1 - e^{i\lambda_j}} e^{-\frac{\pi}{2}\overline{d}_0 i} w_\varepsilon(\lambda_j)^* \right],$$

$$D_6 = \frac{2C(1)}{\sqrt{m}} \frac{\Delta X_n}{2\pi n} \sum_{j=1}^{m} \nu_j \left[\frac{\lambda_j^{2\overline{d}_0} e^{-i\lambda_j}}{\left|1 - e^{i\lambda_j}\right|^2} \widetilde{\varepsilon}_{\lambda_j n}\left(\overline{f}_0\right) + \frac{\lambda_j^{2\overline{d}_0} e^{i\lambda_j}}{\left|1 - e^{i\lambda_j}\right|^2} \widetilde{\varepsilon}_{\lambda_j n}\left(\overline{f}_0\right)^* \right],$$

$$D_7 = \frac{2}{\sqrt{m}} \sum_{j=1}^{m} \nu_j R_{j,n}^a, \quad D_8 = \frac{2}{\sqrt{m}} \sum_{j=1}^{m} \nu_j R_{j,n}^b\left(\overline{d}_0\right), \quad D_9 = \frac{2}{\sqrt{m}} \sum_{j=1}^{m} \nu_j R_{j,n}^c\left(\overline{d}_0\right).$$

It has already shown that $D_1 \to_d N\left(0, 4G_0^2\right)$ and $D_2 + D_4 + D_7 + D_8 + D_9 = o_p(1)$. For D_3, from Lemma 7.18 (c) we have

$$E\left|D_3\right| = E\left(\frac{\Delta X_n^2}{n^{2\overline{d}_0-1}}\right) O\left(\frac{1}{\sqrt{m}} \sum_{j=1}^{m} \nu_j j^{2\overline{d}_0-2}\right) = O\left(\frac{\log m}{\sqrt{m}} \sum_{j=1}^{m} j^{2\overline{d}_0-2}\right) = O\left(m^{2\overline{d}_0-\frac{3}{2}} \log m\right),$$

which is $o(1)$ if $\overline{d}_0 < \frac{3}{4} \Leftrightarrow d_0 < \frac{7}{4}$.

For D_5, rewrite

$$\frac{1}{\sqrt{m}} \frac{\Delta X_n}{\sqrt{n}} \sum_{j=1}^{m} \nu_j w_\varepsilon(\lambda_j) \frac{\lambda_j^{\overline{d}_0} e^{-i\lambda_j}}{1 - e^{-i\lambda_j}}$$

as

$$\frac{\Delta X_n}{n^{\overline{d}_0 - \frac{1}{2}}} \times \frac{n^{\overline{d}_0 - 1}}{\sqrt{m}} \sum_{j=1}^{m} \nu_j w_\varepsilon(\lambda_j) \frac{\lambda_j^{\overline{d}_0} e^{-i\lambda_j}}{1 - e^{-i\lambda_j}} = D_{51} \times D_{52}.$$

In view of the fact that

$$E\left[w_\varepsilon(\lambda_j)^* w_\varepsilon(\lambda_k)\right] = \frac{1}{2\pi n} \sum_{p=1}^{n} \sum_{q=1}^{n} E\left[\varepsilon_p e^{-ip\lambda_j}\right]\left[\varepsilon_q e^{iq\lambda_k}\right] = \frac{\sigma^2}{2\pi n} \sum_{p=1}^{n} e^{i(\lambda_k - \lambda_j)p} = \frac{\sigma^2}{2\pi} \mathbf{1}(j = k),$$

we have

$$E\left|D_{52}\right|^2 = O\left(\frac{(\log m)^2}{m} \sum_{j=1}^{m} j^{2\overline{d}_0 - 2}\right) = O\left((\log m)^2 m^{2\overline{d}_0 - 2}\right) = o(1),$$

hence $D_{52} = o_p(1)$. $D_{51} = O_p(1)$ by Corollary 7.18 (c), and $D_5 = o_p(1)$ follows.

