Econometric Estimation of Models of Fractional Integration

A Dissertation
Presented to the Faculty of the Graduate School
of
Yale University
in Candidacy for the Degree of
Doctor of Philosophy

by
Katsumi Shimotsu

Dissertation Director: Peter C. B. Phillips

December 2000

UMI Number: 9991227

Bell & Howell Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

Contents

List of Tables

List of Illustrations

Chapter 2

Chapter 3

Chapter 4

Acknowledgments

I would like first to express my appreciation to Professor Peter C. B. Phillips who introduced me to the thrilling world of time series econometrics and provided invaluable mentorship throughout my dissertation research. His unlimited knowledge of, insights into, and enthusiasm for econometrics have been a guiding light and source of wonder for me and will continue to be in the future. I want also to thank Professor Donald W. K. Andrews for his pertinent comments and warm support. I am grateful to Professors John Rust and Oliver Linton for their helpful comments and advice.

I also wish to acknowledge my indebtedness to Professors Kazumasa Iwata, Koichi Hamada, and Junko Kato for their help and support before and during my study at Yale.

The support provided by my friends proved invaluable. Special thanks go to Naohito Abe, Elena Krasnokutskaya, Richard Lofthouse, Kay Mansfield, and Brett Marston. They have provided great encouragement and brought joy and pleasure to the otherwise austere life of this graduate student.

Financial support from a Fulbright Fellowship, a Cowles Prize from the Cowles Foundation, and an Alfred P. Sloan Doctoral Dissertation Fellowship are gratefully acknowledged.

And finally, my greatest gratitude is to my parents who never failed to send me my favourite sweets and shrimp crackers, welcomed me with open arms whenever I visited them, and to whom this dissertation is dedicated.

Chapter 1

Introduction

1 General Introduction

This dissertation is based on three papers (Shimotsu and Phillips (2000a, 2000b), Phillips and Shimotsu (2000)) that have come out of joint research with Professor Peter C. B. Phillips during 1998-2000.

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. Fractionally integrated processes accommodate a form of strong dependence in the autocorrelogram that is intermediate in intensity between the short memory displayed by weakly dependent time series (the so-called $I(0)$ processes) and the persistence of unit root time series (or $I(1)$ 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.

Theoretical research on fractional processes has centered on the case where $-\frac{1}{2} < d < \frac{1}{2}$. This is primarily because the process has a stationary and invertible representation when d is within this range. However, in many economic applications the process of interest lies close to the boundary between stationarity and nonstationartiy. Furthermore, many tests of economic hypotheses amount to tests of the stationarity of a process or tests of stationarity of the deviations from equilibrium, as witnessed by the prevalence of the unit root and cointegration approach in economics and econometrics. In order to deal with a nonstationary (i.e. $d > \frac{1}{2}$) $I(d)$ time series, the existing literature defines the process as a partial sum of $I(d-1)$ processes. While this approach is not incorrect, it has several problems, one of which is that it employs different data generating mechanisms depending on the value of d, a parameter which is not known *a priori* and which we have to estimate from the data. This dissertation is based on an alternate model of fractionally integrated processes that is valid for all values of d and aims to provide an inferential apparatus that uniformly covers values of d that are normally encountered in empirical applications.

2 Existing literature

The memory parameter d governs the strength of long range dependence of a process and is often the focus of empirical interest. More specifically, the impulse response function of an $I(d)$ process with $d > 0$ decays at the rate of t^{d-1}. This section provides a brief review of the literature on estimation of d to date. Robinson (1994b) and Baillie (1996) give more exhaustive reviews, although many new methods have been developed and advances has been made since then.

When we can parameterize the autocovariance structure of the process of interest completely, such as in a fractionally integrated autoregressive moving average (ARFIMA(p, d, q)) model, the parameters can be estimated by maximum likelihood procedures. Under Gaussianity and when $d < \frac{1}{2}$, exact maximum likelihood estimation of ARFIMA models was proposed by Sowell (1992). While the exact maximum likelihood estimator is theoretically attractive, it requires the inversion of a $T \times T$ covariance matrix, which is a nonlinear function of hypergeometric functions, at each iteration of evaluation of the likelihood, which is costly even by present computing standards.

In terms of computational burden, Whittle's (1951) approach of approximating the exact likelihood in the frequency domain is appealing. Fox and Taqqu (1986) and Dahlhaus (1989) show consistency and asymptotic normality of the approximate frequency domain maximum likelihood estimator for Gaussian processes with $0 < d < \frac{1}{2}$. Giraitis and Surgailis (1990) prove asymptotic normality of the estimator for general non-Gaussian linear processes, whilst Hosoya (1996) extended the frequency domain estimator to vector linear processes.

Although the two estimators are asymptotically equivalent, Cheung and Diebold (1994) find that their finite sample properties differ. When the mean of the process is known, the exact maximum likelihood estimator is substantially more efficient than the frequency domain estimator. On the other hand, when the mean of the process is unknown, which is typically the case, the performance of the two estimators is very similar. This is because the frequency domain estimator is invariant to the mean of the process and strong dependence makes precise estimation of the mean difficult.

Martin and Wilkins (1999) propose indirect estimation methods for estimating ARFIMA and VARFIMA models. The indirect estimator can be computationally attractive, particularly for multivariate VARFIMA models in which computation of the covariance matrix becomes extremely difficult for general model specifications. It was shown that the indirect estimator generates similar small sample properties to the exact maximum likelihood estimator. Tanaka (1999) proposes another parametric estimator based on an alternate model of $I(d)$ processes. While this model is only asymptotically stationary even for $-\frac{1}{2} < d < \frac{1}{2}$, this estimator has the desirable feature that it provides a valid estimator for all values of d.

In many cases, the persistence of a shock to the process is of central interest to empirical researchers. Then, semiparametric estimators become attractive, because in the parametric approach misspecification of the short-run dynamics, e.g. the form of the ARMA model, leads to inconsistent estimates of d. The semiparametric approach focuses on estimation of d by treating short-run dynamics of the process nonparametrically. Hence it is robust to misspecification of the short-run dynamics.

The most commonly used semiparametric estimator is log periodogram (LP) regression. LP regression was proposed by Geweke and Porter-Hudak (1983). This estimator is based on a linear regression of the ordinates of the log periodogram on the logarithm of a trigonometric function around the origin. It exploits the fact that the shape of the spectral density of an $I(d)$ process behaves like $c\lambda^{-2d}$ for $\lambda \sim 0$ and is dominated by d. While the asymptotic theory of Geweke and Porter-Hudak was incomplete, Robinson (1995a) proved consistency and asymptotic normality of a version of LP regression for $d \in \left(-\frac{1}{2}, \frac{1}{2}\right)$ and under Gaussianity. This version, originally hinted at by Künsch (1986), excludes periodogram ordinates immediately around the origin. Hurvich et al. (1998) further show that the omission of lower frequencies is unnecessary and provide an optimal formula for the choice of the number of periodogram ordinates used in regression. Velasco (1999a) extends Robinson's work to extend the range of d in which valid inference is possible. Kim and Phillips (1999) develop a theory of LP regression for the nonstationary case, i.e. when $d > \frac{1}{2}$.

Another semiparametric estimator, the local Whittle estimator, maximizes the frequency domain Gaussian likelihood function that is localized to the neighborhood of the origin. The local Whittle estimator was proposed by Künsch (1987), and Robinson (1995b) showed its consistency and asymptotic normality for $d \in \left(-\frac{1}{2}, \frac{1}{2}\right)$. While log periodogram regression enjoys its popularity because of the simplicity of its construction as a linear regression estimator, the local Whittle estimator has some strong advantages, including the fact that its validity does not require Gaussianity and that it is more efficient than LP regression. Velasco (1999b) showed the local Whittle estimator is consistent and asymptotically normal for a wider range of d. Lobato (1999) extended local Whittle estimation to the multivariate case and established asymptotic equivalence of the multivariate local Whittle estimator and a two-step estimator, which uses consistent univariate estimates of d for each series as the first step.

Robinson (1994a) proposed another semiparametric estimator based on averages of the periodogram and proved its consistency. Lobato and Robinson (1996) derived its limit distribution, but the estimator is asymptotically normal only for $-\frac{1}{2} < d < \frac{1}{4}$ and it depends on a user-chosen number q in addition to the width of the frequency band m. Recently, Moulines and Soulier (1999) propose another semiparametric estimator that approximates the spectral density of the process for $\lambda \in (0, \pi)$ by λ^{-2d} and means of truncated Fourier series. This broadband estimator is shown to outperform other semiparametric estimators asymptotically, although its advantage in finite sample is not so clear.

An enormous volume of empirical research on fractional integration has now been published. What emerges from this large body of work is that evidence supporting fractional integration has been found for many economic and financial time series, such as inflation, the forward exchange rate premium, and stock return volatility. To give an exhaustive survey is beyond the scope of this introduction. The literature to the mid 1990's is surveyed in

Baillie (1996). Instead, we name here a few contributions that solved, or partially solved, empirical "puzzles" by reconciling the implication of economic theory with the characteristics of the observed data. In consumption function studies, Diebold and Rudebusch (1991) and Haubrich (1993) show that consumption is not "too smooth" if the income process is modeled as a fractionally integrated process, providing a partial solution to the so-called "Deaton paradox." Backus and Zin (1993) show that the term structure implied by commonly used models matches the data if the short term interest rate is fractionally integrated. Maynard and Phillips (1998) demonstrate that the so-called forward premium puzzle is a result of imbalance in regression in which an $I(0)$ process is regressed on an $I(d)$ process. Michelacci and Zaffaroni (2000) show that an augmented Solow model can generate a mean reverting but nonstationary growth pattern for each country, which explains why a unit root in output is often accepted in the time-series literature but a 2% rate of convergence of output is repeatedly found in cross-country studies.

3 Overview of the Dissertation

The second and third chapters of this dissertation are concerned with the semiparametric estimation of the memory parameter in the nonstationary case. An $I(d)$ process has a stationary and invertible representation when $-\frac{1}{2} < d < \frac{1}{2}$. For this range of d, two commonly used semiparametric estimators (log periodogram regression, local Whittle estimation) 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) 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 sum of an $I(-0.3)$ process, and the motivation for this construction of economic data is unclear.

The approach taken in the second chapter is to define fractionally integrated processes as weighted sums of short-memory input variables, which are treated nonparametrically. 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 two chapters use a new representation and approximation theory for the discrete Fourier transform of a fractionally integrated time series (based on Phillips, 1999a) 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.

With this representation theory in hand, the second chapter develops 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 (1999a) without any formal development of its properties or asymptotic behavior. This chapter 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 (1999b), extending further the domain of d where valid inference is possible.

The third chapter studies the asymptotic properties of the local Whittle estimator in the nonstationary case for $d \in (\frac{1}{2}, 2)$, including the unit root case and the case where the process has a linear time trend. These cases are of high importance in empirical work especially with economic time series, which commonly exhibit nonstationary behavior and show some evidence of deterministic trends as well as long range dependence. The asymptotic properties of the local Whittle estimator in the nonstationary case over the region $d \in (\frac{1}{2}, 1)$ were explored in Velasco (1999b). Velasco also showed that, upon adequate tapering of the observations, the region of consistent estimation of d may be extended but with corresponding increases in the variance of the limit distribution. For the region $d \geq 1$, there is presently no theory for the untapered Whittle estimator and, for the region $d \in (\frac{3}{4}, 1)$, no limit distribution theory. The unit root case is of particular interest because it stands as an important special case of an $I(d)$ process with $d = 1$ and it has played a central role in the study of nonstationary economic time series. It is also now known to be the borderline that separates cases of consistent and inconsistent estimation by LP regression (Kim and Phillips, 1999) and, as we shall show here, local Whittle estimation.

This chapter demonstrates that the local Whittle estimator (i) is consistent for $d \in (\frac{1}{2}, 1]$, (ii) is asymptotically normally distributed for $d \in (\frac{1}{2}, \frac{3}{4})$, (iii) is asymptotically distributed as a square of fractional Brownian motion for $d \in (\frac{3}{4}, 1)$, (iv) has a mixed normal limit distribution for $d = 1$, (v) converges to unity in probability for $d \in (1, 2)$, and (vi) converges to unity in probability when the process has a linear time trend. This chapter, therefore, complements the earlier work of Robinson (1995b) and Velasco (1999b) and largely completes the study of the asymptotic properties of the local Whittle estimator for regions of d that are empirically relevant in most applications. This chapter also serves as a counterpart to Phillips (1999b) and Kim and Phillips (1999), which analyze the asymptotics of LP regression for $d \in (\frac{1}{2}, 2)$.

The fourth chapter proposes a new estimation method for the memory parameter of a stationary fractionally integrated process. The most common estimator of the memory parameter in the stationary case is provided by log periodogram regression. The conventional log periodogram regression estimator uses the periodogram ordinates only in a frequency band $\lambda_s = 2\pi/n, ..., 2\pi m/n$ which shrinks to the origin (i.e. $m/n \to 0$) as the sample size increases. This shrinking process allows the estimator to achieve consistency and asymptotic normality while at the same time leaving the short-memory property of the process unspecified. However, the periodogram at higher frequencies λ_s ($s = m+1, ..., [n/2]$) continues to contain some information about the memory parameter. This intuition indicates that the conventional log periodogram regression estimator may discard some information in the data and gains may be achieved by using a wider frequency band while preserving the nonparametric property of the log periodogram regression estimator.

Accordingly, this chapter proposes a procedure for estimating d that builds on this idea. The method is a pooled log periodogram regression that is taken over the wider band of frequencies $\lambda_s = \frac{2\pi s}{n}, s = 1, ..., mL$ with $L \to \infty$ and $mL/n \to 0$. This method corrects for variation in the regression intercept by taking subgroup means in the regression. The estimator of d pools the information about d obtained within each (shrinking) band over which the error spectrum is effectively constant as $n \to \infty$. We therefore call the new estimator a pooled log periodogram regression estimator. The pooled estimator is shown to be consistent and asymptotically normally distributed. The pooled estimator has a smaller asymptotic variance than the conventional log periodogram regression estimator, reflecting the greater number of periodogram ordinates used in the regression, but it also has larger asymptotic bias because of the nonconstancy of the error spectrum.

Simulations show that the pooled estimator has advantages over the conventional log periodogram regression estimator in finite samples, because the use of a wider frequency band ($m(L+1)$ rather than m) makes it less sensitive to the presence of peaks in the underlying spectral density. At the same time, it avoids the extremely large bias that is typical of the log periodogram regression estimator when a wide frequency band is employed. Therefore, it provides us with an alternate way of using a wider frequency band in log periodogram regression and a way to use more information, making the estimator more robust to various shapes in the short memory spectrum.

References

[1] Baillie, R. T. (1996). "Long Memory Processes and Fractional Integration in Econometrics." *Journal of Econometrics* **73**: 5-59.

[2] Backus, D. K. and S. E. Zin (1993). "Long-Memory Inflation Uncertainty: Evidence from the Term Structure of Interest Rates." *Journal of Money, Credit, and Banking* **25**: 681-700.

[3] Cheung, Y.-W. and F. X. Diebold (1994). "On Maximum Likelihood Estimation of the Differencing Parameter of Fractionally-Integrated Noise with Unknown Mean." *Journal of Econometrics* **62**: 301-316.

[4] Dahlhaus, R. (1989). "Efficient Parameter Estimation for Self-Similar Processes." *Annals of Statistics* **17**: 1749-1766.

[5] Diebold, F. X. and G. D. Rudebusch (1991). "Is Consumption Too Smooth? Long Memory and the Deaton Paradox." *Review of Economics and Statistics* **73**: 1-9.

[6] Fox, R. and M. S. Taqqu (1986). "Large-Sample Properties of Parameter Estimates for Strongly Dependent Stationary Gaussian Time Series." *Annals of Statistics* **14**: 517-532.

[7] Geweke, J., and S. Porter-Hudak (1983). "The Estimation and Application of Long Memory Time Series Models." *Journal of Time Series Analysis* **4**: 221-38.

[8] Giraitis, L. and D. Surgailis (1990). "A Central Limit Theorem for Quadratic Forms in Strongly Dependent Linear Variables and its Application to Asymptotical Normality of Whittle's Estimate." *Probability Theory and Related Fields* **86**: 87-104.

[9] Haubrich, J. G. (1993). "Consumption and Fractional Differencing: Old and New Anomalies." *Review of Economics and Statistics* **75**: 767-72.

[10] Hosoya, Y. (1996). "The Quasi-Likelihood Approach to Statistical Inference on Multiple Time-Series with Long-Range Dependence." *Journal of Econometrics* **73**: 217-36.

[11] Hurvich, C. M., R. Deo and J. Brodsky (1998). "The Mean Squared Error of Geweke and Porter-Hudak's Estimator of the Memory Parameter of a Long-Memory Time Series." *Journal of Time Series Analysis* **19**: 19-46.

[12] Kim, C. S. and P. C. B. Phillips (1999). "Log Periodogram Regression: The Nonstationary Case." Cowles Foundation Discussion Paper, Yale University (http://cowles.econ.yale.edu).

[13] Künsch, H. (1986). "Discrimination between Monotonic Trends and Long-Range Dependence." *Journal of Applied Probability* **23**: 1025-1030.

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

[15] Lobato, I. N. (1999). "A Semiparametric Two-Step Estimator in a Multivariate Long Memory Model." *Journal of Econometrics* **90**: 129-53.

[16] Lobato, I. and P. M. Robinson (1996). "Averaged Periodogram Estimation of Long Memory." *Journal of Econometrics* **73**: 303-324.

[17] Martin, V. L. and N. P. Wilkins (1999). "Indirect Estimation of ARFIMA and VARFIMA Models." *Journal of Econometrics* **93**: 149-75.

[18] Maynard, A. and P. C. B. Phillips (1998). "Rethinking an Old Empirical Puzzle: Econometric Evidence on the Forward Discount Anomaly." Yale University.

[19] Michelacci, C. and P. Zaffaroni (2000). "(Fractional) Beta Convergence." *Journal of Monetary Economics* **45**: 129-53.

[20] Moulines, E. and P. Soulier (1999). "Broadband Log-Periodogram Regression of Time Series with Long-Range Dependence." *Annals of Statistics* **27**: 1415-1439.

[21] Phillips, P. C. B. (1999a). "Discrete Fourier Transforms of Fractional Processes." Cowles Foundation Discussion Paper #1243, Yale University (http://cowles.econ.yale.edu).

[22] Phillips, P. C. B. (1999b). "Unit Root Log Periodogram Regression." Cowles Foundation Discussion Paper #1244, Yale University (http://cowles.econ.yale.edu).

[23] Phillips, P. C. B. and K. Shimotsu (2000). "Local Whittle Estimation in Nonstationary and Unit Root Cases." Cowles Foundation Discussion Paper #1266, Yale University (http://cowles.econ.yale.edu).

[24] Robinson, P. M. (1994a). "Semiparametric Analysis of Long-Memory Time Series." *Annals of Statistics* **22**: 515-539.

[25] Robinson, P. M. (1994b). Time Series with Strong Dependence. *Advances in Econometrics, Sixth World Congress*, Cambridge University Press.

[26] Robinson, P. M. (1995a). "Log-Periodogram Regression of Time Series with Long Range Dependence." *Annals of Statistics* **23**: 1048-1072.

[27] Robinson, P. M. (1995b). "Gaussian Semiparametric Estimation of Long Range Dependence." *Annals of Statistics* **23**: 1630-1661.

[28] Shimotsu, K. and P. C. B. Phillips (2000a). "Modified Local Whittle Estimation of the Memory Parameter in the Nonstationary Case." Cowles Foundation Discussion Paper #1265, Yale University (http://cowles.econ.yale.edu).

[29] Shimotsu, K. and P. C. B. Phillips (2000b). "Pooled Log Periodogram Regression." Cowles Foundation Discussion Paper #1267, Yale University (http://cowles.econ.yale.edu).

[30] Sowell, F. (1992). "Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models." *Journal of Econometrics* **53**: 165-188.

[31] Tanaka, K. (1999). "The Nonstationary Fractional Unit Root." *Econometric Theory* **15**: 549-582.

[32] Velasco, C. (1999a). "Non-Stationary Log-Periodogram Regression." *Journal of Econometrics* **91**: 325-371.

[33] Velasco, C. (1999b). "Gaussian Semiparametric Estimation of Non-Stationary Time Series." *Journal of Time Series Analysis* **20**: 87-127.

[34] Whittle, P. (1951). *Hypothesis Testing in Time Series Analysis.* Hafner, New York.

Chapter 2

Modified Local Whittle Estimation of the Memory Parameter in the Nonstationary Case

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 $d \lesseqgtr \frac{1}{2}$, 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 sum of an $I(-0.3)$ process, and the motivation for this construction of economic data is unclear. Appendix C in Section 9 of this chapter 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 chapter 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. This chapter 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

[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.

[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.

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. This chapter 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 of this approach as well.

The remainder of this chapter 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 chapter. 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.

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

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

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

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

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

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

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

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

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 $\widetilde{X}_{\lambda 0}(d)$ and $\widetilde{X}_{\lambda n}(d)$ 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

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

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

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

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

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

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 $\widetilde{U}_{\lambda n}(f)$ and X_n and, consequently, it affects the extent to which the (normalized) dft of the observed data, $w_x(\lambda)$, can provide an approximation of the dft of the component innovations, $w_u(\lambda)$.

2.4 The Modified Discrete Fourier Transform

The representation (8) suggests the use of the quantity

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

and $I_v(\lambda_j) = v_x(\lambda_s) v_x(\lambda_s)^*$ to approximate $v_x(\lambda_s; d)$ and $I_v(\lambda_s; d) = v_x(\lambda_s; d) v_x(\lambda_s; d)^*$. This is done by approximating $\widetilde{X}_{\lambda_s 0}(d)$ and $\widetilde{X}_{\lambda_s n}(d)$ by $D_n(e^{i\lambda_s}; d)(1 - e^{i\lambda_s})^{-1} e^{i\lambda_s} X_0$ and $D_n(e^{i\lambda_s}; d)(1 - e^{i\lambda_s})^{-1} e^{i\lambda_s} X_n$. We call $v_x(\lambda_s)$ and $I_v(\lambda_s)$ the *modified discrete Fourier transform* and *modified periodogram*, respectively. In the following, we confine our attention to the case $d \in (0, 2)$, 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*

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

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*

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

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

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

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 $E\left|R^a_{s,n}\right| = O(\lambda_s)$, $E\left|R^b_{s,n}(d)\right| = O\left(s^{d-1}n^{\frac{1}{2}-d}(\log n)^{\frac{1}{2}}\right)$, *and* $E\left|R^c_{s,n}(d)\right| = O\left(s^{2d-2}n^{1-2d}\log n\right)$ *uniformly in s.*

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

$$\lambda_s^{2d} I_v(\lambda_s) = |C(1)|^2 I_\varepsilon(\lambda_s) + R^a_{s,n} + R^b_{s,n}(d) + R^c_{s,n}(d) + R^e_{s,n}(d) + R^g_{s,n}(d),$$

where

$$E\left|R^a_{s,n}\right| = O(\lambda_s), \quad E\left|R^b_{s,n}\right| = O\left(s^{d-2}(\log n)^{\frac{1}{2}}\right), \quad E\left|R^c_{s,n}\right| = O\left(s^{d-1}n^{-\frac{1}{2}}\right),$$
$$E\left|R^e_{s,n}\right| = O\left(s^{2d-4}\log n\right), \quad E\left|R^g_{s,n}\right| = O\left(s^{2d-2}n^{-1}\right),$$

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

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

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))

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

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.

(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* $E|\varepsilon_t|^p < \infty$ *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 $d \leq 1.0$, but has a large negative bias and larger variance when $d \geq 1.2$ (see also Velasco (1999a)). For the value $0.6 \leq d \leq 1.0$, 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 $0.6 \leq d \leq 1.6$.

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

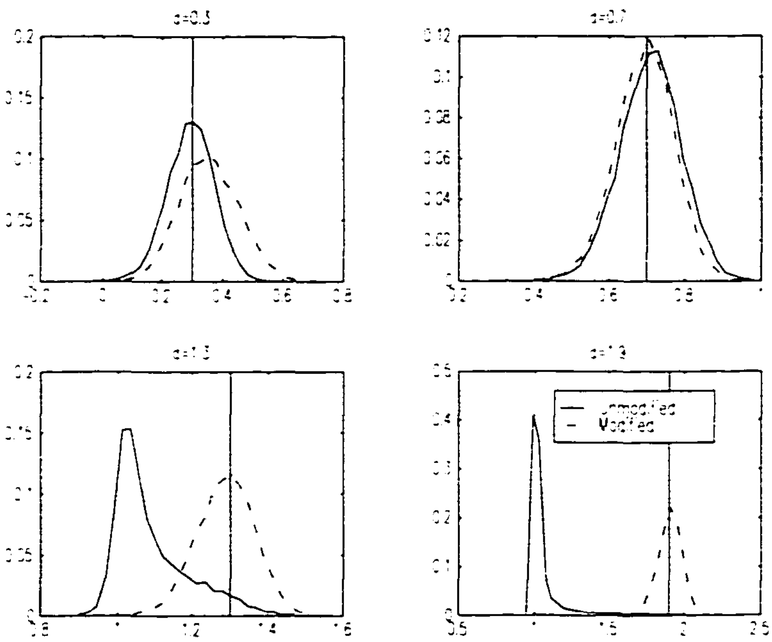

Figure 1: Modified and unmodified local Whittle estimates

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 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 chapter 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

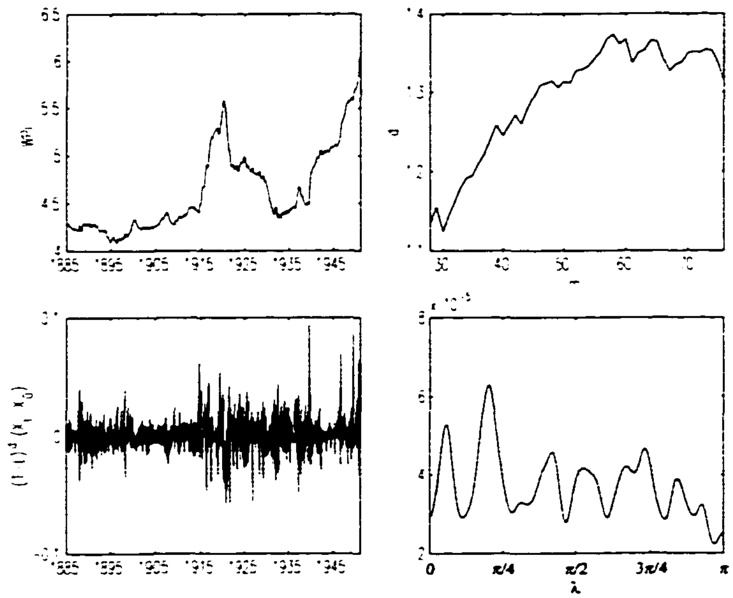

Figure 2: Wholesale price index data and estimates of d

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 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.

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 $\widetilde{f}_{\lambda p}$ in (8). The other gives asymptotic approximations to the term $\widetilde{U}_{\lambda n}(f)$ 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

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

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

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

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

7.4 Lemma

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

$$\begin{aligned} \lambda^{-f}\left(1-e^{i\lambda}\right)^f &= e^{-\frac{\pi}{2}fi} + O(\lambda), \\ \lambda^{-f}\left(1-e^{-i\lambda}\right)^f &= e^{\frac{\pi}{2}fi} + O(\lambda). \end{aligned} \tag{18}$$

7.5 Corollary

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

$$\begin{aligned} \lambda_s^{-f} D_n\left(e^{i\lambda_s}; f\right) &= \lambda_s^{-f}\left(1-e^{i\lambda_s}\right)^f + \lambda_s^{-f} O\left(n^{-f}s^{-1}\right) \\ &= e^{-\frac{\pi}{2}fi} + O(\lambda_s) + O\left(s^{-1-f}\right), \end{aligned} \tag{19}$$

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

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

Taylor expansion yields

$$\begin{aligned}
2\sin\left(\frac{\lambda}{2}\right) &= 2\cos(0)\cdot\frac{\lambda}{2}-\frac{1}{3}\cos\left(\tilde{\lambda}\right)\cdot\left(\frac{\lambda}{2}\right)^3 = \lambda + O\left(\lambda^3\right), \\
\cos\left(\frac{\lambda f}{2}-\frac{\pi f}{2}\right) &= \cos\left(-\frac{\pi f}{2}\right) - \sin\left(\tilde{\lambda}\right)\cdot\left(\frac{\lambda f}{2}\right) = \cos\left(-\frac{\pi f}{2}\right) + O(\lambda), \\
\sin\left(\frac{\lambda f}{2}-\frac{\pi f}{2}\right) &= \sin\left(-\frac{\pi f}{2}\right) + \cos\left(\tilde{\lambda}\right)\cdot\left(\frac{\lambda f}{2}\right) = \sin\left(-\frac{\pi f}{2}\right) + O(\lambda),
\end{aligned}$$

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

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

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

7.7 Lemma

Uniformly in p and s,

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

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

7.8 Proof

The approximation (17) yields

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

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

$$\begin{aligned}
\left|\sum_{k=p+1}^{n} k^{-f-1} e^{2\pi isk/n}\right| &\leq (p+1)^{-f-1} \max_N \left|\sum_{k=p+1}^{p+N} e^{2\pi isk/n}\right|, \\
\left|\sum_{k=p+1}^{p+N} e^{2\pi isk/n}\right| &= O\left(\frac{n}{s}\right), \\
\sum_{k=p+1}^{n} k^{-f-2} &= O\left(p^{-f-1}\right) = O\left(\frac{n}{p^{f+1}s}\right),
\end{aligned}$$

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 $E[\widetilde{\varepsilon}_t]^2 < \infty$. Rewrite the second term as follows:

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

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

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

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

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

From Lemma 7.7,

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

It follows that

$$\begin{aligned}&E\left|\sum_{p=1}^{n-1}\left[\widetilde{f}_{\lambda_s p} e^{-ip\lambda_s} - \widetilde{f}_{\lambda_s (p-1)} e^{-i(p-1)\lambda_s}\right]\widetilde{\varepsilon}_{n-p}\right|^2 \\ &\leq \sum_{p=1}^{n-1}\sum_{q=1}^{n-1} O\left(p^{-f-1}\right)O\left(q^{-f-1}\right)E\left|\widetilde{\varepsilon}_{n-p}\widetilde{\varepsilon}_{n-q}\right| \\ &= \begin{cases} O(1), & \text{for } f > 0, \\ O\left(n^{-2f}\right), & \text{for } f \in (-1,0), \end{cases}\end{aligned}$$

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:

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

7.13 Proof

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

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

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

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

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

$$\begin{aligned}
\sum_{p=0}^{L}\left(\tilde{f}_{\lambda_s p}\right)^2 &= O\left(\sum_{p=1}^{L} p^{-2f}\right) = O\left(L^{1-2f}\right), \\
\sum_{p=L+1}^{n-1}\left(\tilde{f}_{\lambda_s p}\right)^2 &= O\left(\sum_{p=L+1}^{n-1} \frac{1}{p^f}\frac{n}{p^{f+1}s}\right) = O\left(\frac{n}{s}\sum_{p=L+1}^{n-1}\frac{1}{p^{2f+1}}\right) = O\left(\frac{n}{s}L^{-2f}\right).
\end{aligned}$$

For part (c), Minkowski's inequality yields

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

giving the required result. ■

7.14 Lemma

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

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

(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:

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

7.15 Proof

For part (a), using Lemma 7.7 we get

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

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

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

and

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

giving the required result. ■

7.16 Lemma

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

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

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

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

7.17 Proof

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

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

and for part (b),

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

giving the required result. ■

7.18 Lemma

For $d\in\left(\frac{1}{2},2\right)$ and $1\le t\le n$, uniformly in t,

(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*

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

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

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

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

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

Rearrange the second term as follows:

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

where the fourth line follows from the fact that

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

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

$$E\left[\sum_{k=1}^{t-1}k^{d-2}\widetilde{\varepsilon}_{t-k}\right]^2 = E\left[\sum_{k=1}^{t-1}k^{d-2}\widetilde{\varepsilon}_{t-k}\right]\left[\sum_{l=1}^{t-1}l^{d-2}\widetilde{\varepsilon}_{t-l}\right]=\sum_{k=1}^{t-1}k^{d-2}\sum_{l=1}^{t-1}l^{d-2}E\left(\widetilde{\varepsilon}_{t-k}\widetilde{\varepsilon}_{t-l}\right), \tag{27}$$

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

For part (b) and (c), $E\left[X_t^\varepsilon\right]^2$ 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. ■

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). ■

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

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

Using Lemma 7.4 and Corollary 7.5, we have

$$\begin{aligned}\frac{\lambda_s^d D_n\left(e^{i\lambda_s};f\right)}{1-e^{i\lambda_s}} &= \frac{\lambda_s^{-f}D_n\left(e^{i\lambda_s};f\right)}{\lambda_s^{-1}\left(1-e^{i\lambda_s}\right)}=\frac{e^{-\frac{\pi}{2}fi}+O(\lambda_s)+O\left(s^{-1-f}\right)}{e^{-\frac{\pi}{2}i}+O(\lambda_s)}\\ &= e^{\frac{\pi}{2}di}+O(\lambda_s)+O\left(s^{-1-f}\right), \end{aligned} \tag{29}$$

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

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

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

where

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

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

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

where $E\left|r_n^1(\lambda_s)\right|^2 = O\left(\lambda_s^2\right)$. Combining (29) and (31), we obtain the approximation of the first term in (28), viz.

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

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

Now we derive the bound of the second term in (28). It follows from Lemma 7.10 and (30) that

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

where

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

uniformly in s, giving part (a).

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

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

It follows from Minkowski's inequality that

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

uniformly in s, giving the required result.

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

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

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

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

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

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

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

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

Thus

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

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:

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

For arbitrary small $\Delta > 0$, define $\Theta_1 = \left\{d : d_0 - \frac{1}{2} + \Delta < d \le \Delta_2\right\}$ and $\Theta_2 = \left\{d : \Delta_1 \le d \le d_0 - \frac{1}{2} + \Delta\right\}$. 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 $x - \log x \ge 0.5\delta^2$ if $|x| > \delta$, $\widehat{d} \to_p d_0$ if

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

and

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

as $n \to \infty$, 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

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

Note that

$$\begin{aligned}&\frac{\widehat{G}(d)-G(d)}{G(d)}\\ =\ &\frac{\frac{1}{m}\sum_{j=1}^{m}\lambda_j^{2d}I_v(\lambda_j)-G_0\frac{1}{m}\sum_{j=1}^{m}\lambda_j^{2(d-d_0)}}{G_0\frac{1}{m}\sum_{j=1}^{m}\lambda_j^{2(d-d_0)}}\\ =\ &\frac{\frac{1}{m}\sum_{j=1}^{m}\lambda_j^{2d_0}\left(\frac{j}{m}\right)^{2(d-d_0)}\cdot I_v(\lambda_j)-G_0\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2(d-d_0)}}{G_0\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2(d-d_0)}}\\ =\ &\frac{[2(d-d_0)+1]\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0}\left[\lambda_j^{2d_0}I_v(\lambda_j)-G_0\right]}{[2(d-d_0)+1]G_0\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2(d-d_0)}}\\ =\ &\frac{A(d)}{B(d)}.\end{aligned} \tag{34}$$

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)

$$EI_\varepsilon(\lambda_j) = \frac{\sigma^2}{2\pi}, \tag{35}$$

$$Var(I_\varepsilon(\lambda_j)) = O(1), \tag{36}$$

$$Cov(I_\varepsilon(\lambda_j), I_\varepsilon(\lambda_k)) = O(n^{-1}), \quad j \neq k. \tag{37}$$

From (35), (36), and (37) 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 (33)), it follows that

$$\begin{aligned} & E\left[\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0}\left[I_\varepsilon(\lambda_j) - \frac{\sigma^2}{2\pi}\right]\right]^2 \\ &= \frac{1}{m^2}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{4d-4d_0} Var\left[I_\varepsilon(\lambda_j)\right] \\ &\quad + \frac{1}{m^2}\sum_{j\neq k}^{m}\left(\frac{j}{m}\right)^{2d-2d_0}\left(\frac{k}{m}\right)^{2d-2d_0} Cov\left[I_\varepsilon(\lambda_j), I_\varepsilon(\lambda_k)\right] \\ &= O\left(\frac{1}{m^2}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{4d-4d_0}\right) + O\left(\frac{1}{m^2}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0}\sum_{k=1}^{m}\left(\frac{k}{m}\right)^{2d-2d_0} n^{-1}\right) \\ &= \left\{\begin{array}{ll} O(\log m/m) & \text{for } 4d-4d_0 \geq -1 \\ O\left(m^{-2-4d+4d_0}\right) & \text{for } -2+4\Delta < 4d-4d_0 < -1 \end{array}\right\} + O\left(n^{-1}\right) \\ &= O\left(m^{-4\Delta} + m^{-1}\log m + n^{-1}\right) = O\left(m^{-4\Delta} + n^{-1}\right). \end{aligned} \tag{38}$$

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

$$\begin{aligned}
E|A_3(d)| &= O\left(\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0}j^{d_0-1}\right)\\
&= O\left(m^{2d_0-2d-1}\sum_{j=1}^{m}j^{2d-d_0-1}\right)\\
&= \begin{cases} O\left(m^{d_0-1}\right) & \text{for } 2d-d_0>0\\ O\left(m^{2d_0-2d-1}\log m\right) & \text{for } 2d-d_0\le 0\end{cases}\\
&= O\left(m^{d_0-1}+m^{-2\Delta}\log m\right).
\end{aligned}$$

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

$$\begin{aligned}
E|A_3(d)| &= O\left(\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0}j^{d_0-\frac{3}{2}}\right)\\
&= O\left(m^{2d_0-2d-1}\sum_{j=1}^{m}j^{2d-d_0-\frac{3}{2}}\right)\\
&= \begin{cases} O\left(m^{d_0-\frac{3}{2}}\right) & \text{for } 2d-d_0-\frac{3}{2}>-1\\ O\left(m^{2d_0-2d-1}\log m\right) & \text{for } 2d-d_0-\frac{3}{2}\le -1\end{cases}\\
&= O\left(m^{d_0-\frac{3}{2}}+m^{-2\Delta}\log m\right).
\end{aligned}$$

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

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

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

From (40)-(41) 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

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

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

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

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

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

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,

$$\begin{aligned} \Pr\left(\inf_{\Theta_2} S(d) \leq 0\right) &\leq \Pr\left(\frac{1}{m}\sum_{j=1}^{m}(a_j - 1) j^{2d_0} I_v(\lambda_j) \leq 0\right) \\ &= \Pr\left(\frac{1}{m}\sum_{j=1}^{m}(a_j - 1)\frac{I_v(\lambda_j)}{\lambda_j^{-2d_0}} \leq 0\right). \end{aligned} \tag{43}$$

From Corollary 2.6 (b) and (c), (43) is equal to

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

where

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

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

$$\begin{aligned}\sum_{j=1}^{m}a_j &= \sum_{1\le j\le p}a_j+\sum_{p+1\le j\le m}a_j=O(m)+O\left(p^{2d_0+1}\int_p^m x^{-2d_0-1}dx\right)=O(m),\\ \sum_{j=1}^{m}a_j{}^2 &= p^{2-4\Delta}\sum_{j=1}^{p}j^{4\Delta-2}+p^{4d_0+2}\sum_{j=p+1}^{m}j^{-4d_0-2}=O\left(m^{2-4\Delta}+m\right),\end{aligned}$$

we have

$$\begin{aligned}&E\left[\frac{1}{m}\sum_{j=1}^{m}(a_j-1)\left[I_\varepsilon(\lambda_j)-\frac{\sigma^2}{2\pi}\right]\right]^2\\ &= O\left(\frac{1}{m^2}\sum_{j=1}^{m}(a_j-1)^2\right)+O\left(\frac{1}{m^2}\sum_{j=1}^{m}(a_j-1)\sum_{k=1}^{m}(a_k-1)\frac{1}{n}\right)\\ &= O\left(m^{-4\Delta}+m^{-1}\right)+O\left(n^{-1}\right).\end{aligned}\tag{44}$$

