MULTIVARIATE ANALYSIS OF LONG MEMORY SERIES IN THE FREQUENCY DOMAIN

Ignacio Norberto Lobato Garcia

1995

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS OF THE DEGREE OF DOCTOR OF PHILOSOPHY IN ECONOMICS AT THE LONDON SCHOOL OF ECONOMICS

Abstract

This thesis examines some statistical procedures in the frequency domain to analyze long-memory series.

We define a long-memory series and review part of the literature. Then we proceed by analyzing different estimation procedures for H, the parameter that characterizes the existence of long-memory.

Parametric estimates have as a main drawback that they can lead to inconsistent estimates of H if the parametric model is misspecified. Therefore we focus on semiparametric estimates in the frequency domain. In our case, semiparametric means that we only need to assume a parametric model for the spectral density in a neighbourhood of zero frequency.

We focus mainly on a multivariate framework. First we analyze estimates based on the average periodogram. We prove the consistency of the average cross-periodogram for the cumulative cross-spectrum. We also establish the asymptotic distribution in the scalar case. Then we focus on an implicit estimate based on a discrete approximation of the Gaussian likelihood in a neighbourhood of zero frequency. We prove the consistency and asymptotic normality of this estimate. Based on this estimate we establish a Lagrange multiplier test for weak dependence.

We finish with an application of these methods to financial data.

Acknowledgements

I am very grateful to Peter Robinson for being my supervisor. His generous help went far beyond the task of a supervisor. He provided useful suggestions, ongoing encouragement and insightful comments at all stages of the work. I am indebted to Javier Hidalgo for many friendly discussions and encouragement during these years. I thank E.J. Godolphin and P. Newbold for some helpful comments. Financial support from the Bank of Spain is gratefully acknowledged.

TABLE OF CONTENTS

Chapter page

LIST OF TABLES AND FIGURES

Chapter page

6. A LAGRANGE MULTIPLIER TEST FOR I(0)

Chapter 1

General Introduction

1.1 Introduction

ARIMA[1] modelling has been used extensively in applied econometric work with time series data during the last twenty years. Basically, ARIMA analysis removes the nonstationary component of a series by taking a suitable number of differences and explains the stationary component by an ARMA model. Apart from the arguable way of analyzing the nonstationary component, a drawback of this procedure is the limitation of ARMA models to characterize stationary processes.

The Wold decomposition theorem states that any covariance stationary stochastic process can be decomposed as the sum of two uncorrelated processes, one that can be predicted from its own past with zero prediction variance and one that can be expressed as:

$$x_t=\sum_{j=0}^{\infty} \psi_j \epsilon_{t-j} , \tag{1.1.1}$$

where ϵ_t is white noise, $E\epsilon_t=0$, $E\epsilon_t\epsilon_s=\sigma^2$ for t=s, 0 otherwise, with $\psi_0=1$ and $\sum_{j=0}^{\infty} \psi_j^2<\infty$.

As a model, (1.1.1) possesses infinitely many parameters so it cannot be estimated from a finite sample. Box-Jenkins' justification of ARMA models was to approximate (1.1.1) by an ARMA model:

$$\sum_{j=0}^{p} \phi_j x_{t-j}=\sum_{j=0}^{q} \psi_j \epsilon_{t-j} , \tag{1.1.2}$$

[1]ARIMA stands for autoregressive integrated moving average and ARMA for autoregressive moving average.

where (1.1.2) is a parsimonious representation of (1.1.1).

But ARMA models are not general enough to include any covariance stationary process. In fact, an ARMA model can only reflect a process whose autocovariances, eventually, decay exponentially[2]. This is an important limitation of this kind of model, it cannot reflect an eventual hyperbolic decay of the autocovariances, for instance. In fact, this slow hyperbolic decay is a feature of some long-memory models. Long-memory models are a wide class of time series models, in this thesis we will focus on analyzing covariance stationary long-memory models.

In this chapter we will define and show examples of long-memory models and then, we will review some of the literature concerning these processes.