For D_6, rewrite

$$\frac{1}{\sqrt{m}} \frac{\Delta X_n}{n} \sum_{j=1}^{m} \nu_j \frac{\lambda_j^{2\overline{d}_0} e^{-i\lambda_j}}{\left|1 - e^{i\lambda_j}\right|^2} \widetilde{\varepsilon}_{\lambda_j n}\left(\overline{f}_0\right)$$

as

$$\frac{\Delta X_n}{n^{\overline{d}_0 - \frac{1}{2}}} \times \frac{n^{\overline{d}_0 - \frac{3}{2}}}{\sqrt{m}} \sum_{j=1}^{m} \nu_j \frac{\lambda_j^{2\overline{d}_0} e^{-i\lambda_j}}{\left|1 - e^{i\lambda_j}\right|^2} \widetilde{\varepsilon}_{\lambda_j n}\left(\overline{f}_0\right) = D_{61} \times D_{62}.$$

We have $E\left|D_{61}\right|^2 = O(1)$, and from Lemma 7.12 (c), $E\left|D_{62}\right|$ is bounded by

$$\frac{n^{\overline{d}_0 - \frac{3}{2}} \log m}{\sqrt{m}} \sum_{j=1}^{m} \left(\frac{j}{n}\right)^{2\overline{d}_0 - 2} L^{\overline{d}_0 - \frac{1}{2}} + \frac{n^{\overline{d}_0 - \frac{3}{2}} \log m}{\sqrt{m}} \sum_{j=1}^{m} \left(\frac{j}{n}\right)^{2\overline{d}_0 - 2} \left(\frac{n}{j}\right)^{1/2} L^{\overline{d}_0 - 1}$$

$$= O\left(n^{\frac{1}{2} - \overline{d}_0}(\log m) m^{2\overline{d}_0 - \frac{3}{2}} L^{\overline{d}_0 - \frac{1}{2}} + n^{1 - \overline{d}_0}(\log m)^2 m^{-\frac{1}{2}} m^{\max\{2\overline{d}_0 - \frac{3}{2}, 0\}} L^{\overline{d}_0 - 1}\right)$$

$$= O\left(\log m \left(\frac{mL}{n}\right)^{\overline{d}_0 - \frac{1}{2}} m^{\overline{d}_0 - 1} + (\log m)^2 \left(\frac{mL}{n}\right)^{\overline{d}_0 - 1} m^{\max\{\overline{d}_0 - 1, \frac{1}{2} - \overline{d}_0\}}\right) = o(1),$$

by setting $L = \frac{n}{m}$. $D_6 = o_p(1)$ follows from Cauchy-Schwartz inequality. Therefore, $m^{\frac{1}{2}} R'(d_0) \Rightarrow \frac{1}{G_0} N\left(0, 4G_0^2\right)$, giving the required result. ∎

48

8.9 Proof of Theorem 4.4

The condition $\frac{m^{2d_0-2}(\log m)^{12}}{n} \to 0$ implies that $\sup_{\Theta_1}\left|\frac{A(d)}{B(d)}\right| = o_p\left((\log m)^{-6}\right)$, which gives $R''(d^*) = R''(d_0) + o_p(1)$. Recall

$$
\begin{aligned}
m^{4-2d_0}\left(\widehat{d} - d_0\right) &= -\frac{m^{4-2d_0}R'(d_0)}{R''(d^*)} \\
&= -\frac{m^{\frac{7}{2}-2d_0}(D_1 + D_2 + D_4 + D_5 + D_6 + D_7 + D_8 + D_9)}{4G_0 + o_p(1)} \tag{55} \\
&\quad -\frac{m^{\frac{7}{2}-2d_0}D_3}{4G_0 + o_p(1)}, \tag{56}
\end{aligned}
$$