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

$$\begin{aligned}\frac{1}{m}\sum_{j=1}^{m}a_j j^{d_0-1} &= \frac{p^{1-2\Delta}}{m}\sum_{j=1}^{p}j^{2\Delta+d_0-2}+\frac{p^{1+2d_0}}{m}\sum_{j=p+1}^{m}j^{-2-d_0}\\ &= O\left(m^{-2\Delta}\log m+m^{d_0-1}\right),\\ \frac{1}{m}\sum_{j=1}^{m}a_j j^{d_0-\frac{3}{2}} &= \frac{p^{1-2\Delta}}{m}\sum_{j=1}^{p}j^{2\Delta+d_0-\frac{5}{2}}+\frac{p^{1+2d_0}}{m}\sum_{j=p+1}^{m}j^{-\frac{5}{2}-d_0}\\ &= O\left(m^{-2\Delta}\log m+m^{d_0-\frac{3}{2}}\right),\end{aligned}$$

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,

$$\frac{1}{m}\sum_{j=1}^{m}(a_j-1) \geq \frac{1}{m}\sum_{1\leq j\leq p} a_j - 1 \sim \frac{1}{2\Delta e} - 1 > \delta > 0.$$

Hence,

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

It follows that

$$\Pr(B_1 + B_2 + B_3 \leq 0) \to 0 \quad \text{as } m \to \infty. \tag{45}$$

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

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

$$\begin{aligned} A_1(d) &= g\frac{|C(1)|^2}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0}\left[I_\varepsilon(\lambda_j) - \frac{\sigma^2}{2\pi}\right], \\ A_2(d) &= g\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0} R_{j,n}^a, \quad A_3(d) = g\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0} R_{j,n}^b(d_0), \\ A_4(d) &= g\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0} R_{j,n}^c(d_0), \quad A_5(d) = g\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0} R_{j,n}^e(d_0), \\ A_6(d) &= g\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0} R_{j,n}^g(d_0). \end{aligned}$$

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

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

$$
\begin{aligned}
&= O\left((\log n)^{\frac{1}{2}} m^{2d_0-2d-1} \sum_{j=1}^{m} j^{2d-d_0-2}\right) \\
&= O\left((\log n)^{\frac{1}{2}} m^{d_0-2} + (\log n)^{\frac{1}{2}} m^{-2\Delta} \log m\right), \\
E\left|A_4(d)\right| &= O\left(\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0} j^{d_0-1} n^{-\frac{1}{2}}\right) = O\left(n^{-\frac{1}{2}} m^{2d_0-2d-1} \sum_{j=1}^{m} j^{2d-d_0-1}\right) \\
&= O\left(n^{-\frac{1}{2}} m^{d_0-1} + n^{-\frac{1}{2}} m^{-2\Delta} \log m\right), \\
E\left|A_5(d)\right| &= O\left(\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0} j^{2d_0-4} \log n\right) = O\left((\log n)\, m^{2d_0-2d-1} \sum_{j=1}^{m} j^{2d-4}\right) \\
&= O\left((\log n)\, m^{2d_0-4} + (\log n)\, m^{-2\Delta} \log m\right), \\
E\left|A_6(d)\right| &= O\left(\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0} j^{2d_0-2} n^{-1}\right) = O\left(n^{-1} m^{2d_0-2d-1} \sum_{j=1}^{m} j^{2d-2}\right) \\
&= O\left(n^{-1} m^{2d_0-2} + n^{-1} m^{-2\Delta} \log m\right).
\end{aligned}
$$

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

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

$B_3, \ldots, B_6$ converge to zero in probability, because

$$
\begin{aligned}
E\left|B_3\right| &= O\left(\frac{1}{m}\sum_{j=1}^{m} (a_j - 1)\, j^{d_0-2} (\log n)^{\frac{1}{2}}\right) \\
&= O\left((\log n)^{\frac{1}{2}} m^{-2\Delta} \log m + (\log n)^{\frac{1}{2}} m^{d_0-2}\right), \\
E\left|B_4\right| &= O\left(\frac{1}{m}\sum_{j=1}^{m} (a_j - 1)\, j^{d_0-1} n^{-\frac{1}{2}}\right) = O\left(n^{-\frac{1}{2}} m^{-2\Delta} \log m + n^{-\frac{1}{2}} m^{d_0-1}\right), \\
E\left|B_5\right| &= O\left(\frac{1}{m}\sum_{j=1}^{m} (a_j - 1)\, j^{2d_0-4} \log n\right) = O\left((\log n)\, m^{-2\Delta} \log m + (\log n)\, m^{2d_0-4}\right),
\end{aligned}
$$

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

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

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

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

$$\begin{aligned}
E|A_3(d)| &= O\left(\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0} j^{d_0-1}n^{\frac{1}{2}-d_0}(\log n)^{\frac{1}{2}}\right) \\
&= O\left(n^{\frac{1}{2}-d_0}(\log n)^{\frac{1}{2}}\, m^{2d_0-2d-1}\sum_{j=1}^{m} j^{2d-d_0-1}\right) \\
&= \begin{cases} O\left(n^{\frac{1}{2}-d_0}(\log n)^{\frac{1}{2}}\, m^{d_0-1}\right), & \text{for } 2d-d_0>0, \\ O\left(n^{\frac{1}{2}-d_0}(\log n)^{\frac{1}{2}}\, m^{2d_0-2d-1}\log m\right), & \text{for } 2d-d_0\le 0, \end{cases} \\
E|A_4(d)| &= O\left(\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0} j^{2d_0-2}n^{1-2d_0}\log n\right) \\
&= O\left(n^{1-2d_0}(\log n)\, m^{2d_0-2d-1}\sum_{j=1}^{m} j^{2d-2}\right) \\
&= \begin{cases} O\left(n^{1-2d_0}(\log n)\, m^{2d_0-2}\right), & \text{for } d>1/2, \\ O\left(n^{1-2d_0}(\log n)\, m^{2d_0-2d-1}\log m\right), & \text{for } d\le 1/2. \end{cases}
\end{aligned}$$

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

$$\begin{aligned}
n^{\frac{1}{2}-d_0}(\log n)^{\frac{1}{2}}\, m^{2d_0-2d-1}\log m &= n^{d_0-\frac{1}{2}}(\log n)^{-\frac{1}{2}}\left(n^{1-2d_0}(\log n)\, m^{2d_0-2d-1}\log m\right) \\
&= O\left(n^{1-2d_0}(\log n)\, m^{2d_0-2d-1}\log m\right).
\end{aligned}$$

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

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

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

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

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

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

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

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

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 $|d^* - d_0| \leq \left|\widehat{d} - d_0\right|$. 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 (40) and (41), 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

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

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

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

$$\begin{aligned} R''(d_0) &= \frac{4G_0^2\left[\frac{1}{m}\sum_{j=1}^{m}(\log j)^2-\left(\frac{1}{m}\sum_{j=1}^{m}(\log j)\right)^2\right]}{G_0^2}[1+o_p(1)] \\ &= \left\{4\left[\frac{1}{m}\left\{\left(m+\frac{1}{2}\right)(\log m)^2-2m\log m+2m+O(1)\right\}\right]\right. \\ &\quad \left.-4\left[\frac{1}{m^2}\left\{\left(m+\frac{1}{2}\right)(\log m)-m+O(1)\right\}^2\right]\right\}[1+o_p(1)] \\ &= 4+o_p(1). \end{aligned} \tag{47}$$

Next we consider the first term on the right side of (46). 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

$$\begin{aligned}
D_1 &= \frac{2|C(1)|^2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\left[I_\varepsilon(\lambda_j)-\frac{\sigma^2}{2\pi}\right], \\
D_2 &= \frac{2|C(1)|^2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\frac{\lambda_j^{2d_0}}{|1-e^{i\lambda_j}|^2}\frac{|\widetilde{\varepsilon}_{\lambda_j n}(f_0)|^2}{2\pi n}, \\
D_3 &= -\frac{2|C(1)|^2}{\sqrt{m}} \\
&\quad\times\sum_{j=1}^{m}\nu_j\left[e^{\frac{\pi}{2}d_0 i}w_\varepsilon(\lambda_j)\frac{\lambda_j^{d_0}}{1-e^{-i\lambda_j}}\frac{\widetilde{\varepsilon}_{\lambda_j n}(f_0)^*}{\sqrt{2\pi n}}+\frac{\lambda_j^{d_0}}{1-e^{i\lambda_j}}\frac{\widetilde{\varepsilon}_{\lambda_j n}(f_0)}{\sqrt{2\pi n}}e^{-\frac{\pi}{2}d_0 i}w_\varepsilon(\lambda_j)^*\right], \\
D_4 &= \frac{2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j R_{j,n}^a,\quad D_5=\frac{2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j R_{j,n}^b(d_0),\quad D_6=\frac{2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j R_{j,n}^c(d_0).
\end{aligned}$$

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

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

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

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

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

$$\begin{aligned}
&\frac{1}{\sqrt{m}} \sum_{j=1}^{m} |\nu_j| \lambda_j^{2d_0-2} \frac{E\left|\widetilde{\varepsilon}_{\lambda_j n}(f_0)\right|^2}{2\pi n} \\
&= O\left(\frac{\log m}{\sqrt{m}} \sum_{j=1}^{m} \left(\frac{j}{n}\right)^{2d_0-2} \frac{1}{n} \left(L^{2d_0-1} + \frac{n}{j} L^{2d_0-2} \right) \right) \\
&= O\left(\frac{\log m}{\sqrt{m}} \left(m^{2d_0-1} n^{1-2d_0} L^{2d_0-1} + n^{2-2d_0} L^{2d_0-2} \right) \right) \\
&= O\left(\log m \left(m^{-\frac{1}{2}} \left(\frac{mL}{n}\right)^{2d_0-1} + m^{\frac{3}{2}-2d_0} \left(\frac{mL}{n}\right)^{2d_0-2} \right) \right).
\end{aligned}$$

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

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

Therefore, $D_2 = o_p(1)$ 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), $E|D_2|$ is bounded by

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

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

$$\begin{aligned}
&\frac{1}{\sqrt{m}} \sum_{j=1}^{m} |\nu_j| \frac{\lambda_j^{2d_0}}{\left|1 - e^{i\lambda_j}\right|^2} \frac{E\left|\widetilde{\varepsilon}_{\lambda_j n}(f_0)\right|^2}{2\pi n} \\
&= O\left(\frac{\log m}{\sqrt{m}} \sum_{j=1}^{m} \left(\left(\frac{j}{n}\right)^{2d_0-2} \frac{L^{d_0-1} n^{d_0-1}}{j} + \left(\frac{j}{n}\right)^{2d_0-2} \frac{L^{2d_0-3} n}{j^2} \right) \right) \\
&= O\left(\frac{\log m}{\sqrt{m}} \sum_{j=1}^{m} \left(j^{2d_0-3} L^{d_0-1} n^{1-d_0} + j^{2d_0-4} L^{2d_0-3} n^{3-2d_0} \right) \right) \\
&= O\left(\log m \left(\left(\frac{L}{n}\right)^{d_0-1} m^{2d_0-\frac{5}{2}} + \left(\frac{L}{n}\right)^{2d_0-3} m^{-\frac{1}{2}} \right) \right) \\
&= O\left(\log m \left(\left(\frac{Lm}{n}\right)^{d_0-1} m^{d_0-\frac{3}{2}} + \left(\frac{Lm}{n}\right)^{2d_0-3} m^{\frac{5}{2}-2d_0} \right) \right) = o(1),
\end{aligned}$$

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

For D_3, in view of the fact that

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

we obtain the decomposition

$$\begin{aligned}
&\frac{2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\frac{\lambda_j^{d_0}}{1-e^{i\lambda_j}}\frac{\widetilde{\varepsilon}_{\lambda_j n}(f_0)}{\sqrt{2\pi n}}w_\varepsilon(\lambda_j)^* \\
=\ &\frac{2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\frac{\lambda_j^{d_0}}{1-e^{i\lambda_j}}\frac{1}{2\pi n}\left(\sum_{p=0}^{n-1}\widetilde{f}_{\lambda_j p}e^{-ip\lambda_j}\varepsilon_{n-p}\right)\left(\sum_{q=0}^{n-1}e^{iq\lambda_j}\varepsilon_{n-q}\right).
\end{aligned}$$

It follows that

$$\begin{aligned}
&E\left|\frac{2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\frac{\lambda_j^{d_0}}{1-e^{i\lambda_j}}\frac{\widetilde{\varepsilon}_{\lambda_j n}(f_0)}{\sqrt{2\pi n}}w_\varepsilon(\lambda_j)^*\right|^2 \\
=\ &\frac{1}{\pi^2 mn^2}E\left[\sum_{j=1}^{m}\nu_j\frac{\lambda_j^{d_0}}{1-e^{i\lambda_j}}\left(\sum_{p=0}^{n-1}\widetilde{f}_{\lambda_j p}e^{-ip\lambda_j}\varepsilon_{n-p}\right)\left(\sum_{q=0}^{n-1}e^{iq\lambda_j}\varepsilon_{n-q}\right)\right] \\
&\times\left[\sum_{k=1}^{m}\nu_k\frac{\lambda_k^{d_0}}{1-e^{-i\lambda_k}}\left(\sum_{r=0}^{n-1}\widetilde{f}_{-\lambda_k r}e^{ir\lambda_k}\varepsilon_{n-r}\right)\left(\sum_{s=0}^{n-1}e^{-is\lambda_k}\varepsilon_{n-s}\right)\right]. \qquad (50)
\end{aligned}$$

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

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

Therefore, (50) is bounded by

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

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

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

(51) 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 $d_0 \in \left(\frac{1}{2}, 1\right)$, (52) is bounded by

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

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

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

$$\begin{aligned}
&\frac{\log m}{\sqrt{m}}\sum_{j=1}^{m}\left(\frac{j}{n}\right)^{d_0-1}\frac{1}{n}\left(\sum_{p=0}^{n-1}\left|\tilde{f}_{\lambda_j p}\right|\right)\\
&= O\left(\frac{\log m}{\sqrt{m}}\sum_{j=1}^{m}\left(j^{d_0-1}n^{-d_0}L^{d_0}+j^{d_0-2}n^{1-d_0}L^{d_0-1}\right)\right)\\
&= O\left(\frac{\log m}{\sqrt{m}}m^{d_0}\left(\frac{L}{n}\right)^{d_0}+\frac{\log m}{\sqrt{m}}\left(\frac{n}{L}\right)^{1-d_0}\right).
\end{aligned}$$

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

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

thus (52) is $o(1)$. For $d_0\in\left(1,\frac{3}{2}\right)$, (52) is bounded by

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

Using Lemma 7.7, we have

$$\begin{aligned}
&\frac{\log m}{\sqrt{m}}\sum_{j=1}^{m}\left(\frac{j}{n}\right)^{d_0-1}\frac{1}{n}\left(\sum_{p=0}^{n-1}\left|\tilde{f}_{\lambda_j p}\right|\right)\\
&= O\left(\frac{\log m}{\sqrt{m}}\sum_{j=1}^{m}\left(\frac{j}{n}\right)^{d_0-1}\frac{1}{n}\sum_{p=1}^{n-1}\frac{n}{p^{f_0+1}j}\right)\\
&= O\left(\frac{\log m}{\sqrt{m}}\sum_{j=1}^{m}\left(\frac{j}{n}\right)^{d_0-1}\frac{n^{-f_0}}{j}\right)\\
&= O\left(\frac{\log m}{\sqrt{m}}\sum_{j=1}^{m}j^{d_0-2}\right)=O\left(m^{d_0-\frac{3}{2}}\log m\right),
\end{aligned}$$

and, hence, (52) is $o(1)$.

In view of the fact that $\sum_{q=0}^{n-1}e^{iq(\lambda_j-\lambda_k)}=n\mathbf{1}(j=k)$, (53) is bounded by, for $d_0\in\left(\frac{1}{2},1\right)$,

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

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

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

$$
\begin{aligned}
\frac{(\log m)^2}{mn}\sum_{j=1}^{m}\lambda_j^{2d_0-2}\sum_{p=0}^{n-1}\left|\widetilde{f}_{\lambda_j p}\right|^2 &= O\left(\frac{(\log m)^2}{mn}\sum_{j=1}^{m}\left(\frac{j}{n}\right)^{2d_0-2}\left(\sum_{p=1}^{n-1} n^{-f_0}\frac{n}{p^{f_0+1}j}\right)\right) \\
&= O\left(\frac{(\log m)^2}{mn}\sum_{j=1}^{m}\left(\frac{j}{n}\right)^{2d_0-2} n^{2d_0-1}j^{-1}\right) \\
&= O\left(\frac{(\log m)^2}{m}\sum_{j=1}^{m} j^{2d_0-3}\right) \\
&= O\left((\log m)^2 m^{2d_0-3}\right).
\end{aligned}
$$

Therefore, (50) converges to zero, and thus $D_3 = o_p(1)$.

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

$$
\begin{aligned}
E\left|D_4\right| &= O\left(\frac{1}{\sqrt{m}}\sum_{j=1}^{m}\left|\nu_j\right|\lambda_j\right) = O\left(\frac{\log m}{\sqrt{m}}\sum_{j=1}^{m}\frac{j}{n}\right) = O\left(\frac{m^{\frac{3}{2}}\log m}{n}\right), \\
E\left|D_5\right| &= O\left(\frac{1}{\sqrt{m}}\sum_{j=1}^{m}\left|\nu_j\right| j^{d_0-2}\right) = \begin{cases} O\left(\frac{\log m}{\sqrt{m}}\right), & \text{for } d_0 \in \left(\frac{1}{2}, 1\right), \\ O\left(m^{d_0-\frac{3}{2}}\log m\right), & \text{for } d_0 \in \left(1, \frac{3}{2}\right), \end{cases} \\
E\left|D_6\right| &= \begin{cases} O\left(\frac{2}{\sqrt{m}}\sum_{j=1}^{m}\left|\nu_j\right| j^{d_0-1}n^{\frac{1}{2}-d_0}\right) = O\left(n^{\frac{1}{2}-d_0}m^{d_0-\frac{1}{2}}\log m\right), & \text{for } d_0 \in \left(\frac{1}{2}, 1\right), \\ O\left(\frac{2}{\sqrt{m}}\sum_{j=1}^{m}\left|\nu_j\right| j^{d_0-1}n^{-\frac{1}{2}}\right) = O\left(n^{-\frac{1}{2}}m^{d_0-\frac{1}{2}}\log m\right), & \text{for } d_0 \in \left(1, \frac{3}{2}\right). \end{cases}
\end{aligned}
$$

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

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

It follows from (46), (47) and (54) that

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

giving the required result. ■

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

$$\begin{aligned}
E|C_3| &= O\left(\frac{1}{m}\sum_{j=1}^{m}(\log j)^k j^{d_0-2}(\log n)^{\frac{1}{2}}\right) = O\left((\log m)^k m^{d_0-2}(\log n)^{\frac{1}{2}}\right),\\
E|C_4| &= O\left(\frac{1}{m}\sum_{j=1}^{m}(\log j)^k j^{d_0-1}n^{-\frac{1}{2}}\right) = O\left((\log m)^k m^{d_0-1}n^{-\frac{1}{2}}\right),\\
E|C_5| &= O\left(\frac{1}{m}\sum_{j=1}^{m}(\log j)^k j^{2d_0-4}\log n\right) = O\left((\log m)^k m^{2d_0-4}\log n\right),\\
E|C_6| &= O\left(\frac{1}{m}\sum_{j=1}^{m}(\log j)^k j^{2d_0-2}n^{-1}\right) = O\left((\log m)^k m^{2d_0-2}n^{-1}\right),
\end{aligned}$$

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

Before we evaluate $\frac{2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\left[\lambda_j^{2d_0}I_v(\lambda_j) - G_0\right]$, we derive the approximation of $\lambda_j^{2d_0}I_v(\lambda_j)$ 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 $\bar{d}_0 = d_0 - 1 \in \left(\frac{1}{2}, 1\right)$ and $\bar{f}_0 = 1 - \bar{d}_0$. From Lemma 2.5 (a) we have

$$\begin{aligned} & \lambda_s^{\bar{d}_0} w_{\Delta x}(\lambda_s) \\ = \; & e^{\frac{\pi}{2}\bar{d}_0 i} C(1) w_\varepsilon(\lambda_s) - \frac{\lambda_s^{\bar{d}_0} C(1)}{1-e^{i\lambda_s}} \frac{\tilde{\varepsilon}_{\lambda_s n}(\bar{f}_0)}{\sqrt{2\pi n}} - \frac{\lambda_s^{\bar{d}_0} e^{i\lambda_s}}{1-e^{i\lambda_s}} \frac{\Delta X_n}{\sqrt{2\pi n}} + r_{s,n}^a + r_{s,n}^b(\bar{d}_0) + r_{s,n}^c(\bar{d}_0), \end{aligned}$$

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

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

hence we obtain

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

where $E\left|R_{s,n}^a\right| = O(\lambda_s)$, $E\left|R_{s,n}^b(\bar{d}_0)\right| = O\left(s^{\bar{d}_0-2}\right)$, and $E\left|R_{s,n}^c(\bar{d}_0)\right| = O\left(s^{\bar{d}_0-1} n^{\frac{1}{2}-\bar{d}_0}\right)$ 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 $E\left|\lambda_s^{2\bar{d}_0} I_{\Delta x}(\lambda_s)\right| = O(1)$, it follows that

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

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

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

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

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

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

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

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

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

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

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

For D_5, rewrite

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

as

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

In view of the fact that

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

we have

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

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

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

as

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

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

$$\begin{aligned}
&\frac{n^{\bar{d}_0-\frac{3}{2}}\log m}{\sqrt{m}}\sum_{j=1}^{m}\left(\frac{j}{n}\right)^{2\bar{d}_0-2}L^{\bar{d}_0-\frac{1}{2}} + \frac{n^{\bar{d}_0-\frac{3}{2}}\log m}{\sqrt{m}}\sum_{j=1}^{m}\left(\frac{j}{n}\right)^{2\bar{d}_0-2}\left(\frac{n}{j}\right)^{1/2}L^{\bar{d}_0-1} \\
= \; & O\left(n^{\frac{1}{2}-\bar{d}_0}(\log m)\, m^{2\bar{d}_0-\frac{3}{2}}L^{\bar{d}_0-\frac{1}{2}} + n^{1-\bar{d}_0}(\log m)^2 m^{-\frac{1}{2}}m^{\max\left\{2\bar{d}_0-\frac{3}{2},0\right\}}L^{\bar{d}_0-1}\right) \\
= \; & O\left(\log m\left(\frac{mL}{n}\right)^{\bar{d}_0-\frac{1}{2}}m^{\bar{d}_0-1} + (\log m)^2\left(\frac{mL}{n}\right)^{\bar{d}_0-1}m^{\max\left\{\bar{d}_0-1,\frac{1}{2}-\bar{d}_0\right\}}\right) = o(1),
\end{aligned}$$

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. ■

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

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

where

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

$$= \frac{2(2\pi)^{2\bar{d}_0-2}}{\sqrt{m}} \frac{\Delta X_n^2}{2\pi n^{2\bar{d}_0-1}} \sum_{j=1}^{m} \nu_j j^{2\bar{d}_0-2} + \frac{\Delta X_n^2}{n^{2\bar{d}_0-1}} O\left(\frac{\log m}{n\sqrt{m}} \sum_{j=1}^{m} j^{2\bar{d}_0-1}\right).$$

The second term is

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

Now evaluate the summation:

$$\begin{aligned}
\sum_{j=1}^{m} \nu_j j^{2\bar{d}_0-2} &= \sum_{j=1}^{m} \left(\log j - \frac{1}{m}\sum_{j=1}^{m} \log j\right) j^{2\bar{d}_0-2} \\
&= \sum_{j=1}^{m} j^{2\bar{d}_0-2} \log j - \sum_{j=1}^{m} j^{2\bar{d}_0-2} \left(\frac{1}{m}\sum_{j=1}^{m} \log j\right).
\end{aligned}$$

For the first term, we have

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

and, for the second term,

$$\begin{aligned}
\sum_{j=1}^{m} j^{2\bar{d}_0-2}\left(\frac{1}{m}\sum_{j=1}^{m} \log j\right) &= \left(\int_1^m x^{2\bar{d}_0-2} dx + \frac{1}{2}\left(m^{2\bar{d}_0-2}+1\right) + O\left(\int_1^m x^{2\bar{d}_0-3} dx\right)\right) \\
&\quad \times \left(\log m - 1 + O\left(\frac{\log m}{m}\right)\right) \\
&= \left(\frac{m^{2\bar{d}_0-1}}{2\bar{d}_0-1} + O(1)\right)\left(\log m - 1 + O\left(\frac{\log m}{m}\right)\right) \\
&= \frac{m^{2\bar{d}_0-1} \log m}{2\bar{d}_0-1} - \frac{m^{2\bar{d}_0-1}}{2\bar{d}_0-1} + O(\log m).
\end{aligned}$$

Therefore,

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

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

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

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

$$\begin{aligned} m^{\frac{7}{2}-2d_0} D_3 &= 2\left(2\pi\right)^{2\bar{d}_0-2} \frac{\Delta X_n^2}{2\pi n^{2\bar{d}_0-1}} m^{1-2\bar{d}_0} \sum_{j=1}^{m} \nu_j j^{2\bar{d}_0-2} \\ &= 2\left(2\pi\right)^{2\bar{d}_0-2} \left[\frac{C\left(1\right)^2 \left(\Delta X_n^{\varepsilon}\right)^2}{2\pi n^{2\bar{d}_0-1}} + o_p\left(1\right)\right] m^{1-2\bar{d}_0} \sum_{j=1}^{m} \nu_j j^{2\bar{d}_0-2} \\ &\to {}_d 2\left(2\pi\right)^{2\bar{d}_0-2} \frac{C\left(1\right)^2 \sigma^2}{2\pi} B_{\bar{d}_0-1}\left(1\right)^2 \frac{2\bar{d}_0 - 2}{\left(2\bar{d}_0 - 1\right)^2}. \end{aligned}$$

For $d_0 = \frac{7}{4}$, (55) converges to $N\left(0, \frac{1}{4}\right)$ and

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

For $d_0 \in \left(\frac{3}{4}, 1\right)$, (55) is $o_p\left(1\right)$ and

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

giving the required result. ■

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

Two main approaches to defining a nonstationary $I\left(d\right)$ 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 $I\left(d-1\right)$ process and satisfies

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

where ε_t is a short-memory stationary process. Combining (57) and (58), we obtain

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

leading to a definition of the operator equation

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

in terms of (57) and (58). 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

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

and again X_t is integrated of order d. The two definitions (59) and (61) are different, however, because the stationary input formulation (58) 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

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

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 (62) 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 (60). 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 (62). 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 (60) 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). This chapter does not deal with this additional level of difficulty, but works from the definition (60) 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." Cowles Foundation Discussion Paper #1243, Yale University (http://cowles.econ.yale.edu).

[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.

Chapter 3

Local Whittle Estimation in Nonstationary and Unit Root Cases

1 Introduction

Semiparametric estimation of the memory parameter (d) in fractionally integrated ($I(d)$) time series has attracted much recent study and is attractive in empirical applications because of its general treatment of the short memory component. Two commonly used semiparametric estimators are log periodogram (LP) regression and local Whittle estimation. Log periodogram regression is popular mainly because of the simplicity of its construction as a linear regression estimator. Local Whittle estimation involves numerical methods but is more efficient than LP regression. The local Whittle estimator was proposed by Künsch (1987), and Robinson (1995) showed its consistency and asymptotic normality for $d \in (-\frac{1}{2}, \frac{1}{2})$. Velasco (1999) extended Robinson's results to show that the estimator is consistent for $d \in (-\frac{1}{2}, 1)$ and asymptotically normally distributed for $d \in (-\frac{1}{2}, \frac{3}{4})$.

This chapter studies the asymptotic properties of the local Whittle estimator in the nonstationary case for $d \in (\frac{1}{2}, 2)$, including the unit root case and the case where the process has a linear time trend. These cases are of high importance in empirical work especially with economic time series, which commonly exhibit nonstationary behavior and show some evidence of deterministic trends as well as long range dependence. The asymptotic properties of the local Whittle estimator in the nonstationary case over the region $d \in (\frac{1}{2}, 1)$ were explored in Velasco (1999). Velasco also showed that, upon adequate tapering of the observations, the region of consistent estimation of d may be extended but with corresponding increases in the variance of the limit distribution. For the region $d \geq 1$, there is presently no theory for the untapered Whittle estimator and, for the region $d \in (\frac{3}{4}, 1)$, no limit distribution theory. The unit root case is of particular interest because it stands as an important special case of an $I(d)$ process with $d = 1$ and it has played a central role in the study of nonstationary economic time series. It is also now known to be the borderline that separates cases of consistent and inconsistent estimation by LP regression (Kim and Phillips, 1999) and, as we shall show here, local Whittle estimation.

This chapter demonstrates that the local Whittle estimator (i) is consistent for $d \in (\frac{1}{2}, 1]$, (ii) is asymptotically normally distributed for $d \in (\frac{1}{2}, \frac{3}{4})$, (iii) is asymptotically distributed as a square of fractional Brownian motion for $d \in (\frac{3}{4}, 1)$, (iv) has a mixed normal limit distribution for $d = 1$, (v) converges to unity in probability for $d \in (1, 2)$, and (vi) converges to unity in probability when the process has a linear time trend. The present chapter, therefore, complements the earlier work of Robinson (1995) and Velasco (1999) and largely completes the study of the asymptotic properties of the local Whittle estimator for regions of d that are empirically relevant in most applications. This chapter also serves as a counterpart to Phillips (1999b) and Kim and Phillips (1999), which analyze the asymptotics of LP regression for $d \in (\frac{1}{2}, 2)$.

The approach in the present chapter draws on an exact representation and approximation theory for the discrete Fourier transform (dft) of fractionally integrated processes. The theory was developed by Phillips (1999a) and chapter 2 of this dissertation and provides an apparatus for analyzing the asymptotic behavior of the dft's of fractionally integrated processes. The study of the limit distribution in the unit root case is based on an embedding and conditioning argument that uses an asymptotic representation of the dft of a short memory time series in terms of Brownian motion. The technique was developed in Phillips (1999b) to derive the asymptotic distribution of LP regression in the unit root case.

The remainder of this chapter is organized as follows. Section 2 briefly reviews the representation and approximation theory and tailors it for our analysis in subsequent sections. Consistency of the local Whittle estimator for $d \in (\frac{1}{2}, 1]$ and its inconsistency for $1 < d < 2$ are demonstrated in Section 3. Section 4 derives the limit distributions. Results for fractionally integrated processes with a linear time trend are given in Section 5. Section 6 reports some simulation results and gives an empirical application using economic data. Some technical results are collected in Appendix A in Section 7. Proofs are given in Appendix B in Section 8.

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 in the case where X_t is nonstationary and $\frac{1}{2} < d < 2$, 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 a valid representation for all values of d

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

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

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

for all t and with $\varepsilon_t = iid(0, \sigma^2)$ and $E\varepsilon_t^4 = \mu_4 < \infty$. Under (4), the spectrum of u_t is $f_u(\lambda) = \frac{\sigma^2}{2\pi} \left|\sum_{j=0}^{\infty} c_j e^{ij\lambda}\right|^2$ and this specificity certainly involves a loss of generality in comparison with local assumptions on the short memory spectrum that are used in other work (Robinson, 1995, and Velasco, 1999). However, (4) is satisfied by a wide class of parametric and nonparametric models for u_t and it enables the use of the techniques in Phillips and Solo (1992) and embedding arguments that are used later in this chapter.

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

$$
\begin{aligned}
w_a(\lambda_s) &= \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^{n} a_t e^{it\lambda_s}, \quad \lambda_s = \frac{2\pi s}{n}, s = 1, ..., n, \\
I_a(\lambda_s) &= w_a(\lambda_s) w_a(\lambda_s)^* .
\end{aligned} \tag{5}
$$

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 (1999a) leads to an exact expression that we can use for the model in frequency domain form.

2.2 Lemma

(a) If X_t follows (1), then

$$
w_x(\lambda)\left(1 - e^{i\lambda}\right) = D_n\left(e^{i\lambda}; f\right) w_u(\lambda) - \frac{e^{in\lambda}}{\sqrt{2\pi n}} \widetilde{U}_{\lambda n}(f) - \frac{e^{i\lambda}}{\sqrt{2\pi n}} \left(e^{in\lambda} X_n - X_0\right), \tag{6}
$$

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

$$
\widetilde{U}_{\lambda n}(f) = \widetilde{D}_{n\lambda}\left(e^{-i\lambda}L; f\right) u_n = \sum_{p=0}^{n-1} \widetilde{f}_{\lambda p} e^{-ip\lambda} u_{n-p}, \quad \widetilde{f}_{\lambda p} = \sum_{k=p+1}^{n} \frac{(-f)_k}{k!} e^{ik\lambda}. \tag{7}
$$

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

$$
w_x(\lambda)\left(1 - e^{i\lambda}\right) = w_u(\lambda) - \frac{e^{i\lambda}}{\sqrt{2\pi n}} \left(e^{in\lambda} X_n - X_0\right). \tag{8}
$$

Without loss of generality and to simplify formulae, we shall hereafter assume $X_0 = 0$.

2.3 Approximation of $w_x(\lambda_s)$ and $I_x(\lambda_s)$

Dividing both sides of (6) by $\left(1 - e^{i\lambda_s}\right)$, we obtain the following expression for $w_x(\lambda_s)$:

$$
w_x(\lambda_s) = \frac{D_n\left(e^{i\lambda_s}; f\right)}{1 - e^{i\lambda_s}} w_u(\lambda_s) - \frac{e^{i\lambda_s}}{1 - e^{i\lambda_s}} \frac{X_n}{\sqrt{2\pi n}} - \frac{1}{1 - e^{i\lambda_s}} \frac{\widetilde{U}_{\lambda_s n}(f)}{\sqrt{2\pi n}}. \tag{9}
$$

Neglecting the third term of (9) as a remainder, $w_x(\lambda_s)$ is seen to comprise two terms – a function of the dft of u_t and a function of X_n. These two components have very different characteristics. The first component is asymptotically uncorrelated for different frequencies λ_s, while the second component is perfectly correlated across all λ_s. As the value of d changes, the stochastic magnitude of the two components changes, and this influences the asymptotic behavior of $w_x(\lambda_s)$. As shown below, when $d < 1$ the first term dominates the second term and $w_x(\lambda_s)$ and $w_x(\lambda_r)$ are asymptotically uncorrelated for $s \neq r$. When $d > 1$, the second term becomes dominant and the $w_x(\lambda_s)$ are asymptotically perfectly correlated across all λ_s. This switching behavior of $w_x(\lambda_s)$ at $d = 1$ is a key determinant of the asymptotic properties of the local Whittle estimator. When $d = 1$, the two terms have the same stochastic order and this leads to a form of asymptotic behavior that is particular to this case. The next lemmas establish this relationship and are used as the basis of the analysis in the following sections.

2.4 Lemma

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

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

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

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* $E\left|r_{s,n}^c(d)\right|^2 = O\left(s^{2d-2} n^{1-2d}\right)$ *uniformly in* s.

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

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

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

(c) For $d = 1$,

$$\lambda_s w_x(\lambda_s) = iC(1) w_\varepsilon(\lambda_s) - i\frac{X_n}{\sqrt{2\pi n}} + r_{s,n}^a,$$

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

(d) For $d \in (1, 2)$,

$$s^{1-d} \lambda_s^d w_x(\lambda_s) = -\frac{s^{1-d}\lambda_s^d}{1-e^{i\lambda_s}} \frac{e^{i\lambda_s} X_n}{\sqrt{2\pi n}} + r_{s,n}^a(d) + r_{s,n}^b,$$

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

2.5 Corollary

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

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

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* $E\left|R_{s,n}^c(d)\right| = O\left(s^{d-1} n^{\frac{1}{2}-d}\right)$ *uniformly in s.*

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

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

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

(c) For $d = 1$,

$$\lambda_s^2 I_x(\lambda_s) = \left| C(1) w_\varepsilon(\lambda_s) - \frac{X_n}{\sqrt{2\pi n}} \right|^2 + R_{s,n}^a,$$

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

(d) For $d \in (1, 2)$,

$$s^{2-2d} \lambda_s^{2d} I_x(\lambda_s) = \frac{s^{2-2d} \lambda_s^{2d}}{\left|1 - e^{i\lambda_s}\right|^2} \frac{X_n^2}{2\pi n} + R_{s,n}^a(d) + R_{s,n}^b,$$

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

3 Local Gaussian Estimation: Consistency for $d \leq 1$ and Inconsistency for $d > 1$

We set up the local Whittle likelihood as in Künsch (1987) and Robinson (1995). Specifically, we start with the following Gaussian objective function, defined in terms of the parameter d and G

$$Q_m(G, d) = \frac{1}{m} \sum_{j=1}^{m} \left[\log\left(G \lambda_j^{-2d}\right) + \frac{\lambda_j^{2d}}{G} I_x(\lambda_j) \right], \tag{10}$$

where m is some integer less than n. The local Whittle procedure estimates G and d by minimising $Q_m(G, d)$, so that

$$\left(\widehat{G}, \widehat{d}\right) = \arg \min_{0 < G < \infty,\ -\frac{1}{2} < d < M < \infty} Q_m(G, d),$$

which involves numerical optimization. Hence, it will be convenient in what follows to distinguish the true values of the parameters by the notation $G_0 = f_{uu}(0)$ and d_0.

Concentrating (10) with respect to G, we find that $\widehat{d}$ satisfies

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

where

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

Velasco (1999) shows that $\widehat{d}$ is consistent for $d_0 \in \left(\frac{1}{2}, 1\right)$. Theorem 3.1 below establishes that $\widehat{d}$ is consistent for $d_0 \in \left(\frac{1}{2}, 1\right]$ and hence consistency carries over to the unit root case. However, Theorem 3.1 shows that, while $\widehat{G}$ is consistent for $d_0 \in \left(\frac{1}{2}, 1\right)$, it is inconsistent and tends to a random quantity when $d_0 = 1$.

3.1 Theorem

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

3.2 Theorem

If $\frac{1}{m} + \frac{m}{n} \to 0$ *as* $n \to \infty$, *then,*

$$\widehat{G}\left(\widehat{d}\right) \to_d \begin{cases} G_0, & \text{for } d_0 \in \left(\frac{1}{2}, 1\right), \\ G_0 + \omega^2 B(1)^2/2\pi, & \text{for } d_0 = 1. \end{cases}$$

When $d_0 > 1$, $\widehat{d}$ manifests very different behavior. It converges to unity in probability, and the local Whittle estimator becomes inconsistent. Therefore, the local Whittle estimator is biased downward whenever the true value of d is greater than unity. Kim and Phillips (1999) show that the log periodogram regression estimator converges to unity when $d > 1$.

3.3 Theorem

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

4 Local Gaussian Estimation: Asymptotic Distribution

The following theorems establish the asymptotic distribution of the local Whittle estimator for $d_0 \in \left(\frac{1}{2}, 1\right]$. When $d_0 \in \left(\frac{1}{2}, \frac{3}{4}\right)$, $\widehat{d}$ is asymptotically normally distributed (c.f. Velasco, 1999).

4.1 Theorem

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

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

For $d_0 \geq \frac{3}{4}$, $\widehat{d}$ has a non-normal distribution. This phenomenon occurs because, when d_0 is large, the stochastic magnitude of X_n in the representation (6) becomes so large that it dominates the behavior of $\widehat{d}$.