1.2 Definitions

Let x_t be a covariance stationary stochastic process with mean μ and autocovariances γ_j, that is:

$$\mu = Ex_1 \ , \quad \gamma_j = E(x_1-\mu)(x_{1+j}-\mu)$$

and assume it has a spectral density function $f(\lambda)$ defined by:

$$\gamma_j = \int_{-\pi}^{\pi} f(\lambda)\cos j\lambda d\lambda.$$

Introduce[3] also condition 1.A:

$$f(\lambda) \sim C_1\lambda^{1-2H} \quad as \quad \lambda \to 0^+ \ , \quad C_1>0 \ , \quad 0<H<1 \ , \quad H\neq 1/2 \qquad (1.2.1)$$

[2]In an ARMA model the eventual behaviour of the autocovariances is determined by a stable finite order difference equation, Yule-Walker, whose solution is a linear combination of exponentials.

[3]Where the symbol $\sim$ means that, as $\lambda \to 0^+$, the ratio of the left hand side (LHS), $f(\lambda)$, and the right hand side (RHS), $C_1 \lambda^{1-2H}$, tends to 1.

and condition 1.B:

$$\gamma_j \sim C_2\, j^{2H-2}\ ,\quad as\quad j\to\infty\ ,\ |C_2|<\infty\ ,\quad 0<H<1\ ,\ H\neq 1/2. \tag{1.2.2}$$

We say that x_t has long-memory, broadly speaking, when condition 1.A or condition 1.B are satisfied. Other names used in the literature instead of long-memory are long-range dependence and strong dependence.

These definitions include two different cases: a) when $H\in(1/2,1)$ and b) when $H\in(0,1/2)$. The first is called, strictly speaking, long-memory[4], the spectral density will tend to infinity as it is evaluated at frequencies approaching 0. In case b) the spectral density will be zero at zero frequency, this case has been called antipersistent (Mandelbrot, 1969), in practice it occurs when a series is overdifferenced. In (1.2.1) the case $H=1/2$ reflects the usual weakly dependent case in which the spectral density of x_t at zero frequency is bounded and bounded away from zero. In this case we write $x_t \sim I(0)$. We need $H<1$ in order to have covariance stationarity[5].

Conditions 1.A and 1.B are not always equivalent but if $1/2<H<1$ and if we suppose that γ_j are quasi-monotonically convergent to zero, i.e., for some $B\geq 0$:

---

[4]In this situation we have that, either condition 1.A or condition 1.B, imply condition 1.C:

$$\sum_{j=0}^{\infty} |\gamma_j| = \infty\ ,$$

this condition has been used as a definition of long-memory also.

[5]We can see that the spectral density is only integrable, and therefore, the variance is defined, if $H<1$:

$$\int_o^{\pi} f(\lambda)d\lambda = \int_0^{\epsilon} f(\lambda)d\lambda + \int_{\epsilon}^{\pi} f(\lambda)d\lambda = C + \int_0^{\epsilon} C_1 \lambda^{1-2H} d\lambda = C + \frac{C_1 \epsilon^{2-2H}}{2-2H} < \infty\ ,$$

if $H<1$, as ϵ approaches to zero.

$$\gamma_{n+1} \le \gamma_n \left(1+\frac{B}{n}\right) \quad for \quad all \quad n \ge n_0(B) \tag{1.2.3}$$

then (1.2.1) and (1.2.2) are satisfied with:

$$C_1 = \frac{\pi}{2\Gamma(2H-2)\cos\left(\frac{(2H-2)\pi}{2}\right)} C_2 .$$

This is a particular case of theorem III-14 in Yong (1974) where a proof can be found.

1.3 Examples of Long-Memory

In this section we analyze the most studied examples of long-memory: the fractional Gaussian noise and the fractional ARIMA.

A) Fractional Gaussian noise. This has been analyzed in Mandelbrot and Van Ness (1968) and Sinai (1976). It is a Gaussian stochastic process with zero mean and autocovariances:

$$\gamma_j = \frac{Ex_1^2}{2}\left(|j+1|^{2H} - 2|j|^{2H} + |j-1|^{2H}\right) ,$$

so that, as $j \to \infty$, $\gamma_j \sim Kj^{2H-2}$, where $K = H(2H-1)Ex^2_1$ the spectral density is:

$$f(\lambda) = c|e^{i\lambda}-1|^2 \sum_{j=-\infty}^{\infty} \frac{1}{|\lambda + 2\pi j|^{2H+1}} ,$$

with c some positive finite constant, so for low frequencies:

$$f(\lambda) \sim c\lambda^{1-2H} \quad as \quad \lambda \to 0^+ .$$

This process can be derived in the following way: denote by Z(t) a Gaussian continuous self-similar process[6], with parameter H and with stationary increments[7], this

[6]This means that, for any $a>0$ and any $t_1,\dots,t_n$, the joint distribution of $Z(t_1),\dots,Z(t_n)$ is the same as a^{-H} times the joint distribution of $Z(at_1),\dots Z(at_n)$.