where

$$
\begin{aligned}
D_3 &= \frac{2}{\sqrt{m}}\frac{\Delta X_n^2}{2\pi n}\sum_{j=1}^{m}\nu_j\frac{\lambda_j^{2\overline{d}_0}}{\left|1 - e^{i\lambda_j}\right|^2} \\
&= \frac{2(2\pi)^{2\overline{d}_0-2}}{\sqrt{m}}\frac{\Delta X_n^2}{2\pi n^{2\overline{d}_0-1}}\sum_{j=1}^{m}\nu_j j^{2\overline{d}_0-2}\left(1 + O(\lambda_j)\right) \\
&= \frac{2(2\pi)^{2\overline{d}_0-2}}{\sqrt{m}}\frac{\Delta X_n^2}{2\pi n^{2\overline{d}_0-1}}\sum_{j=1}^{m}\nu_j j^{2\overline{d}_0-2} + \frac{\Delta X_n^2}{n^{2\overline{d}_0-1}}O\left(\frac{\log m}{n\sqrt{m}}\sum_{j=1}^{m}j^{2\overline{d}_0-1}\right).
\end{aligned}
$$

The second term is

$$
O_p(1)O\left(\frac{m^{2\overline{d}_0-\frac{1}{2}}\log m}{n}\right) = o_p(1).
$$

Now evaluate the summation:

$$
\sum_{j=1}^{m}\nu_j j^{2\overline{d}_0-2} = \sum_{j=1}^{m}\left(\log j - \frac{1}{m}\sum_{j=1}^{m}\log j\right)j^{2\overline{d}_0-2} = \sum_{j=1}^{m}j^{2\overline{d}_0-2}\log j - \sum_{j=1}^{m}j^{2\overline{d}_0-2}\left(\frac{1}{m}\sum_{j=1}^{m}\log j\right).
$$

For the first term, we have

$$
\begin{aligned}
\sum_{j=1}^{m}j^{2\overline{d}_0-2}\log j &= \int_1^m x^{2\overline{d}_0-2}\log x\, dx + \frac{1}{2}\left(m^{2\overline{d}_0-2}\log m\right) + O\left(\log m\int_1^m x^{2\overline{d}_0-3}dx\right) \\
&= \left[\frac{x^{2\overline{d}_0-1}\log x}{2\overline{d}_0 - 1}\right]_1^m - \int_1^m\frac{x^{2\overline{d}_0-2}}{2\overline{d}_0 - 1}dx + O(\log m) \\
&= \frac{m^{2\overline{d}_0-1}\log m}{2\overline{d}_0 - 1} - \frac{m^{2\overline{d}_0-1}}{(2\overline{d}_0-1)^2} + O(\log m),
\end{aligned}
$$

and, for the second term,

$$
\sum_{j=1}^{m}j^{2\overline{d}_0-2}\left(\frac{1}{m}\sum_{j=1}^{m}\log j\right) = \left(\int_1^m x^{2\overline{d}_0-2}dx + \frac{1}{2}\left(m^{2\overline{d}_0-2}+1\right) + O\left(\int_1^m x^{2\overline{d}_0-3}dx\right)\right)
$$

$$\times \left(\log m - 1 + O\left(\frac{\log m}{m} \right) \right)$$

$$= \left(\frac{m^{2\overline{d}_0 - 1}}{2\overline{d}_0 - 1} + O\left(1\right) \right) \left(\log m - 1 + O\left(\frac{\log m}{m} \right) \right)$$

$$= \frac{m^{2\overline{d}_0 - 1} \log m}{2\overline{d}_0 - 1} - \frac{m^{2\overline{d}_0 - 1}}{2\overline{d}_0 - 1} + O\left(\log m \right).$$

Therefore,

$$\sum_{j=1}^{m} \nu_j j^{2\overline{d}_0 - 2} = -\frac{m^{2\overline{d}_0 - 1}}{\left(2\overline{d}_0 - 1\right)^2} + \frac{m^{2\overline{d}_0 - 1}}{2\overline{d}_0 - 1} + O\left(\log m\right) = \frac{\left(2\overline{d}_0 - 2\right) m^{2\overline{d}_0 - 1}}{\left(2\overline{d}_0 - 1\right)^2} + O\left(\log m\right).$$