4.2 Theorem

If $\frac{1}{m} + \frac{m^{\frac{3}{2}} \log m}{n} \to 0$ *as* $n \to \infty$, *then,*

(a) For $d_0 = \frac{3}{4}$, *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 $d_0 \in \left(\frac{3}{4}, 1\right)$,

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

When $d_0 = 1$, the two main components of $w_x(\lambda_s)$, i.e. $w_u(\lambda_s)$ and $X_n/\sqrt{2\pi n}$, have the same stochastic magnitude. This poses difficulties in applying central limit theory to a weighted average of $I_x(\lambda_s)$. In order to circumvent this technical difficulty, we follow Phillips (1999b) and use a direct approximation (Komlós et al. (1976), Csörgő and Horváth (1993)) for the partial sum $S_k = \sum_{j=1}^{k} \varepsilon_j$. This approach gives a uniform approximation to S_k over $0 \leq k \leq n$ in terms of a Brownian motion $B(\cdot)$ with variance σ^2. As a result, $w_x(\lambda_s)$ can be approximated by a sum of two independent Gaussian random variables ξ_s and η, where ξ_s is independent across s and η is common to all s. Then, conditional on η, we can apply the central limit theorem to the weighted sum of $w_x(\lambda_s)$ to establish the asymptotic normality of $\widehat{d}$. Because the application of the central limit theorem is conditional on η, the unconditional limit distribution of the local Whittle estimator turns out to be mixed normal (denoted as MN below). Intriguingly, the variance of $\widehat{d}$ becomes smaller than the case where $d_0 < 1$, as was found in the corresponding case for LP regression (Phillips, 1999b).

4.3 Theorem

(a) If $E|\varepsilon_t|^p < \infty$ *for* $p \geq 4$, $m \to \infty$, *and* $\frac{m^{\frac{3}{2}} \log m}{n^{\frac{1}{2}-\frac{1}{p}}} = O(1)$ *as* $n \to \infty$, *then, for* $d_0 = 1$, *we have*

$$m^{\frac{1}{2}}\left(\widehat{d} - d_0\right) \to_d MN\left(0, \sigma^2(\eta)\right) \equiv \int_{-\infty}^{\infty} N\left(0, \sigma^2(\eta)\right)\phi(\eta)\, d\eta,$$

where η *is* $N(0,1)$, $\phi(\cdot)$ *is standard normal pdf, and*

$$\sigma^2(\eta) = \frac{1}{4}\frac{1 + 2\eta^2}{1 + 2\eta^2 + \eta^4}.$$

(b) If ε_t *is Gaussian and* $\frac{1}{m} + \frac{m^{\frac{3}{2}} \log m}{n} \to 0$ *as* $n \to \infty$, *then, for* $d_0 = 1$, *we have*

$$m^{\frac{1}{2}}\left(\widehat{d} - d_0\right) \to_d MN\left(0, \sigma^2(\eta)\right) \equiv \int_{-\infty}^{\infty} N\left(0, \sigma^2(\eta)\right)\phi(\eta)\, d\eta.$$

4.4 Remarks

(a) When $d_0 = 1$, the variance of the limit distribution of $m^{\frac{1}{2}}\left(\widehat{d} - d_0\right)$ is less than $\frac{1}{4}$ since $\sigma^2(\eta) \leq \frac{1}{4}$ *a.s.*. Numerical evaluation gives

$$\sigma_{\widehat{d}}^2 = \frac{1}{4}\int_{-\infty}^{\infty} \frac{1 + 2\eta^2}{1 + 2\eta^2 + \eta^4}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{\eta^2}{2}\right) d\eta = 0.2028.$$

Thus, the limit distribution of the local Whittle estimator has less dispersion than in the stationary case. A similar phenomena applies in the limit theory for LP regression (Phillips, 1999b).

(b) The condition on the expansion rate of m becomes strong in the unit root case, but may be a consequence of the method of proof. When all moments of ε_t are finite, $m = o\left(n^{\frac{1}{3}}\right)$ is needed. This condition weakens substantially when ε_t is Gaussian.

5 Fractional Integration with a Linear Time Trend

In many applications, a nonstationary process is accompanied by a linear time trend. Accordingly, this section extends the analysis above to fractional processes with a linear time trend. Specifically, the process X_t is generated by the model

$$(1-L)^d (X_t - X_0 - \mu t) = u_t, \quad t = 0, 1, 2, \ldots, \quad \mu \neq 0 \tag{11}$$

where X_0 and u_t are defined as above. Inversion of (11) gives an alternate form for X_t, e.g. an $I(d)$ process plus a linear trend

$$X_t = (1-L)^{-d} u_t + X_0 + \mu t = \sum_{k=0}^{t-1} \frac{(d)_k}{k!} u_{t-k} + X_0 + \mu t. \tag{12}$$

Now we obtain a representation of the dft of this process. By straightforward calculation[1], we have

$$\frac{1}{\sqrt{2\pi n}} \sum_{t=1}^{n} t e^{i\lambda_s t} = -\frac{e^{i\lambda_s}}{1-e^{i\lambda_s}} \frac{\sqrt{n}}{\sqrt{2\pi}}, \tag{13}$$

and it follows that (assuming $X_0 = 0$ hereafter)

$$w_x(\lambda_s)\left(1-e^{i\lambda_s}\right) = -\frac{\mu e^{i\lambda_s}\sqrt{n}}{\sqrt{2\pi}} + D_n\left(e^{i\lambda_s}; f\right) w_u(\lambda_s) - \frac{\widetilde{U}_{\lambda_s n}(f)}{\sqrt{2\pi n}} - \frac{e^{i\lambda_s} X_n}{\sqrt{2\pi n}}. \tag{14}$$

We use (14) to examine the asymptotics of the dft and the periodogram of an $I(d)$ process with a linear trend. The following lemmas give approximations. For all the cases below, the behavior of $w_x(\lambda_s)$ is dominated by that of the trend or the final observation X_n. When $d < \frac{3}{2}$, the effect of the trend dominates. When $d > \frac{3}{2}$, the effect of X_n becomes dominant because $X_n = O_p\left(n^{d-\frac{1}{2}}\right)$.

5.1 Lemma

Suppose X_t follows (11). Then,

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

$$n^{d-\frac{3}{2}} s^{1-d} \lambda_s^d w_x(\lambda_s) = -\mu \frac{s^{1-d}\lambda_s^d}{1-e^{i\lambda_s}} \frac{e^{i\lambda_s} n^{d-1}}{\sqrt{2\pi}} + r_{s,n}(d),$$

where $E\left|r_{s,n}(d)\right|^2 = O\left(n^{2d-3} s^{2-2d}\right)$ uniformly in s.

[1] See also Corbae, Ouliaris and Phillips (1999) who give (13) and recursive formulae for dft's of higher order trends.

(b) For $d = 1$,

$$n^{-\frac{1}{2}}\lambda_s w_x(\lambda_s) = -\mu\frac{\lambda_s}{1-e^{i\lambda_s}}\frac{e^{i\lambda_s}}{\sqrt{2\pi}} + r_{s,n},$$

where $E|r_{s,n}|^2 = O(n^{-1})$ *uniformly in* s.

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

$$n^{d-\frac{3}{2}}s^{1-d}\lambda_s^d w_x(\lambda_s) = -\mu\frac{s^{1-d}\lambda_s^d}{1-e^{i\lambda_s}}\frac{e^{i\lambda_s}n^{d-1}}{\sqrt{2\pi}} + r_{s,n}(d),$$

where $E|r_{s,n}(d)|^2 = O\left(n^{2d-3}\right)$ *uniformly in* s.

(d) For $d = \frac{3}{2}$,

$$s^{1-d}\lambda_s^d w_x(\lambda_s) = -\mu\frac{s^{1-d}\lambda_s^d}{1-e^{i\lambda_s}}\frac{e^{i\lambda_s}\sqrt{n}}{\sqrt{2\pi}} - \frac{s^{1-d}\lambda_s^d}{1-e^{i\lambda_s}}\frac{e^{i\lambda_s}X_n}{\sqrt{2\pi n}} + r_{s,n},$$

where $E|r_{s,n}|^2 = O(s^{-1})$ *uniformly in* s.

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

$$s^{1-d}\lambda_s^d w_x(\lambda_s) = -\frac{s^{1-d}\lambda_s^d}{1-e^{i\lambda_s}}\frac{e^{i\lambda_s}X_n}{\sqrt{2\pi n}} + r_{s,n}^a + r_{s,n}^b(d),$$

where $E\left|r_{s,n}^a\right|^2 = O(s^{-1})$ *and* $E\left|r_{s,n}^b(d)\right|^2 = O\left(n^{3-2d}\right)$ *uniformly in* s.

5.2 Corollary

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

$$n^{2d-3}s^{2-2d}\lambda_s^{2d}I_x(\lambda_s) = \mu^2\frac{s^{2-2d}\lambda_s^{2d}}{|1-e^{i\lambda_s}|^2}\frac{n^{2d-2}}{2\pi} + R_{s,n}(d),$$

where $E|R_{s,n}(d)| = O\left(n^{d-\frac{3}{2}}s^{1-d}\right)$ *uniformly in* s.

(b) For $d = 1$,

$$n^{-1}\lambda_s^2 I_x(\lambda_s) = \mu^2\frac{\lambda_s^2}{|1-e^{i\lambda_s}|^2}\frac{1}{2\pi} + R_{s,n},$$

where $E|R_{s,n}| = O\left(n^{-\frac{1}{2}}\right)$ *uniformly in* s.

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

$$n^{2d-3}s^{2-2d}\lambda_s^{2d}I_x(\lambda_s) = \mu^2\frac{s^{2-2d}\lambda_s^{2d}}{|1-e^{i\lambda_s}|^2}\frac{n^{2d-2}}{2\pi} + R_{s,n}(d),$$

where $E|R_{s,n}(d)| = O\left(n^{d-\frac{3}{2}}\right)$ *uniformly in* s.

(d) For $d = \frac{3}{2}$,

$$s^{2-2d}\lambda_s^{2d} I_x(\lambda_s) = \frac{s^{2-2d}\lambda_s^{2d}}{\left|1-e^{i\lambda_s}\right|^2}\left|\mu\frac{\sqrt{n}}{\sqrt{2\pi}} + \frac{X_n}{\sqrt{2\pi n}}\right|^2 + R_{s,n},$$

where $E\left|R_{s,n}\right| = O\left(s^{-\frac{1}{2}}\right)$ *uniformly in* s.

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

$$s^{2-2d}\lambda_s^{2d} I_x(\lambda_s) = \frac{s^{2-2d}\lambda_s^{2d}}{\left|1-e^{i\lambda_s}\right|^2}\frac{X_n^2}{2\pi n} + R_{s,n}^a + R_{s,n}^b(d),$$

where $E\left|R_{s,n}^a\right| = O\left(s^{-\frac{1}{2}}\right)$ *and* $E\left|R_{s,n}^b(d)\right| = O\left(n^{\frac{3}{2}-d}\right)$ *uniformly in* s.

5.3 Local Gaussian Estimation: Inconsistency in the Trend Case

The following theorem shows that $\widehat{d}$ converges to unity in probability when X_t follows (11). Therefore, the local Whittle estimator is inconsistent except when $d_0 = 1$. In consequence, some caution is needed in applying the Whittle estimator to investigate the degree of long range dependence when a time series exhibits trending behavior and the nature of the trend is uncertain.

5.4 Theorem

Suppose X_t *follows (11) and* $\frac{1}{m} + \frac{m}{n} \to 0$ *as* $n \to \infty$. *Then , for* $d_0 \in \left(\frac{1}{2}, 2\right)$, $\widehat{d} \to_p 1$ *as* $n \to \infty$.

6 Simulations and an Empirical Application

First, we report simulations that were conducted to examine the finite sample performance of the local Whittle estimator using (1) with $u_t \equiv iidN(0,1)$.[2] All the results are based on 10,000 replications.

Table 1. Simulation Results for $d = 0.7$ and $d = 1.0$

	$d = 0.7$			$d = 1.0$		
n	bias	s.d.	t.s.d.	bias	s.d.	t.s.d.
200	0.0002	0.1977	0.1336	-0.0235	0.1779	0.1204
500	0.0093	0.1451	0.1066	-0.0129	0.1280	0.0960
1,000	0.0101	0.1162	0.0898	-0.0102	0.1019	0.0809

note: t.s.d signifies theoretical standard deviation.

[2] We also conducted simulations with u_t following a t_v distribution with $v = 5$ degrees of freedom. The results were very similar.

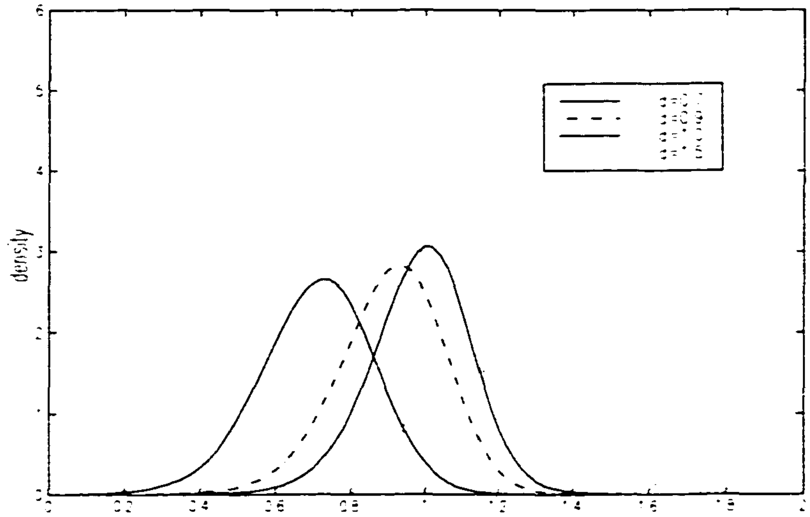

Figure 1: Densities of the local Whittle estimator: $n = 500$, $m = n^{0.5}$

Table 1 shows the simulation results for $d = 0.7$ and $d = 1.0$. The sample size and m were chosen to be $n = 200$, 500, 1000 and $m = [n^{0.5}]$. The estimator is seen to have smaller standard deviation when $d = 1.0$, corroborating the theoretical result.

Figure 1 plots the empirical distribution of the estimator for $d = 0.7$, 0.9, 1.0, 1.5 when $n = 500$ and $m = [n^{0.5}]$. The estimator appears to have a symmetric distribution when $d \leq 1$, and the positive bias and skewness of the limit distribution for $d = 0.9$ is not evident for this sample size. When $d > 1$, distribution of the estimator is concentrated around unity, again corroborating the asymptotic result.

Figure 2 displays the empirical distribution of the estimator when the process has a linear time trend. The parameter values were chosen to be $d = 0.7$, $\mu = 0.00$, 0.02, 0.05. As expected from the theory, when the value of μ increases, the distribution shifts toward unity.

As an empirical illustration, the local Whittle estimator was applied to the historical economic times series considered in Nelson and Plosser (1982). We also estimate d by the modified local Whittle estimator (see chapter 2), which is known to be consistent for $0 < d < 2$ and invariant to a linear trend[3]. Table 2 shows the estimates based on both

[3]The modified Whittle estimator adjusts the dft to take into account the second term of (9). As discussed in chapter 2, the procedure is close to one in which the local Whittle procedure is applied to differenced data and unity is added to the resulting estimate.

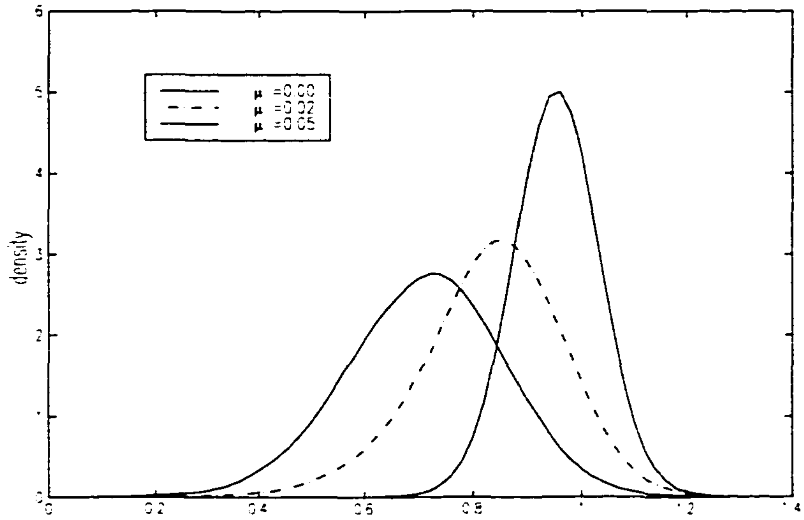

Figure 2: Densities of the local Whittle estimator with trending data: $n = 500$, $m = n^{0.5}$

$m = n^{0.5}$ and $m = n^{0.6}$. These series produce long memory estimates over a wide interval that ranges from around 0.5 for the unemployment rate to 1.38 for the bond yield. For the unemployment rate, the local Whittle estimator ($\widehat{d}_{LW}$) and the modified estimator ($\widehat{d}_{MLW}$) have values which are very close together, both indicating only marginal nonstationarity in the data. For the bond yield, the modified estimator ($\widehat{d}_{MLW}$) is very different from $\widehat{d}_{LW}$. Especially for the GNP measures, industrial production and employment, the presence of a linear trend component in the data (which is supported by much of the empirical work with this data set following Nelson and Plosser, 1982) appears to bias $\widehat{d}_{LW}$ heavily toward unity. These results, in particular, suggest that, although the local Whittle estimator is consistent for $0.5 < d \leq 1$, the use of the modified estimator may be preferable, unless the time series clearly does not involve a deterministic trend or data detrending is conducted prior to the estimation of the long memory parameter.

Table 2. Estimates of d for US Economic Data

		$m = n^{0.5}$		$m = n^{0.6}$	
	n	$\widehat{d}_{LW}$	$\widehat{d}_{MLW}$	$\widehat{d}_{LW}$	$\widehat{d}_{MLW}$
Real GNP	62	0.990	0.644	0.946	0.692
Nominal GNP	62	0.983	0.919	0.930	0.868
Real per capita GNP	62	0.976	0.650	0.912	0.700
Industrial production	111	0.918	0.488	0.968	0.570
Employment	81	1.001	0.655	0.977	0.685
Unemployment rate	81	0.507	0.520	0.705	0.718
GNP deflator	82	1.143	0.987	1.049	1.086
CPI	111	1.020	1.276	0.828	1.136
Nominal wage	71	1.080	1.080	1.015	0.973
Real wage	71	1.105	0.813	1.030	0.805
Money stock	82	1.042	0.930	0.993	1.202
Velocity of money	102	1.055	0.925	0.970	0.776
Bond yield	71	0.676	1.235	0.740	1.384
Stock prices	100	0.914	0.920	0.984	0.734

7 Appendix A: Technical Lemmas

This section lists some technical lemmas from chapter 2, to which the reader is referred for proofs.

7.1 Component Approximations (deterministic part)

The following lemmas give approximate representations of the sinusoidal polynomials $D_n\left(e^{i\lambda_s}; d\right)$ and $\widetilde{f}_{\lambda p}$ in (6).

7.2 Lemma

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

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

7.3 Lemma

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

$$\begin{aligned} \lambda^{-f}\left(1 - e^{i\lambda}\right)^f &= e^{-\frac{\pi}{2} f i} + O(\lambda), \\ \lambda^{-f}\left(1 - e^{-i\lambda}\right)^f &= e^{\frac{\pi}{2} f i} + O(\lambda). \end{aligned} \tag{16}$$

7.4 Corollary

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

$$\begin{aligned} \lambda_s^{-f} D_n\left(e^{i\lambda_s}; f\right) &= \lambda_s^{-f}\left(1 - e^{i\lambda_s}\right)^f + \lambda_s^{-f} O\left(n^{-f} s^{-1}\right) \\ &= e^{-\frac{\pi}{2} f i} + O\left(\lambda_s\right) + O\left(s^{-1-f}\right). \end{aligned} \tag{17}$$

7.5 Lemma

Uniformly in p and s,

$$(a) \quad \widetilde{f}_{\lambda_s p} = \begin{cases} O\left(p^{-f}\right), & \text{for } f > 0, \\ O\left(n^{-f}\right), & \text{for } f \in (-1, 0), \end{cases} \tag{18}$$

$$(b) \quad \widetilde{f}_{\lambda_s p} = O\left(\frac{n}{p^{f+1} s}\right), \text{ for } f > -1. \tag{19}$$

7.6 Component Approximations (stochastic part)

The following lemmas give asymptotic approximations to the terms $\widetilde{U}_{\lambda n}(f)$ and X_n in (6).

7.7 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.

7.8 Lemma

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

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

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

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

7.9 Lemma

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

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

7.10 Lemma

For $d\in\left(\frac{1}{2},2\right)$ and $1\le t\le n$, uniformly in t,

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

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

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

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

8 Appendix B: Proofs

8.1 Proof of Lemma 2.2

See Theorems 2.2 and 2.7 of Phillips (1999a). ■

8.2 Proof of Lemma 2.4

Multiplying both sides of (6) by $\lambda_s^d\left(1-e^{i\lambda_s}\right)^{-1}$ yields

$$\lambda_s^d w_x(\lambda_s) = \frac{\lambda_s^d D_n\left(e^{i\lambda_s};f\right)}{1-e^{i\lambda_s}}w_u(\lambda_s) - \frac{\lambda_s^d}{1-e^{i\lambda_s}}\left[\frac{\widetilde{U}_{\lambda_s n}(f)}{\sqrt{2\pi n}}+\frac{e^{i\lambda_s}X_n}{\sqrt{2\pi n}}\right]. \tag{20}$$

In the proof of Lemma 2.5 of chapter 2, it is shown that

$$\begin{aligned} \frac{\lambda_s^d D_n\left(e^{i\lambda_s};f\right)}{1-e^{i\lambda_s}}w_u(\lambda_s) &= e^{\frac{\pi}{2}di}C(1)w_\varepsilon(\lambda_s)+r_n^a(\lambda_s)+r_n^2(\lambda_s), \qquad (21) \\ \frac{\lambda_s^d}{1-e^{i\lambda_s}}\frac{\widetilde{U}_{\lambda_s n}(f)}{\sqrt{2\pi n}} &= \frac{\lambda_s^d}{1-e^{i\lambda_s}}\frac{C(1)\widetilde{\varepsilon}_{\lambda_s n}(f)}{\sqrt{2\pi n}}+r_n^3(\lambda_s), \end{aligned}$$

where $E\left|r_n^a(\lambda_s)\right|^2 = O\left(\lambda_s^2\right)$, $E\left|r_n^2(\lambda_s)\right|^2 = O\left(s^{2d-4}\right)$, and $E\left|r_n^3(\lambda_s)\right|^2 = O\left(s^{2d-2}n^{1-2d}\right)$, giving part (a).

For part (b), it follows from Lemma 7.7, Lemma 7.8 (a), and Lemma 7.10 (c) that

$$E\left|\frac{\lambda_s^d}{1-e^{i\lambda_s}}\left[\frac{\tilde{U}_{\lambda_s n}(f)}{\sqrt{2\pi n}} + \frac{e^{i\lambda_s}X_n}{\sqrt{2\pi n}}\right]\right|^2 = O\left(\lambda_s^{2d-2}n^{2d-2}\right) = O\left(s^{2d-2}\right), \tag{22}$$

giving the required result.

For part (c), from Lemma 2.5 of chapter 2 we have

$$\lambda_s v_x(\lambda_s) \equiv \lambda_s\left[w_x(\lambda_s) + \frac{e^{i\lambda_s}}{1-e^{i\lambda_s}}\frac{X_n}{\sqrt{2\pi n}}\right] = iC(1)\,w_\varepsilon(\lambda_s) + r_{s,n},$$

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

$$\begin{aligned}
\lambda_s w_x(\lambda_s) &= iC(1)\,w_\varepsilon(\lambda_s) - \frac{\lambda_s}{1-e^{i\lambda_s}}\frac{e^{i\lambda_s}X_n}{\sqrt{2\pi n}} + r_{s,n} \qquad (23)\\
&= iC(1)\,w_\varepsilon(\lambda_s) - \left(e^{\frac{\pi}{2}i} + O(\lambda_s)\right)(1+O(\lambda_s))\frac{X_n}{\sqrt{2\pi n}} + r_{s,n}\\
&= iC(1)\,w_\varepsilon(\lambda_s) - i\frac{X_n}{\sqrt{2\pi n}} + r_{s,n}^a,
\end{aligned}$$

where we use the fact that $EX_n^2 = O(n)$.

For part (d), multiplying both sides of (20) by s^{1-d} yields

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

From Lemma 7.7, Lemma 7.9, and Lemma 7.10, we have

$$\begin{aligned}
E\left|\frac{s^{1-d}\lambda_s^d D_n\left(e^{i\lambda_s}; f\right)}{1-e^{i\lambda_s}}w_u(\lambda_s)\right|^2 &= O\left(s^{2-2d}\right),\\
E\left|\frac{s^{1-d}\lambda_s^d}{1-e^{i\lambda_s}}\frac{\tilde{U}_{\lambda_s n}(f)}{\sqrt{2\pi n}}\right|^2 &= O\left(s^{-1}\right), \quad E\left|\frac{s^{1-d}\lambda_s^d}{1-e^{i\lambda_s}}\frac{e^{i\lambda_s}X_n}{\sqrt{2\pi n}}\right|^2 = O(1),
\end{aligned}$$

giving the required result. ■

8.3 Proof of Theorem 3.1

We follow the approach of chapter 2 and refer the reader to chapter 2 for details when they are not provided here. 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:

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

For arbitrary small $\Delta > 0$, define $\Theta_1 = \left\{ d : d_0 - \frac{1}{2} + \Delta < d < M \right\}$ and $\Theta_2 = \left\{ d : -\frac{1}{2} < d \leq d_0 - \frac{1}{2} + \Delta \right\}$. Without loss of generality, we assume $\Delta < \frac{1}{4}$ hereafter. In view of the arguments in Robinson (1995), $\widehat{d} \rightarrow_p d_0$ 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

$$\begin{aligned} T(d) &= \log \frac{\widehat{G}(d_0)}{G_0} - \log \frac{\widehat{G}(d)}{G(d)} - \log \left(\frac{1}{m} \sum_{j=1}^{m} j^{2d-2d_0} / \frac{m^{2(d-d_0)}}{2(d-d_0)+1} \right) \\ & \quad +(2d - 2d_0) \left[\frac{1}{m} \sum_{j=1}^{m} \log j - (\log m - 1) \right] . \end{aligned}$$

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

$$\begin{aligned} \frac{1}{m} \sum_{j=1}^{m} \log j - (\log m - 1) &= O\left(\frac{\log m}{m} \right), \\ \frac{2(d-d_0)+1}{m} \sum_{j=1}^{m} \left(\frac{j}{m} \right)^{2d-2d_0} - 1 &= O\left(\frac{1}{m^{2\Delta}} \right). \end{aligned} \tag{24}$$

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

$$\sup_{\Theta_1} \left| \log \frac{\widehat{G}(d_0)}{G_0} - \log \frac{\widehat{G}(d)}{G(d)} \right| \rightarrow_p 0.$$

Note that

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

$$
\begin{aligned}
&= \frac{[2(d-d_0)+1]\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0}\left[\lambda_j^{2d_0}I_x(\lambda_j)-G_0\right]}{[2(d-d_0)+1]G_0\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2(d-d_0)}} \\
&= \frac{A(d)}{B(d)}, \qquad (25)
\end{aligned}
$$

and

$$
\begin{aligned}
&\log\frac{\widehat{G}(d_0)}{G_0}-\log\frac{\widehat{G}(d)}{G(d)} \\
&= \log\left(\frac{B(d_0)+A(d_0)}{B(d_0)}\right)-\log\left(\frac{B(d)+A(d)}{B(d)}\right) \\
&= \log\left(\frac{B(d)}{B(d_0)}\right)+\log\left(\frac{B(d_0)+A(d_0)}{B(d)+A(d)}\right) \\
&= \log\left(1+\frac{B(d)-B(d_0)}{B(d_0)}\right)+\log\left(1+\frac{B(d_0)-B(d)+A(d_0)-A(d)}{B(d)+A(d)}\right).
\end{aligned}
$$

Therefore, by the fact that $\Pr(|\log Y|\geq\varepsilon)\leq\Pr(|Y-1|\geq\varepsilon/2)$ for any nonnegative random variable Y and $\varepsilon\leq 1$, $\sup_{\Theta_1}\left|\log\left(\widehat{G}(d_0)/G_0\right)-\log\left(\widehat{G}(d)/G(d)\right)\right|\to_p 0$ if

$$
\sup_{\Theta_1}\left|\frac{B(d)-B(d_0)}{B(d_0)}\right|\to_p 0 \quad\text{and}\quad \sup_{\Theta_1}\left|\frac{B(d_0)-B(d)+A(d_0)-A(d)}{B(d)+A(d)}\right|\to_p 0.
$$

From Corollary 2.5 (b) and (c), we have

$$
\lambda_j^{2d_0}I_x(\lambda_j)=\begin{cases}|C(1)|^2 I_\varepsilon(\lambda_j)+R_{j,n}^a+R_{j,n}^b(d_0), & \text{for } d_0\in\left(\frac{1}{2},1\right),\\ \left|C(1)w_\varepsilon(\lambda_j)-\frac{X_n}{\sqrt{2\pi n}}\right|^2+R_{j,n}^c, & \text{for } d_0=1,\end{cases}
$$

where $E\left|R_{j,n}^a\right|=O(\lambda_j)$, $E\left|R_{j,n}^b(d)\right|=O\left(j^{d-1}\right)$, and $E\left|R_{j,n}^c\right|=O(\lambda_j)$. 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

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

where $g=2(d-d_0)+1$ and

$$
\begin{aligned}
A_1(d) &= g\frac{|C(1)|^2}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0}\left[I_\varepsilon(\lambda_j)-\frac{\sigma^2}{2\pi}\right], \\
A_2(d) &= g\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0}R_{j,n}^a, \quad A_3(d)=g\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2d_0}R_{j,n}^b(d_0), \\
A_4(d) &= -g\frac{C(1)}{m}\frac{X_n}{\sqrt{2\pi n}}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2}\left[w_\varepsilon(\lambda_j)+w_\varepsilon(\lambda_j)^*\right], \\
A_5(d) &= g\frac{X_n^2}{2\pi n}\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2}, \quad A_6(d)=g\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2}R_{j,n}^c.
\end{aligned}
$$

We proceed to consider the successive terms $A_i(d)$ $i = 1, ..., 6$. Chapter 2 shows that for all $d \in \Theta_1$ we have $A_1(d) = O_p\left(m^{-2\Delta} + n^{-1/2}\right)$, $A_2(d) = O_p\left(n^{-1}m\right)$, $A_6(d) = O_p\left(n^{-1}m\right)$, and $A_3(d) = O_p\left(m^{d_0-1} + m^{-2\Delta}\log m\right)$.

Next consider $A_4(d)$. Using the fact that

$$\begin{aligned} E\left[w_\varepsilon(\lambda_j)\, w_\varepsilon(\lambda_k)^*\right] &= \frac{1}{2\pi n}\sum_{p=1}^{n}\sum_{q=1}^{n} E\left[\varepsilon_p e^{-ip\lambda_j}\right]\left[\varepsilon_q e^{iq\lambda_k}\right] \\ &= \frac{\sigma^2}{2\pi n}\sum_{p=1}^{n} e^{i(\lambda_k-\lambda_j)p} = \begin{cases} \sigma^2/2\pi, & \text{if } j = k, \\ 0, & \text{if } j \neq k, \end{cases} \end{aligned} \tag{26}$$

we have

$$\begin{aligned} & E\left|\left[\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2} w_\varepsilon(\lambda_j)\right]\left[\frac{1}{m}\sum_{k=1}^{m}\left(\frac{k}{m}\right)^{2d-2} w_\varepsilon(\lambda_k)^*\right]\right| \\ &= \frac{1}{m^2}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{4d-4}\frac{\sigma^2}{2\pi} = O\left(m^{-4\Delta}\right). \end{aligned}$$

Therefore, $E\left|A_4(d)\right|$ is bounded by $O\left(m^{-2\Delta}\right)$ for all $d \in \Theta_1$ because $E\left|X_n/\sqrt{n}\right|^2 = O(1)$.

For $A_5(d)$, from Phillips and Solo (1992) we have

$$A_5(d) = \frac{gX_n^2}{2\pi n}\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2} \to_d \frac{\omega^2 B(1)^2}{2\pi}\left[\lim_{m\to\infty}\frac{g}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2}\right] = \frac{\omega^2 B(1)^2}{2\pi},$$

where $\omega = \sigma C(1)$ and $B(r)$ is a standard Brownian motion.

In sum, $A(d)$ is bounded uniformly for all $d \in \Theta_1$ as follows: for $d_0 \in \left(\frac{1}{2}, 1\right)$,

$$A(d) = O_p\left(m^{-2\Delta}\log m + n^{-1/2} + n^{-1}m + m^{d_0-1}\right), \tag{27}$$

and for $d_0 = 1$,

$$A(d) = O_p\left(m^{-2\Delta} + n^{-1/2} + n^{-1}m\right) + \frac{X_n^2}{2\pi n}\frac{g}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2}. \tag{28}$$

It follows that, for $d_0 \in \left(\frac{1}{2}, 1\right)$,

$$\sup_{\Theta_1}\left|A(d) - A(d_0)\right| = O_p\left(m^{-2\Delta}\log m + n^{-1/2} + n^{-1}m + m^{d_0-1}\right), \tag{29}$$

and for $d_0 = 1$,

$$\begin{aligned} \sup_{\Theta_1}\left|A(d) - A(d_0)\right| &= O_p\left(m^{-2\Delta} + n^{-1/2} + n^{-1}m\right) + \frac{X_n^2}{2\pi n}\left(\frac{g}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2} - 1\right) \\ &= O_p\left(m^{-2\Delta} + n^{-1/2} + n^{-1}m\right). \end{aligned} \tag{30}$$

Finally, observe that

$$B(d) = [2(d-d_0)+1]\, G_0 \frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2(d-d_0)} = G_0 + O\left(m^{-2\Delta}\right), \tag{31}$$

uniformly for all $d \in \Theta_1$.

From (27)-(31) we deduce that

$$\sup_{\Theta_1}\left|\frac{B(d)-B(d_0)}{B(d_0)}\right| = o_p(1), \tag{32}$$

$$\sup_{\Theta_1}\left|\frac{B(d_0)-B(d)+A(d_0)-A(d)}{B(d)+A(d)}\right| = o_p(1). \tag{33}$$

Also we have established

$$\frac{\widehat{G}(d)}{G(d)} = 1 + \frac{\widehat{G}(d)-G(d)}{G(d)} \to_d 1 + \frac{\xi(d_0)}{G_0},$$

where

$$\xi(d_0) = \begin{cases} 0, & \text{for } d_0 \in \left(\frac{1}{2}, 1\right), \\ \omega^2 B(1)^2/2\pi, & \text{for } d_0 = 1. \end{cases}$$

Now we consider $\Theta_2 = \left\{d : -\frac{1}{2} < d \leq d_0 - \frac{1}{2} + \Delta\right\}$. From chapter 2, we have

$$\Pr\left(\inf_{\Theta_2} S(d) \leq 0\right) \leq \Pr\left(\frac{1}{m}\sum_{j=1}^{m}(a_j-1)\frac{I_x(\lambda_j)}{\lambda_j^{-2d_0}} \leq 0\right), \tag{34}$$

where $p = \exp\left(m^{-1}\sum_1^m \log j\right)$, and (see chapter 2 p.44)

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

From Corollary 2.5 (b) and (c), (34) is equal to

$$\begin{cases} \Pr(B_1+B_2+B_3 \leq 0), & \text{for } d_0 \in \left(\frac{1}{2}, 1\right), \\ \Pr(B_1+B_4+B_5+B_6 \leq 0), & \text{for } d_0 = 1. \end{cases}$$

where

$$\begin{aligned} B_1 &= \frac{|C(1)|^2}{m}\sum_{j=1}^{m}(a_j-1)\, I_\varepsilon(\lambda_j), \\ B_2 &= \frac{1}{m}\sum_{j=1}^{m}(a_j-1)\, R_{j,n}^a, \quad B_3 = \frac{1}{m}\sum_{j=1}^{m}(a_j-1)\, R_{j,n}^b(d_0), \\ B_4 &= -\frac{X_n}{\sqrt{2\pi n}}\frac{1}{m}\sum_{j=1}^{m}(a_j-1)\left[w_\varepsilon(\lambda_j)+w_\varepsilon(\lambda_j)^*\right], \\ B_5 &= \frac{X_n^2}{2\pi n}\frac{1}{m}\sum_{j=1}^{m}(a_j-1), \quad B_6 = \frac{1}{m}\sum_{j=1}^{m}(a_j-1)\, R_{j,n}^c. \end{aligned}$$

We proceed to consider the successive terms as above. Chapter 2 shows that

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

For B_4, the fact that $m^{-2} \sum_{j=1}^{m} (a_j - 1)^2 = O\left(m^{-4\Delta} + m^{-1}\right)$ (see chapter 2) and (26) yield

$$E \left| \frac{1}{m} \sum_{j=1}^{m} (a_j - 1) w_\varepsilon (\lambda_j) \right|^2 = O\left(m^{-4\Delta} + m^{-1}\right),$$

giving $B_4 = o_p(1)$. For B_5, we have

$$B_5 = \frac{X_n^2}{2\pi n} \frac{1}{m} \sum_{j=1}^{m} (a_j - 1) \to_d \frac{\omega^2 B(1)^2}{2\pi} \frac{1}{m} \sum_{j=1}^{m} (a_j - 1).$$

$B_6 = o_p(1)$ follows from $B_2 = o_p(1)$. From chapter 2, for sufficiently large m we have $\frac{1}{m} \sum_{j=1}^{m} (a_j - 1) > \delta > 0$. Thus,

$$\begin{aligned} B_1 + B_2 + B_3 &\to_p \quad G_0 \frac{1}{m} \sum_{j=1}^{m} (a_j - 1) \geq G_0 \delta > 0, \\ B_1 + B_4 + B_5 + B_6 &\to_d \quad \left(G_0 + \frac{\omega^2 B(1)^2}{2\pi}\right) \left[\frac{1}{m} \sum_{j=1}^{m} (a_j - 1)\right] \geq \left(G_0 + \frac{\omega^2 B(1)^2}{2\pi}\right) \delta \overset{a.s.}{>} 0. \end{aligned}$$

It follows that

$$\left. \begin{array}{l} \Pr(B_1 + B_2 + B_3 \leq 0) \to 0 \\ \Pr(B_1 + B_4 + B_5 + B_6 \leq 0) \to 0 \end{array} \right\} \text{ as } m \to \infty. \tag{35}$$

Therefore, $\hat{d} \to_p d_0$, giving the stated result. ■

8.4 Proof of Theorem 3.2

Since $\hat{d} \to_p d_0$ and $\widehat{G}(d)$ is a continuous function of d, we may analyse $\widehat{G}(d_0)$. We have

$$\frac{\widehat{G}(d_0) - G(d_0)}{G(d_0)} = \frac{\frac{1}{m} \sum_{j=1}^{m} \lambda_j^{2d_0} I_x(\lambda_j) - G_0}{G_0} = \frac{A(d_0)}{B(d_0)} \to_d \frac{\xi(d_0)}{G_0}.$$

We can write the final expression in the form stated for $\xi(d_0)$ and we have

$$\widehat{G}(d_0) \to_d G_0 + \xi(d_0),$$

which gives the required result. ■

8.5 Proof of Theorem 3.3

Define $G(d) = G_0 \frac{1}{m} \sum_{j=1}^{m} \lambda_j^{2(d-1)}$ and $S(d) = R(d) - R(1)$. Rewrite $S(d)$ as follows:

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

For arbitrary small $\Delta > 0$, define $\Theta_1' = \left\{ d : \frac{1}{2} + \Delta < d < M \right\}$ and $\Theta_2' = \left\{ d : -\frac{1}{2} < d \leq \frac{1}{2} + \Delta \right\}$. Without loss of generality, we assume $\Delta < \frac{1}{4}$ hereafter. Then, by the same argument as above, $\widehat{d} \to_p 1$ if

$$\sup_{\Theta_1'} \left| \log \frac{\widehat{G}(d)}{G(d)} - \log \frac{\widehat{G}(1)}{G_0} \right| \to_p 0,$$

and

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

as $n \to \infty$. Note that

$$\begin{aligned}
\frac{\widehat{G}(d)}{G(d)} &= \frac{(2d-1)\frac{1}{m}\sum_{j=1}^{m} \lambda_j^{2d} I_x(\lambda_j)}{(2d-1) G_0 \frac{1}{m} \sum_{j=1}^{m} \lambda_j^{2(d-1)}} \\
&= \frac{(2d-1)(2\pi)^{2d-2d_0} m^{2d-2} n^{2d_0-2d} \frac{1}{m} \sum_{j=1}^{m} \left(\frac{j}{m}\right)^{2d-2} j^{2-2d_0} \lambda_j^{2d_0} I_x(\lambda_j)}{(2d-1)(2\pi)^{2d-2} G_0 m^{2d-2} n^{2-2d} \frac{1}{m} \sum_{j=1}^{m} \left(\frac{j}{m}\right)^{2d-2}}
\end{aligned}$$

$$
\begin{aligned}
&= (2\pi)^{2-2d_0} n^{2d_0-2} \frac{(2d-1)\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2} j^{2-2d_0}\lambda_j^{2d_0} I_x(\lambda_j)}{(2d-1)G_0\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2}} \\
&= (2\pi)^{2-2d_0} n^{2d_0-2} \frac{A(d)}{B(d)},
\end{aligned}
$$

and

$$
\log\frac{\widehat{G}(d)}{G(d)} - \log\frac{\widehat{G}(1)}{G_0} = \log\left(\frac{A(d)}{A(1)}\right) - \log\left(\frac{B(d)}{G_0}\right).
$$

Therefore, $\sup_{\Theta_1'}\left|\log\left(\widehat{G}(d)/G(d)\right) - \log\left(\widehat{G}(1)/G_0\right)\right| \to_p 0$ if

$$
\sup_{\Theta_1'}\left|\frac{A(d)-A(1)}{A(1)}\right| \to_p 0 \quad \text{and} \quad \sup_{\Theta_1'}\left|\frac{B(d)-G_0}{G_0}\right| \to_p 0. \tag{36}
$$

From Corollary 2.5 (d), we have

$$
j^{2-2d_0}\lambda_j^{2d_0} I_x(\lambda_j) = \frac{j^{2-2d_0}\lambda_j^{2d_0}}{|1-e^{i\lambda_j}|^2}\frac{X_n^2}{2\pi n} + R_{j,n}^a(d_0) + R_{j,n}^b,
$$

where $E\left|R_{j,n}^a(d)\right| = O\left(j^{1-d}\right)$ and $E\left|R_{j,n}^b\right| = O\left(j^{-\frac{1}{2}}\right)$. Thus, $A(d)$ can be written as

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

where

$$
\begin{aligned}
A_1(d) &= \frac{2d-1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2}\frac{j^{2-2d_0}\lambda_j^{2d_0}}{|1-e^{i\lambda_j}|^2}\frac{X_n^2}{2\pi n}, \\
A_2(d) &= \frac{2d-1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2} R_{j,n}^a(d_0), \quad A_3(d) = \frac{2d-1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2} R_{j,n}^b.
\end{aligned}
$$

For $A_1(d)$, from Lemma 7.3 we have

$$
\frac{j^{2-2d_0}\lambda_j^{2d_0}}{|1-e^{i\lambda_j}|^2} = \frac{j^{2-2d_0}\lambda_j^{2d_0-2}}{\lambda_j^{-2}|1-e^{i\lambda_j}|^2} = \frac{(2\pi)^{2d_0-2}n^{2-2d_0}}{1+O(\lambda_j)}, \tag{37}
$$

and it follows that

$$
\frac{j^{2-2d_0}\lambda_j^{2d_0}}{|1-e^{i\lambda_j}|^2}\frac{X_n^2}{2\pi n} = \frac{(2\pi)^{2d_0-3}X_n^2}{n^{2d_0-1}} + r_{n,j}, \tag{38}
$$

where $E|r_{n,j}| = O(\lambda_j)$. Thus, $A_1(d)$ can be written as

$$
A_1(d) = \frac{(2\pi)^{2d_0-3}X_n^2}{n^{2d_0-1}}\frac{2d-1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2} + \frac{2d-1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2} r_{n,j}.
$$

Furthermore, the fact that $\frac{2d-1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2} = 1 + O\left(m^{-2\Delta}\right)$ and

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

yields

$$A_1(d) = n^{1-2d_0}(2\pi)^{2d_0-3} X_n^2 + O_p\left(m^{-2\Delta} + n^{-1}m\right).$$

For $A_2(d)$ and $A_3(d)$, we obtain

$$\begin{aligned} E\left|A_2(d)\right| &= O\left(\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2} j^{1-d_0}\right) \\ &= O\left(m^{1-2d}\sum_{j=1}^{m} j^{2d-d_0-1}\right) \\ &= \begin{cases} O\left(m^{1-d_0}\right) & \text{for } 2d-d_0>0 \\ O\left(m^{1-2d}\log m\right) & \text{for } 2d-d_0\le 0 \end{cases} \\ &= O\left(m^{1-d_0} + m^{-2\Delta}\log m\right), \end{aligned}$$

$$E\left|A_3(d)\right| = O\left(\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2} j^{-\frac{1}{2}}\right) = O\left(m^{1-2d}\sum_{j=1}^{m} j^{2d-\frac{5}{2}}\right) = O\left(m^{-\frac{1}{2}} + m^{-2\Delta}\log m\right). \tag{39}$$

Therefore, we deduce that, uniformly over $d \in \Theta_1'$,

$$\begin{aligned} A(d) &= n^{1-2d_0}(2\pi)^{2d_0-3} X_n^2 + O_p\left(n^{-1}m + m^{1-d_0} + m^{-2\Delta}\log m\right), \\ B(d) &= G_0\left(1 + O\left(m^{-2\Delta}\right)\right), \end{aligned}$$

which gives

$$\sup_{\Theta_1'}\left|\frac{B(d)-G_0}{G_0}\right| \to 0.$$

For $A(d)$, from Lemma 7.10 and Akonom and Gourieroux (1987), we have

$$n^{1-2d_0}(2\pi)^{2d_0-3} X_n^2 \to_d (2\pi)^{2d_0-3} C(1)^2 \sigma^2 B_{d_0-1}(1)^2,$$

where $B_d(1) = \Gamma(d+1)^{-1}\int_0^1 (1-s)^d dB(s)$. It follows that, for $d \in \Theta_1'$,

$$\frac{A(d)-A(1)}{A(1)} = \frac{O_p\left(n^{-1}m + m^{1-d_0} + m^{-2\Delta}\log m\right)}{n^{1-2d_0}(2\pi)^{2d_0-3} X_n^2 + O_p\left(n^{-1}m + m^{1-d_0} + m^{-2\Delta}\log m\right)}.$$

Because the denominator is $o_p\left((\log m)^{-1}\right)$ and $\Pr\left(A(1)(\log m) < \varepsilon\right) \to 0$ for all $\varepsilon > 0$, it follows that

$$\sup_{\Theta_1'}\left|\frac{A(d)-A(1)}{A(1)}\right| \to_p 0,$$

and we establish (36).

Now we consider $\Theta_2' = \left\{d : -\frac{1}{2} < d \leq \frac{1}{2} + \Delta\right\}$. Let $p = \exp\left(m^{-1} \sum_1^m \log j\right)$. Then, $S(d) = \log\left\{\widehat{D}(d) / \widehat{D}(1)\right\}$, where

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

It follows that

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

where

$$a_j = \begin{cases} \left(\frac{j}{p}\right)^{2\Delta - 1}, & \text{for } 1 \leq j \leq p, \\ \left(\frac{j}{p}\right)^{-3}, & \text{for } p < j \leq m. \end{cases}$$

Then,

$$\begin{aligned} \Pr\left(\inf_{\Theta_2'} S(d) \leq 0\right) &\leq \Pr\left(\frac{1}{m} \sum_{j=1}^{m} (a_j - 1) j^2 I_x(\lambda_j) \leq 0\right) \\ &= \Pr\left(\frac{1}{m} \sum_{j=1}^{m} (a_j - 1) j^{2-2d_0} \lambda_j^{2d_0} I_x(\lambda_j) \leq 0\right). \end{aligned} \tag{40}$$

From Corollary 2.5 (d), (40) is equal to

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

where

$$\begin{aligned} B_1 &= \frac{1}{m} \sum_{j=1}^{m} (a_j - 1) \frac{j^{2-2d_0} \lambda_j^{2d_0}}{|1 - e^{i\lambda_j}|^2} \frac{X_n^2}{2\pi n}, \\ B_2 &= \frac{1}{m} \sum_{j=1}^{m} (a_j - 1) R_{j,n}^a(d_0), \quad B_3 = \frac{1}{m} \sum_{j=1}^{m} (a_j - 1) R_{j,n}^b, \end{aligned}$$

with $E\left|R_{j,n}^a(d)\right| = O\left(j^{1-d}\right)$ and $E\left|R_{j,n}^b\right| = O\left(j^{-\frac{1}{2}}\right)$. We proceed to consider the successive terms as above. For B_1, it follows from (38) that

$$B_1 = \frac{(2\pi)^{2d_0 - 3} X_n^2}{n^{2d_0 - 1}} \frac{1}{m} \sum_{j=1}^{m} (a_j - 1) + O_p\left(\frac{1}{m} \sum_{j=1}^{m} (a_j - 1) \frac{j}{n}\right),$$

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 fact that

$$\sum_{j=1}^{m} a_j = \sum_{1 \leq j \leq p} a_j + \sum_{p+1 \leq j \leq m} a_j = O(m) + O\left(p^3 \int_p^m x^{-3} dx\right) = O(m),$$

we have

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

$E|B_2|$ and $E|B_3|$ are $o(1)$, because

$$\begin{aligned}\frac{1}{m}\sum_{j=1}^{m}a_j j^{1-d_0} &= \frac{p^{1-2\Delta}}{m}\sum_{j=1}^{p}j^{2\Delta-d_0}+\frac{p^3}{m}\sum_{j=p+1}^{m}j^{-2-d_0}=O\left(m^{-2\Delta}\log m+m^{1-d_0}\right),\\ \frac{1}{m}\sum_{j=1}^{m}a_j j^{-\frac{1}{2}} &= \frac{p^{1-2\Delta}}{m}\sum_{j=1}^{p}j^{2\Delta-\frac{3}{2}}+\frac{p^3}{m}\sum_{j=p+1}^{m}j^{-\frac{7}{2}}=O\left(m^{-2\Delta}\log m+m^{-\frac{1}{2}}\right),\end{aligned}$$

and $m^{-1}\sum_1^m j^{1-d_0}=O\left(m^{1-d_0}\right)$, $m^{-1}\sum_1^m j^{-\frac{1}{2}}=O\left(m^{-\frac{1}{2}}\right)$.

Choose $\Delta<1/(2e)<1/4$ with no loss of generality, then for sufficiently large m,

$$\frac{1}{m}\sum_{j=1}^{m}(a_j-1)\geq\frac{1}{m}\sum_{1\leq j\leq p}a_j-1\sim\frac{1}{2\Delta e}-1>\delta>0.$$

Hence,

$$B_1+B_2+B_3=\frac{(2\pi)^{2d_0-3}X_n^2}{n^{2d_0-1}}\frac{1}{m}\sum_{j=1}^{m}(a_j-1)+o_p(1),$$

and it follows that

$$\Pr(B_1+B_2+B_3\leq 0)\to 0\quad\text{as } m\to\infty.$$

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

8.6 Proof of Theorem 4.1

Because the proof follows the same argument as Theorems 4.1 and 4.2 of chapter 2, we provide here only the relevant parts. We work from the first order conditions for $\widehat{d}$, viz.

$$0=R'\left(\widehat{d}\right)=R'(d_0)+R''(d^*)\left(\widehat{d}-d_0\right),\tag{41}$$

where $|d^*-d_0|\leq\left|\widehat{d}-d_0\right|$, and

$$R''(d)=\frac{4\left[\widehat{F}_2(d)\widehat{F}_0(d)-\widehat{F}_1(d)^2\right]}{\widehat{F}_0(d)^2},\quad \widehat{F}_k(d)=\frac{1}{m}\sum_{j=1}^{m}(\log j)^k\lambda_j^{2d}I_x(\lambda_j)$$

The condition on m and n implies that $\widehat{d}$ is consistent, and from (29) and (31) we have

$$\begin{aligned}\sup_{\Theta_1}\left|\frac{B(d)-B(d_0)}{B(d_0)}\right| &= o_p\left((\log m)^{-6}\right),\\ \sup_{\Theta_1}\left|\frac{B(d_0)-B(d)+A(d_0)-A(d)}{B(d)+A(d)}\right| &= o_p\left((\log m)^{-6}\right),\end{aligned}$$

which gives $R''(d^*) = R''(d_0) + o_p(1)$. Now, from Corollary 2.5 (b), we find

$$\widehat{F}_k(d_0) = \frac{1}{m}\sum_{j=1}^{m}(\log j)^k \lambda_j^{2d_0} I_x(\lambda_j) = C_1 + C_2 + C_3,$$

where

$$C_1 = \frac{|C(1)|^2}{m}\sum_{j=1}^{m}(\log j)^k I_\varepsilon(\lambda_j),$$
$$C_2 = \frac{1}{m}\sum_{j=1}^{m}(\log j)^k R_{j,n}^a, \quad C_3 = \frac{1}{m}\sum_{j=1}^{m}(\log j)^k R_{j,n}^b(d_0).$$

From the proof of Theorem 4.1 of chapter 2, we have $C_1 = \frac{1}{m}\sum_{j=1}^{m}(\log j)^k (G_0 + o_p(1))$, $E|C_2| = o\left((\log m)^k\right)$, and $E|C_3| = o\left((\log m)^k\right)$. Hence

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

and it follows that

$$R''(d_0) = 4 + o_p(1). \tag{42}$$

Next we consider the first term on the right side of (41). It follows from chapter 2 that

$$m^{\frac{1}{2}}R'(d_0) = \frac{\frac{2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\left[\lambda_j^{2d_0}I_x(\lambda_j) - G_0\right]}{\widehat{G}(d_0)}, \tag{43}$$

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 3.2 we have

$$\widehat{G}(d_0) \to_p G_0. \tag{44}$$

By Corollary 2.5 (a), the numerator can be decomposed as follows:

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

where

$$D_1 = \frac{2|C(1)|^2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\left[I_\varepsilon(\lambda_j) - \frac{\sigma^2}{2\pi}\right],$$
$$D_2 = \frac{2|C(1)|^2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\frac{\lambda_j^{2d_0}}{|1-e^{i\lambda_j}|^2}\frac{\left|\widetilde{\varepsilon}_{\lambda_j n}(f_0)\right|^2}{2\pi n}, \quad D_3 = \frac{2}{\sqrt{m}}\frac{X_n^2}{2\pi n}\sum_{j=1}^{m}\nu_j\frac{\lambda_j^{2d_0}}{|1-e^{i\lambda_j}|^2},$$

$$
\begin{aligned}
D_4 &= -\frac{2|C(1)|^2}{\sqrt{m}} \\
&\quad \times \sum_{j=1}^{m} \nu_j \left[e^{\frac{\pi}{2}d_0 i} w_\varepsilon(\lambda_j) \frac{\lambda_j^{d_0}}{1-e^{-i\lambda_j}} \frac{\widetilde{\varepsilon}_{\lambda_j n}(f_0)^*}{\sqrt{2\pi n}} + \frac{\lambda_j^{d_0}}{1-e^{i\lambda_j}} \frac{\widetilde{\varepsilon}_{\lambda_j n}(f_0)}{\sqrt{2\pi n}} e^{-\frac{\pi}{2}d_0 i} w_\varepsilon(\lambda_j)^* \right], \\
D_5 &= -\frac{2C(1)}{\sqrt{m}} \frac{X_n}{\sqrt{2\pi n}} \sum_{j=1}^{m} \nu_j \left[e^{\frac{\pi}{2}d_0 i} w_\varepsilon(\lambda_j) \frac{\lambda_j^{d_0} e^{-i\lambda_j}}{1-e^{-i\lambda_j}} + \frac{\lambda_j^{d_0} e^{i\lambda_j}}{1-e^{i\lambda_j}} e^{-\frac{\pi}{2}d_0 i} w_\varepsilon(\lambda_j)^* \right], \\
D_6 &= \frac{2C(1)}{\sqrt{m}} \frac{X_n}{2\pi n} \sum_{j=1}^{m} \nu_j \left[\frac{\lambda_j^{2d_0} e^{-i\lambda_j}}{|1-e^{i\lambda_j}|^2} \widetilde{\varepsilon}_{\lambda_j n}(f_0) + \frac{\lambda_j^{2d_0} e^{i\lambda_j}}{|1-e^{i\lambda_j}|^2} \widetilde{\varepsilon}_{\lambda_j n}(f_0)^* \right], \\
D_7 &= \frac{2}{\sqrt{m}} \sum_{j=1}^{m} \nu_j R_{j,n}^a, \quad D_8 = \frac{2}{\sqrt{m}} \sum_{j=1}^{m} \nu_j R_{j,n}^b(d_0), \quad D_9 = \frac{2}{\sqrt{m}} \sum_{j=1}^{m} \nu_j R_{j,n}^c(d_0).
\end{aligned}
$$

From the proof of Theorem 4.1 of chapter 2, we have $D_1 \to_d N(0, 4G_0^2)$ and $D_2 + D_4 + D_7 + D_8 + D_9 = o_p(1)$. For D_3, D_5 and D_6, we can apply the proof of Theorem 4.2 of chapter 2 with replacing $\overline{d}_0$ in chapter 2 by d_0 to show that $D_3 + D_5 + D_6 = o_p(1)$.

Therefore, we obtain

$$
m^{\frac{1}{2}} R'(d_0) \Rightarrow \frac{1}{G_0} N\left(0, 4G_0^2\right). \tag{45}
$$

It follows from (41), (42) and (45) that

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

giving the required result. ■

8.7 Proof of Theorem 4.2

The proof follows from the proof of Theorem 4.4 of chapter 2, replacing $\overline{d}_0$ by d_0. First, the conditions on m and n imply that $\widehat{d}$ is consistent and

$$
\begin{aligned}
\sup_{\Theta_1} \left| \frac{B(d) - B(d_0)}{B(d_0)} \right| &= o_p\left((\log m)^{-6}\right), \\
\sup_{\Theta_1} \left| \frac{B(d_0) - B(d) + A(d_0) - A(d)}{B(d) + A(d)} \right| &= o_p\left((\log m)^{-6}\right),
\end{aligned}
$$

which gives $R''(d^*) = R''(d_0) + o_p(1)$. Recall

$$
\begin{aligned}
m^{2-2d_0}\left(\widehat{d} - d_0\right) &= -\frac{m^{\frac{3}{2}-2d_0}(D_1 + D_2 + D_4 + D_5 + D_6 + D_7 + D_8 + D_9)}{4G_0 + o_p(1)} \qquad (46) \\
&\quad -\frac{m^{\frac{3}{2}-2d_0} D_3}{4G_0 + o_p(1)}, \qquad (47)
\end{aligned}
$$

where

$$D_3 = \frac{2}{\sqrt{m}} \frac{X_n^2}{2\pi n} \sum_{j=1}^{m} \nu_j \frac{\lambda_j^{2d_0}}{\left|1 - e^{i\lambda_j}\right|^2} = \frac{2\,(2\pi)^{2d_0-2}}{\sqrt{m}} \frac{X_n^2}{2\pi n^{2d_0-1}} \sum_{j=1}^{m} \nu_j j^{2d_0-2} + o_p(1).$$

From the proof of Theorem 4.4 of chapter 2, we have

$$\sum_{j=1}^{m} \nu_j j^{2d_0-2} = \frac{(2d_0 - 2)\, m^{2d_0-1}}{(2d_0 - 1)^2} + O(\log m),$$

and, for $d_0 \in \left[\frac{3}{4}, 1\right)$,

$$m^{\frac{3}{2}-2d_0} D_3 \to_d 2\,(2\pi)^{2d_0-2} \frac{C(1)^2 \sigma^2}{2\pi} B_{d_0-1}(1)^2 \frac{2d_0 - 2}{(2d_0 - 1)^2}.$$

For $d_0 = \frac{3}{4}$, (46) converges to $N\left(0, \frac{1}{4}\right)$ and

$$-\frac{m^{\frac{3}{2}-2d_0} D_3}{4G_0 + o_p(1)} \to_d \frac{(1 - d_0)\,(2\pi)^{2d_0-2}}{(2d_0 - 1)^2} B_{d_0-1}(1)^2 \equiv (2\pi)^{-\frac{1}{2}} B_{-\frac{1}{4}}(1)^2.$$

For $d_0 \in \left(\frac{3}{4}, 1\right)$, (46) is $o_p(1)$ and

$$-\frac{m^{\frac{3}{2}-2d_0} D_3}{4G_0 + o_p(1)} \to_d \frac{(1 - d_0)\,(2\pi)^{2d_0-2}}{(2d_0 - 1)^2} B_{d_0-1}(1)^2,$$

giving the required result. ■

8.8 Proof of Theorem 4.3

The condition on m and n implies that $\widehat{d}$ is consistent and $R''(d^*) = R''(d_0) + o_p(1)$. Now, from Corollary 2.5 (c), we find

$$\widehat{F}_k(d_0) = \frac{1}{m} \sum_{j=1}^{m} (\log j)^k \lambda_j^{2d_0} I_x(\lambda_j) = C_1 + C_2 + C_3 + C_4,$$

where

$$\begin{aligned} C_1 &= \frac{|C(1)|^2}{m} \sum_{j=1}^{m} (\log j)^k I_\varepsilon(\lambda_j), \quad C_2 = -\frac{C(1)}{m} \frac{X_n}{\sqrt{2\pi n}} \sum_{j=1}^{m} (\log j)^k \left[w_\varepsilon(\lambda_j) + w_\varepsilon(\lambda_j)^*\right], \\ C_3 &= \frac{1}{m} \frac{X_n^2}{2\pi n} \sum_{j=1}^{m} (\log j)^k, \quad C_4 = \frac{1}{m} \sum_{j=1}^{m} (\log j)^k R_{j,n}^a, \end{aligned}$$

and $E\left|R_{j,n}^a\right| = O(\lambda_j)$. From the proof of Theorem 4.1 of chapter 2, we have $C_1 = \frac{1}{m}\sum_{j=1}^{m} (\log j)^k (G_0 + o_p(1))$ and $E|C_4| = o\left((\log m)^k\right)$. Next consider C_2. In view of the fact that $E\left[w_\varepsilon(\lambda_j)^* w_\varepsilon(\lambda_k)\right] = \frac{\sigma^2}{2\pi}\mathbf{1}\{j=k\}$, we have

$$\begin{aligned} E\left|\left[\frac{1}{m}\sum_{j=1}^{m}(\log j)^k w_\varepsilon(\lambda_j)\right]\left[\frac{1}{m}\sum_{l=1}^{m}(\log l)^k w_\varepsilon(\lambda_l)^*\right]\right| &= O\left(\frac{1}{m^2}\sum_{j=1}^{m}(\log j)^{2k}\right) \\ &= o\left((\log m)^{2k}\right). \end{aligned}$$

It follows that

$$C_2 = O_p(1) \times o_p\left((\log m)^k\right) = o_p\left((\log m)^k\right).$$

Hence

$$\widehat{F}_k(d_0) = \left(G_0 + \frac{X_n^2}{2\pi n}\right)\left[\frac{1}{m}\sum_{j=1}^{m}(\log j)^k\right][1+o_p(1)],$$

and we obtain

$$\begin{aligned} R''(d_0) &= \frac{4\left[\widehat{F}_2(d_0)\widehat{F}_0(d_0) - \widehat{F}_1(d_0)^2\right]}{\widehat{F}_0(d_0)^2} \\ &= \frac{4\left(G_0 + \frac{X_n^2}{2\pi n}\right)^2\left[\frac{1}{m}\sum_{j=1}^{m}(\log j)^2 - \left(\frac{1}{m}\sum_{j=1}^{m}\log j\right)^2\right]}{\left(G_0 + \frac{X_n^2}{2\pi n}\right)^2}[1+o_p(1)] \\ &= 4 + o_p(1). \end{aligned} \tag{48}$$

Recall that (note that $\sum_{j=1}^{m}\nu_j = 0$)

$$m^{\frac{1}{2}}R'(d_0) = \frac{\frac{2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\left[\lambda_j^{2d_0}I_x(\lambda_j) - G_0\right]}{\widehat{G}(d_0)} = \frac{\frac{2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\lambda_j^{2d_0}I_x(\lambda_j)}{\widehat{G}(d_0)}. \tag{49}$$

From Lemma 7.10, we have

$$\frac{X_n}{\sqrt{2\pi n}} = \frac{C(1)X_n^\varepsilon}{\sqrt{2\pi n}} + r_n = C(1)w_\varepsilon(\lambda_0) + r_n, \tag{50}$$

where $E|r_n|^2 = O(n^{-1})$. It follows from (50) and Corollary 2.5 (c) that

$$\frac{2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\lambda_j^{2d_0}I_x(\lambda_j) = D_1 + D_2 + D_3,$$

where

$$\begin{aligned} D_1 &= \frac{2|C(1)|^2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\left|w_\varepsilon(\lambda_j) - w_\varepsilon(\lambda_0)\right|^2, \\ D_2 &= \frac{2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j R_{j,n}^a, \quad D_3 = \frac{2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j R_{j,n}^b, \end{aligned}$$

and $E\left|R^{a}_{j,n}\right| = O(\lambda_j)$ and $E\left|R^{b}_{j,n}\right| = O\left(n^{-\frac{1}{2}}\right)$. It is straightforward to show that $E|D_2| = O\left(n^{-1}m^{\frac{3}{2}}\log m\right)$ and $E|D_3| = O\left(n^{-\frac{1}{2}}m^{\frac{1}{2}}\log m\right)$, and hence both D_2 and D_3 are $o_p(1)$.

From Theorem 3.2 of Phillips (1999b), we have

$$\begin{aligned} w_\varepsilon(\lambda_s) &= \frac{1}{\sqrt{2\pi}}\int_0^1 e^{2\pi i s r}dB(r) + o_p\left(\frac{m}{n^{\frac{1}{2}-\frac{1}{p}}}\right), \\ w_\varepsilon(\lambda_0) &= \frac{1}{\sqrt{2\pi}}B(1) + o_p\left(\frac{m}{n^{\frac{1}{2}-\frac{1}{p}}}\right), \end{aligned}$$

where the error magnitude holds uniformly in $s \leq m$. Let us define

$$\xi_s = \int_0^1 e^{2\pi i s r}dW(r), \quad \eta = W(1),$$

where $W(\cdot)$ is a standard Brownian motion. The variates $\{\xi_s\}_{s=1}^m$ are independent complex Gaussian $N_c(0,1)$ and are independent of η, which is real Gaussian $N(0,1)$. Then, it follows that

$$\begin{aligned} |w_\varepsilon(\lambda_j) - w_\varepsilon(\lambda_0)|^2 &\equiv \frac{\sigma^2}{2\pi}\left|\xi_j - \eta + o_p\left(\frac{m}{n^{\frac{1}{2}-\frac{1}{p}}}\right)\right|^2 \\ &= \frac{\sigma^2}{2\pi}\left|\xi_j - \eta\right|^2 + \left|\xi_j - \eta\right| o_p\left(\frac{m}{n^{\frac{1}{2}-\frac{1}{p}}}\right) + o_p\left(\frac{m^2}{n^{1-\frac{2}{p}}}\right), \end{aligned} \quad (51)$$

where $\equiv$ signifies equality in distribution, and the error magnitude holds uniformly in $j \leq m$. Applying the Cauchy inequality

$$\sum |x_j y_j| \leq \left(\sum |x_j|^2\right)^{1/2}\left(\sum |y_j|^2\right)^{1/2},$$

to $x_j = \left|\xi_j - \eta\right|/\sqrt{m}$ and $y_j = \nu_j o_p\left(m/n^{\frac{1}{2}-\frac{1}{p}}\right)$, we obtain

$$\begin{aligned} \left(\frac{1}{\sqrt{m}}\sum_{j=1}^m \nu_j\left|\xi_j - \eta\right| o_p\left(\frac{m}{n^{\frac{1}{2}-\frac{1}{p}}}\right)\right)^2 &\leq \frac{1}{m}\sum_{j=1}^m\left|\xi_j - \eta\right|^2\sum_{j=1}^m \nu_j^2 o_p\left(\frac{m^2}{n^{1-\frac{2}{p}}}\right) \\ &= O_p(1) \times o_p\left(\frac{m^3(\log m)^2}{n^{1-\frac{2}{p}}}\right) = o_p(1), \end{aligned}$$

by the fact that $\frac{m^{\frac{3}{2}}\log m}{n^{\frac{1}{2}-\frac{1}{p}}} = O(1)$. It follows that the other remainder term in (51) is $o_p(1)$.

Therefore, we have

$$D_1 \equiv \frac{|C(1)|^2}{\sqrt{m}}\frac{\sigma^2}{2\pi}\sum_{j=1}^m \nu_j 2\left|\xi_j - \eta\right|^2 + o_p(1).$$

Using the argument in Phillips (1999b), let us write $\xi_j = \zeta_{1j} + \zeta_{2j} i$. The components ζ_{1j}, ζ_{2j} are independent and each is $N\left(0, \frac{1}{2}\right)$. Then

$$2\left|\xi_j - \eta\right|^2 = 2\left[\left(\zeta_{1j} - \eta\right)^2 + \zeta_{2j}^2\right] = G_{j\eta}.$$

Conditional on η, $\zeta_{1j} - \eta$ is $N\left(-\eta, \frac{1}{2}\right)$, and so, conditional on η,

$$G_{j\eta} = \frac{\left(\zeta_{1j} - \eta\right)^2 + \zeta_{2j}^2}{1/2} \equiv \chi_2^2(\delta).$$

Thus, conditional on η, the family $\{G_{j\eta}\}_1^m$ are independent and identically distributed non-central chi-squared variates with two degrees of freedom and noncentrality parameter δ where

$$\delta = \left(\frac{-\eta}{1/\sqrt{2}}\right)^2 = 2\eta^2.$$

Let $E(G_{j\eta}|\eta) = \mu_\eta$ and $Var(G_{j\eta}|\eta) = \sigma_\eta^2$. From the moments of non-central chi-squared random variables, we get

$$\mu_\eta = 2 + 2\eta^2, \quad \sigma_\eta^2 = 2\left(2 + 4\eta^2\right).$$

Thus, conditional on η, $G_{j\eta}$ is $iid\left(\mu_\eta, \sigma_\eta^2\right)$ and it follows that

$$\frac{1}{\sqrt{m}}\sum_{j=1}^{m} \nu_j G_{j\eta} = \frac{1}{\sqrt{m}}\sum_{j=1}^{m} \nu_j\left(G_{j\eta} - \mu_\eta\right) \rightarrow_d N\left(0, \sigma_\eta^2\left(\lim_{m\rightarrow\infty}\frac{1}{m}\sum_{j=1}^{m}\nu_j^2\right)\right) \equiv N\left(0, \sigma_\eta^2\right),$$

by the Lindeberg-Feller central limit theorem (c.f. Robinson, 1995, p. 1070). From (28), the denominator of (49) is (note that $G_0 = \frac{\sigma^2}{2\pi}\left|C(1)\right|^2$)

$$\widehat{G}(d_0) = A(d_0) + G_0 = G_0 + \frac{X_n^2}{2\pi n} + o_p(1) = G_0 + \frac{C(1)^2 X_n^{\varepsilon 2}}{2\pi n} + o_p(1) \equiv G_0\left(1 + \eta^2\right) + o_p(1).$$

Therefore, conditionally on η, we have

$$\begin{aligned}
\frac{\frac{2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\lambda_j^{2d_0} I_x(\lambda_j)}{\widehat{G}(d_0)} &\equiv \frac{G_0\frac{1}{\sqrt{m}}\sum_{j=1}^{m}\nu_j G_{j\eta} + o_p(1)}{G_0(1+\eta^2) + o_p(1)} \\
&= \frac{\frac{1}{\sqrt{m}}\sum_{j=1}^{m}\nu_j G_{j\eta}}{1+\eta^2} + o_p(1) \rightarrow_d N\left(0, \frac{4(1+2\eta^2)}{(1+\eta^2)^2}\right).
\end{aligned}$$

It follows that, conditionally on η,

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

Unconditionally, we therefore obtain

$$\begin{aligned} m^{\frac{1}{2}}\left(\widehat{d}-d_0\right) \rightarrow_d & \quad MN\left(0, \frac{1}{4}\frac{1+2\eta^2}{1+2\eta^2+\eta^4}\right) \\ & \equiv \int_{-\infty}^{\infty} N\left(0, \frac{1}{4}\frac{1+2\eta^2}{1+2\eta^2+\eta^4}\right) \mathrm{pdf}(\eta)\, d\eta, \end{aligned}$$

giving the required result.

When ε_t is Gaussian,

$$\left|w_\varepsilon(\lambda_j) - w_\varepsilon(\lambda_0)\right|^2 \equiv \frac{\sigma^2}{2\pi}\left|\xi_j - \eta\right|^2,$$

holds exactly. It follows that

$$\begin{aligned} \frac{2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\lambda_j^{2d_0}I_x(\lambda_j) & \equiv \frac{|C(1)|^2}{\sqrt{m}}\frac{\sigma^2}{2\pi}\sum_{j=1}^{m}\nu_j 2\left|\xi_j-\eta\right|^2 + o_p(1), \\ \widehat{G}(d_0) & \equiv G_0\left(1+\eta^2\right) + o_p(1), \end{aligned}$$

under $n^{-1}m^{\frac{3}{2}}\log m \rightarrow 0$, giving the required result. ■

8.9 Proof of Lemma 5.1

For part (a), multiplying both sides of (14) by $n^{d-\frac{3}{2}}s^{1-d}\lambda_s^d\left(1-e^{i\lambda_s}\right)^{-1}$ yields

$$\begin{aligned} & n^{d-\frac{3}{2}}s^{1-d}\lambda_s^d w_x(\lambda_s) \\ = & -\mu\frac{s^{1-d}\lambda_s^d}{1-e^{i\lambda_s}}\frac{e^{i\lambda_s}n^{d-1}}{\sqrt{2\pi}} + \frac{n^{d-\frac{3}{2}}s^{1-d}\lambda_s^d}{1-e^{i\lambda_s}}\left[D_n\left(e^{i\lambda_s};f\right)w_u(\lambda_s) - \frac{\widetilde{U}_{\lambda_s n}(f)}{\sqrt{2\pi n}} - \frac{e^{i\lambda_s}X_n}{\sqrt{2\pi n}}\right]. \end{aligned}$$

From (21) and (22), we have

$$E\left|\frac{n^{d-\frac{3}{2}}s^{1-d}\lambda_s^d}{1-e^{i\lambda_s}}\left[D_n\left(e^{i\lambda_s};f\right)w_u(\lambda_s) - \frac{\widetilde{U}_{\lambda_s n}(f)}{\sqrt{2\pi n}} - \frac{e^{i\lambda_s}X_n}{\sqrt{2\pi n}}\right]\right|^2 = O\left(n^{2d-3}s^{2-2d}\right),$$

and Corollary (a) follows from the fact that

$$\frac{s^{1-d}\lambda_s^d}{1-e^{i\lambda_s}}\frac{e^{i\lambda_s}n^{d-1}}{\sqrt{2\pi}} = O(1).$$

Part (b) follows from (14) and Lemma 2.4 (c). Part (c), part (d), and part(e) follow from (14) and Lemma 2.4 (d). ■

8.10 Proof of Theorem 5.4

We can apply the technique used in the proof of Theorem 3.3 with modifying definition of $A(d)$ and $B_k(d)$. First consider the case $d_0 \in (\frac{1}{2}, \frac{3}{2})$. For $d \in \Theta_1' = \left\{d : \frac{1}{2} + \Delta < d < M\right\}$, let $A(d)$ be defined as

$$A(d) = n^{2d_0-3}(2d-1)\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2} j^{2-2d_0}\lambda_j^{2d_0} I_x(\lambda_j).$$

Rewrite $A(d)$ as $A(d) = A_1(d) + A_2(d)$, where

$$\begin{aligned} A_1(d) &= \mu^2 \frac{2d-1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2} \frac{j^{2-2d_0}\lambda_j^{2d_0}}{|1-e^{i\lambda_j}|^2}\frac{n^{2d_0-2}}{2\pi}, \\ A_2(d) &= \frac{2d-1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2} R_{j,n}(d_0), \end{aligned}$$

and

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

For $A_1(d)$, from (37) we have $j^{2-2d_0}\lambda_j^{2d_0}\left|1-e^{i\lambda_j}\right|^{-2} = (2\pi)^{2d_0-2} n^{2-2d_0}(1+O(\lambda_j))$, and it follows that

$$\frac{j^{2-2d_0}\lambda_j^{2d_0}}{|1-e^{i\lambda_j}|^2}\frac{n^{2d_0-2}}{2\pi} = (2\pi)^{2d_0-3} + o(1), \tag{52}$$

Hence, $A_1(d) = \mu^2(2\pi)^{2d_0-3} + o(1)$. For $A_2(d)$,

$$\begin{aligned} E|A_2(d)| &= \begin{cases} O\left(n^{d_0-\frac{3}{2}} m^{-1}\sum_1^m (j/m)^{2d-2} j^{1-d_0}\right) & \text{for } d_0 \in \left(\frac{1}{2}, 1\right] \\ O\left(n^{d_0-\frac{3}{2}} m^{-1}\sum_1^m (j/m)^{2d-2}\right) & \text{for } d_0 \in \left(1, \frac{3}{2}\right) \end{cases} \\ &= \begin{cases} O\left(n^{d_0-\frac{3}{2}} m^{1-2d}\sum_1^m j^{2d-1-d_0}\right) & \text{for } d_0 \in \left(\frac{1}{2}, 1\right] \\ O\left(n^{d_0-\frac{3}{2}} m^{1-2d}\sum_1^m j^{2d-2}\right) & \text{for } d_0 \in \left(1, \frac{3}{2}\right) \end{cases} \\ &= \begin{cases} O\left(n^{d_0-\frac{3}{2}} m^{1-d_0}\right) & \text{for } d_0 \in \left(\frac{1}{2}, 1\right] \\ O\left(n^{d_0-\frac{3}{2}}\right) & \text{for } d_0 \in \left(1, \frac{3}{2}\right) \end{cases} \\ &= o(1), \end{aligned}$$

because $n^{d_0-\frac{3}{2}} m^{1-d_0} = (m/n)^{1-d_0} n^{-\frac{1}{2}} \to 0$. It follows that $\sup_{\Theta_1'} |(A(d) - A(1))/A(1)| \to_p 0$. For $d \in \Theta_2' = \left\{d : -\frac{1}{2} < d \le \frac{1}{2} + \Delta\right\}$, let B_k defined as

$$\begin{aligned} B_1 &= \mu^2 \frac{1}{m}\sum_{j=1}^{m}(a_j - 1)\frac{j^{2-2d_0}\lambda_j^{2d_0}}{|1-e^{i\lambda_j}|^2}\frac{n^{2d_0-2}}{2\pi}, \\ B_2 &= \frac{1}{m}\sum_{j=1}^{m}(a_j - 1) R_{j,n}(d_0). \end{aligned}$$

As $n \to \infty$, $B_1 > \delta > 0$ by (52). For $d_0 \in \left(\frac{1}{2}, 1\right]$, $E|B_2|$ is $o(1)$ because

$$\begin{aligned} \frac{1}{m}\sum_{j=1}^{m} a_j n^{d_0-\frac{3}{2}} j^{1-d_0} &= n^{d_0-\frac{3}{2}} \frac{p^{1-2\Delta}}{m} \sum_{j=1}^{p} j^{2\Delta-d_0} + n^{d_0-\frac{3}{2}} \frac{p^3}{m} \sum_{j=p+1}^{m} j^{-2-d_0} \\ &= O\left(n^{d_0-\frac{3}{2}} m^{1-d_0}\right), \\ \frac{1}{m}\sum_{j=1}^{m} n^{d_0-\frac{3}{2}} j^{1-d_0} &= O\left(n^{d_0-\frac{3}{2}} m^{1-d_0}\right). \end{aligned}$$

For $d_0 \in \left(1, \frac{3}{2}\right)$, $B_2 = o_p(1)$ is straightforward, giving the required result.