[7]That is, the finite-dimensional distributions of $\{Z(t+s)-Z(t)\}$ do not depend on s.

process is called a Fractional Brownian Motion, then the 1-step increments, $x_t=Z(t+1)-Z(t)$ with t integer, follows a fractional Gaussian noise.

B) Fractional ARIMA. This has been studied in Granger and Joyeux (1980) and Hosking (1981). It is a generalization of ARIMA models. Instead of considering d an integer we allow it to be fractional in:

$$\phi(L)(1-L)^d x_t=\theta(L)\epsilon_t ,$$

where L is the lag operator ($Lx_t=x_{t-1}$), ϵ_t is a white noise process and

$$\phi(L)=1-\phi_1 L-\phi_2 L^2...-\phi_p L^p \quad and \quad \theta(L)=1+\theta_1 L+\theta_2 L^2+...+\theta_q L^q$$

have all zeros outside the unit circle. We say that x_t is integrated of order d and denote it by $x_t \sim I(d)$. The expression $(1-L)^d$ is defined by the binomial expansion:

$$(1-L)^d=\sum_{j=0}^{\infty}\pi_j L^j , \tag{1.3.1}$$

where

$$\pi_j=\frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)}=\prod_{k=1}^{j}\frac{k-1-d}{k} , \quad j=0,1,2..., \tag{1.3.2}$$

where $\Gamma(\ )$ is the gamma function:

$$\Gamma(x)=\begin{cases}\int_0^{\infty} t^{x-1}e^{-t}dt & x>0\\ \infty & x=0\\ x^{-1}\Gamma(1+x) & x<0 .\end{cases}$$

The fractional Gaussian noise characterizes all the autocorrelations by just one parameter, H. This is a drawback in empirical applications, where we would prefer to model the short run behaviour with more flexibility. Fractional ARIMA allows this, because the eventual behaviour of the autocovariances is determined by d, which corresponds to H-1/2, but the short run behaviour can be modelled with an ARMA process.

We consider first the properties of a fractional ARIMA(0,d,0). Then:

$$x_t=(1-L)^{-d}\epsilon_t \ , \ i.e. \ , \quad x_t=\sum_{j=0}^{\infty}\psi_j\epsilon_{t-j} \ , \quad where \quad \psi_j=\prod_{k=1}^{j}\frac{k-1+d}{k}$$

will be properly defined (have finite variance) when

$$\sum_{j=0}^{\infty}\psi_j^2<\infty \ ; \tag{1.3.3}$$

using Sterling's formula we can approximate ψ_j by:

$$\frac{1}{\Gamma(d)}j^{d-1} \ ,$$

so (1.3.3) will be equivalent to $d<1/2$.

Similarly we will have an infinite autoregressive representation:

$$\sum_{j=0}^{\infty}\pi_j x_{t-j}=\epsilon_t \ ,$$

with the coefficients defined in (1.3.2) when

$$\sum_{j=0}^{\infty}\pi_j^2<\infty \ ,$$

and this corresponds with $d>-1/2$.

Therefore the process x_t is covariance stationary and invertible when $-1/2<d<1/2$; we assume that this is the case. The autocovariances satisfy:

$$\gamma_j\sim C_d\, j^{2d-1} \quad as \quad j\to\infty \ , \quad where \quad C_d=\frac{1}{\pi}\Gamma(1-2d)\sin\pi d \ ,$$

and the spectral density is:

$$f(\lambda)=\frac{\sigma^2}{2\pi}|1-e^{-i\lambda}|^{-2d}=\frac{\sigma^2}{2\pi}(2\sin\frac{\lambda}{2})^{-2d} \ ,$$

so

$$f(\lambda) \sim \frac{\sigma^2}{2\pi} \lambda^{-2d} , \quad as \quad \lambda \to 0^+ .$$