>From Akonom and Gourieroux (1987), if $E\left|\varepsilon_t\right|^p < \infty$ for $p > \max\left\{ \frac{1}{\overline{d}_0 - \frac{1}{2}}, 2 \right\}$, we have

$$\frac{\Delta X_n^{\varepsilon}}{n^{\overline{d}_0 - \frac{1}{2}}} \to_d \sigma B_{\overline{d}_0 - 1}\left(1\right) = \frac{\sigma}{\Gamma\left(\overline{d}_0\right)} \int_0^1 \left(1 - s\right)^{\overline{d}_0 - 1} dB\left(s\right).$$

When $d_0 = \frac{7}{4}$, $\frac{1}{\overline{d}_0 - \frac{1}{2}} = 4$ and we need an additional moment condition $E\left|\varepsilon_t\right|^p < \infty$ for $p > 4$ for convergence. When $d_0 > \frac{7}{4}$, $\frac{1}{\overline{d}_0 - \frac{1}{2}} < 4$ and the condition $E\left|\varepsilon_t\right|^4 < \infty$ suffice. It follows that, for $d_0 \in \left[\frac{7}{4}, 2\right)$,

$$m^{\frac{7}{2} - 2d_0} D_3 = 2\left(2\pi\right)^{2\overline{d}_0 - 2} \frac{\Delta X_n^2}{2\pi n^{2\overline{d}_0 - 1}} m^{1 - 2\overline{d}_0} \sum_{j=1}^{m} \nu_j j^{2\overline{d}_0 - 2}$$

$$= 2\left(2\pi\right)^{2\overline{d}_0 - 2} \left[\frac{C\left(1\right)^2 \left(\Delta X_n^{\varepsilon}\right)^2}{2\pi n^{2\overline{d}_0 - 1}} + o_p\left(1\right) \right] m^{1 - 2\overline{d}_0} \sum_{j=1}^{m} \nu_j j^{2\overline{d}_0 - 2}$$

$$\to_d 2\left(2\pi\right)^{2\overline{d}_0 - 2} \frac{C\left(1\right)^2 \sigma^2}{2\pi} B_{\overline{d}_0 - 1}\left(1\right)^2 \frac{2\overline{d}_0 - 2}{\left(2\overline{d}_0 - 1\right)^2}.$$

For $d_0 = \frac{7}{4}$, (??) converges to $N\left(0, \frac{1}{4}\right)$ and

$$-\frac{m^{\frac{7}{2} - 2d_0} D_3}{4G_0 + o_p\left(1\right)} \to_d \frac{\left(1 - \overline{d}_0\right) \left(2\pi\right)^{2\overline{d}_0 - 2}}{\left(2\overline{d}_0 - 1\right)^2} B_{\overline{d}_0 - 1}\left(1\right)^2 \equiv \left(2\pi\right)^{-\frac{1}{2}} B_{-\frac{1}{4}}\left(1\right)^2.$$

For $d_0 \in \left(\frac{3}{4}, 1\right)$, (??) is $o_p\left(1\right)$ and

$$-\frac{m^{\frac{7}{2} - 2d_0} D_3}{4G_0 + o_p\left(1\right)} \to_d \frac{\left(1 - \overline{d}_0\right) \left(2\pi\right)^{2\overline{d}_0 - 2}}{\left(2\overline{d}_0 - 1\right)^2} B_{\overline{d}_0 - 1}\left(1\right)^2 = \frac{\left(2 - d_0\right) \left(2\pi\right)^{2d_0 - 4}}{\left(2d_0 - 3\right)^2} B_{d_0 - 2}\left(1\right)^2,$$

giving the required result. ∎

50

9 Appendix C: Different Characterizations of Nonstationary $I(d)$ Processes

Two main approaches to defining a nonstationary $I(d)$ process have been used in the literature to date. They are by no means exhaustive. The first, which is used in Hurvich and Ray (1995) and Velasco (1999a, 1999b), is to define the observed process X_t as the partial sum of a stationary fractionally integrated process, viz.

$$X_t = X_0 + \sum_{j=1}^{t} z_j, \quad t \geq 1, \tag{57}$$

where z_j is a stationary $I(d-1)$ process and satisfies

$$z_t = (1-L)^{1-d} \varepsilon_t = \sum_{j=0}^{\infty} \frac{(d-1)_j}{j!} \varepsilon_{t-j}, \tag{58}$$

where ε_t is a short-memory stationary process. Combining (??) and (??), we obtain

$$(1-L)(X_t - X_0) = (1-L)^{1-d} \varepsilon_t,$$

leading to a definition of the operator equation

$$(1-L)^d (X_t - X_0) = \varepsilon_t, \quad t \geq 1, \tag{59}$$

in terms of (??) and (??). X_t is said to be integrated of order d.