Next consider the case $d_0 = \frac{3}{2}$. Similarly to the above, define

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

and

$$\begin{aligned} A_1(d) &= \frac{2d-1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2} \frac{j^{2-2d_0}\lambda_j^{2d_0}}{|1-e^{i\lambda_j}|^2}\left|\mu\frac{\sqrt{n}}{\sqrt{2\pi}} + \frac{X_n}{\sqrt{2\pi n}}\right|^2, \\ A_2(d) &= \frac{2d-1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2d-2} R_{j,n}, \end{aligned}$$

where $E|R_{j,n}| = O\left(j^{-\frac{1}{2}}\right)$. From (37), we have

$$A_1(d) = \left|\mu + \frac{X_n}{n}\right|^2 + o_p(1).$$

$A_2(d) = o_p(1)$ is shown in (39). Thus $\sup_{\Theta_1'} |(A(d) - A(1))/A(1)| \to_p 0$. For $d \in \Theta_2'$, let B_k defined as

$$\begin{aligned} B_1 &= \frac{1}{m}\sum_{j=1}^{m}(a_j - 1)\frac{j^{2-2d_0}\lambda_j^{2d_0}}{|1-e^{i\lambda_j}|^2}\left|\mu\frac{\sqrt{n}}{\sqrt{2\pi}} + \frac{X_n}{\sqrt{2\pi n}}\right|^2, \\ B_2 &= \frac{1}{m}\sum_{j=1}^{m}(a_j - 1) R_{j,n}. \end{aligned}$$

It follows that $B_1 > \delta > 0$ as $n \to \infty$ and $B_2 = o_p(1)$ (see B_3 in proof of Theorem 3.3).

For $d_0 \in \left(\frac{3}{2}, 2\right)$, the proof follows from the proof of Theorem 3.3. ■

References

[1] Akonom, J. and C. Gourieroux (1987). "A Functional Limit Theorem for Fractional Processes." Technical Report #8801, CEPREMAP, Paris.

[2] Corbae, D., S. Ouliaris and P. C. B. Phillips (1999). "Band Spectral Regression with Trending Data." Yale University, mimeographed.

[3] Csörgő, M. and L. Horváth (1993). *Weighted Approximations in Probability and Statistics*. New York, Wiley.

[4] Kim, C. S. and P. C. B. Phillips (1999). "Log Periodogram Regression: The Nonstationary Case". Cowles Foundation Discussion Paper, Yale University (http://cowles.econ.yale.edu).

[5] Komlós, J., P. Major and G. Tusnády (1976). "An Approximation of Partial Sums of Independent RV's, and the Sample DF. II." *Zeitschrift für Wahrscheinlichkeitsheorie und verwandte Gebiete* **34**: 33-58.

[6] 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.

[7] Nelson, C. R. and C. I. Plosser (1982). "Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications." *Journal of Monetary Economics* **10**: 139-62.

[8] Phillips, P. C. B. (1999a). "Discrete Fourier Transforms of Fractional Processes." Cowles Foundation Discussion Paper #1243, Yale University (http://cowles.econ.yale.edu).

[9] Phillips, P. C. B. (1999b). "Unit Root Log Periodogram Regression." Cowles Foundation Discussion Paper #1244, Yale University (http://cowles.econ.yale.edu).

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

[11] Robinson, P. M. (1995). "Gaussian Semiparametric Estimation of Long Range Dependence." *Annals of Statistics* **23**: 1630-1661.

[12] Shimotsu, K. and P. C. B. Phillips (2000). "Modified Local Whittle Estimation of the Memory Parameter in the Nonstationary Case". Cowles Foundation Discussion Paper #1265, Yale University (http://cowles.econ.yale.edu).

[13] Velasco, C. (1999). "Gaussian Semiparametric Estimation of Non-Stationary Time Series." *Journal of Time Series Analysis* **20**: 87-127.

Chapter 4

Pooled Log Periodogram Regression

1 Introduction

The model we work with is a stationary Gaussian long-memory process X_t whose spectral density has the form

$$f_{XX}(\lambda) = \left|1 - e^{i\lambda}\right|^{-2d} f_{uu}(\lambda), \tag{1}$$

where $-1/2 < d < 1/2$ and $f_{uu}(\lambda)$ is a symmetric, periodic (with period 2π), positive, and continuous function bounded above and away from zero with a finite third derivative. Our objective is to estimate the parameter d in (1), which governs the long-memory property of X_t. The time domain version of (1) has the form $(1-L)^d X_t = u_t$, where u_t is a covariance stationary time series with spectral density $f_{uu}(\lambda)$. It is often preferable to leave the precise generating mechanism of u_t unspecified, so that the treatment of u_t is nonparametric. The estimation of d then falls within the framework of semiparametric methods. The most common estimator for d in this framework is provided by log periodogram regression, which was proposed by Geweke and Porter-Hudak (1983) and is sometimes called the GPH estimator. Rigorous analysis by Künsch (1986), Robinson (1995), and, most recently, Hurvich, Deo, and Brodsky (1998) followed the earlier work and established asymptotic properties of the estimator, including consistency and asymptotic normality and an optimal formula for the choice of the number of periodogram ordinates used in the regression. There is now a large and growing literature on the subject, the estimator is commonly used in empirical work, especially in economics, and it offers the computational convenience of least squares regression.

In view of (1), we have the following relation between the spectral density of X_t and u_t in logarithmic form

$$\ln(f_{XX}(\lambda)) = -2d \ln\left|1 - e^{i\lambda}\right| + \ln(f_{uu}(\lambda)).$$

Using periodogram ordinates in place of the actual spectra and evaluating these at the fundamental frequencies $\lambda_s = \frac{2\pi s}{n}, s = 1, ..., n-1$ leads to the 'regression' relationship

$$\ln\left(I_{XX}\left(\lambda_s\right)\right) = -2d\ln\left|1-e^{i\lambda_s}\right| + \ln\left(f_{uu}\left(\lambda_s\right)\right) + U\left(\lambda_s\right), \tag{2}$$

where $I_{XX}(\lambda_s) = w_X(\lambda_s)w_X(\lambda_s)^*$ and $w_X(\lambda_s)$ is the discrete Fourier transform $w_X(\lambda_s) = \frac{1}{\sqrt{2\pi n}}\sum_{t=1}^{n} X_t e^{it\lambda_s}$. The error in (2) is

$$U\left(\lambda_s\right) = \ln\left[\frac{I_{XX}\left(\lambda_s\right)}{f_{XX}\left(\lambda_s\right)}\right].$$

By virtue of the continuity of f_{uu}, $f_{uu}\left(\lambda_s\right)$ is effectively constant for frequencies in a shrinking band around the origin. This motivates the log-periodogram regression estimator of d, which is based on a linear least squares regression over frequencies $s = \ell + 1, ..., m$ (with ℓ a trimming number and m a truncation number) leading to

$$\ln\left(I_{XX}\left(\lambda_s\right)\right) = -2\widehat{d}(\ell)\ln\left|1-e^{i\lambda_s}\right| + \widehat{\mu} + \text{error} \tag{3}$$

and a class of estimates $\widehat{d}(\ell)$ and $\widehat{\mu}$ that depend on a subset of $m - \ell$ frequencies. Under the rate condition $\frac{\ell}{m} + \frac{m}{n} \longrightarrow 0$, the regression effectively uses $O(m)$ periodogram ordinates as $n \longrightarrow \infty$.

The heuristic motivation for this regression is based on the idea that the errors $U\left(\lambda_s\right)$ in (2) would be asymptotically independent across frequencies if the spectrum $f_{XX}(\lambda)$ were bounded. But, in the present case that is not so, and the errors $U(\lambda_s)$ are asymptotically correlated as shown by Künsch (1986), a feature that suggests the trimming of some (ℓ) observations away from the origin. Robinson (1995) proved that $\widehat{d}(\ell)$ is consistent and asymptotically normally distributed under some additional conditions on ℓ, m, and n. Hurvich, Deo, and Brodsky (1998) derived the asymptotic bias, variance, and the mean squared error of $\widehat{d}(0)$, the original GPH estimator, and showed under some stronger conditions on m, n, and $f_{uu}(\lambda)$ that

$$\sqrt{m}\left(\widehat{d}(0) - d\right) \rightarrow_d N\left(0, \frac{\pi^2}{24}\right).$$

The GPH estimator achieves consistency and asymptotic normality by using only m periodogram ordinates at frequencies $\lambda_s = 2\pi/n, ..., 2\pi m/n$ with $m/n \rightarrow 0$. The truncation at λ_m implies that as n increases, the estimator uses a smaller and smaller proportion of the full frequency band $[0, \pi]$, so that the effective band shrinks to the origin. The shrinking process is deliberate in the design of the GPH estimator because, given the nonparametric specification of $f_{uu}(\lambda)$, it is natural to confine attention in the regression to an immediate neighbourhood of the origin $\lambda \sim 0$, because in this case (1) has the simpler asymptotic form $f_{XX}(\lambda) \sim \lambda^{-2d}G$ as $\lambda \to 0+$, with $G = f_{uu}(0)$ constant. However, as is apparent from (2), the periodogram at higher frequencies λ_s $(s = m+1, .., [n/2])$ continues to contain some information about d, although the intercept involves $f_{uu}(\lambda_s)$ and will now vary over frequency bands to the extent that $f_{uu}(\lambda)$ is not constant. This intuition indicates that conventional log periodogram regression may discard some information in the data and gains may be achieved by using more frequency bands while at the same time allowing for variation in $f_{uu}(\lambda)$.

Accordingly, we now propose a new procedure for estimating d that builds on this idea. The method is a pooled log periodogram regression that is taken over the wider band of frequencies $\lambda_s = \frac{2\pi s}{n}, s = 1, ..., mL$ with $L \to \infty$ and $mL/n \to 0$. This method corrects for variation in the regression intercept by taking subgroup means in the regression, so that it allows that the error spectrum $f_{uu}(\lambda)$ may be nonconstant across bands. The new estimator treats $\ln(f_{uu}(\lambda_s))$ in (2) as an infinite dimensional nuisance parameter appearing in the regression intercept. The approach taken is then analogous to the treatment of fixed effects in panel data regression. The estimator of d pools the information about d obtained within each (shrinking) band over which the error spectrum is effectively constant as $n \to \infty$. We therefore call the new estimator a pooled log periodogram regression estimator.

The pooled estimator is shown to be consistent and asymptotically normally distributed. The pooled estimator has a smaller asymptotic variance than the GPH estimator, reflecting the greater number of periodogram ordinates used in the regression, but it also has larger asymptotic bias because of the nonconstancy of $f_{uu}(\lambda)$. Simulations show that in finite samples the pooled estimator performs substantially better than the GPH estimator when $f_{uu}(\lambda)$ has spectral peaks near the origin. On the other hand, the pooled estimator generally performs worse than the GPH estimator when $f_{uu}(\lambda)$ changes monotonically from $\lambda = 0$ to $\lambda = \pi$, although the difference is small.

This chapter is organized as follows. The new estimator is constructed in Section 2.

Section 3 gives assumptions and derives some preliminary asymptotic results. Section 4 proves consistency of the pooled estimator and derives its asymptotic mean squared error. Section 5 demonstrates asymptotic normality. Section 6 discusses the simulation results and gives an empirical illustration. Proofs are collected in Section 7.

2 Pooling Log Periodogram Ordinates in Regression

The idea of pooling ordinates in log periodogram regression can be explained as follows. First, we use an alternate form of the log periodogram representation, viz.

$$\begin{aligned} \ln\left(I_{XX}\left(\lambda_s\right)\right) &= \ln\left(f_{uu}\left(\lambda_s\right)\right) + \ln\left|1-e^{i\lambda_s}\right|^{-2d} + \ln\left(\frac{I_{XX}\left(\lambda_s\right)}{f_{XX}\left(\lambda_s\right)}\right) \\ &= \ln\left(f_{uu}\left(\omega_j\right)\right) - 2d\ln\left|1-e^{i\lambda_s}\right| + \ln\left(\frac{I_{XX}\left(\lambda_s\right)}{f_{XX}\left(\lambda_s\right)}\right) + \ln\left(\frac{f_{uu}\left(\lambda_s\right)}{f_{uu}\left(\omega_j\right)}\right), \end{aligned} \tag{4}$$

which allows for periodogram ordinates λ_s in the neighbourhood of a set of frequencies ω_j for $j = 0, 1, ...M-1$, where M is a parameter that determines the total number of distinct bands.

The implied log periodogram relation is now

$$\ln\left(I_{XX}\left(\lambda_s\right)\right) = \ln\left(f_{uu}\left(\omega_j\right)\right) - 2d\ln\left|1-e^{i\lambda_s}\right| + V\left(\lambda_s\right), \quad \lambda_s \in B_j \tag{5}$$

where

$$V\left(\lambda_s\right) = \ln\left(\frac{I_{XX}\left(\lambda_s\right)}{f_{XX}\left(\lambda_s\right)}\right) + \ln\left(\frac{f_{uu}\left(\lambda_s\right)}{f_{uu}\left(\omega_j\right)}\right),$$

and

$$B_j = \begin{cases} \left\{\lambda_s | \omega_j - \frac{\pi}{2M} < \lambda_s \le \omega_j + \frac{\pi}{2M}\right\}, & \omega_j = \frac{(2j+1)\pi}{2M}, j = 1, ..., M-1 \\ \left\{\lambda_s | 0 < \lambda_s \le \frac{\pi}{M}\right\}, & \omega_0 = 0, j = 0 \end{cases}$$

are the frequency bands, which are of width $\frac{\pi}{M}$. We compute the regressor sequence in (5) using $\left|1-e^{i\lambda_s}\right|^2 = 4\sin^2\left(\frac{\lambda_s}{2}\right)$ and do not use the conventional replacement $\left|1-e^{i\lambda_s}\right|^2 \sim \lambda^2$, which is appropriate only for λ_s in the vicinity of the zero frequency.

We propose to estimate the parameter d in (5) by linear regression using $(L+1)$ bands $B_0, ..., B_L$ where L is a number such that $L \to \infty$ and $L/M \to 0$. Thereby we remove the intercept in the regression by pooling observations over bands to fit d. Write (5) as

$$Y_{sj} = \mu_j + dX_{sj} + \eta_{sj} + \varepsilon_{sj} \quad s = 1, ..., m; j = 0, 1, ..., L \tag{6}$$

with

$$\begin{aligned}
Y_{sj} &= \ln\left(I_{XX}\left(\lambda_s\right)\right), \quad \lambda_s \in B_j \\
X_{sj} &= -2\ln\left|1-e^{i\lambda_s}\right| = -\ln\left[4\sin^2\left(\frac{\lambda_s}{2}\right)\right] = -\ln\left(2-2\cos\lambda_s\right), \quad \lambda_s \in B_j \\
\eta_{sj} &= \ln\left(\frac{f_{uu}\left(\lambda_s\right)}{f_{uu}\left(\omega_j\right)}\right) = \ln f_{uu}\left(\lambda_s\right) - \ln f_{uu}\left(\omega_j\right), \quad \lambda_s \in B_j \\
\varepsilon_{sj} &= \ln\left(\frac{I_{XX}\left(\lambda_s\right)}{f_{XX}\left(\lambda_s\right)}\right) - \psi(1), \quad \lambda_s \in B_j \\
\mu_j &= \ln\left(f_{uu}\left(\omega_j\right)\right) + \psi(1),
\end{aligned}$$

where $\psi(1) = \Gamma'(1) = -\gamma$ and $\gamma = 0.57721566...$ is Euler's constant.

The pooled estimator $\widehat{d}$ is given by the formula

$$\widehat{d} = \frac{\sum_{j=0}^{L}\sum_{\{s:\lambda_s\in B_j\}}\left(Y_{sj}-\overline{Y}_{\cdot j}\right)\left(X_{sj}-\overline{X}_{\cdot j}\right)}{\sum_{j=0}^{L}\sum_{\{s:\lambda_s\in B_j\}}\left(X_{sj}-\overline{X}_{\cdot j}\right)^2} = \frac{\sum_{j=0}^{L}\sum_{\{s:\lambda_s\in B_j\}}Y_{sj}\left(X_{sj}-\overline{X}_{\cdot j}\right)}{\sum_{j=0}^{L}\sum_{\{s:\lambda_s\in B_j\}}\left(X_{sj}-\overline{X}_{\cdot j}\right)^2}, \tag{7}$$

where

$$\begin{aligned}
\overline{Y}_{\cdot j} &= \frac{1}{m}\sum_{\{s:\lambda_s\in B_j\}} Y_{sj} = \frac{1}{m}\sum_{\{s:\lambda_s\in B_j\}} \ln\left(I_{XX}\left(\lambda_s\right)\right), \\
\overline{X}_{\cdot j} &= \frac{1}{m}\sum_{\{s:\lambda_s\in B_j\}} X_{sj} = -\frac{1}{m}\sum_{\{s:\lambda_s\in B_j\}} \ln\left[4\sin^2\left(\frac{\lambda_s}{2}\right)\right].
\end{aligned}$$

Note that the estimator $\widehat{d}$ uses data over an increasing number of frequency bands, not just those frequencies in B_0. The estimator still uses frequencies only in the vicinity of the origin because $mL/n \to 0$. In other words, the pooled estimator retains semiparametric nature of the log periodogram regression while using increasing number of bands. Subgroup means are subtracted in order to allow for the fact that the intercept μ_j may change over frequency bands $j = 0, ..., L$.

Combining equations (6) and (7) gives the estimation error

$$\widehat{d} - d = \frac{\sum_{j=0}^{L}\sum_{\{s:\lambda_s\in B_j\}}\eta_{sj}\left(X_{sj}-\overline{X}_{\cdot j}\right) + \sum_{j=0}^{L}\sum_{\{s:\lambda_s\in B_j\}}\varepsilon_{sj}\left(X_{sj}-\overline{X}_{\cdot j}\right)}{\sum_{j=0}^{L}\sum_{\{s:\lambda_s\in B_j\}}\left(X_{sj}-\overline{X}_{\cdot j}\right)^2}. \tag{8}$$

The idea of pooling ordinates over the bands B_j while allowing for variation in the spectrum across bands can be applied to other semiparametric estimators of the long memory parameter d. In particular, it is readily implemented in the case of the local Whittle estimator suggested by Künsch (1986) to give a pooled Whittle estimator. Our attention in the present chapter, however, will be confined to the pooled log periodogram procedure.

3 Assumptions and Asymptotic Results

To establish a limit theory for the pooled estimator, many of the results in Robinson (1995) and Hurvich, Deo, and Brodsky (1998) are relevant, and our approach draws substantially on their earlier work. We start by introducing the following assumptions.

Assumption 1 $m \to \infty$, $n \to \infty$, $\frac{m}{n} = \frac{1}{2M} \to 0$.

Assumption 2 $\frac{M}{m} + \frac{m \ln m}{n} + \frac{\ln^2 n}{m} \to 0$.

Assumption 3 $f'_{uu}(0) = 0$, $f_{uu}(\omega) > B_0 > 0, |f'_{uu}(\omega)| < B_1 < \infty, |f''_{uu}(\omega)| < B_2 < \infty$, $|f'''_{uu}(\omega)| < B_3 < \infty\ \forall \omega \in [0, \pi]$.

3.1 Remark In what follows, it will be taken as a convention that $n = 2mM$ holds exactly with both m and M integers (so that n is even). The convention is convenient, but not essential in what follows, and M is simply defined by the ratio $M = \frac{n}{2m}$ when n is odd. Any particular choice of m and expansion rate for m affect the bandwidth $\frac{\pi}{M}$, its rate of contraction, and the number of bands in the regression. The rate condition in Assumption 2 controls the relative rates at which m, M, and $n \to \infty$. Assumption 3 implies that $f_{uu}(\omega)$ is bounded away from zero and smooth with finite third derivative, much as in Hurvich et al. (1998).

3.2 Lemma *For a number ℓ such that $\ell \to \infty$ and $\ell^5/M^4 \to 0$, the following results hold:*

(a) $\sum_{\{s:\lambda_s \in B_0\}} \left(X_{sj} - \overline{X}_{\cdot j}\right)^2 = 4m + o(m)$.

(b) $\sum_{j=1}^{\ell} \sum_{\{s:\lambda_s \in B_j\}} \left(X_{sj} - \overline{X}_{\cdot j}\right)^2 = 4m\Xi + o(m)$,

where $\Xi = \sum_{j=1}^{\infty} \left[-j(j+1)(\ln(j+1) - \ln j)^2 + 1\right] \doteqdot 0.0803$.

3.3 Lemma *For a number ℓ such that $\ell/M \to 0$,*

$$\sum_{j=0}^{\ell} \sum_{\{s:\lambda_s \in B_j\}} \eta_{sj} \left(X_{sj} - \overline{X}_{\cdot j}\right) = -\frac{2\pi^2}{3} \frac{f''_{uu}(0)}{f_{uu}(0)} \frac{m^3 \ell}{n^2} + o\left(\frac{m^3 \ell}{n^2}\right).$$

3.4 Remarks

(a) The GPH estimator involves regression only over the band B_0. In place of (8) it satisfies

$$\widehat{d}_{GPH} - d = \frac{\sum_{\{s:\lambda_s \in B_0\}} \eta_{sj} \left(X_{sj} - \overline{X}_{.j}\right) + \sum_{\{s:\lambda_s \in B_0\}} \varepsilon_{sj} \left(X_{sj} - \overline{X}_{.j}\right)}{\sum_{\{s:\lambda_s \in B_0\}} \left(X_{sj} - \overline{X}_{.j}\right)^2}. \tag{9}$$

The denominator of $\widehat{d} - d$

$$\sum_{j=0}^{L} \sum_{\{s:\lambda_s \in B_j\}} \left(X_{sj} - \overline{X}_{.j}\right)^2 = 4\left(1 + \Xi\right) m + o\left(m\right), \tag{10}$$

is larger than the denominator of the GPH estimator

$$\sum_{\{s:\lambda_s \in B_0\}} \left(X_{sj} - \overline{X}_{.j}\right)^2 = 4m + o\left(m\right),$$

as $m \to \infty$. Roughly speaking, the denominator measures the excitation level of the regressors and indicates the information content in the regressors about the coefficient (d) in the regression (6). From (10), it is apparent that this information content is larger when the frequency band $B_0, ..., B_L$ is employed than when the immediate band around the zero frequency B_0 is used. As we will see, this increase in information content reduces the asymptotic variance of the pooled estimator relative to that of the GPH estimator.

(b) The optimal expansion rate of m for the GPH estimator is known from Hurvich et al. (1998) to be $O(n^{\frac{4}{5}})$, whereas the optimal rate for the pooled estimator is, as we will see later, $O(n^{\frac{4}{5}} L^{-\frac{2}{5}})$. Thus, if optimal rates were chosen the variance gains of the pooled estimator would vanish as $n \to \infty$. Issues of a joint optimal choice of m and L have not been considered by the authors.

(c) Lemma 3.3 shows that the nonrandom bias of the pooled estimator that arises from the presence of the first term in the numerator of (8) is $O\left(\frac{m^2 L}{n^2}\right)$ when $f''_{uu}(\lambda) \neq 0$ and hence this bias tends to zero as $n \to \infty$.

4 MSE and Consistency

We start with the following theorem, which is a variant of theorem 2 in Robinson (1995).

4.1 Theorem *Let Assumption 3 hold. Then, for any sequences of positive integers $j = j(n)$ and $k = k(n)$ such that $0 < k < j < n/2$, as $n \to \infty$*

(a) $E\left[w(\lambda_j)\overline{w}(\lambda_j)/f_{XX}(\lambda_j)\right] = 1 + O\left(j^{-1}\ln n\right)$,

(b) $E\left[w(\lambda_j)w(\lambda_j)/f_{XX}(\lambda_j)\right] = O\left(j^{-1}\ln n\right)$,

(c) $E\left[w(\lambda_j)\overline{w}(\lambda_k)/\left(f_{XX}(\lambda_j)f_{XX}(\lambda_k)\right)^{1/2}\right] = O\left(k^{-1}\ln n\right)$,

(d) $E\left[w(\lambda_j)w(\lambda_k)/\left(f_{XX}(\lambda_j)f_{XX}(\lambda_k)\right)^{1/2}\right] = O\left(k^{-1}\ln n\right)$,

where $\lambda_j = 2\pi j/n$ *and* $w(\lambda) = (2\pi n)^{-1/2}\sum_1^n X_t e^{it\lambda}$.

Define the quantities

$$A_s = \frac{1}{\sqrt{2\pi n}}\sum_{t=0}^{n-1} X_t \cos\lambda_s t, \quad C_s = \frac{1}{\sqrt{2\pi n}}\sum_{t=0}^{n-1} X_t \sin\lambda_s t,$$

and

$$\begin{aligned}\alpha_{st} &= \max\left\{\left|\operatorname{cov}\left(A_s/f_{XX}^{1/2}(\lambda_s), A_t/f_{XX}^{1/2}(\lambda_t)\right)\right|, \left|\operatorname{cov}\left(A_s/f_{XX}^{1/2}(\lambda_s), C_t/f_{XX}^{1/2}(\lambda_t)\right)\right|,\right.\\ &\left.\left|\operatorname{cov}\left(C_s/f_{XX}^{1/2}(\lambda_s), A_t/f_{XX}^{1/2}(\lambda_t)\right)\right|, \left|\operatorname{cov}\left(C_s/f_{XX}^{1/2}(\lambda_s), C_t/f_{XX}^{1/2}(\lambda_t)\right)\right|\right\}.\end{aligned}$$

Theorem 4.1, combined with the Gaussianity of X_t, enables us to evaluate the means, variances, and covariances of

$$\varepsilon_{sj} = \ln\left(\frac{I_{XX}(\lambda_s)}{f_{XX}(\lambda_s)}\right) - \psi(1) = \ln\left(\frac{A_j^2}{f_{XX}(\lambda_j)} + \frac{C_j^2}{f_{XX}(\lambda_j)}\right) - \psi(1),$$

because the distribution of the normal vector

$$\left(A_j/\left(f_{XX}(\lambda_j)\right)^{1/2},\ C_j/\left(f_{XX}(\lambda_j)\right)^{1/2},\ A_k/\left(f_{XX}(\lambda_k)\right)^{1/2},\ C_k/\left(f_{XX}(\lambda_k)\right)^{1/2}\right)$$

is solely determined by its covariance matrix. In particular, it follows directly from theorem 4.1 that

$$\begin{aligned}E\left(\frac{A_j^2}{f_{XX}(\lambda_j)}\right) &= \frac{1}{2} + O\left(\frac{\ln n}{j}\right), \quad E\left(\frac{C_j^2}{f_{XX}(\lambda_j)}\right) = \frac{1}{2} + O\left(\frac{\ln n}{j}\right),\\ E\left(\frac{A_jC_j}{f_{XX}(\lambda_j)}\right) &= O\left(\frac{\ln n}{j}\right),\end{aligned}$$

uniformly for $1 \le t < s < n/2$.

The following lemma is also a consequence of theorem 4.1. Because the proofs of its four component parts are very similar to those of lemmas 2,3,6, and 7 in Hurvich, Deo, and Brodsky (1998), they are omitted here.

4.2 Lemma

(a) $\alpha_{st} = O(\ln n/t)$, *uniformly for* $1 \le t < s < n/2$.

(b) $Cov(\varepsilon_{sj}, \varepsilon_{tk}) = O(\alpha_{st}^2)$, *uniformly for* $\ln^2 n \le t < s < n/2$.

(c) $E(\varepsilon_{sj}) = O(\ln n/s)$, *uniformly for* $\ln^2 n \le s < n/2$.

(d) $Var(\varepsilon_{sj}) = \pi^2/6 + O(\ln n/s)$, *uniformly for* $\ln^2 n \le s < n/2$.

Lemma 4.2 shows that $\varepsilon_{sj}, \varepsilon_{tk}$ are asymptotically mean zero and independent and identically distributed for $\ln^2 n \le s$, $t \le n/2$. We now proceed to derive asymptotic representations of the bias and variances and covariances of $\sum_{\{s:\lambda_s \in B_j\}} \varepsilon_{sj}\left(X_{sj} - \overline{X}_{.j}\right)$ over different frequency bands.

4.3 Lemma (Bias) *For a number* $\ell < M$,

$$\sum_{j=1}^{\ell} \sum_{\{s:\lambda_s \in B_j\}} E(\varepsilon_{sj})\left(X_{sj} - \overline{X}_{.j}\right) = O(\ln m).$$

4.4 Lemma (Variance and covariances between bands B_j, B_k, $1 \le k < j < M$)

For $1 \le k < j < M$,

$$Var\left[\sum_{\{s:\lambda_s \in B_j\}} \varepsilon_{sj}\left(X_{sj} - \overline{X}_{.j}\right)\right] = \frac{\pi^2}{6} \sum_{\{s:\lambda_s \in B_j\}} \left(X_{sj} - \overline{X}_{.j}\right)^2 + O\left(\frac{\ln^2 m}{j^3}\right),$$

and

$$Cov\left[\sum_{\{s:\lambda_s \in B_j\}} \varepsilon_{sj}\left(X_{sj} - \overline{X}_{.j}\right), \sum_{\{t:\lambda_t \in B_k\}} \varepsilon_{tk}\left(X_{tk} - \overline{X}_{.k}\right)\right] = O\left(\frac{\ln^2 m}{jk^2}\right).$$

4.5 Lemma (Covariances between bands B_j, B_0, $1 \le j < M$)

$$Cov\left[\sum_{\{s:\lambda_s \in B_j\}} \varepsilon_{sj}\left(X_{sj} - \overline{X}_{.j}\right), \sum_{\{t:\lambda_t \in B_0\}} \varepsilon_{t0}\left(X_{t0} - \overline{X}_{.0}\right)\right] = O\left(\frac{m^{1/2}\ln^7 m}{j}\right) + O\left(\frac{m}{j\ln^3 m}\right).$$

4.6 Lemma (Asymptotic variance)

$$Var\left[\sum_{j=0}^{\ell}\sum_{\{s:\lambda_s\in B_j\}}\varepsilon_{sj}\left(X_{sj}-\overline{X}_{\cdot j}\right)\right]=\frac{4\pi^2 m}{6}(1+\Xi)+o(m).$$

4.7 Remark We can express $\widehat{d}$ as

$$\begin{aligned}
\widehat{d} &= \frac{\sum_{j=0}^{L}\sum_{\{s:\lambda_s\in B_j\}}\ln\left(I_{XX}(\lambda_s)\right)\left(X_{sj}-\overline{X}_{\cdot j}\right)}{\sum_{j=0}^{L}\sum_{\{s:\lambda_s\in B_j\}}\left(X_{sj}-\overline{X}_{\cdot j}\right)^2}\\
&= \sum_{j=0}^{L}\left(\frac{\sum_{\{s:\lambda_s\in B_j\}}\left(X_{sj}-\overline{X}_{\cdot j}\right)^2}{\sum_{j=0}^{\ell}\sum_{\{s:\lambda_s\in B_j\}}\left(X_{sj}-\overline{X}_{\cdot j}\right)^2}\frac{\sum_{\{s:\lambda_s\in B_j\}}\ln\left(I_{XX}(\lambda_s)\right)\left(X_{sj}-\overline{X}_{\cdot j}\right)}{\sum_{\{s:\lambda_s\in B_j\}}\left(X_{sj}-\overline{X}_{\cdot j}\right)^2}\right)\\
&= \sum_{j=0}^{L}\left(\frac{\sum_{\{s:\lambda_s\in B_j\}}\left(X_{sj}-\overline{X}_{\cdot j}\right)^2}{\sum_{j=0}^{L}\sum_{\{s:\lambda_s\in B_j\}}\left(X_{sj}-\overline{X}_{\cdot j}\right)^2}\right)\widehat{d}_j,
\end{aligned}$$

where $\widehat{d}_j$ is the estimator of d obtained from using the band B_j only. Lemmas 4.4, 4.5, and 4.6 imply that the $\widehat{d}_j$ are asymptotically independent. Therefore, $\widehat{d}$ is a weighted average of asymptotically independent component estimators $\widehat{d}_j$, and we may therefore anticipate that the variance of $\widehat{d}$ is smaller than that of $\widehat{d}_0\equiv\widehat{d}_{GPH}$.

Specifically, lemmas 3.2, 3.3, 4.3, and 4.6 yield an asymptotic representation of the mean squared error of $\widehat{d}$, which is given in the next theorem.

4.8 Theorem *Let assumptions 1-3 hold. Then*

$$E\left(\widehat{d}-d\right) = -\frac{\pi^2}{6(1+\Xi)}\frac{f''_{uu}(0)}{f_{uu}(0)}\frac{m^2L}{n^2}+o\left(\frac{m^2L}{n^2}\right)+O\left(\frac{\ln^3 m}{m}\right), \tag{11}$$

$$Var\left(\widehat{d}\right) = \frac{\pi^2}{24(1+\Xi)m}+o\left(\frac{1}{m}\right),$$

$$\begin{aligned}MSE\left(\widehat{d}\right) &= \frac{\pi^4}{36(1+\Xi)^2}\left\{\frac{f''_{uu}(0)}{f_{uu}(0)}\right\}^2\frac{m^4L^2}{n^4}+\frac{\pi^2}{24(1+\Xi)m}\\ &\quad+o\left(\frac{m^4L^2}{n^4}\right)+O\left(\frac{\ln^6 m}{m^2}\right)+O\left(\frac{mL\ln^3 m}{n^2}\right)+o\left(\frac{1}{m}\right).\end{aligned} \tag{12}$$

4.9 Remarks

(a) The mean squared error tends to zero as $n\to\infty$ and $\widehat{d}$ is consistent.

(b) Hurvich, Deo, and Brodsky (1998) derive the following formulae for the asymptotic bias and variance of the GPH estimator:

$$E\left(\widehat{d}_{GPH}-d\right) = -\frac{2\pi^2}{9}\frac{f''_{uu}(0)}{f_{uu}(0)}\frac{m^2}{n^2}+O\left(\frac{\ln^3 m}{m}\right)+o\left(\frac{m^2}{n^2}\right),$$

$$Var\left(\widehat{d}_{GPH}\right) = \frac{\pi^2}{24m} + o\left(\frac{1}{m}\right).$$

Compared with the GPH estimator, the pooled estimator has a larger bias but smaller variance. Which effect dominates in finite samples will depend on the sample size and the shape of the error spectrum $f_{uu}(\omega)$. In the extreme case where the error spectrum is constant, $\eta_{sj} = 0$, and both estimators are unbiased (the first term in the numerator of both (8) and (9) is zero).

5 Asymptotic Normality

To establish the asymptotic normality of $\widehat{d}$, we prove that the standardized quantity

$$m^{-1/2} \sum_{j=0}^{L} \sum_{\{s:\lambda_s \in B_j\}} \varepsilon_{sj} \left(X_{sj} - \overline{X}_{.j}\right) \tag{13}$$

with $L = O(\ln M)$ has a limiting normal distribution.

The following lemma gives the basis of the limiting distribution theory. Its proof draws heavily on the derivations in Theorem 3 of Robinson (1995) and applies the approach developed in that article to a case in which there are ℓm rather than m observations.

5.1 Lemma *Let $a_{kn} = a_k$ be a triangular array for which*

$$\max_k |a_k| = O(m), \qquad \sum_{k=1+m^{0.5+\delta}}^{\ell m} a_k^2 \sim m, \qquad \sum_{k=1+m^{0.5+\delta}}^{\ell m} |a_k|^p = O(m \ln \ell), \tag{14}$$

for all $p \geq 1$, and let ℓ be a number that satisfies $\ell \to \infty$, $\ell^2 m^2 m^{0.5+\Delta}/n^2 \to 0$ for some $0 < \Delta < \delta < 0.5$ and $\ln^K \ell / m^{\Delta} \to 0$ for any $K > 0$. Then,

$$\frac{1}{\sqrt{m}} \sum_{k=1+m^{0.5+\delta}}^{\ell m} a_k U_k \xrightarrow{d} N(0, \Omega) = N\left(0, \psi'(1)\right),$$

where

$$\begin{aligned} U_k &= \log\left[\left(v^R(\lambda_k)\right)^2 + \left(v^I(\lambda_k)\right)^2\right] - \psi(1) \\ &= \log\left(\frac{I_{XX}(\lambda_k)}{C_g \lambda_k^{-2d}}\right) - \psi(1) = \log\left(\frac{I_{XX}(\lambda_k)}{f_{uu}(0)\, \lambda_k^{-2d}}\right) - \psi(1). \end{aligned}$$

With this lemma in hand, we extract a limiting distributional result for the quantity (13). In particular, we have the following.

5.2 Lemma *Let assumptions 1-3 hold and additionally require that* $m = O\left(n^{\frac{4}{5}-\varepsilon}\right)$ *for some* $\varepsilon > 0$ *and* $L = O(\ln M)$. *Then, we have*

$$m^{-1/2}\sum_{j=0}^{L}\sum_{\{s:\lambda_s\in B_j\}}\varepsilon_{sj}\left(X_{sj}-\overline{X}_{\cdot j}\right)\xrightarrow{d}N\left(0,4\pi^2\left(1+\Xi\right)/6\right).$$

These preliminary results lead to the asymptotic normality of the estimator $\widehat{d}$.

5.3 Theorem *Let assumptions 1-3 hold. Moreover, if* $m = O\left(n^{\frac{4}{5}-\varepsilon}\right)$ *for some* $\varepsilon > 0$ *and* $L = O(\ln M)$, *we have*

$$m^{1/2}\left(\widehat{d}-d\right)\xrightarrow{d}N\left(0,\frac{\pi^2}{24\left(1+\Xi\right)}\right).$$

6 Simulations and Empirical Illustration

This section reports some simulations that were conducted to compare the finite sample performance of the two estimators $\widehat{d}_{GPH}$ and $\widehat{d}_{pooled}$. Because both estimators treat u_t nonparametrically, it seems desirable to examine their finite sample properties over models that allow for a variety of spectral shapes for $f_{uu}(\lambda)$. With this objective in mind, we used the following AR(2) generating mechanism for u_t

$$u_t - a_1u_{t-1} - a_2u_{t-2} = \varepsilon_t, \quad \varepsilon_t \sim iidN(0,1),$$

which permits a range of spectral shapes, including some with spectral peaks away from the origin. We generate the process $X_t = (1-L)^{-d}u_t$ by the algorithm of Davies and Harte (1987).

We set $m = n^{0.65}$ in the simulations. This amounts to using the frequency band $(0, 0.31\pi)$, $(0, 0.22\pi)$, and $(0, 0.09\pi)$ for sample size $n = 200$, 500, and 1000. We set $L = 2$ in the construction of the pooled estimator. Although the pooled estimator requires a slightly stronger condition on m for asymptotic normality than the GPH estimator (specifically, $m = O(n^{\frac{4}{5}-\varepsilon})$ rather than $m = o(n^{\frac{4}{5}})$), we use the same m for comparison because m is rarely chosen to be as large as $n^{\frac{4}{5}}$ in practice.

6.1 Simulations over the (a_1, a_2) plane

First, we report some comprehensive simulations over the (a_1, a_2) parameter space, so that the effect of spectral shape on performance can be assessed. We take the region of the (a_1, a_2) plane for which u_t is stationary and use a grid with a step size of 0.1 in this plane. The bias, variance, and mean squared error (MSE) were computed using 1,000 replications. Sample size and long memory parameter were chosen to be $n = 500$ and $d = 0.3$, respectively. A second experiment, reported below, looks at performance for different values of n.

For the GPH estimator, we used the regressor $X_s = -2\ln(\lambda_s)$, instead of the exact regressor $-\ln(4\sin^2(\lambda_s/2))$, as is common practice, and in our simulations the former regressor generally gave better results for the parameter values considered. The variances of the estimators were very similar and seemed to vary little across the different parameter values. Hence, most of the variation that appears in the MSE is due to differences in bias.

Figures 1 and 2 plot the MSE's. When a_1 and a_2 are close to the line $a_1 + a_2 = 1$, the MSE of both estimators becomes quite large. The MSE of the GPH estimator has a particularly large spike when a_1 is large and a_2 is small. The MSE of the pooled estimator also has a spike, although the magnitude of the spike is substantially smaller than that of the MSE of the GPH estimator. The MSE of the GPH estimator decreases monotonically as a_1 decreases, whereas the MSE of the pooled estimator has small bumps, especially when a_2 is small and negative.