Similarly for a fractional ARIMA(p,d,q):

$$\gamma_j \sim C j^{2d-1} , \quad as \quad j \to \infty ,$$

for C a positive finite constant when d > 0 and a negative finite constant when d < 0, and

$$f(\lambda) = \frac{\sigma^2}{2\pi} \frac{|\theta(e^{-i\lambda})|^2}{|\phi(e^{-i\lambda})|^2} |1 - e^{-i\lambda}|^{-2d} \sim K \lambda^{-2d} , \ as \ \lambda \to 0^+ , \ with \ K = \frac{\sigma^2}{2\pi} \frac{\theta(1)^2}{\phi(1)^2}.$$

As $(1-L)^{d_0} = (1-L)^n (1-L)^{d_1}$ with n integer and $d_1 \in (-1/2,1/2)$ for any d_0 real, it is interesting to analyze the properties of a fractional ARIMA with $d \in (-1/2,1/2)$.

So far we have only analyzed the univariate case. The multivariate case is very relevant in order to analyze the interrelationships between different variables. For instance, we can have series with different short correlation structure, but with a similar correlation pattern for higher lags, i.e., with the same H; in this case we can estimate more efficiently this parameter if we take into account the cross-information. In several disciplines, e.g. economics, it is interesting to analyze the impact (both, in the short and in the long run, and the way it happens) that some variables have on others. For example, this can be relevant in terms of economic policy. In general, if we try to build a model of several variables, it is crucial to study the co-movements of these variables.

The extension of the preceding definitions and models to the multivariate framework is not difficult. We assume that the conditions that characterize the long memory behaviour are fulfilled for every series, i.e., if we consider an r-dimensional vector process and a is the typical series (a=1,...r) and we assume that its spectral density exists and denote it by

$f_{aa}(\lambda)$ and by γ_j^a its autocovariance at lag j:

$$f_{aa}(\lambda) \sim g_{aa}\ \lambda^{1-2H_a}\ ,\ as\ \ \lambda \to 0^+\ ,\ \ g_{aa}>0\ \ for\ \ a=1,...r\ ,$$

$$\gamma_j^a\ \sim\ K_a\, j^{2H_a-2}\ ,\ \ as\ \ j\to\infty\ ,\ \ for\ \ a=1,...r\ ,$$

g_{aa} and K_a denoting general constants and $H_a \in (0,1)$, $H_a \neq 1/2$. In chapter 4 we will introduce carefully our multivariate framework.

When analyzing a multivariate fractional ARIMA one consideration worth noticing is that, while in the univariate case the models[8]:

$$(I)\quad \phi(L)(1-L)^d x_t = \theta(L)\epsilon_t$$

and

$$(II)\quad (1-L)^d \phi(L) x_t = \theta(L)\epsilon_t$$

are equivalent, in the sense that the general linear process that both models implied are the same:

$$(I)\!:\ x_t = \sum_{j=0}^{\infty} b_{1j}\epsilon_{t-j}\ ,$$

$$(II)\!:\ x_t = \sum_{j=0}^{\infty} b_{2j}\epsilon_{t-j}\ ,$$

with $\{b_{1j}\} = \{b_{2j}\}$, in the multivariate framework this does not happen. The models:

$$\Phi(L)\ diag\{(1-L)^{d_a}\} X_t = \Theta(L)\epsilon_t$$

and

[8]Where we assume the necessary conditions for identification, in particular, ϕ(L) and θ(L) have all their roots outside the unit circle.

$$diag\{(1-L)^{d_a}\}\Phi(L)X_t=\Theta(L)\epsilon_t$$

are not equivalent. Consider, for instance, a simple example[9]:

$$(I):\ \begin{pmatrix}\phi_{11}(L) & \phi_{12}(L)\\ \phi_{21}(L) & \phi_{22}(L)\end{pmatrix}\begin{pmatrix}(1-L)^{d_1} & 0\\ 0 & (1-L)^{d_2}\end{pmatrix}\begin{pmatrix}x_{1t}\\ x_{2t}\end{pmatrix}=\begin{pmatrix}\epsilon_{1t}\\ \epsilon_{2t}\end{pmatrix},$$