A second definition (Phillips, 1999), corresponding to that in (2) above, defines the nonstationary fractionally integrated process X_t directly in terms of the short memory inputs by using a finite order expansion of the operator $(1-L)^{-d}$, viz.

$$X_t = X_0 + \sum_{k=0}^{t-1} \frac{(d)_k}{k!} \varepsilon_{t-k}, \tag{60}$$

where ε_t is a short-memory stationary process. This leads to the operator expression

$$(1-L)^d (X_t - X_0) = \varepsilon_t, \quad t \geq 1, \tag{61}$$

and again X_t is integrated of order d. The two definitions (??) and (??) are different, however, because the stationary input formulation (??) implies that, by the first definition, $X_t \left(= X_0 + \sum_{j=1}^{t} (1-L)^{1-d} \varepsilon_j \right)$ involves inputs ε_s with $s \leq 0$. In fact, for each t we have

$$(1-L)^{1-d} \varepsilon_t = \frac{(d-1)_0}{0!} \varepsilon_t + \frac{(d-1)_1}{1!} \varepsilon_{t-1} + ... + \frac{(d-1)_{t+k}}{(t+k)!} \varepsilon_{-k} + ...$$

so that the infinite past history of the short memory stationary inputs ε_s figures in X_t.

Some further comparisons involving the impulse responses may be helpful. When $d \in \left(\frac{1}{2}, 1\right)$, according to the first definition, X_t is integrated of order $d < 1$ and the increments z_t constitute an $I(f)$ process with negative $f = 1 - d$. In other words, the increments have negative correlation and are often described as antipersistent. On the other hand, according to the second definition, X_t is integrated of order $d < 1$ because the coefficients of ε_{t-k} are not unity but decay slowly, too slowly for the process to be stationary and have finite variance. Thus, the second definition gives the anticipated slow decay of the impulse responses directly, and as such is more apparently intermediate in form between a unit root process and a stationary long memory process or a short-memory process (but see (??) below).

In some cases, the empirical context may be helpful in motivating the formative process. Suppose that $d \in \left(1, \frac{3}{2}\right)$. Then, according to the first definition, X_t is integrated of order $d > 1$ because it is the accumulation of stationary increments z_t that have long memory with $f = 1 - d > 0$. According to the second definition, X_t is integrated of order $d > 1$ because the coefficients of ε_{t-k} increase as k increases. When it is known that the process of interest is the result of an accumulation of past long-memory shocks (perhaps, like the diameter of a tree), the first definition would seem to be appropriate. However, when it is expected that the shocks each period have short memory but may have increasing impulse responses over time on the observed variable, then the second definition seems more appropriate. For instance, in seeking to characterize a time series like GDP as a nonstationary $I(d)$ process with $d > 1$, the first definition posits GDP as the sum of past shocks which have long memory, whereas the second definition posits that the shocks to GDP each period have short memory but the cumulative effect of these shocks is allowed to increase over time, perhaps by way of some internal feedback mechanism.

Whether the first or the second definition is used, it will often be useful to extract the impulse responses from the short memory components to the observed series. In the second definition these appear directly as the coefficients $\frac{(d)_k}{k!}$ in (??). By rearrangement of the series in the first definition, one finds that the impulse responses are the same in this case as well. In particular, it can be shown that an $I(d)$ process by the first definition can be written as

$$X_t = X_0 + \xi_0(d) + \sum_{k=0}^{t-1} \frac{(d)_k}{k!} \varepsilon_{t-k}, \tag{62}$$

where the term $\xi_0(d)$ has an order of magnitude that is dominated by that of the third term asymptotically. Thus, the essential difference between the definitions can be interpreted as one relating to initialization.