To obtain a better idea of the differences between the two estimators, a contour plot of the MSE difference (MSE(GPH)−MSE(pooled)) is displayed in Figure 3. In general, the difference is small when it is negative, except in the area near the line $a_1 + a_2 = 1$. As expected from Figures 1 and 2, the difference is large when a_1 is large and a_2 is small. Figure 4 shows a contour plot of the logarithm of the relative efficiency ($= \log_2 MSE(GPH) - \log_2 MSE(pooled)$). The area of the (a_1, a_2) plane in which the MSE of the pooled estimator is smaller than that of the GPH estimator (i.e., the logarithm of the relative efficiency is greater than 0) is not very large. Because the pooled estimator has a larger MSE than the GPH estimator when the MSE of both estimators is small, the magnitudes of positive and negative relative efficiency are roughly equal to each other.

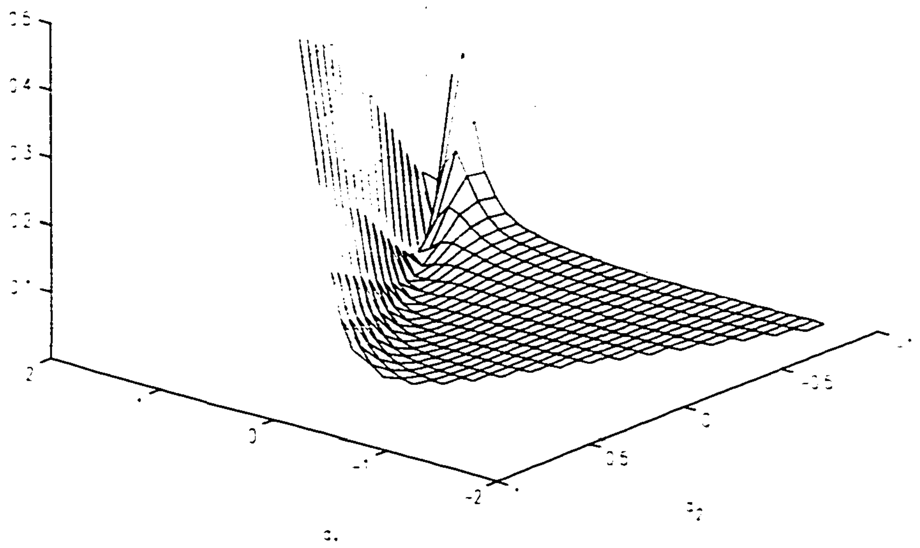

Figure 1: MSE of the GPH estimator ($d = 0.3, n = 500, m = 56$)

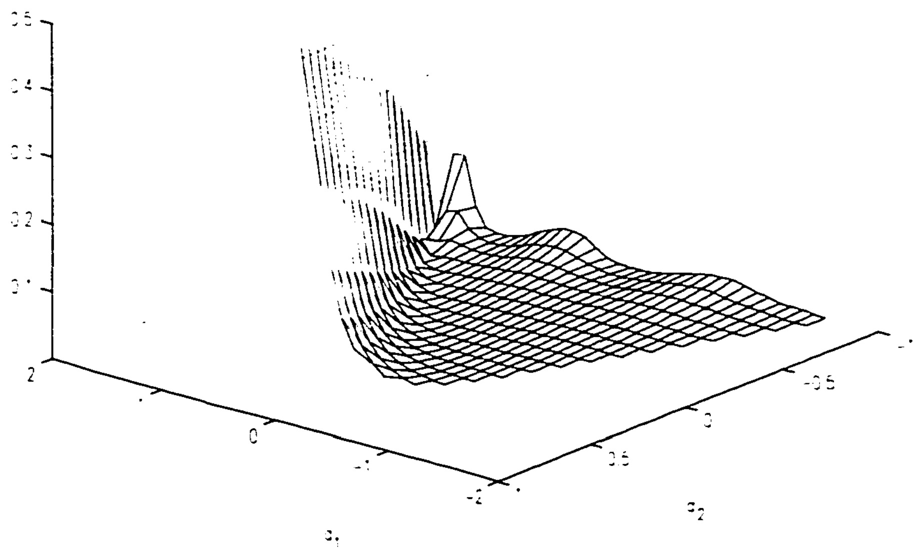

Figure 2: MSE of the pooled estimator ($d = 0.3, n = 500, m = 56$)

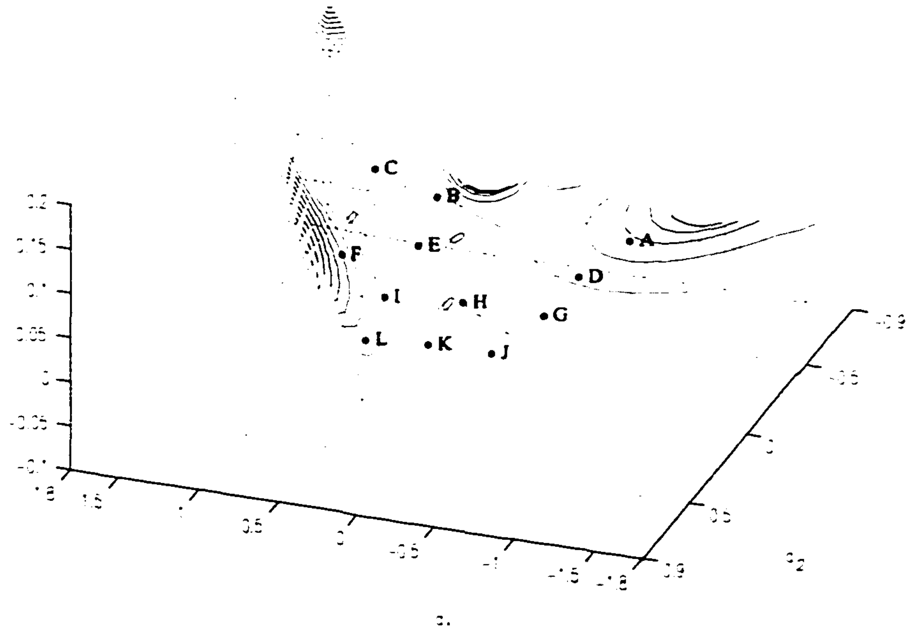

Figure 3: MSE(GPH)−MSE(pooled) ($d = 0.3, n = 500, m = 56$)

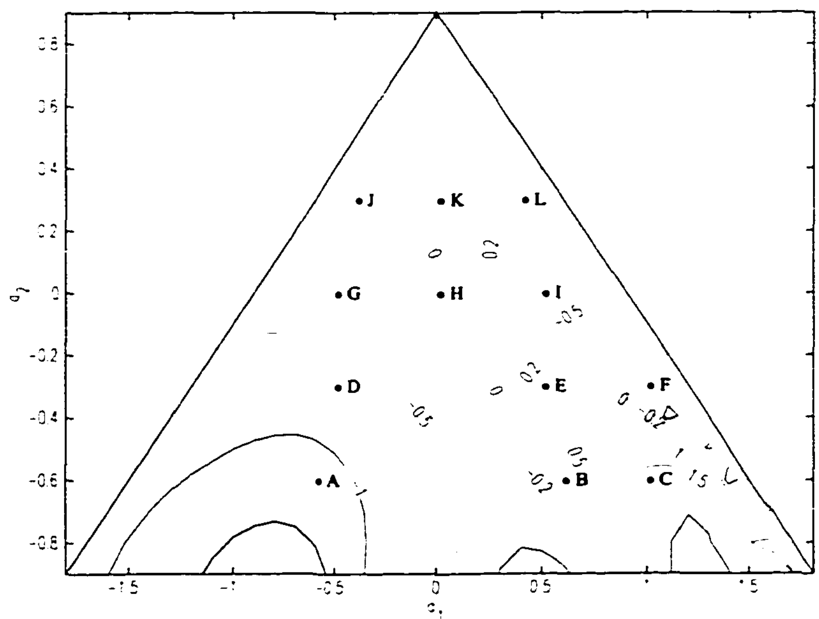

Figure 4: $\log_2 (MSE\,(GPH)\,/MSE\,(pooled))$ ($d = 0.3, n = 500, m = 56$)

Figure 5 plots the spectral density of u_t for some of the more important parameter combinations in Figures 3 and 4. The values of a_1 and a_2 at these points are given in Table 1. At the points where the pooled estimator has a substantially smaller MSE than the GPH estimator, such as at B, C, and E, the spectral density has a peak near the origin. At the points where the spectral density has a peak away from the origin or changes monotonically, such as at A, D, and I, the GPH estimator has a smaller MSE than the pooled estimator. At the points F and L, where both estimators have quite large MSE, the spectral density decreases sharply from the origin. Because the slope of the periodogram around the origin contains the strongest signal for both the GPH estimator and the pooled estimator, this result is hardly surprising, and neither procedure can be expected to work well. At point H, where u_t is a white noise, both estimators have similar MSE. The shape of the spectral density far away from the origin does not affect the MSE substantially, as indicated by the results on the points G and J.

Table 1. The values of a_1 and a_2 at the points A-L

	A	B	C	D	E	F	G	H	I	J	K	L
a_1	-0.6	0.6	1.0	-0.5	0.5	1.0	-0.5	0.0	0.5	-0.4	0.0	0.4
a_2	-0.6	-0.6	-0.6	-0.3	-0.3	-0.3	0.0	0.0	0.0	0.3	0.3	0.3

In sum, the pooled estimator has the advantage of being robust to the presence of a peak in the error spectrum $f_{uu}(\lambda)$ and produces substantial reductions in both bias and MSE when the peak is close to the regression frequency band. In such cases, of course, it is the peak in the short memory spectrum that exacerbates the bias in the GPH estimator.

6.2 Detailed simulation for several pairs of parameter values

For several points in Figures 3 and 4, we conducted a more detailed simulation covering different sample sizes[1]. Tables 2-4 show the simulation results for parameter combinations G, H, and I from Table 1, for which $a_2 = 0$ and u_t follows an AR(1) process. In general, for AR(1) processes with $a_1 \neq 0$, $\widehat{d}_{pooled}$ (the second row) is more biased than $\widehat{d}_{GPH}$ (the first row). The increase in the bias occurs because $f_{uu}(\lambda)$ is monotonically increasing (decreasing), and the bias effect of the nonzero slope of $f_{uu}(\lambda)$ on the estimator is accumulated across bands in the case of the pooled estimator $\widehat{d}_{pooled}$, as is apparent from the asymptotic

[1] We report only the case $d = 0.3$, because the results for different memory parameter values are very similar.

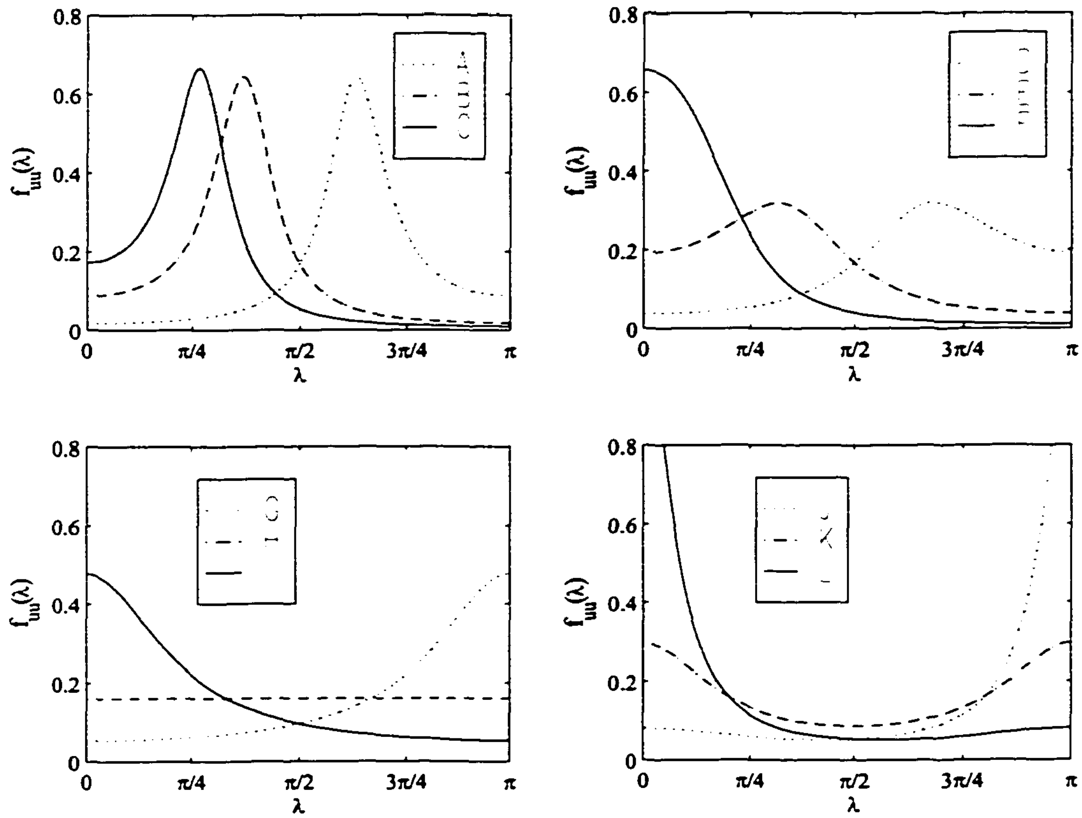

Figure 5: Spectral densities for several AR(2) processes

formula (11). The variance of the two estimators are almost equal. MSE($\widehat{d}_{pooled}$) is similar to MSE($\widehat{d}_{GPH}$) when $a_1 = 0$, whereas the pooled estimator has a larger MSE than $\widehat{d}_{GPH}$ when $a_1 = \pm 0.5$. The difference in the MSE is smaller when a_1 is negative because the slope of $f_{uu}(\lambda)$ changes primarily where λ is far from the origin, and the effect of the shape of $f_{uu}(\lambda)$ far away from the origin is smaller. Thus, for AR(1) errors it appears that the GPH estimator is generally better than the pooled estimator in finite samples. The difference in the MSE is small when $n = 1000$, however, because the frequency band $[0, \lambda_m]$ becomes narrow relative to $[0, 2\pi]$.

Tables 5-7 show simulation results for the parameter combinations A, B, and C, for which the spectral density of u_t has a peak at the frequency $\lambda_p = \arccos\left\{-a_1 \frac{1-a_2}{4a_2}\right\}$. The values of λ_p are 1.98 ($\simeq 0.63\pi$), 1.16 ($\simeq 0.37\pi$), and 0.84 ($\simeq 0.27\pi/$) at A, B, and C, respectively. The simulation results for these parameter combinations are very different from the case of AR(1) errors. Now, the performance of the estimators depends very much on the location of the peak in the spectrum $f_{uu}(\lambda)$ relative to the frequency band being used in the regression. When the peak in the spectrum is near the frequency band $[\lambda_1, \lambda_m]$, as it is for B and C with $n = 200$ and 500, the estimator $\widehat{d}_{GPH}$ appears to be severely biased, and the pooled estimator $\widehat{d}_{pooled}$ has much smaller bias than $\widehat{d}_{GPH}$ (see Tables 6 and 7). On the other hand, at A, the peak in the spectrum of $f_{uu}(\lambda)$ is far from the origin, and $f_{uu}(\lambda)$ is close to constant around the origin. In this case (see Table 5), the bias of $\widehat{d}_{GPH}$ is relatively small, and $\widehat{d}_{pooled}$ has a larger bias. In all cases, although the variance of $\widehat{d}_{pooled}$ is smaller than that of $\widehat{d}_{GPH}$, the difference is not substantial. In terms of MSE, at parameter combinations B and C and for $n = 200$ and 500, MSE($\widehat{d}_{pooled}$) is decisively smaller than MSE($\widehat{d}_{GPH}$) due to the bias reduction. However, at A, $\widehat{d}_{pooled}$ has a larger MSE than $\widehat{d}_{GPH}$ because of its larger bias. When $n = 1000$, the difference between the two estimators becomes smaller.

In the simulations above, the pooled estimator uses a total $m(L+1) = 3m$ frequencies, while the GPH estimator uses m frequencies. We examine the effects of using the wider frequency band on the GPH estimator and see how this affects the comparison of the two estimators. The third row of Tables 2-7 reports results for $\widehat{d}_{GPH}(3m)$, the GPH estimator using $3m$ frequencies. Using $3m$ frequencies in GPH leads to a dramatic increase in bias except for the parameter combinations H (where $f_{uu}(\lambda)$ is constant) and B with $m = 200$ and 500. The results in Tables 2, 4, 5, and 7 reveal that the GPH estimator based on the wider frequency band is very sensitive to the shape of the spectrum $f_{uu}(\lambda)$ around the origin and is generally much inferior to the pooled estimator. On the other hand, using only m frequencies, dividing the frequency band $[\lambda_1, \lambda_m]$ into bands and applying the pooled estimator does not provide a superior pooled estimator. The fourth row of Tables 2-7 shows results for $\widehat{d}_{pooled}(m)$, the pooled estimator that uses only m frequencies in total and two blocks, each block containing $[m/2]$ frequencies. Evidently, the increase in variance in this case more than offsets the reduction of bias.

In sum, the pooled estimator has advantages over the GPH estimator in finite samples, because the use of a wider frequency band ($m(L+1)$ rather than m) makes it less sensitive to the presence of peaks in the underlying spectral density $f_{uu}(\lambda)$. At the same time, it avoids the extremely large bias that is typical of the GPH estimator when a wide frequency band is employed. Therefore, it provides us with an alternate way of using a wider frequency band in log periodogram regression and a way to use more information, making the estimator more robust to various shapes in the short memory spectrum. In so doing, it can lead to substantial bias and MSE reductions when $f_{uu}(\lambda)$ has peaks that are close to the regression frequency band. On the other hand, it suffers from a mild bias increase when the error spectrum $f_{uu}(\lambda)$ changes monotonically, as it does in the case of AR(1) errors.

6.3 Empirical illustration

The methods were applied to US inflation series and stock returns. The inflation series constituted 624 observations of the monthly CPI inflation rate over the period 1947:1-1999:2; and the stock return series involved 3600 observations of the absolute value of returns on the daily S&P500 stock index from January 1979 to October 1992.[2] The first panel of Figure 6 graphs each series. The second panel of Figure 6 plots $\widehat{d}_{GPH}$ and $\widehat{d}_{pooled}$ for different values of m. (Specifically, $m = n^{0.5}, ..., n^{0.65}$ were used).

The value of the both estimator changes as m increases, although the pooled estimator

[2]The inflation series were computed as $X_t = 100\Delta[\log(x_t)]$, where x_t is the US monthly CPI, over the period from January 1947 to February 1999. The stock return series were computed as $X_t = 100|\Delta(\log(x_t))|$, where x_t is the S&P 500 stock price index from January 1979 to October 1992.

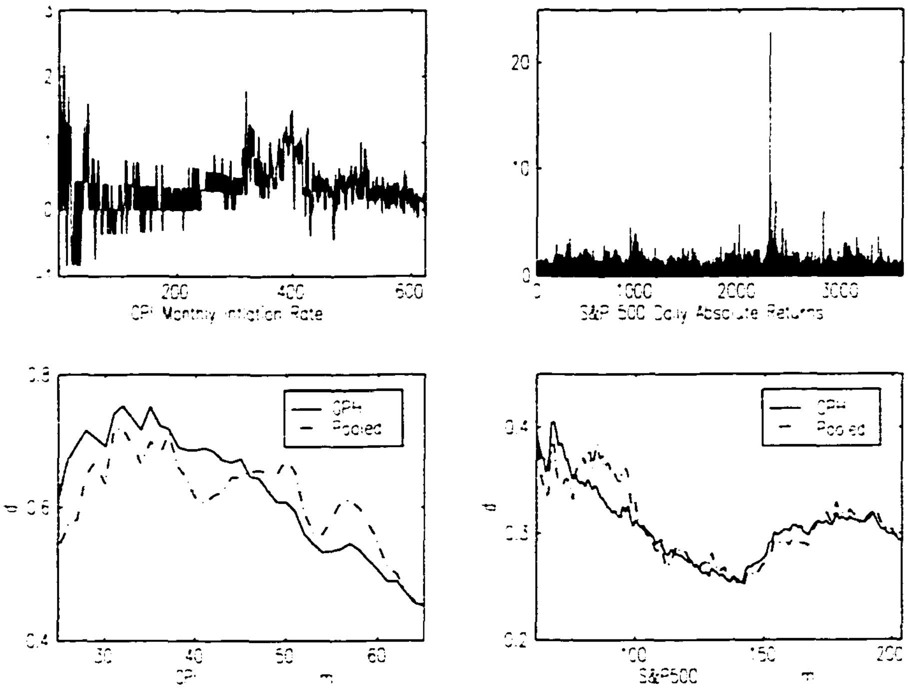

Figure 6: Inflation rate and stock return data and estimates of d

appears to have a less sharp peak. The estimates of the memory parameter for inflation are in the region $(0.4, 0.8)$, indicating marginal nonstationarity of inflation. Those for stock return magnitudes are around 0.3, indicating stationary long range dependence.

The first panel of Figure 7 shows the residual fractionally differenced series $\widehat{u}_t = (1-L)^{\widehat{d}} X_t$, where $\widehat{d}$ is the pooled estimate with $m = n^{0.55}$. The spectral density estimates of $\widehat{u}_t$ are displayed in the second panel of Figure 7, using $\widehat{d}_{GPH}$ and $\widehat{d}_{pooled}$ estimates calculated with $m = n^{0.55}$.

In both cases, the empirical estimates of the spectral density of u_t appear to have more power concentrated at higher frequencies. This is partly explained by the natural tendency of the GPH estimator to attribute power in the periodogram at lower frequencies to the long memory parameter d. The estimates of $f_{uu}(0)$ implied by the pooled estimator of d are higher than those obtained from the GPH estimator for the inflation series and lower for the stock return series. In other respects, the spectral densities estimates are very close.

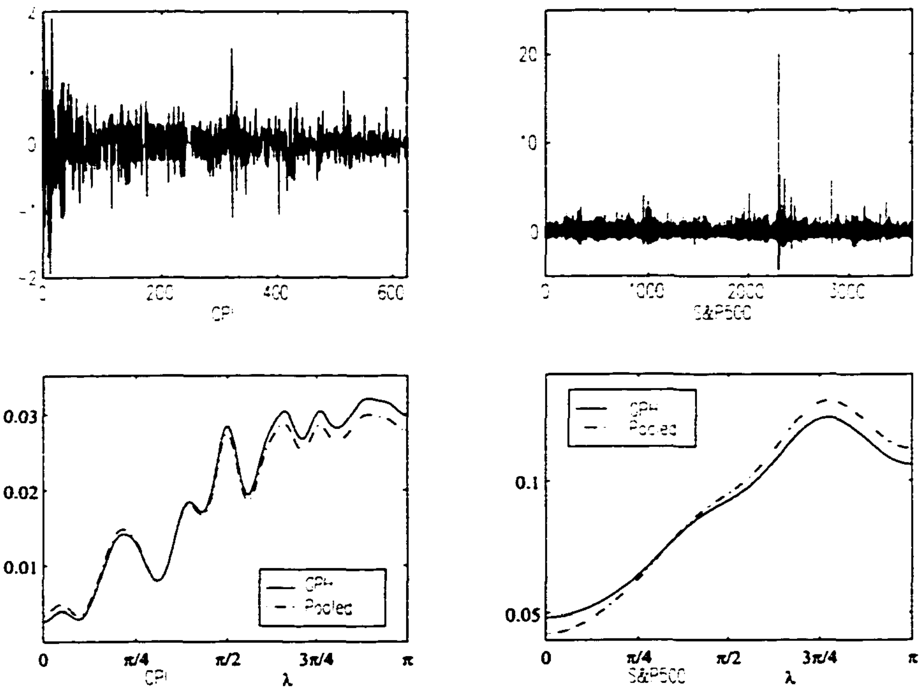

Figure 7: $(1-L)^{\hat{d}} X_t$ and spectral density estimates

7 Appendix: Proofs

7.1 Proof of Lemma 3.2

(a) $j = 0$

Assumptions 1 and 2 yield

$$\sum_{\{s:\lambda_s \in B_0\}} (X_{sj} - \overline{X}_{.j})^2 = 4 \sum_{s=1}^{m} \left(\ln \left| 1 - e^{i\lambda_s} \right| - \frac{1}{m} \sum_{s=1}^{m} \ln \left| 1 - e^{i\lambda_s} \right| \right)^2 = 4m + o(m),$$

as shown in lemma 1 of Hurvich and Beltrao (1994).

(b) $1 \le j \le \ell$

Taylor expansion gives, as in Hurvich and Beltrao (1994)

$$\left| 1 - e^{i\lambda_s} \right|^2 = 2(1 - \cos \lambda_s) = \lambda_s^2 \cos \xi_s, \quad 0 < \xi_s < \lambda_s,$$

and

$$-\frac{1}{2} X_{sj} \;=\; \ln \lambda_s + \frac{1}{2} \ln \cos \xi_s, \quad \lambda_s \in B_j$$

$$
\begin{aligned}
&= \ln \lambda_s - \frac{\xi_s^2}{4\cos^2 \Omega_s}, \quad 0 < \Omega_s < \xi_s \\
&= \ln \lambda_s + O\left(\xi_s^2\right) = \ln \lambda_s + O\left(\frac{j^2}{M^2}\right),
\end{aligned}
$$

for a sufficiently large n, because $\ell/M \to 0$ implies $\Omega_s \to 0$, and $\left(\cos^2 \Omega_s\right)^{-1} = O(1)$ follows. Also, the order $O\left(j^2/M^2\right)$ is uniform in j.

Note that

$$
\omega_j = \frac{(2j+1)\pi}{2M} = \frac{\pi m(2j+1)}{n}, \quad \text{(using } n = 2mM\text{)}
$$

and it follows that

$$
\begin{aligned}
\lambda_s \in B_j &\Leftrightarrow -\frac{\pi}{2M} < \lambda_s - \omega_j \le \frac{\pi}{2M} \\
&\Leftrightarrow -\frac{\pi m}{n} < \frac{2\pi s}{n} - \frac{(2j+1)\pi m}{n} \le \frac{\pi m}{n} \\
&\Leftrightarrow -m < 2s - m(2j+1) \le m \\
&\Leftrightarrow 0 < 2s - 2mj \le 2m \\
&\Leftrightarrow 1 \le s - mj \le m.
\end{aligned}
$$

For $mj + 1 \le s, s' \le mj + m$, by the mean value theorem

$$
\begin{aligned}
\ln s - \ln s' &= \frac{1}{\bar{s}}\left(s - s'\right) \quad \bar{s} \in \left[s, s'\right] \\
&= O(1/mj)\, O(m) = O(1/j),
\end{aligned}
$$

and

$$
\ln s - \frac{1}{m}\sum_{s=mj+1}^{mj+m} \ln s = \frac{1}{m}\sum_{s'=mj+1}^{mj+m}\left(\ln s - \ln s'\right) = O(1/j).
$$

It follows that

$$
\begin{aligned}
&\sum_{\{s:\lambda_s \in B_j\}} \left(X_{sj} - \overline{X}_{\cdot j}\right)^2 \\
&= 4\sum_{\{s:\lambda_s \in B_j\}}\left[\ln \lambda_s - \frac{1}{m}\sum_{\{s:\lambda_s \in B_j\}} \ln \lambda_s + O\left(\frac{j^2}{M^2}\right)\right]^2 \\
&= 4\sum_{\{s:\lambda_s \in B_j\}}\left[\ln s - \frac{1}{m}\sum_{\{s:\lambda_s \in B_j\}} \ln s + O\left(\frac{j^2}{M^2}\right)\right]^2 \\
&= 4\sum_{s=mj+1}^{mj+m}\left(\ln s - \frac{1}{m}\sum_{s=mj+1}^{mj+m} \ln s\right)^2 + \sum_{\{s:\lambda_s \in B_j\}} O\left(\frac{j}{M^2}\right) + O\left(\frac{mj^4}{M^4}\right)
\end{aligned}
$$

$$
\begin{aligned}
&= 4\sum_{s/m=j+1/m}^{j+1}\left(\ln\frac{s}{m}-\frac{1}{m}\sum_{s/m=j+1/m}^{j+1}\ln\frac{s}{m}\right)^2+O\left(\frac{mj}{M^2}\right)+O\left(\frac{mj^4}{M^4}\right)\\
&\sim 4m\int_j^{j+1}(\ln x-c_j)^2\,dx+O\left(\frac{mj}{M^2}\right)+O\left(\frac{mj^4}{M^4}\right),
\end{aligned}
$$

where

$$
\begin{aligned}
c_j &= \int_j^{j+1}\ln x dx\\
&= [x(\ln x-1)]_j^{j+1}\\
&= (j+1)(\ln(j+1)-1)-j(\ln j-1)\\
&= (j+1)\ln(j+1)-j\ln j-1,\\
c_j+1 &= (j+1)\ln(j+1)-j\ln j,
\end{aligned}
$$

which gives

$$
\begin{aligned}
&\int_j^{j+1}(\ln x-c_j)^2\,dx\\
&= \left[x\left((\ln x-c_j-1)^2+1\right)\right]_j^{j+1}\\
&= (j+1)(\ln(j+1)-c_j-1)^2+j+1-j(\ln j-c_j-1)^2-j\\
&= (j+1)j^2(\ln(j+1)-\ln j)^2-j(j+1)^2(\ln(j+1)-\ln j)^2+1\\
&= -j(j+1)(\ln(j+1)-\ln j)^2+1.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
&\sum_{j=1}^{\ell}\sum_{\{s:\lambda_s\in B_j\}}(X_{sj}-\overline{X}_{.j})^2\\
&\sim 4m\sum_{j=1}^{\ell}\int_j^{j+1}(\ln x-c_j)^2\,dx+\sum_{j=1}^{\ell}\left(O\left(\frac{mj}{M^2}\right)+O\left(\frac{mj^4}{M^4}\right)\right)\\
&= 4m\sum_{j=1}^{\ell}\left[-j(j+1)(\ln(j+1)-\ln j)^2+1\right]+O\left(\frac{m\ell^2}{M^2}\right)+O\left(\frac{m\ell^5}{M^4}\right)\\
&= 4m\sum_{j=1}^{\ell}\left[-j(j+1)(\ln(j+1)-\ln j)^2+1\right]+o(m).
\end{aligned}
$$

There is no explicit numerical value for the quantity $\sum_1^{\ell}\left(-j(j+1)(\ln(j+1)-\ln j)^2+1\right)$ or its limit as $\ell\to\infty$. Nevertheless, the sum converges since

$$
\int_j^{j+1}(\ln x-c_j)^2\,dx
$$

$$\leq \left(\ln(j+1) - \ln j\right)^2 \text{ (by the mean value theorem)}$$
$$\leq 1/j^2,$$

which implies that $\sum_{j=1}^{\infty} \int_j^{j+1} (\ln x - c_j)^2 \, dx < \infty$. Direct calculations using Mathematica produce the approximate numerical value $\sum_{j=1}^{\infty} \left(-j(j+1)(\ln(j+1) - \ln j)^2 + 1\right) \doteqdot 0.0803$.

7.2 Proof of Lemma 3.3

For $j = 0$, by lemma 1 of Hurvich, Deo, and Brodsky (1998) (hereafter HDB), we have

$$\sum_{\{s:\lambda_s \in B_0\}} \eta_{sj} \left(X_{sj} - \overline{X}_{.j}\right) = \sum_{\{s:\lambda_s \in B_0\}} \left(\ln f_{uu}(\lambda_s) - \ln f_{uu}(0)\right)\left(X_{sj} - \overline{X}_{.j}\right)$$
$$= \sum_{\{s:\lambda_s \in B_0\}} \ln f_{uu}(\lambda_s)\left(X_{sj} - \overline{X}_{.j}\right) = -\frac{8\pi^2}{9} \frac{f''_{uu}(0)}{f_{uu}(0)} \frac{m^3}{n^2} + o\left(\frac{m^3}{n^2}\right).$$

For $j \geq 1$, we first collect together some useful technicalities in the following Sub-lemma.

7.2.1 Sub-lemma

1. $\frac{f'_{uu}(\omega_j)}{f_{uu}(\omega_j)} = \frac{f''_{uu}(0)\omega_j + O(\omega_j^2)}{f_{uu}(0) + O(\omega_j)} = \frac{f''_{uu}(0)}{f_{uu}(0)} \omega_j \left(1 + O(\omega_j)\right) = O(\omega_j)$ *uniformly in* j.
2. $O\left(\frac{M}{nj}\right) = O\left(\frac{1}{j^2}\right)$ *because* $j = o(m)$.
3. $\frac{\sin \omega_j}{1 - \cos \omega_j} = \frac{\omega_j + O(\omega_j^2)}{0.5\omega_j^2 + O(\omega_j^4)} = \frac{2}{\omega_j}\left(1 + O(\omega_j)\right)$ *uniformly in* j.

■

First, note that

$$\sum_{\{s:\lambda_s \in B_j\}} (\lambda_s - \omega_j) = \sum_{s=mj+1}^{mj+m} \left(\frac{2\pi s}{n} - \frac{(2j+1)\pi m}{n}\right) = \frac{2\pi}{n} \sum_{s=mj+1}^{mj+m} \left(s - (j + 1/2)m\right)$$
$$= \frac{2\pi}{n} \sum_{s=mj+1}^{mj+m} (s - mj - m/2) = \frac{2\pi}{n} \sum_{k=1}^{m} (k - m/2)$$
$$= \frac{2\pi}{n}\left(\frac{m(m+1)}{2} - \frac{m^2}{2}\right) = \frac{2\pi}{n}\frac{m}{2},$$

because positive and negative terms cancel out as we sum over $s - jm$ from 1 to m. It follows that

$$-\overline{X}_{.j} = -\frac{1}{m} \sum_{\{s:\lambda_s \in B_j\}} X_{sj} = \frac{1}{m} \sum_{\{s:\lambda_s \in B_j\}} \ln(2 - 2\cos\lambda_s)$$

$$
\begin{aligned}
&= \frac{1}{m}\sum_{\{s:\lambda_s\in B_j\}}\left[\ln(2-2\cos\omega_j)+\frac{\sin\omega_j}{1-\cos\omega_j}(\lambda_s-\omega_j)+\frac{1}{2}\frac{1}{\cos\widetilde{\lambda}_s-1}(\lambda_s-\omega_j)^2\right]\\
&= \ln(2-2\cos\omega_j)+\frac{\sin\omega_j}{1-\cos\omega_j}\frac{1}{m}\sum_{\{s:\lambda_s\in B_j\}}(\lambda_s-\omega_j)+O\left(\frac{1}{\omega_j^2M^2}\right)\\
&= \ln(2-2\cos\omega_j)+\frac{\sin\omega_j}{1-\cos\omega_j}\frac{\pi}{n}+O\left(\frac{1}{j^2}\right),
\end{aligned}
$$

where $\widetilde{\lambda}_s\in(\omega_j,\lambda_s)$, and the term $O(1/j^2)$ is uniform in j.

By Taylor expansion

$$
\begin{aligned}
\ln f_{uu}(\lambda_s) &= \ln f_{uu}(\omega_j)+\frac{f'_{uu}(\omega_j)}{f_{uu}(\omega_j)}(\lambda_s-\omega_j)+\left(\frac{f''_{uu}(\widetilde{\lambda}_s)}{f_{uu}(\widetilde{\lambda}_s)}-\left(\frac{f'_{uu}(\widetilde{\lambda}_s)}{f_{uu}(\widetilde{\lambda}_s)}\right)^2\right)\frac{(\lambda_s-\omega_j)^2}{2}\\
&= \ln f_{uu}(\omega_j)+\frac{f'_{uu}(\omega_j)}{f_{uu}(\omega_j)}(\lambda_s-\omega_j)+O\left(\frac{1}{M^2}\right).
\end{aligned}
$$

Thus, in view of Sub-lemma 7.2.1,

$$
\begin{aligned}
&\sum_{\{s:\lambda_s\in B_j;j\ge1\}}\eta_{sj}\left(X_{sj}-\overline{X}_{\cdot j}\right)\\
&= \sum_{\{s:\lambda_s\in B_j;j\ge1\}}\left[\frac{f'_{uu}(\omega_j)}{f_{uu}(\omega_j)}(\lambda_s-\omega_j)+O\left(\frac{1}{M^2}\right)\right]\\
&\quad\times\left[-\left(\frac{\sin\omega_j}{1-\cos\omega_j}\right)\left(\lambda_s-\omega_j-\frac{\pi}{n}\right)+O\left(\frac{1}{j^2}\right)\right]\\
&= \sum_{\{s:\lambda_s\in B_j;j\ge1\}}\left[\frac{f'_{uu}(\omega_j)}{f_{uu}(\omega_j)}(\lambda_s-\omega_j)+O\left(\frac{1}{M^2}\right)\right]\\
&\quad\times\left[-\left(\frac{\sin\omega_j}{1-\cos\omega_j}\right)(\lambda_s-\omega_j)+O\left(\frac{M}{nj}\right)+O\left(\frac{1}{j^2}\right)\right]\\
&= -\left(\frac{f'_{uu}(\omega_j)}{f_{uu}(\omega_j)}\right)\left(\frac{\sin\omega_j}{1-\cos\omega_j}\right)\sum_{\{s:\lambda_s\in B_j;j\ge1\}}(\lambda_s-\omega_j)^2\\
&\quad+\left(\frac{f'_{uu}(\omega_j)}{f_{uu}(\omega_j)}\right)\sum_{\{s:\lambda_s\in B_j;j\ge1\}}(\lambda_s-\omega_j)\cdot O\left(\frac{1}{j^2}\right)\\
&\quad-\left(\frac{\sin\omega_j}{1-\cos\omega_j}\right)\sum_{\{s:\lambda_s\in B_j;j\ge1\}}(\lambda_s-\omega_j)\cdot O\left(\frac{1}{M^2}\right)+O\left(\frac{m}{M^2j^2}\right)\\
&= -\left(\frac{f'_{uu}(\omega_j)}{f_{uu}(\omega_j)}\right)\left(\frac{\sin\omega_j}{1-\cos\omega_j}\right)\left(\frac{2\pi}{n}\right)^2\sum_{k=1}^{m}\left(k-\frac{m}{2}\right)^2\\
&\quad+O\left(\frac{j}{M}\right)\sum_{\{s:\lambda_s\in B_j\}}O\left(\frac{1}{Mj^2}\right)-O\left(\frac{M}{j}\right)\sum_{\{s:\lambda_s\in B_j\}}O\left(\frac{1}{M^3}\right)+O\left(\frac{m}{M^2j^2}\right)\\
&= -\left(\frac{f'_{uu}(\omega_j)}{f_{uu}(\omega_j)}\right)\left(\frac{\sin\omega_j}{1-\cos\omega_j}\right)\frac{\pi^2}{m^2M^2}\left(\frac{m^3}{12}+O(m^2)\right)+O\left(\frac{m}{M^2j}\right)+O\left(\frac{m}{M^2j^2}\right)
\end{aligned}
$$

$$= -\left(\frac{f'_{uu}(\omega_j)}{f_{uu}(\omega_j)}\right)\left(\frac{\sin\omega_j}{1-\cos\omega_j}\right)\frac{\pi^2}{M^2}\frac{m}{12}+O\left(\frac{1}{M^2}\right)+O\left(\frac{m}{M^2 j}\right)$$
$$= -\left(\frac{f'_{uu}(\omega_j)}{f_{uu}(\omega_j)}\right)\left(\frac{\sin\omega_j}{1-\cos\omega_j}\right)\frac{\pi^2}{M^2}\frac{m}{12}+O\left(\frac{m}{M^2 j}\right),$$

where we use the fact that

$$\sum_{k=1}^{m}\left(k-\frac{m}{2}\right)^2=\frac{m(m+1)(2m+1)}{6}-\frac{m^2(m+1)}{2}+\frac{m^3}{4}=\frac{m^3}{12}+O\left(m^2\right).$$

Furthermore, from Sub-lemma 7.2.1 we have

$$\left(\frac{f'_{uu}(\omega_j)}{f_{uu}(\omega_j)}\right)\left(\frac{\sin\omega_j}{1-\cos\omega_j}\right)=\frac{2f''_{uu}(0)}{f_{uu}(0)}\left(1+O(\omega_j)\right).$$

Therefore,

$$\sum_{j=1}^{\ell}\sum_{\{s:\lambda_s\in B_j\}}\eta_{sj}\left(X_{sj}-\overline{X}_{\cdot j}\right)$$
$$= -\frac{2f''_{uu}(0)}{f_{uu}(0)}\frac{\pi^2}{M^2}\frac{m}{12}+O\left(\sum_{j=1}^{\ell}\frac{j}{M}\frac{m}{M^2}\right)+O\left(\sum_{j=1}^{\ell}\frac{m}{M^2 j}\right)$$
$$= -\frac{2\pi^2}{3}\frac{f''_{uu}(0)}{f_{uu}(0)}\frac{m^3\ell}{n^2}+O\left(\sum_{j=1}^{\ell}\frac{m^4 j}{n^3}\right)+O\left(\sum_{j=1}^{\ell}\frac{m^3}{n^2 j}\right)$$
$$= -\frac{2\pi^2}{3}\frac{f''_{uu}(0)}{f_{uu}(0)}\frac{m^3\ell}{n^2}+O\left(\frac{m^4\ell^2}{n^3}\right)+O\left(\frac{m^3}{n^2}\ln\ell\right)$$
$$= -\frac{2\pi^2}{3}\frac{f''_{uu}(0)}{f_{uu}(0)}\frac{m^3\ell}{n^2}+o\left(\frac{m^3\ell}{n^2}\right),$$

giving the required result. ■

7.3 Proof of Theorem 4.1

Again, it is helpful to start by collecting some useful preliminary results in the following Sub-lemmas.