As with the definition of unit root processes, there are alternative ways of dealing with initial conditions for nonstationary fractional processes and these may or may not affect large sample behavior. If X_0 is taken to be any $O_p(1)$ random variable then its value has no affect on large sample behavior. Similar considerations apply to the term $\xi_0(d)$ in (??). However, when X_0 has the same stochastic order as X_t for $t = O(n)$ then initializations do matter, as indeed has been found to be the case for unit root time series (e.g., Phillips and Lee, 1996, and Canjels and Watson, 1997). In

the present case, the generalization might involve a distant past initialization of the form

$$X_0 = X_0^\kappa = \sum_{k=0}^{[n\kappa]} \frac{(d)_k}{k!} \varepsilon_{-k},$$

or one might extend (**??**) directly by writing

$$X_t = \sum_{k=0}^{t+[n\kappa]} \frac{(d)_k}{k!} \varepsilon_{-k}.$$

In both these cases, the effective initialization is pushed into the distant past and is parameterized by κ, which measures the extent of the pre-sample history on the current data X_t. While κ is not consistently estimable, in general, it will figure in the asymptotic theory, just as it does in the case of unit root asymptotics (Phillip and Lee, 1996). The present paper does not deal with this additional level of difficulty, but works from the definition (**??**) with $X_0 = O_p(1)$.

References

[1] Akonom, J. and C. Gourieroux (1987). "A functional limit theorem for fractional processes." Technical Report #8801, CEPREMAP, Paris.

[2] Canjels, E. and M. W. Watson (1997). "Estimating deterministic trends in the presence of serially correlated errors." *Review of Economics and Statistics* **79**: 184-200.

[3] Corbae, D., S. Ouliaris and P. C. B. Phillips (1999). "Band spectral regression with trending data." Yale University, mimeographed.

[4] Erdélyi, A. (1953). *Higher Transcendental Functions: Volume 1*. Malabar, Krieger.

[5] Davidson, J. (1994). *Stochastic Limit Theory*. Oxford University Press.

[6] Hannan, E. J. (1970). *Multiple Time Series*. New York, Wiley.

[7] Hurvich, C. M. and B. K. Ray (1995). "Estimation of the memory parameter for nonstationary or noninvertible fractionally integrated processes." *Journal of Time Series Analysis* **16**: 17-42.

[8] Kim, C. S. and P. C. B. Phillips (1999). "Modified Log Periodogram Regression". Yale University, mimeographed.

[9] Künsch, H. (1987). "Statistical Aspects of Self-Similar Processes." In *Proceedings of the First World Congress of the Bernoulli Society*. Y. Prokhorov and V. V. Sazanov (Eds.). Utrecht, VNU Science Press. **1**: 67-74.

[10] Phillips, P. C. B. (1999). "Discrete Fourier transforms of fractional processes." Yale University, mimeographed.

[11] Phillips, P. C. B. and C. C. Lee (1996). "Efficiency gains from quasi-differencing under nonstationarity." *Essays in Memory of E. J. Hannan.* P. M. Robinson and M. Rosenblatt (Eds.). New York, Springer.

[12] Phillips, P. C. B. and V. Solo (1992). "Asymptotics for linear processes." *Annals of Statistics* **20**: 971-1001.

[13] Priestley, M. B. (1981). *Spectral Analysis and Time Series.* New York, Academic Press.

[14] Robinson, P. M. (1995a). "Log-periodogram regression of time series with long range dependence." *Annals of Statistics* **23**: 1048-1072.

[15] Robinson, P. M. (1995b). "Gaussian semiparametric estimation of long range dependence." *Annals of Statistics* **23**: 1630-1661.

[16] Tanaka, K. (1999). "The nonstationary fractional unit root." *Econometric Theory* **15**: 549-582.

[17] Velasco, C. (1999a). "Gaussian semiparametric estimation of non-stationary time series." *Journal of Time Series Analysis* **20**: 87-127.

[18] Velasco, C. (1999b). "Non-stationary log-periodogram regression." *Journal of Econometrics* **91**: 325-371.

[19] Zygmund, A. (1959). *Trigonometric Series.* Cambridge, Cambridge University Press.