7.3.1 Sub-Lemma

1. $\int_{-\pi}^{\pi}|D(\lambda)|\,d\lambda=O(\ln n)$
2. $K(\lambda)=O\left(\left(n\lambda^2\right)^{-1}\right),\quad 0<|\lambda|\leq\pi$
3. $|D(\lambda)|\leq 2|\lambda|^{-1},\quad 0<|\lambda|\leq\pi.$

where $K(\lambda)=(2\pi n)^{-1}\left|\sum_{t=1}^{n}\sum_{s=1}^{n}e^{i(t-s)\lambda}\right|=(2\pi n)^{-1}|D(\lambda)|^2,\quad D(\lambda)=\sum_{t=1}^{n}e^{it\lambda}.$

Proof See Zygmund (1959) (p. 67 for 1, pp.89-90 for 2, pp.49-51 for 3). ■

7.3.2 Sub-Lemma

For some $C < \infty$ and for $\lambda \in [-\pi, 0) \cup (0, \pi]$,

1. $|f_{XX}(\lambda)| \leq C |\lambda|^{-2d}$
2. $|f_{XX}(\lambda)|^{-1} \leq C |\lambda|^{2d}$
3. $|f'_{XX}(\lambda)| \leq C |\lambda|^{-2d-1}$.

Proof First, the inequality $|x/2| \leq |\sin x| \leq |x|$ for $x \in [-\pi/2, \pi/2]$ implies that, for both positive and negative d, $\left|\sin^2(\lambda/2)\right|^{-d} \leq C_1 |\lambda|^{-2d}$ holds for $\lambda \in [-\pi, 0) \cup (0, \pi]$ and for some $C_1 < \infty$. It follows that

$$\begin{aligned}
|f_{XX}(\lambda)| &= \left|4\sin^2(\lambda/2)\right|^{-d} |f_{uu}(\lambda)| \leq 4^{-d} C_1 |\lambda|^{-2d} \sup_{\lambda \in (-\pi,\pi)} |f_{uu}(\lambda)| \leq C |\lambda|^{-2d}, \\
|f_{XX}(\lambda)|^{-1} &= \left|4\sin^2(\lambda/2)\right|^{d} |f_{uu}(\lambda)|^{-1} \leq 4^{d} C_1 |\lambda|^{2d} \sup_{\lambda \in (-\pi,\pi)} |f_{uu}(\lambda)|^{-1} \leq C |\lambda|^{2d},
\end{aligned}$$

because $\sup_{\lambda \in (-\pi,\pi)} |f_{uu}(\lambda)| < \infty$ and $\inf_{\lambda \in (-\pi,\pi)} |f_{uu}(\lambda)| > 0$. The bound for $|f'_{XX}(\lambda)|$ uses the fact that $\sup_{\lambda \in (-\pi,\pi)} |f'_{uu}(\lambda)| < \infty$ and then

$$\begin{aligned}
&|f'_{XX}(\lambda)| \\
\leq\ & \left|4^{-d}(-d)\left(\sin^2(\lambda/2)\right)^{-d-1} \sin(\lambda/2)\cos(\lambda/2) f_{uu}(\lambda)\right| + \left|\left(4\sin^2(\lambda/2)\right)^{-d} f'_{uu}(\lambda)\right| \\
\leq\ & C_2 |\sin(\lambda/2)|^{-2d-1} + C_3 |\lambda|^{-2d} \leq C |\lambda|^{-2d-1}. \blacksquare
\end{aligned}$$

7.3.3 Sub-Lemma

For $0 < \lambda + \lambda_j < 2\pi$ and $0 \leq \lambda, \lambda_j \leq \pi$,

$$|\lambda - \lambda_j| \, |D(\lambda + \lambda_j)| \leq 2, \tag{15}$$

Proof Because $|D(\lambda)| = |D(2\pi - \lambda)|$ for $\pi \leq \lambda \leq 2\pi$, and in view of Sub lemma 7.3.1 (3),

$$\begin{aligned} & |\lambda - \lambda_j| \, |D(\lambda + \lambda_j)| \\ \leq \; & 2\frac{|\lambda - \lambda_j|}{\lambda + \lambda_j}, \quad 0 < \lambda + \lambda_j \leq \pi \\ \leq \; & 2\frac{|\lambda - \lambda_j|}{2\pi - \lambda - \lambda_j}, \quad \pi < \lambda + \lambda_j < 2\pi, \end{aligned}$$

and

$$\begin{aligned} \frac{|\lambda - \lambda_j|}{\lambda + \lambda_j} &= \frac{(\lambda - \lambda_j)}{2\lambda_j + (\lambda - \lambda_j)} \leq 1, \quad \lambda \geq \lambda_j \\ &= \frac{(\lambda_j - \lambda)}{2\lambda + (\lambda_j - \lambda)} \leq 1, \quad \lambda < \lambda_j \\ \frac{|\lambda - \lambda_j|}{2\pi - \lambda - \lambda_j} &= \frac{(\lambda - \lambda_j)}{2\pi - 2\lambda + (\lambda - \lambda_j)} \leq 1, \quad \lambda \geq \lambda_j \\ &= \frac{(\lambda_j - \lambda)}{2\pi - 2\lambda_j + (\lambda_j - \lambda)} \leq 1. \quad \lambda < \lambda_j \blacksquare \end{aligned}$$

7.3.4 Sub-Lemma

If $\kappa_1 \lambda_j \leq \lambda_j + \lambda \leq 2\pi - \kappa_2 \lambda_j$ for $\kappa_1, \kappa_2 > 0$, $\sup \lambda_j |D(\lambda_j + \lambda)| < \infty$.
If $\kappa_1 \lambda_j \leq \lambda_j - \lambda \leq 2\pi - \kappa_2 \lambda_j$ for $\kappa_1, \kappa_2 > 0$, $\sup \lambda_j |D(\lambda_j - \lambda)| < \infty$.

Proof

$$\begin{aligned} \lambda_j |D(\lambda_j + \lambda)| \leq \frac{2\lambda_j}{\lambda_j + \lambda} &\leq \frac{2}{\kappa_1}, \quad \kappa_1 \lambda_j \leq \lambda_j + \lambda \leq \pi \\ \lambda_j |D(\lambda_j + \lambda)| \leq \frac{2\lambda_j}{2\pi - \lambda_j - \lambda} &\leq \frac{2}{\kappa_2}, \quad \pi < \lambda_j + \lambda \leq 2\pi - \kappa_2 \lambda_j \end{aligned}$$

similarly for $\lambda_j |D(\lambda_j - \lambda)|$. ■

With these technicalities in hand, the proof of theorem 4.1 is almost identical to that in Robinson (1995). The main difference is that $f_{XX}(\lambda)$ is bounded by $|\lambda|^{-2d}$ over the full range of λ, and the evaluation of $|D(\lambda)|$ becomes complicated because λ_j is not restricted to a neighborhood of the origin. We proceed with each part in turn.

7.3.5 Proof of (a)

We start by showing

$$E[w(\lambda_j)\overline{w}(\lambda_j)] - f_{XX}(\lambda_j) = \int_{-\pi}^{\pi} \{f_{XX}(\lambda) - f_{XX}(\lambda_j)\} K(\lambda - \lambda_j) \, d\lambda = O\left(j^{-1}\lambda_j^{-2d} \ln n\right).$$

where $K(\lambda) = (2\pi n)^{-1} \left|\sum_{t=1}^{n} \sum_{s=1}^{n} e^{i(t-s)\lambda}\right|$ is Fejer's kernel. By sub-lemmas 7.3.1 and 7.3.2,

$$\begin{aligned}
\left|\int_{\lambda_j/2}^{\pi}\right| &\leq \left\{\max_{\lambda_j/2 \leq \lambda \leq \pi} |f'_{XX}(\lambda)|\right\} \int_{\lambda_j/2}^{\pi} |\lambda - \lambda_j| K(\lambda - \lambda_j)\, d\lambda \\
&= O\left(\lambda_j^{-1-2d} n^{-1} \int_{\lambda_j/2}^{\pi} |D(\lambda - \lambda_j)|\, d\lambda\right), \\
\text{because } \lambda - \lambda_j &\in (-\lambda_j/2, \pi - \lambda_j) \subseteq [-\pi, \pi] \\
&= O\left(\lambda_j^{-1-2d} n^{-1} \int_{-\pi}^{\pi} |D(\lambda)|\, d\lambda\right) \\
&= O\left(\lambda_j^{-1-2d} n^{-1} \ln n\right) = O\left(j^{-1} \lambda_j^{-2d} \ln n\right).
\end{aligned}$$

The symmetry of $f_{XX}(\lambda)$ and sub-lemma 7.3.3 (applicable since $0 < \lambda + \lambda_j < 2\pi$) yields

$$\begin{aligned}
\left|\int_{-\pi}^{-\lambda_j/2}\right| &= \left|\int_{\lambda_j/2}^{\pi} \{f_{XX}(\lambda) - f_{XX}(\lambda_j)\} K(\lambda + \lambda_j)\, d\lambda\right| \\
&\leq \left\{\max_{\lambda_j/2 \leq \lambda \leq \pi} |f'_{XX}(\lambda)|\right\} \int_{\lambda_j/2}^{\pi} |\lambda - \lambda_j| K(\lambda + \lambda_j)\, d\lambda \\
&= O\left(\left\{\max_{\lambda_j/2 \leq \lambda \leq \pi} |f'_{XX}(\lambda)|\right\} n^{-1} \int_{\lambda_j/2}^{\pi} |\lambda - \lambda_j| \, |D(\lambda + \lambda_j)|^2\, d\lambda\right) \\
&= O\left(\left\{\max_{\lambda_j/2 \leq \lambda \leq \pi} |f'_{XX}(\lambda)|\right\} n^{-1} \int_{\lambda_j/2}^{\pi} |D(\lambda + \lambda_j)|\, d\lambda\right) \\
&= O\left(\lambda_j^{-1-2d} n^{-1} \ln n\right) = O\left(j^{-1} \lambda_j^{-2d} \ln n\right).
\end{aligned}$$

For the remaining part, as in Robinson (1995),

$$\begin{aligned}
\left|\int_{-\lambda_j/2}^{\lambda_j/2}\right| &\leq \max_{|\lambda| \leq \lambda_j/2} K(\lambda - \lambda_j) \int_{-\lambda_j/2}^{\lambda_j/2} \{|f_{XX}(\lambda)| + |f_{XX}(\lambda_j)|\}\, d\lambda \qquad (16) \\
&= O\left(n^{-1} \lambda_j^{-1-2d}\right) = O\left(j^{-1} \lambda_j^{-2d}\right).
\end{aligned}$$

Finally, by sub-lemma 7.3.2,

$$E[w(\lambda_j) \overline{w}(\lambda_j) / f_{XX}(\lambda_j)] = 1 + O\left(j^{-1} \ln n \lambda_j^{-2d} / f_{XX}(\lambda_j)\right) = 1 + O\left(j^{-1} \ln n\right).$$

7.3.6 Proof of (b)

By Robinson (1995), we have

$$E[w(\lambda_j) w(\lambda_j)] = \int_{-\pi}^{\pi} \{f_{XX}(\lambda) - f_{XX}(\lambda_j)\} E_{j,-j}(\lambda)\, d\lambda, \qquad (17)$$

for $0 < k < j < n/2$, where $E_{jk}(\lambda) = (2\pi n)^{-1} D(\lambda_j - \lambda) D(\lambda - \lambda_k)$. The calculations are similar to those given before, but we have to check the range of integration for each subinterval.

$$\begin{aligned}
\left| \int_{\lambda_j/2}^{\pi} \right| &= \left| \int_{\lambda_j/2}^{\pi} (2\pi n)^{-1} \{ f_{XX}(\lambda) - f_{XX}(\lambda_j) \} D(\lambda_j - \lambda) D(\lambda + \lambda_j) \, d\lambda \right| \\
&\leq \left\{ \max_{\lambda_j/2 \leq \lambda \leq \pi} |f'_{XX}(\lambda)| \right\} (2\pi n)^{-1} \int_{\lambda_j/2}^{\pi} |\lambda_j - \lambda| \, |D(\lambda_j - \lambda)| \, |D(\lambda + \lambda_j)| \, d\lambda \\
&= O\left(\lambda_j^{-1-2d} n^{-1} \int_{\lambda_j/2}^{\pi} |D(\lambda + \lambda_j)| \, d\lambda \right) \\
&= O\left(\lambda_j^{-1-2d} n^{-1} \ln n \right) = O\left(j^{-1} \lambda_j^{-2d} \ln n \right),
\end{aligned}$$

because $-\pi < -\pi + \lambda_j \leq \lambda_j - \lambda \leq \lambda_j/2 < \pi$. Similarly,

$$\begin{aligned}
&\left| \int_{-\pi}^{-\lambda_j/2} \{ f_{XX}(\lambda) - f_{XX}(\lambda_j) \} (2\pi n)^{-1} D(\lambda_j - \lambda) D(\lambda + \lambda_j) \, d\lambda \right| \\
&= \left| \int_{\lambda_j/2}^{\pi} \{ f_{XX}(\lambda) - f_{XX}(\lambda_j) \} (2\pi n)^{-1} D(\lambda_j - \lambda) D(\lambda + \lambda_j) \, d\lambda \right| \\
&= O\left(j^{-1} \lambda_j^{-2d} \ln n \right).
\end{aligned}$$

For the remaining part, $\lambda j/2 \leq \lambda_j - \lambda \leq 3\lambda_j/2 \leq 2\pi - \lambda_j/2$ and sub-lemma 7.3.4 yield

$$\begin{aligned}
&\int_{-\lambda_j/2}^{-\lambda_j/2} \{ f_{XX}(\lambda) - f_{XX}(\lambda_j) \} E_{j,-j}(\lambda) \, d\lambda \\
&= O\left(\max_{|\lambda| \leq \lambda_j/2} |E_{j,-j}(\lambda)| \int_0^{\lambda_j} \{ |f_{XX}(\lambda)| + |f_{XX}(\lambda_j)| \} \, d\lambda \right) \\
&= O\left(n^{-1} \lambda_j^{-2} \lambda_j^{1-2d} \right) = O\left(j^{-1} \lambda_j^{-2d} \right).
\end{aligned}$$

7.3.7 Proof of (c)

Similar to Robinson (1995), we expand the integral as follows, for $0 < k < j < n/2$,

$$\begin{aligned}
E\left[w(\lambda_j) \overline{w}(\lambda_k) \right] &= \int_{-\pi}^{\pi} f_{XX}(\lambda) E_{jk}(\lambda) \, d\lambda \\
&= \int_{(\lambda_j + \lambda_k)/2}^{\pi} \{ f_{XX}(\lambda) - f_{XX}(\lambda_j) \} E_{jk}(\lambda) \, d\lambda && (18) \\
&\quad + \int_{\lambda_k/2}^{(\lambda_j + \lambda_k)/2} \{ f_{XX}(\lambda) - f_{XX}(\lambda_k) \} E_{jk}(\lambda) \, d\lambda && (19) \\
&\quad - \{ f_{XX}(\lambda_j) - f_{XX}(\lambda_k) \} \int_{\lambda_k/2}^{(\lambda_j + \lambda_k)/2} E_{jk}(\lambda) \, d\lambda && (20) \\
&\quad + \int_{-\pi}^{\lambda_k/2} \{ f_{XX}(\lambda) - f_{XX}(\lambda_j) \} E_{jk}(\lambda) \, d\lambda. && (21)
\end{aligned}$$

First, because $-\pi < \lambda_j - \pi \leq \lambda_j - \lambda \leq (\lambda_j - \lambda_k)/2 < \pi$, (18) is bounded by

$$\begin{aligned}
&\left\{\max_{(\lambda_j+\lambda_k)/2\leq\lambda\leq\pi}\left|f'_{XX}(\lambda)\right|\right\} n^{-1}\int_{(\lambda_j+\lambda_k)/2}^{\pi}|\lambda_j-\lambda|\,|D(\lambda_j-\lambda)|\,|D(\lambda-\lambda_k)|\,d\lambda \\
=& \left\{\max_{(\lambda_j+\lambda_k)/2\leq\lambda\leq\pi}\left|f'_{XX}(\lambda)\right|\right\} n^{-1}\int_{(\lambda_j+\lambda_k)/2}^{\pi}|D(\lambda-\lambda_k)|\,d\lambda \qquad (22)\\
=& O\left(\lambda_j^{-1-2d}n^{-1}\ln n\right)\\
=& O\left(j^{-1}\lambda_j^{-2d}\ln n\right)\\
=& O\left(k^{-1}\lambda_j^{-d}\lambda_k^{-d}\ln n\left(\frac{k}{j}\right)^{1+d}\right)\\
=& O\left(k^{-1}\lambda_j^{-d}\lambda_k^{-d}\ln n\right).
\end{aligned}$$

Next, when $k \geq j/2$, (19) is bounded by,

$$\begin{aligned}
&\left\{\max_{\lambda_k/2\leq\lambda\leq(\lambda_j+\lambda_k)/2}\left|f'_{XX}(\lambda)\right|\right\} n^{-1}\int_{\lambda_k/2}^{(\lambda_j+\lambda_k)/2}|\lambda-\lambda_k|\,|D(\lambda_j-\lambda)|\,|D(\lambda-\lambda_k)|\,d\lambda \\
=& \left\{\max_{\lambda_k/2\leq\lambda\leq(\lambda_j+\lambda_k)/2}\left|f'_{XX}(\lambda)\right|\right\} n^{-1}\int_{\lambda_k/2}^{(\lambda_j+\lambda_k)/2}|D(\lambda_j-\lambda)|\,d\lambda \\
=& O\left(\lambda_k^{-1-2d}n^{-1}\ln n\right)\\
=& O\left(k^{-1}\lambda_j^{-d}\lambda_k^{-d}\ln n\left(\frac{k}{j}\right)^{-d}\right)\\
=& O\left(k^{-1}\lambda_j^{-d}\lambda_k^{-d}\ln n\right),
\end{aligned}$$

because $-\pi < -\lambda_k/2 \leq \lambda - \lambda_k \leq (\lambda_j - \lambda_k)/2 < \pi$, and, when $k < j/2$, by

$$\begin{aligned}
&\left\{\max_{\lambda_k/2\leq\lambda\leq\lambda_j}|f_{XX}(\lambda)| + |f_{XX}(\lambda_k)|\right\}\int_{\lambda_k/2}^{(\lambda_j+\lambda_k)/2}|E_{jk}(\lambda)|\,d\lambda \\
=& O\left(\left(\lambda_j^{-2d}+\lambda_k^{-2d}\right)(j-k)^{-1}\int_{\lambda_k/2}^{(\lambda_j+\lambda_k)/2}|D(\lambda-\lambda_k)|\,d\lambda\right)\\
=& O\left(k^{-1}\lambda_j^{-d}\lambda_k^{-d}\ln n\left(\frac{k}{j-k}\right)\left\{\left(\frac{\lambda_j}{\lambda_k}\right)^{-d}+\left(\frac{\lambda_k}{\lambda_j}\right)^{-d}\right\}\right)\\
=& O\left(k^{-1}\lambda_j^{-d}\lambda_k^{-d}\ln n\left(\frac{j}{j-k}\right)\left\{\left(\frac{k}{j}\right)^{1+d}+\left(\frac{k}{j}\right)^{1-d}\right\}\right)\\
=& O\left(k^{-1}\lambda_j^{-d}\lambda_k^{-d}\ln n\right),
\end{aligned}$$

because $0 < (\lambda_j - \lambda_k)/2 \leq \lambda_j - \lambda \leq \lambda_j < \pi$.

Similarly, (20) is bounded by

$$(\lambda_j-\lambda_k)\left\{\max_{\lambda_k\leq\lambda\leq\lambda_j}\left|f'_{XX}(\lambda)\right|\right\}\int_{\lambda_k/2}^{(\lambda_j+\lambda_k)/2}|E_{jk}(\lambda)|\,d\lambda$$

$$
\begin{aligned}
&= O\left(\frac{j-k}{n}\lambda_k^{-1-2d}(j-k)^{-1}\int_{\lambda_k/2}^{(\lambda_j+\lambda_k)/2}|D(\lambda-\lambda_k)|\,d\lambda\right) \\
&= O\left(\lambda_k^{-1-2d}n^{-1}\ln n\right) = O\left(k^{-1}\lambda_j^{-d}\lambda_k^{-d}\ln n\right).
\end{aligned}
$$

when $k \geq j/2$, and by

$$
\begin{aligned}
&\{|f_{XX}(\lambda_j)| + |f_{XX}(\lambda_k)|\}(j-k)^{-1}\int_{\lambda_k/2}^{(\lambda_j+\lambda_k)/2}|D(\lambda-\lambda_k)|\,d\lambda \\
&= O\left(\left(\lambda_j^{-2d}+\lambda_k^{-2d}\right)(j-k)^{-1}\int_{-\pi}^{\pi}|D(\lambda)|\,d\lambda\right) \\
&= O\left(k^{-1}\lambda_j^{-d}\lambda_k^{-d}\ln n\right).
\end{aligned}
$$

when $k < j/2$.

Finally, (21) is calculated by dividing the interval into $[-\pi, -\lambda_j/2]$, $[-\lambda_j/2, -\lambda_k/2]$,and $[-\lambda_k/2, \lambda_k/2]$. First, for the integral on $[-\pi, -\lambda_j/2]$, we use $|D(-\lambda)| = |D(\lambda)|$ and sub-lemma 7.3.3 to derive

$$
\begin{aligned}
&\int_{-\pi}^{-\lambda_j/2}(2\pi n)^{-1}\{f_{XX}(\lambda)-f_{XX}(\lambda_j)\}D(\lambda_j-\lambda)D(\lambda-\lambda_k)\,d\lambda \\
&\leq \int_{\lambda_j/2}^{\pi}(2\pi n)^{-1}\{f_{XX}(\lambda)-f_{XX}(\lambda_j)\}|D(\lambda_j+\lambda)||D(\lambda+\lambda_k)|\,d\lambda \\
&= O\left(\left\{\max_{\lambda_j/2\leq\lambda\leq\pi}|f'_{XX}(\lambda)|\right\}\int_{\lambda_j/2}^{\pi}n^{-1}|\lambda-\lambda_j||D(\lambda_j+\lambda)||D(\lambda+\lambda_k)|\,d\lambda\right) \\
&= O\left(\lambda_j^{-1-2d}n^{-1}\int_{\lambda_j/2}^{\pi}|D(\lambda+\lambda_k)|\,d\lambda\right) \\
&= O\left(\lambda_j^{-1-2d}n^{-1}\ln n\right) = O\left(k^{-1}\lambda_j^{-d}\lambda_k^{-d}\ln n\right).
\end{aligned}
$$

For the integral on $[-\lambda_k/2, \lambda_k/2]$, note that both $\lambda_j|D(\lambda_j-\lambda)|$ and $\lambda_k|D(\lambda_k-\lambda)|$ are bounded (see sub-Lemma 7.3.4) because $\lambda_j/2 < \lambda_j-\lambda < 2\pi-\lambda_j/2$ and $\lambda_k/2 \leq \lambda_k-\lambda < 2\pi-\lambda_k/2$, and then

$$
\int_{-\lambda_k/2}^{\lambda_k/2}(2\pi n)^{-1}\{f_{XX}(\lambda)-f_{XX}(\lambda_j)\}|D(\lambda_j-\lambda)||D(\lambda-\lambda_k)|\,d\lambda \tag{23}
$$

$$
\begin{aligned}
&= \int_{-\lambda_k/2}^{\lambda_k/2}(2\pi n)^{-1}\{f_{XX}(\lambda)-f_{XX}(\lambda_j)\}|D(\lambda_j-\lambda)||D(\lambda_k-\lambda)|\,d\lambda \\
&= O\left(\frac{1}{n\lambda_j\lambda_k}\int_{-\lambda_k/2}^{\lambda_k/2}\{|f_{XX}(\lambda)|+|f_{XX}(\lambda_j)|\}\,d\lambda\right) \\
&= O\left(\frac{1}{n\lambda_j\lambda_k}\left(\lambda_k^{1-2d}+\lambda_j^{-2d}\lambda_k\right)\right)
\end{aligned}
$$

$$
\begin{aligned}
&= O\left(n^{-1}\lambda_j^{-1}\lambda_k^{-2d} + n^{-1}\lambda_j^{-1-2d}\right) \\
&= O\left(k^{-1}\lambda_j^{-1}\lambda_k^{1-2d} + n^{-1}\lambda_j^{-1-2d}\right) \\
&= O\left(k^{-1}\lambda_j^{-d}\lambda_k^{-d}\left(\frac{\lambda_k}{\lambda_j}\right)^{1-d} + n^{-1}\lambda_j^{-1-2d}\right) = O\left(k^{-1}\lambda_j^{-d}\lambda_k^{-d}\ln n\right).
\end{aligned}
$$

In evaluating the integral on $[-\lambda_j/2, -\lambda_k/2]$, we use sub-lemma 7.3.4 and $\lambda_j < \lambda_j + \lambda < 2\pi - \lambda_j/2$ for $\lambda_k/2 \leq \lambda \leq \lambda_j/2$, giving

$$
\begin{aligned}
&\int_{-\lambda j/2}^{-\lambda_k/2} (2\pi n)^{-1}\{f_{XX}(\lambda) - f_{XX}(\lambda_j)\} D(\lambda_j - \lambda) D(\lambda - \lambda_k)\, d\lambda \\
\leq\ & \int_{\lambda_k/2}^{\lambda_j/2} (2\pi n)^{-1}\{f_{XX}(\lambda) - f_{XX}(\lambda_j)\} |D(\lambda_j + \lambda)|\,|D(\lambda + \lambda_k)|\, d\lambda \\
=\ & O\left(\max_{\lambda_k/2\leq\lambda\leq\lambda_j} |f_{XX}(\lambda)| \int_{\lambda_k/2}^{\lambda_j/2} n^{-1}|D(\lambda_j + \lambda)|\,|D(\lambda + \lambda_k)|\, d\lambda\right) \qquad (24) \\
=\ & O\left(\left(\lambda_k^{-2d} + \lambda_j^{-2d}\right)\frac{\ln n}{j}\right) \\
=\ & O\left(n^{-1}\lambda_j^{-1}\lambda_k^{-2d}\ln n + n^{-1}\lambda_j^{-1-2d}\ln n\right) = O\left(k^{-1}\lambda_j^{-d}\lambda_k^{-d}\ln n\right).
\end{aligned}
$$

7.3.8 Proof of (d)

$$
\begin{aligned}
E\left[w(\lambda_j)\, w(\lambda_k)\right] &= \int_{-\pi}^{\pi} \{f_{XX}(\lambda) - f_{XX}(\lambda_j)\} E_{j,-k}(\lambda)\, d\lambda \\
&= \int_{-\pi}^{\pi} (2\pi n)^{-1}\{f_{XX}(\lambda) - f_{XX}(\lambda_j)\} D(\lambda_j - \lambda) D(\lambda + \lambda_k)\, d\lambda
\end{aligned}
$$

The evaluations are similar to those in the proof of (c). In particular,

$$
\begin{aligned}
&\int_{\lambda_j/2}^{\pi} (2\pi n)^{-1}\{f_{XX}(\lambda) - f_{XX}(\lambda_j)\} D(\lambda_j - \lambda) D(\lambda + \lambda_k)\, d\lambda \\
=\ & O\left(\max_{\lambda_j/2\leq\lambda\leq\pi} |f'_{XX}(\lambda)|\, n^{-1}\int_{\lambda_j/2}^{\pi} |D(\lambda + \lambda_k)|\, d\lambda\right) \\
=\ & O\left(\lambda_j^{-1-2d} n^{-1}\ln n\right) \\
=\ & O\left(k^{-1}\lambda_j^{-d}\lambda_k^{-d}\ln n\right),
\end{aligned}
$$

because $-\pi < \lambda_j - \pi \leq \lambda_j - \lambda \leq \lambda_j/2 < \pi$. Next,

$$
\begin{aligned}
&\int_{-\pi}^{-\lambda_j/2} (2\pi n)^{-1}\{f_{XX}(\lambda) - f_{XX}(\lambda_j)\} D(\lambda_j - \lambda) D(\lambda + \lambda_k)\, d\lambda \\
\leq\ & \int_{\lambda_j/2}^{\pi} (2\pi n)^{-1}\{f_{XX}(\lambda) - f_{XX}(\lambda_j)\} |D(\lambda_j + \lambda)|\,|D(\lambda - \lambda_k)|\, d\lambda
\end{aligned}
$$

$$= O\left(\max_{\lambda_j/2\le\lambda\le\pi}\left|f'_{XX}(\lambda)\right| n^{-1}\int_{\lambda_j/2}^{\pi}\left|D(\lambda-\lambda_k)\right| d\lambda\right)$$
$$= O\left(\lambda_j^{-1-2d} n^{-1}\ln n\right)$$
$$= O\left(k^{-1}\lambda_j^{-d}\lambda_k^{-d}\ln n\right),$$

by $0<\lambda_j+\lambda<2\pi$ and sub-lemma 7.3.3. The same argument as used in (24) yields

$$\int_{-\lambda_j/2}^{-\lambda_k/2}(2\pi n)^{-1}\left\{f_{XX}(\lambda)-f_{XX}(\lambda_j)\right\} D(\lambda_j-\lambda) D(\lambda+\lambda_k)\, d\lambda$$
$$\le \int_{\lambda_k/2}^{\lambda_j/2}(2\pi n)^{-1}\left\{f_{XX}(\lambda)-f_{XX}(\lambda_j)\right\}\left|D(\lambda_j+\lambda)\right|\left|D(\lambda-\lambda_k)\right| d\lambda$$
$$= O\left(\max_{\lambda_k/2\le\lambda\le\lambda_j}\left|f_{XX}(\lambda)\right|\int_{\lambda_k/2}^{\lambda_j/2} n^{-1}\left|D(\lambda_j+\lambda)\right|\left|D(\lambda-\lambda_k)\right| d\lambda\right)$$
$$= O\left(\left(\lambda_k^{-2d}+\lambda_j^{-2d}\right)\frac{\ln n}{j}\right) = O\left(k^{-1}\lambda_j^{-d}\lambda_k^{-d}\ln n\right).$$

Further, the same argument as in (23) yields

$$\int_{-\lambda_k/2}^{\lambda_k/2}(2\pi n)^{-1}\left\{f_{XX}(\lambda)-f_{XX}(\lambda_j)\right\} D(\lambda_j-\lambda) D(\lambda+\lambda_k)\, d\lambda$$
$$= O\left(\frac{1}{n\lambda_j\lambda_k}\int_{-\lambda_k/2}^{\lambda_k/2}\left\{\left|f_{XX}(\lambda)\right|+\left|f_{XX}(\lambda_j)\right|\right\} d\lambda\right)$$
$$= O\left(k^{-1}\lambda_j^{-d}\lambda_k^{-d}\ln n\right).$$

Finally,

$$\int_{\lambda_k/2}^{\lambda_j/2}(2\pi n)^{-1}\left\{f_{XX}(\lambda)-f_{XX}(\lambda_j)\right\} D(\lambda_j-\lambda) D(\lambda+\lambda_k)\, d\lambda$$
$$= O\left(\max_{\lambda_k/2\le\lambda\le\lambda_j}\left|f_{XX}(\lambda)\right|\int_{\lambda_k/2}^{\lambda_j/2} n^{-1}\left|D(\lambda_j-\lambda)\right|\left|D(\lambda+\lambda_k)\right| d\lambda\right)$$
$$= O\left(\left(\lambda_k^{-2d}+\lambda_j^{-2d}\right)\frac{\ln n}{j}\right) = O\left(k^{-1}\lambda_j^{-d}\lambda_k^{-d}\ln n\right),$$

by $0<\lambda_j/2<\lambda_j-\lambda<\pi$ and sub-lemma 7.3.4. ■

7.4 Proof of Lemma 4.3

The following elementary result is useful.

7.4.1 Sub-Lemma

1. $\ln M = O(\ln m)$,

2. $\ln n = O(\ln m)$,

3. $\ln^{-1} n = O(\ln^{-1} m)$.

Proof $n = 2mM$ and $M/m \to 0$ implies

$$\begin{aligned} \ln M &= O(\ln m), \\ \ln n &= \ln 2 + \ln m + \ln M = O(\ln m), \\ \frac{\ln m}{\ln n} &= \frac{\ln m}{\ln 2 + \ln m + \ln M} = O(1). \blacksquare \end{aligned}$$

From the proof of lemma 3.3, we have

$$X_{sj} - \overline{X}_{.j} = O(1/j),$$

uniformly in $1 \leq j < M$. Also, $E(\varepsilon_{sj}) = O(\ln n/s)$ uniformly for $m \leq s < 2/n$ because $\ln^2 n = o(m)$. It follows that

$$\sum_{\{s:\lambda_s \in B_j\}} E(\varepsilon_{sj})(X_{sj} - \overline{X}_{.j}) = O\left(\frac{1}{j}\sum_{s=mj+1}^{mj+m}\frac{\ln n}{s}\right) = O\left(\frac{\ln n}{j^2}\right),$$

because $1/s = O(1/mj)$ in $\{s : \lambda_s \in B_j\}$. Therefore,

$$\sum_{j=1}^{\ell}\sum_{\{s:\lambda_s \in B_j\}} E(\varepsilon_{sj})(X_{sj} - \overline{X}_{.j}) = O\left(\sum_{j=1}^{\ell}\frac{\ln n}{j^2}\right) = O(\ln n) = O(\ln m). \blacksquare$$

7.5 Proof of Lemma 4.4

$\mathrm{Var}(\varepsilon_{sj}) = \pi^2/6 + O(\ln n/s)$ and $\mathrm{Cov}(\varepsilon_{sj}, \varepsilon_{tk}) = O(\ln^2 n/t^2)$ uniformly for $m \leq t < s < n/2$ because $\ln^2 n = o(m)$. Hence

$$\begin{aligned} &\mathrm{Var}\left[\sum_{\{s:\lambda_s \in B_j\}} \varepsilon_{sj}(X_{sj} - \overline{X}_{.j})\right] \\ = &\sum_{\{s:\lambda_s \in B_j\}} (X_{sj} - \overline{X}_{.j})^2 \mathrm{Var}(\varepsilon_{sj}) + 2\sum_{t=mj+1}^{mj+m}\sum_{s=t+1}^{mj+m} (X_{sj} - \overline{X}_{.j})(X_{tj} - \overline{X}_{.j})\,\mathrm{cov}(\varepsilon_{sj}, \varepsilon_{tj}) \end{aligned}$$

$$= \sum_{\{s:\lambda_s \in B_j\}} (X_{sj} - \overline{X}_{.j})^2 \left(\frac{\pi^2}{6} + O\left(\frac{\ln n}{s}\right)\right) + O\left(\frac{1}{j^2} \sum_{t=mj+1}^{mj+m} \sum_{s=t+1}^{mj+m} \frac{\ln^2 n}{t^2}\right)$$

$$= \frac{\pi^2}{6} \sum_{\{s:\lambda_s \in B_j\}} (X_{sj} - \overline{X}_{.j})^2 + O\left(\frac{1}{j^2} \sum_{s=mj+1}^{mj+m} \frac{\ln n}{s}\right) + O\left(\frac{1}{j^2} \sum_{t=mj+1}^{mj+m} \frac{m \ln^2 n}{t^2}\right)$$

$$= \frac{\pi^2}{6} \sum_{\{s:\lambda_s \in B_j\}} (X_{sj} - \overline{X}_{.j})^2 + O\left(\frac{\ln n}{j^3}\right) + O\left(\frac{1}{j^2} \frac{m \ln^2 n}{mj}\right)$$

$$= \frac{\pi^2}{6} \sum_{\{s:\lambda_s \in B_j\}} (X_{sj} - \overline{X}_{.j})^2 + O\left(\frac{\ln^2 n}{j^3}\right)$$

$$= \frac{\pi^2}{6} \sum_{\{s:\lambda_s \in B_j\}} (X_{sj} - \overline{X}_{.j})^2 + O\left(\frac{\ln^2 m}{j^3}\right),$$

because $1/s = O(1/mj)$ and $\sum_{t=mj+1}^{mj+m} \frac{1}{t^2} \sim \int_{mj+1}^{mj+m} \frac{dt}{t^2} = \frac{1}{mj+1} - \frac{1}{mj+m} = O(1/mj)$.

For the covariance, we have for $j > k$

$$\text{Cov}\left[\sum_{\{s:\lambda_s \in B_j\}} \varepsilon_{sj} (X_{sj} - \overline{X}_{.j}), \sum_{\{t:\lambda_t \in B_k\}} \varepsilon_{tk} (X_{tk} - \overline{X}_{.k})\right]$$

$$= \sum_{\{s:\lambda_s \in B_j\}} \sum_{\{t:\lambda_t \in B_k\}} (X_{sj} - \overline{X}_{.j}) (X_{tk} - \overline{X}_{.k}) \text{Cov}(\varepsilon_{sj}, \varepsilon_{tk})$$

$$= \sum_{\{s:\lambda_s \in B_j\}} \sum_{\{t:\lambda_t \in B_k\}} O\left(\frac{1}{jk}\right) O\left(\frac{\ln^2 n}{t^2}\right)$$

$$= O\left(\frac{1}{jk} \sum_{s=mj+1}^{mj+m} \sum_{t=mk+1}^{mk+m} \frac{\ln^2 n}{t^2}\right)$$

$$= O\left(\frac{1}{jk} \frac{m \ln^2 n}{mk}\right) = O\left(\frac{\ln^2 n}{jk^2}\right) = O\left(\frac{\ln^2 m}{jk^2}\right). \blacksquare$$

7.6 Proof of Lemma 4.5

$$\text{Cov}\left[\sum_{\{s:\lambda_s \in B_j\}} \varepsilon_{sj} (X_{sj} - \overline{X}_{.j}), \sum_{\{t:\lambda_t \in B_0\}} \varepsilon_{t0} (X_{t0} - \overline{X}_{.0})\right]$$

$$= \text{Cov}\left[\sum_{\{s:\lambda_s \in B_j\}} \varepsilon_{sj} (X_{sj} - \overline{X}_{.j}), \sum_{t=1}^{\ln^6 n} \varepsilon_{t0} (X_{t0} - \overline{X}_{.0})\right]$$

$$+\text{Cov}\left[\sum_{\{s:\lambda_s \in B_j\}} \varepsilon_{sj} (X_{sj} - \overline{X}_{.j}), \sum_{t=\ln^6 n+1}^{m} \varepsilon_{t0} (X_{t0} - \overline{X}_{.0})\right].$$

First,

$$\begin{aligned}
&\operatorname{Cov}\left[\sum_{\{s:\lambda_s\in B_j\}} \varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right), \sum_{t=1}^{\ln^6 n} \varepsilon_{t0}\left(X_{t0}-\overline{X}_{.0}\right)\right] \\
\leq\ & Var\left[\left(\sum_{\{s:\lambda_s\in B_j\}} \varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right)\right)\right]^{1/2} \left[Var\left(\sum_{t=1}^{\ln^6 n} \varepsilon_{t0}\left(X_{t0}-\overline{X}_{.0}\right)\right)\right]^{1/2} \\
=\ & O\left(\frac{\sqrt{m}\ln^7 m}{j}\right),
\end{aligned}$$

because

$$Var\left(\sum_{t=1}^{\ln^6 n} \varepsilon_{t0}\left(X_{t0}-\overline{X}_{.0}\right)\right) = O\left(\ln^{14} m\right),$$

by HDB p.25, and as shown in the proof of lemma 4.3,

$$\begin{aligned}
Var\left(\sum_{\{s:\lambda_s\in B_j\}} \varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right)\right) &= \frac{\pi^2}{6}\sum_{\{s:\lambda_s\in B_j\}} \left(X_{sj}-\overline{X}_{.j}\right)^2 + O\left(\frac{\ln^2 n}{j^3}\right) \\
&= \frac{\pi^2}{6}\sum_{\{s:\lambda_s\in B_j\}} O\left(\frac{1}{j^2}\right) + O\left(\frac{\ln^2 n}{j^3}\right) \\
&= O\left(\frac{m}{j^2}\right).
\end{aligned}$$

Next,

$$\begin{aligned}
&\operatorname{Cov}\left[\sum_{\{s:\lambda_s\in B_j\}} \varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right), \sum_{t=\ln^6 n+1}^{m} \varepsilon_{t0}\left(X_{t0}-\overline{X}_{.0}\right)\right] \\
=\ & \sum_{\{s:\lambda_s\in B_j\}} \sum_{t=\ln^6 n+1}^{m} \left(X_{sj}-\overline{X}_{.j}\right)\left(X_{t0}-\overline{X}_{.0}\right)\operatorname{Cov}\left(\varepsilon_{sj},\varepsilon_{t0}\right) \\
=\ & O\left(\frac{\ln m}{j}\sum_{t=\ln^6 n+1}^{m}\ \sum_{s=mj+1}^{mj+m} \frac{\ln^2 n}{t^2}\right) \\
=\ & O\left(\frac{\ln m}{j}\sum_{t=\ln^6 n+1}^{m} \frac{m\ln^2 n}{t^2}\right) \\
=\ & O\left(\frac{m\ln m}{j\ln^4 n}\right) = O\left(\frac{m}{j\ln^3 m}\right),
\end{aligned}$$

by $\sum_{t=\ln^6 n+1}^{m}\frac{1}{t^2} \sim \int_{\ln^6 n+1}^{m}\frac{dt}{t^2} = \frac{1}{\ln^6 n+1} - \frac{1}{m} = O\left(\frac{1}{\ln^6 n}\right)$. ∎

7.7 Proof of Lemma 4.6

$\mathrm{Var}\left[\sum_{j=0}^{\ell}\sum_{\{s:\lambda_s\in B_j\}}\varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right)\right]$ can be decomposed into the following parts:

$$\begin{aligned}
&\mathrm{Var}\left[\sum_{j=0}^{\ell}\sum_{\{s:\lambda_s\in B_j\}}\varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right)\right]\\
=&\sum_{j=0}^{\ell}\mathrm{Var}\left[\sum_{\{s:\lambda_s\in B_j\}}\varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right)\right]\\
&+\sum_{j\neq k}^{\ell}\mathrm{Cov}\left[\sum_{\{s:\lambda_s\in B_j\}}\varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right),\sum_{\{t:\lambda_t\in B_k\}}\varepsilon_{tk}\left(X_{tk}-\overline{X}_{.k}\right)\right]\\
=&\sum_{j=0}^{\ell}\mathrm{Var}\left[\sum_{\{s:\lambda_s\in B_j\}}\varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right)\right]\\
&+2\sum_{j=1}^{\ell}\mathrm{Cov}\left[\sum_{\{s:\lambda_s\in B_j\}}\varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right),\sum_{\{t:\lambda_t\in B_0\}}\varepsilon_{t0}\left(X_{t0}-\overline{X}_{.0}\right)\right]\\
&+2\sum_{k=1}^{\ell}\sum_{j=k+1}^{\ell}\mathrm{Cov}\left[\sum_{\{s:\lambda_s\in B_j\}}\varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right),\sum_{\{t:\lambda_t\in B_k\}}\varepsilon_{tk}\left(X_{tk}-\overline{X}_{.k}\right)\right].
\end{aligned}$$

By theorem 1 of HDB,

$$\mathrm{Var}\left[\sum_{\{s:\lambda_s\in B_0\}}\varepsilon_{s0}\left(X_{s0}-\overline{X}_{.0}\right)\right]=\frac{4\pi^2 m}{6}+o(m).$$

For the remaining parts,

$$\begin{aligned}
\sum_{j=1}^{\ell}\mathrm{Var}\left[\sum_{\{s:\lambda_s\in B_j\}}\varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right)\right]&=\sum_{j=1}^{\ell}\left(\frac{\pi^2}{6}\sum_{\{s:\lambda_s\in B_j\}}\left(X_{sj}-\overline{X}_{.j}\right)^2+O\left(\frac{\ln^2 m}{j^3}\right)\right)\\
&=\frac{\pi^2}{6}4m\Xi+o(m)+O\left(\ln^2 m\right)\\
&=\frac{\pi^2}{6}4m\Xi+o(m),
\end{aligned}$$

$$\begin{aligned}
&\sum_{j=1}^{\ell}\mathrm{Cov}\left[\sum_{\{s:\lambda_s\in B_j\}}\varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right),\sum_{\{t:\lambda_t\in B_0\}}\varepsilon_{t0}\left(X_{t0}-\overline{X}_{.0}\right)\right]\\
=&\ O\left(\sqrt{m}\ln^7 m\ln M\right)+O\left(\frac{m\ln M}{\ln^3 m}\right)=o(m),
\end{aligned}$$

and

$$\sum_{k=1}^{\ell}\sum_{j=k+1}^{\ell} \mathrm{Cov}\left[\sum_{\{s:\lambda_s\in B_j\}} \varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right), \sum_{\{t:\lambda_t\in B_k\}} \varepsilon_{tk}\left(X_{tk}-\overline{X}_{.k}\right)\right]$$
$$= \sum_{k=1}^{\ell}\sum_{j=k+1}^{\ell} O\left(\frac{\ln^2 m}{jk^2}\right) = \sum_{k=1}^{\ell} O\left(\frac{\ln^2 m \ln M}{k^2}\right) = O\left(\ln^3 m\right) = o\left(m\right). \blacksquare$$

7.8 Proof of Theorem 4.8

Recall

$$\widehat{d}-d = \frac{\sum_{j=0}^{L}\sum_{\{s:\lambda_s\in B_j\}} \eta_{sj}\left(X_{sj}-\overline{X}_{.j}\right)}{\sum_{j=0}^{L}\sum_{\{s:\lambda_s\in B_j\}}\left(X_{sj}-\overline{X}_{.j}\right)^2} + \frac{\sum_{j=0}^{L}\sum_{\{s:\lambda_s\in B_j\}} \varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right)}{\sum_{j=0}^{L}\sum_{\{s:\lambda_s\in B_j\}}\left(X_{sj}-\overline{X}_{.j}\right)^2}.$$

By lemma 8 of HDB,

$$E\left(\sum_{\{s:\lambda_s\in B_0\}} \varepsilon_{s0}\left(X_{s0}-\overline{X}_{.0}\right)\right) = O\left(\ln^3 m\right).$$

Using this result and lemmas 3.2, 3.3, 4.3 and 4.6, we obtain

$$\begin{aligned}
E\left(\widehat{d}-d\right) &= \frac{-\frac{2\pi^2}{3}\frac{f''_{uu}(0)}{f_{uu}(0)}\frac{m^3 L}{n^2}+o\left(\frac{m^3L}{n^2}\right)}{4m\left(1+\Xi\right)+o\left(m\right)} + \frac{O\left(\ln^3 m\right)+O\left(\ln m\right)}{4m\left(1+\Xi\right)+o\left(m\right)}\\
&= -\frac{\pi^2}{6\left(1+\Xi\right)}\frac{f''_{uu}\left(0\right)}{f_{uu}\left(0\right)}\frac{m^2L}{n^2}+o\left(\frac{m^2L}{n^2}\right)+O\left(\frac{\ln^3 m}{m}\right),\\
Var\left(\widehat{d}\right) &= \left[4m\left(1+\Xi\right)+o\left(m\right)\right]^{-2}\left[\frac{4\pi^2 m}{6}\left(1+\Xi\right)+o\left(m\right)\right]\\
&= \frac{\pi^2}{24\left(1+\Xi\right)m}+o\left(\frac{1}{m}\right).
\end{aligned}$$

For the mean squared error, we have

$$\begin{aligned}
MSE\left(\widehat{d}\right) &= \left\{-\frac{\pi^2}{6\left(1+\Xi\right)}\frac{f''_{uu}\left(0\right)}{f_{uu}\left(0\right)}\frac{m^2L}{n^2}+o\left(\frac{m^2L}{n^2}\right)+O\left(\frac{\ln^3 m}{m}\right)\right\}^2\\
&\quad+\frac{\pi^2}{24\left(1+\Xi\right)m}+o\left(\frac{1}{m}\right)\\
&= \frac{\pi^4}{36\left(1+\Xi\right)^2}\left\{\frac{f''_{uu}\left(0\right)}{f_{uu}\left(0\right)}\right\}^2\frac{m^4L^2}{n^4}+\frac{\pi^2}{24\left(1+\Xi\right)m}\\
&\quad+o\left(\frac{m^2L}{n^2}\right)\left(\frac{m^2L}{n^2}\right)+O\left(\frac{\ln^3 m}{m}\right)\left(\frac{\ln^3 m}{m}\right)+O\left(\frac{m^2L}{n^2}\frac{\ln^3 m}{m}\right)+o\left(\frac{1}{m}\right)\\
&= \frac{\pi^4}{36\left(1+\Xi\right)^2}\left\{\frac{f''_{uu}\left(0\right)}{f_{uu}\left(0\right)}\right\}^2\frac{m^4L^2}{n^4}+\frac{\pi^2}{24\left(1+\Xi\right)m}\\
&\quad+o\left(\frac{m^4L^2}{n^4}\right)+O\left(\frac{\ln^6 m}{m^2}\right)+O\left(\frac{mL\ln^3 m}{n^2}\right)+o\left(\frac{1}{m}\right).
\end{aligned}$$

7.9 Proof of Lemma 5.1

First, we modify the propositions in Theorem 2 of Robinson (1995, p.1056). The assumptions $f'_{uu}(0) = 0$ and $|f''_{uu}(\omega)| < B_2 < \infty$ imply $\alpha = 2$. The univariate model in our case implies there is no β here. Therefore, (4.2) on p.1060 in Robinson changes to

$$\rho(\lambda_j) - C_g \lambda_j^{-\delta} = O\left[\left(\frac{j}{n}\right)^2 \lambda_j^{-\delta}\right]$$

and (a)-(d) in his Theorem 2 become

$$\begin{aligned}
(a') \quad E\left[v(\lambda_j)\bar{v}(\lambda_j)\right] &= 1 + O\left[\frac{\ln j}{j} + \left(\frac{j}{n}\right)^2\right] \\
(b') \quad E\left[v(\lambda_j)v(\lambda_j)\right] &= O\left(\frac{\ln j}{j}\right) \\
(c') \quad E\left[v(\lambda_j)\bar{v}(\lambda_k)\right] &= O\left(\frac{\ln j}{k}\right), \quad j > k \\
(d') \quad E\left[v(\lambda_j)v(\lambda_k)\right] &= O\left(\frac{\ln j}{k}\right), \quad j > k
\end{aligned}$$

where

$$\lambda_k = 2\pi k/n, \quad v(\lambda) = w(\lambda)/\left(C_g^{1/2}\lambda^{-d}\right), \quad w(\lambda) = (2\pi n)^{-1/2}\sum\nolimits_1^n X_t e^{it\lambda}.$$

In the following, we repeat Robinson's argument (1995, pp.1067-70) with corresponding modifications, although we try to keep the derivations here as self-contained as possible. Write $\chi_k = m^{-1/2}a_k U_k$. Fix an integer N. Then, $E\left(\sum_k \chi_k\right)^N$ is a sum of finitely many terms of the form

$$\sum_{k_1} \cdots \sum_{k_K} E\left(\prod_{i=1}^{K} \chi_{k_i}^{N_{k_i}}\right), \tag{25}$$

where $N_{k_1}, ..., N_{k_K}$ are all positive and sum to N and $1 \le K \le N$. Fix such a K and $N_{k_1}, ..., N_{k_K}$. Introduce the 2−vector $v_k^* = \left(v^R(\lambda_k), v^I(\lambda_k)\right)'$ and $2K \times 1$ vector $v^* = \left(v_{k_1}^{*\prime}, ..., v_{k_K}^{*\prime}\right)$, which is normally distributed with zero mean.

Theorem 2 of Robinson (1995) implies that

$$\begin{aligned}
E\left(v_j^* v_k^*\right) &= R + O\left(\frac{\ln j}{j} + \left(\frac{j}{n}\right)^2\right), \quad j = k, \quad R = \frac{1}{2}I_2, \\
&= O\left(\frac{\ln j}{k}\right), \quad j > k,
\end{aligned}$$

as $n \to \infty$. It follows that, $\Sigma^* = E(v^* v^{*\prime})$ satisfies

$$\Sigma^* = I_K \otimes R + O\left(\frac{\ln \ell m}{m^{0.5+\delta}} + \left(\frac{\ell m}{n}\right)^2\right) = I_K \otimes R + o\left(m^{-1/2-\Delta}\right), \tag{26}$$

as $n \to \infty$.

For n sufficiently large, $\Psi = \Sigma^{*-1}$ exists by (26). Denote by Ψ_{ij} the (i,j) th 2×2 submatrix of Ψ and write

$$\widetilde{\Psi} = \begin{bmatrix} \Psi_{11} & & 0 \\ & \ddots & \\ 0 & & \Psi_{KK} \end{bmatrix}, \quad \overline{\Psi} = \Psi - \widetilde{\Psi}.$$

It follows that

$$\widetilde{\Psi} = I_K \otimes R^{-1} + o\left(m^{-1/2-\Delta}\right), \quad \overline{\Psi} = o\left(m^{-1/2-\Delta}\right) \quad \text{as} \quad n \to \infty. \tag{27}$$

Now denote by φ_p the density function of a p-dimensional standard normal variate. Then (25) is

$$\sum_{k_1} \cdots \sum_{k_K} |\Psi|^{1/2} \int \left(\prod_{i=1}^{K} \chi_{k_i}^{N_{k_i}}\right) \varphi_{2K}\left(\Psi^{1/2} v^*\right) dv^*, \tag{28}$$

for n sufficiently large. Consider

$$\sum_{k_1} \cdots \sum_{k_K} |\Psi|^{1/2} \prod_{i=1}^{K} \left\{ \int \chi_{k_i}^{N_{k_i}} \varphi_2\left(\Psi_{ii}^{1/2} v_{k_i}^*\right) dv_{k_i}^* \right\}. \tag{29}$$

The difference between (28) and (29) is

$$\sum_{k_1} \cdots \sum_{k_K} |\Psi|^{1/2} \int \left(\prod_{i=1}^{K} \chi_{k_i}^{N_{k_i}}\right) \varphi_{2K}\left(\widetilde{\Psi}^{1/2} v^*\right) \left\{\exp\left(-\frac{1}{2} v^{*\prime} \overline{\Psi} v^*\right) - 1\right\} dv^*. \tag{30}$$

For any positive integer r, the mean value theorem indicates that $\left|e^u - \sum_{t=0}^{r-1} u^t/t!\right| \leq |u|^r e^{|u|}/r!$, for all u. For all $\varepsilon > 0$ there exists $C_\varepsilon < \infty$ such that $|u|^r \leq C_\varepsilon e^{|u|}$ for all u. Following (27), choose n so large that $\|\overline{\Psi}\| < \varepsilon$, where $\|\cdot\|$ is the Euclidean norm, that is, $\exp\left(\frac{1}{2}\left|v^{*\prime}\overline{\Psi} v^*\right|\right) \leq \exp\left(\frac{1}{2}\varepsilon \|v^*\|^2\right)$. Again using (27), $\left|v^{*\prime}\overline{\Psi} v^*\right| = o\left(m^{-1/2-\Delta}\|v^*\|^2\right)$ uniformly in v^*. Thus

$$\left|\exp\left(-\frac{1}{2} v^{*\prime}\overline{\Psi} v^*\right) - \sum_{t=0}^{r-1} \frac{\left(-\frac{1}{2} v^{*\prime}\overline{\Psi} v^*\right)^t}{t!}\right| = o\left(m^{-r/2-r\Delta} \exp\left(\varepsilon \|v^*\|^2\right)\right),$$

as $n \to \infty$, uniformly in v^*. Thus the difference between (30) and

$$\sum_{k_1} \cdots \sum_{k_K} |\Psi|^{1/2} \int \left(\prod_{i=1}^{K} \chi_{k_i}^{N_{k_i}}\right) \varphi_{2K}\left(\widetilde{\Psi}^{1/2} v^*\right) \sum_{t=1}^{r-1} \frac{\left(-\frac{1}{2} v^{*\prime}\overline{\Psi} v^*\right)^t}{t!} dv^* \tag{31}$$

is

$$o\left(m^{-r/2-r\Delta}\sum_{k_1}\cdots\sum_{k_K}|\Psi|^{1/2}\int\prod_{i=1}^{K}\left|\chi_{k_i}^{N_{k_i}}\right|\exp\left\{-\frac{1}{2}v^{*\prime}\left(\widetilde{\Psi}-\varepsilon I_{2K}\right)v^{*}\right\}dv^{*}\right). \qquad (32)$$

In view of (26), $|\Psi| = O(1)$, while $\frac{1}{2}v^{*\prime}\left(\widetilde{\Psi}-\varepsilon I_{2K}\right)v^{*} > \frac{1}{2}\eta\|v^{*}\|^{2}$ for some $\eta > 0$. Because $\|\chi_k\| \leq m^{-1/2}|a_k|\,\|U_k\|$, we deduce from the finiteness of the moments of all orders of the log of a chi-squared variate that (32) is $o\left(\ln^{K}\ell\cdot m^{K-N/2-r/2-r\Delta}\right) \to 0$ on choosing $r = \max(2K-N, 1)$.

Now (31) makes a contribution only when such $r \geq 2$, which occurs only when $2K-N \geq 2$. Let D be the number of N_{k_i} which equal 1. Clearly $D \geq 2K-N$, that is, $D > t$ for $t = 1,...,r-1 = 2K-N-1$ in (31). Note that $v^{*\prime}\widetilde{\Psi}v^{*}$ is bilinear in the $v_{k_i}^{*}$ and for each $t = 1,...,r-1$, hence $\left(v^{*\prime}\widetilde{\Psi}v^{*}\right)^{t}$ cannot involve more than t of the $v_{k_i}^{*}$. The corresponding t or fewer k_i can overlap with the Dk_i for which $N_{k_i} = 1$, but because $D > r-1$, the $(k_1,...,k_K,t)'$th summand in (31) can be written

$$|\Psi|^{1/2}\prod_{i=1}^{D-t}\left(\int\chi_{k_i}\varphi_2\left(\Psi_{ii}^{1/2}v_{k_i}^{*}\right)dv_{k_i}^{*}\right) \qquad (33)$$

$$\times\int\frac{\left(-1/2v^{*\prime}\widetilde{\Psi}v^{*}\right)^{t}}{t!}\left(\prod_{i=D-t+1}^{K}\chi_{k_i}^{N_{k_i}}\varphi_2\left(\Psi_{ii}^{1/2}v_{k_i}^{*}\right)dv_{k_i}^{*}\right). \qquad (34)$$

From (27),

$$\varphi_2\left(\Psi_{ii}^{1/2}v_{k_i}^{*}\right) = \varphi_2\left(R^{-1/2}v_{k_i}^{*}\right)\left(1+o\left(m^{-1/2-\Delta}\right)\|v_{k_i}^{*}\|^{2}\right), \qquad (35)$$

uniformly in $v_{k_i}^{*}$ and $\int\chi_k\varphi_2\left(R^{-1/2}v_k^{*}\right)dv_k^{*} = 0$. For all positive p and q, uniformly in k,

$$\int\left\|m^{1/2}\chi_k\right\|^{p}\|v_k^{*}\|^{q}\varphi_2\left(R^{-1/2}v_k^{*}\right)dv_k^{*} = O\left(|a_k|^{p}\right).$$

Thus (33) is $o\left(m^{-(D-t)-\Delta(D-t)}\prod_{i=1}^{D-t}|a_{k_i}|\right)$ and (34) is $o\left(m^{-t-(N-D)/2-t\Delta}\prod_{i=D-t+1}^{K}|a_{k_i}|^{N_{k_i}}\right)$. It follows from the third part of (14) that (31) is $o\left(\ln^{K}\ell\cdot m^{K-N/2-D/2-\Delta D}\right) \to 0$ as $n\to\infty$. Thus, we have shown that (30) $\to 0$ as $n \to \infty$. Now from (27), $|\Psi| = |R|^{-K} + o\left(m^{-1/2-\Delta}\right)$ and

$$|R|^{-1/2}\int\chi_{k_i}^{N_{k_i}}\varphi_2\left(\Psi_{ii}^{1/2}v_{k_i}^{*}\right)dv_{k_i}^{*} = \mu_{k_i}^{(N_{k_i})}\left(1+o\left(m^{-1/2-\Delta}\right)\right),$$

where $\mu_k^{(p)} = |R|^{-1/2}\int\chi_k^{p}\varphi_2\left(R^{-1/2}v_k^{*}\right)dv_k^{*}$. The difference between (29) and

$$\sum_{k_1}\cdots\sum_{k_K}\prod_{i=1}^{K}\mu_{k_i}^{(N_{k_i})} \qquad (36)$$

is readily seen to be $o\left(\ln^K \ell \cdot m^{K-N/2-(1/2)\max(1,D)-\Delta}\right)$ using (35), and using $K - N/2 - D/2 \leq 0$ when $D \geq 1$ and $K \leq N/2$ when $D = 0$. However, (36) is

$$E\left[\sum_{k_1}\cdots\sum_{k_K}\prod_{i=1}^{K}\left(\frac{a_{k_i}W_{k_i}}{m^{1/2}}\right)^{N_{k_i}}\right].$$

Therefore, we have shown that the moments of $\sum_k \chi_k$ differ negligibly from those of the variate $m^{-1/2}\sum_k a_k W_k$, which converges in distribution to $N\left(0, \psi'(1)\right)$ upon applying (14), $W_k \sim iid\left(0, \psi'(1)\right)$ and the Lindeberg-Feller CLT. ■

7.10 Proof of Lemma 5.2

We adopt the argument from HDB, theorem 2. Let

$$m^{-1/2}\sum_{j=0}^{L}\sum_{\{s:\lambda_s\in B_j\}}\varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right)\equiv T_1+T_2+T_3,$$

where

$$\begin{aligned} T_1 &= \frac{1}{m^{1/2}}\sum_{s=1}^{\ln^8 m}\varepsilon_{s0}\left(X_{s0}-\overline{X}_{.0}\right),\\ T_2 &= \frac{1}{m^{1/2}}\sum_{s=1+\ln^8 m}^{m^{0.5+\delta}}\varepsilon_{s0}\left(X_{s0}-\overline{X}_{.0}\right),\\ T_3 &= \frac{1}{m^{1/2}}\sum_{s=1+m^{0.5+\delta}}^{m}\varepsilon_{s0}\left(X_{s0}-\overline{X}_{.0}\right)+\frac{1}{m^{1/2}}\sum_{j=1}^{L}\sum_{\{s:\lambda_s\in B_j\}}\varepsilon_{sj}\left(X_{sj}-\overline{X}_{.j}\right), \end{aligned}$$

where $0 < \delta < 0.5$. Note that the condition required for theorem 2 of HDB, i.e., $m = o\left(n^{4/5}\right)$, is satisfied by our assumption. HDB show that $T_1 = o_p(1)$ and $T_2 = o_p(1)$.

We now prove that T_3 is asymptotically normal. Let

$$U_s = \ln I_s - \ln f_{uu}(0) - \psi'(1) + 2d\ln\lambda_s,$$

as defined in Equation (2,4) of Robinson (1995). Then we have

$$U_s = \varepsilon_s + \ln\left\{\frac{f_{uu}(\lambda_s)}{f_{uu}(0)}\right\} - 2d\ln\left\{\frac{|1-\exp(-i\lambda_s)|}{\lambda_s}\right\},$$

where $\varepsilon_s = \varepsilon_s$. (we drop the second subscript of ε_{sj}). Hence,

$$T_3 \equiv T_{31} + T_{32} + T_{33},$$

where

$$\begin{aligned}
T_{31} &= \frac{1}{m^{1/2}}\sum_{s=1+m^{0.5+\delta}}^{m} U_s\left(X_{s0}-\overline{X}_{.0}\right)+\frac{1}{m^{1/2}}\sum_{j=1}^{L}\sum_{\{s:\lambda_s\in B_j\}} U_{sj}\left(X_{sj}-\overline{X}_{.j}\right),\\
T_{32} &= -\frac{1}{m^{1/2}}\sum_{s=1+m^{0.5+\delta}}^{m}\ln\left\{\frac{f_{uu}(\lambda_s)}{f_{uu}(0)}\right\}\left(X_{s0}-\overline{X}_{.0}\right)\\
&\quad -\frac{1}{m^{1/2}}\sum_{j=1}^{L}\sum_{\{s:\lambda_s\in B_j\}}\ln\left\{\frac{f_{uu}(\lambda_s)}{f_{uu}(0)}\right\}\left(X_{sj}-\overline{X}_{.j}\right),\\
T_{33} &= \frac{2d}{m^{1/2}}\sum_{s=1+m^{0.5+\delta}}^{m}\ln\left\{\frac{|1-\exp(-i\lambda_s)|}{\lambda_s}\right\}\left(X_{s0}-\overline{X}_{.0}\right)\\
&\quad +\frac{2d}{m^{1/2}}\sum_{j=1}^{L}\sum_{\{s:\lambda_s\in B_j\}}\ln\left\{\frac{|1-\exp(-i\lambda_s)|}{\lambda_s}\right\}\left(X_{sj}-\overline{X}_{.j}\right).
\end{aligned}$$

HDB show that the first term of T_{32} is $o(1)$ and $\ln\{f_{uu}(\lambda_s)/f_{uu}(0)\}=O\left(s^2/n^2\right)$ uniformly for $1\leq s\leq\ m\ln M$. Hence the second term is, by $s=O(mj)$,

$$\begin{aligned}
O\left(\frac{1}{m^{1/2}}\sum_{j=1}^{L}\sum_{\{s:\lambda_s\in B_j\}}\frac{s^2}{n^2}\left(X_{sj}-\overline{X}_{.j}\right)\right) &= O\left(\frac{1}{m^{1/2}}\sum_{j=1}^{L}\sum_{\{s:\lambda_s\in B_j\}}\frac{s^2}{n^2}\frac{1}{j}\right)\\
&= O\left(\frac{1}{m^{1/2}}\sum_{j=1}^{L}\sum_{\{s:\lambda_s\in B_j\}}\frac{jm^2}{n^2}\right)\\
&= O\left(\frac{m^{2.5}L^2}{n^2}\right),
\end{aligned}$$

which is $o(1)$ by $m^5=O\left(n^{4-5\epsilon}\right)$ and $L=O(\ln M)$. Hence, $T_{32}=o(1)$.

HDB show that the first term of T_{33} is $o(1)$. For the second term, from the proof of lemma 3.2(b), we have for $m\leq s\leq m(L+1)$

$$\ln\left|1-e^{i\lambda_s}\right|=-\frac{1}{2}X_{sj}=\ln\lambda_s+O\left(\frac{j^2}{M^2}\right),$$

uniformly in s. Hence,

$$\ln\left\{\frac{|1-\exp(-i\lambda_s)|}{\lambda_s}\right\}=O\left(\frac{j^2}{M^2}\right),$$

uniformly in s. Thus, the second term

$$\begin{aligned}
&\frac{2d}{m^{1/2}}\sum_{j=1}^{L}\sum_{\{s:\lambda_s\in B_j\}}\ln\left\{\frac{|1-\exp(-i\lambda_s)|}{\lambda_s}\right\}\left(X_{sj}-\overline{X}_{.j}\right)\\
&= O\left(\frac{1}{m^{1/2}}\sum_{j=1}^{L}\sum_{\{s:\lambda_s\in B_j\}}\frac{j^2}{M^2}\frac{1}{j}\right)
\end{aligned}$$

$$= O\left(\frac{mL^2}{m^{1/2}M^2}\right)$$
$$= O\left(\frac{m^{2.5}L^2}{n^2}\right) = o(1).$$

Thus, $T_{33} = o(1)$.

Finally, we need to prove that

$$T_{31} = \frac{1}{m^{1/2}} \sum_{s=1+m^{0.5+\delta}}^{m} U_s \left(X_{s0} - \overline{X}_{.0}\right) + \frac{1}{m^{1/2}} \sum_{j=1}^{L} \sum_{\{s:\lambda_s \in B_j\}} U_{sj} \left(X_{sj} - \overline{X}_{.j}\right)$$
$$\xrightarrow{d} N\left(0, 4\pi^2 (1+\Xi)/6\right).$$

For $s = 1 + m^{0.5+\delta}, ..., m$, HDB show that $\left(X_{sj} - \overline{X}_{.j}\right)/4$ satisfies the condition (14) of lemma 5.2. For $s = m+1, ..., m(L+1)$, we can use the argument in the proof of lemma 3.2 to obtain

$$\left|X_{sj} - \overline{X}_{.j}\right| = O(1/j) = o(m), \quad \frac{1}{4\Xi} \sum_{j=1}^{L} \sum_{\{s:\lambda_s \in B_j\}} \left(X_{sj} - \overline{X}_{.j}\right)^2 \sim m,$$

and

$$\sum_{j=1}^{L} \sum_{\{s:\lambda_s \in B_j\}} \left|X_{sj} - \overline{X}_{.j}\right|^p = \sum_{j=1}^{L} \sum_{\{s:\lambda_s \in B_j\}} O\left(\frac{1}{j^p}\right) = O\left(\sum_{j=1}^{L} \frac{m}{j^p}\right) = O(m \ln L).$$

Hence $\left(X_{sj} - \overline{X}_{.j}\right)/4(1+\Xi)$, $s = 1 + m^{0.5+\delta}, ..., m(L+1)$, satisfies the condition (14) of lemma 5.2. Also $\ln M$ satisfies the condition for ℓ in Lemma 5.2, because $\ell^2 m^2 m^{0.5+\Delta}/n^2 = O\left(n^{-2.5\epsilon} m^{\Delta} \ln^2 M\right) \to 0$ by a proper choice of Δ and $\ln^K \ln M / m^{\Delta} \to 0$ for any $K > 0$. Therefore,

$$\frac{1}{m^{1/2}} \sum_{s=1+m^{0.5+\delta}}^{m} U_s \left(X_{s0} - \overline{X}_{.0}\right) + \frac{1}{m^{1/2}} \sum_{j=1}^{L} \sum_{\{s:\lambda_s \in B_j\}} U_{sj} \left(X_{sj} - \overline{X}_{.j}\right) \xrightarrow{d} N\left(0, 4\pi^2 (1+\Xi)/6\right).$$

■

7.11 Proof of Theorem 5.3

Let

$$m^{1/2}\left(\widehat{d} - d\right) = \frac{m^{1/2} \sum_{j=0}^{L} \sum_{\{s:\lambda_s \in B_j\}} \eta_{sj} \left(X_{sj} - \overline{X}_{.j}\right)}{\sum_{j=0}^{L} \sum_{\{s:\lambda_s \in B_j\}} \left(X_{sj} - \overline{X}_{.j}\right)^2}$$
$$+ \frac{m}{\sum_{j=0}^{L} \sum_{\{s:\lambda_s \in B_j\}} \left(X_{sj} - \overline{X}_{.j}\right)^2} \frac{\sum_{j=0}^{L} \sum_{\{s:\lambda_s \in B_j\}} \varepsilon_{sj} \left(X_{sj} - \overline{X}_{.j}\right)}{m^{1/2}}$$
$$= V_1 + V_2.$$

Now

$$\begin{aligned} V_1 &= O\left(\frac{m^{1/2}}{m}\frac{m^3}{n^2}\right) = O\left(\frac{m^{5/2}}{n^2}\right) = o(1), \\ V_2 &= \frac{m}{4(1+\Xi)m + o(m)} \frac{\sum_{j=0}^{L}\sum_{\{s:\lambda_s \in B_j\}}\left(X_{sj} - \overline{X}_{.j}\right)\varepsilon_{sj}}{m^{1/2}} \xrightarrow{d} N\left(0, \frac{\pi^2}{24(1+\Xi)}\right), \end{aligned}$$

giving the required result. ■

8 Monte Carlo simulation results

Table 2. G $((a_1, a_2) = (-0.5, 0.0))$, $d = 0.3$

	BIAS	VAR	MSE	BIAS	VAR	MSE	BIAS	VAR	MSE
	$n = 200$ $(m = 31)$			$n = 500$ $(m = 56)$			$n = 1000$ $(m = 89)$		
$\widehat{d}_{GPH}$	-0.0221	0.0156	0.0161	-0.0190	0.0073	0.0077	0.0252	0.0065	0.0071
$\widehat{d}_{pooled}$	-0.0580	0.0148	0.0182	-0.0400	0.0071	0.0087	0.0167	0.0068	0.0070
$\widehat{d}_{GPH}(3m)$	-0.3242	0.0051	0.1102	-0.1439	0.0020	0.0228	-0.0800	0.0014	0.0078
$\widehat{d}_{pooled}(m)$	0.0918	0.0647	0.0731	0.0246	0.0176	0.0182	0.0207	0.0092	0.0096

Table 3. H $((a_1, a_2) = (0.0, 0.0))$, $d = 0.3$

	BIAS	VAR	MSE	BIAS	VAR	MSE	BIAS	VAR	MSE
	$n = 200$ $(m = 31)$			$n = 500$ $(m = 56)$			$n = 1000$ $(m = 89)$		
$\widehat{d}_{GPH}$	0.0026	0.0154	0.0154	-0.0035	0.0073	0.0073	0.0343	0.0065	0.0077
$\widehat{d}_{pooled}$	0.0019	0.0148	0.0148	-0.0046	0.0071	0.0071	0.0374	0.0066	0.0080
$\widehat{d}_{GPH}(3m)$	-0.0478	0.0050	0.0073	-0.0118	0.0021	0.0022	-0.0015	0.0014	0.0014
$\widehat{d}_{pooled}(m)$	0.1174	0.0708	0.0846	0.0323	0.0177	0.0187	0.0278	0.0096	0.0104

Table 4. I $((a_1, a_2) = (0.5, 0.0))$, $d = 0.3$

	BIAS	VAR	MSE	BIAS	VAR	MSE	BIAS	VAR	MSE
	$n = 200$ $(m = 31)$			$n = 500$ $(m = 56)$			$n = 1000$ $(m = 89)$		
$\widehat{d}_{GPH}$	0.1697	0.0158	0.0446	0.0932	0.0075	0.0162	0.0992	0.0065	0.0163
$\widehat{d}_{pooled}$	0.1936	0.0152	0.0527	0.1203	0.0072	0.0217	0.1281	0.0066	0.0230
$\widehat{d}_{GPH}(3m)$	0.3160	0.0049	0.1047	0.2888	0.0021	0.0855	0.2461	0.0014	0.0620
$\widehat{d}_{pooled}(m)$	0.2059	0.0712	0.1136	0.0793	0.0174	0.0237	0.0593	0.0097	0.0132

Table 5. A $((a_1, a_2) = (-0.6, -0.6))$, $d = 0.3$

	BIAS	VAR	MSE	BIAS	VAR	MSE	BIAS	VAR	MSE
	$n = 200$ $(m = 31)$			$n = 500$ $(m = 56)$			$n = 1000$ $(m = 89)$		
$\widehat{d}_{GPH}$	-0.0988	0.0151	0.0249	-0.0539	0.0073	0.0102	0.0063	0.0067	0.0067
$\widehat{d}_{pooled}$	-0.1629	0.0148	0.0413	-0.1240	0.0070	0.0224	-0.0376	0.0068	0.0083
$\widehat{d}_{GPH}(3m)$	-0.5242	0.0042	0.2789	-0.4990	0.0020	0.2510	-0.2886	0.0015	0.0848
$\widehat{d}_{pooled}(m)$	0.0352	0.0539	0.0552	0.0052	0.0170	0.0170	0.0096	0.0089	0.0090

Table 6. B $((a_1, a_2) = (0.6, -0.6))$, $d = 0.3$

	BIAS	VAR	MSE	BIAS	VAR	MSE	BIAS	VAR	MSE
	$n = 200$ $(m = 31)$			$n = 500$ $(m = 56)$			$n = 1000$ $(m = 89)$		
$\widehat{d}_{GPH}$	-0.2376	0.0152	0.0717	-0.1079	0.0074	0.0190	-0.0261	0.0066	0.0073
$\widehat{d}_{pooled}$	-0.1286	0.0147	0.0313	-0.1003	0.0071	0.0171	-0.0384	0.0067	0.0082
$\widehat{d}_{GPH}(3m)$	0.1692	0.0042	0.0328	-0.0821	0.0021	0.0088	-0.2384	0.0014	0.0582
$\widehat{d}_{pooled}(m)$	0.0127	0.0731	0.0732	-0.0156	0.0181	0.0183	-0.0011	0.0094	0.0094

Table 7. C $((a_1, a_2) = (1.0, -0.6))$, $d = 0.3$

	BIAS	VAR	MSE	BIAS	VAR	MSE	BIAS	VAR	MSE
	$n = 200$ $(m = 31)$			$n = 500$ $(m = 56)$			$n = 1000$ $(m = 89)$		
$\widehat{d}_{GPH}$	-0.2478	0.0153	0.0767	-0.1574	0.0074	0.0322	-0.0585	0.0066	0.0101
$\widehat{d}_{pooled}$	-0.1350	0.0148	0.0330	-0.0604	0.0072	0.0108	-0.0179	0.0066	0.0069
$\widehat{d}_{GPH}(3m)$	0.4783	0.0051	0.2339	0.2561	0.0021	0.0677	0.0635	0.0015	0.0055
$\widehat{d}_{pooled}(m)$	0.0018	0.0728	0.0728	-0.0353	0.0187	0.0200	-0.0155	0.0094	0.0096

References

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

[2] Geweke, J., and S. Porter-Hudak (1983). "The Estimation and Application of Long Memory Time Series Models." *Journal of Time Series Analysis* **4**: 221-38.

[3] Gradshteyn, I. S., and I. M. Ryzhik (1965). *Tables of Integrals, Series and Products.* New York, Academic Press.

[4] Hurvich, C. M., and K. I. Beltrao (1994). "Automatic Semiparametric Estimation of the Memory Parameter of a Long-Memory Time Series." *Journal of Time Series Analysis* **15**: 285-302.

[5] Hurvich, C. M., R. Deo, and J. Brodsky (1998). "The Mean Squared Error of Geweke and Porter-Hudak's Estimator of the Memory Parameter of a Long-Memory Time Series." *Journal of Time Series Analysis* **19**: 19-46.

[6] Künsh, H. (1986). "Discrimination between Monotonic Trends and Long-Range Dependence." *Journal of Applied Probability* **23**: 1025-30.

[7] Robinson, P. M. (1995). "Log-Periodogram Regression of Time Series with Long Range Dependence." *Annals of Statistics* **23**: 1048-72.

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