$$(II):\ \begin{pmatrix}(1-L)^{d_1} & 0\\ 0 & (1-L)^{d_2}\end{pmatrix}\begin{pmatrix}\phi_{11}(L) & \phi_{12}(L)\\ \phi_{21}(L) & \phi_{22}(L)\end{pmatrix}\begin{pmatrix}x_{1t}\\ x_{2t}\end{pmatrix}=\begin{pmatrix}\epsilon_{1t}\\ \epsilon_{2t}\end{pmatrix},$$

then we can express x_{1t} in (I) as:

$$(\phi_{11}(L)\phi_{22}(L)-\phi_{12}(L)\phi_{21}(L))(1-L)^{d_1}x_{1t}=\phi_{22}(L)\epsilon_{1t}+\phi_{12}(L)\epsilon_{2t}$$

and in (II) as:

$$(\phi_{11}(L)\phi_{22}(L)-\phi_{12}(L)\phi_{21}(L))(1-L)^{d_1+d_2}x_{1t}=\phi_{22}(L)(1-L)^{d_2}\epsilon_{1t}+\phi_{12}(L)(1-L)^{d_1}\epsilon_{2t},$$

so, only if $d_1=d_2$, models (I) and (II) are equivalent.

In case (I) we have that $x_{1t}\sim I(d_1)$ and $x_{2t}\sim I(d_2)$, while in case (II): $x_{1t}\sim I(d_2)$ if $d_1<d_2$ and $\phi_{12}(1)\neq 0$ and $x_{1t}\sim I(d_1)$ otherwise, while $x_{2t}\sim I(d_1)$ when $d_2<d_1$ and $\phi_{21}(1)\neq 0$ and $x_{2t}\sim I(d_2)$ otherwise.

These considerations are relevant when analyzing the subject of fractional cointegration[10]. In chapter 4 we will give further insight in these representations, but the subject of fractional cointegration is not in the scope of this thesis.

What is important to stress is that the multivariate framework provides a more detailed insight of the properties and behaviour of stochastic processes than a univariate framework and this is why much of this thesis will focus on it.

---

[9]Where we assume the necessary conditions for identification, see Hannan (1976), one of them is that all the roots of $|\Phi(L)|$ are outside the unit circle.

[10] That is, when a linear combination of the series possesses an order of integration that is less than the maximum of the different orders of integration of the original series.

1.4 Review of the Literature

The long-memory phenomenon has been noticed by applied statisticians in several fields for a long time. Two main areas, in which it has been detected, has been hydrology and economics, but there are references for biology, geophysics and meteorology in Mandelbrot and Wallis (1969), agriculture in Whittle (1962) just to cite a few. In hydrology, Hurst (1951) analyzed the flow of the river Nile and proposed the widely used R/S statistic to detect long-memory. In economics, Mandelbrot (1969) and Granger (1966) are earlier references. Granger called the "typical spectral shape" of an economic variable to a spectrum that exhibits a comparatively high mass close to the zero frequency. Robinson (1994c) and Beran (1994) provide an extensive bibliography.

Although detection of the phenomenon was quite earlier, the formal analysis didn't start until much later, and it was linked in great part to the analysis of self-similar processes. In this field, Taqqu (1985) and Vervaat (1987) are the usual references for the bibliography up to 1987.

As the literature is very active in several directions this section does not pretend to be comprehensive but will review some aspects on a few topics as justification, simulation, generalizations, testing and convergence results with long-memory series to point out the variety of results that can be found.

Long-memory models have been **justified** in terms of aggregation. Robinson (1978) and Granger (1980) show that if individual series follow AR(1) processes:

$$x_{it}=\alpha_i x_{it-1}+u_{it}\ , \quad i=1,\dots N, \quad t=1,\dots T\ ,$$

then, the aggregate series:

$$x_t=\sum_{i=1}^{N} x_{it} \ , \quad t=1,...T \ ,$$

will exhibit long-memory if, for instance, α_i are drawn from a beta, B(p,q), distribution for certain values of p and q.

A key feature of long-memory processes (and what makes this literature distinguished from the standard) is that, because these processes are not strong mixing, at least with fast enough rates in the Gaussian case, the usual central limit theorem cannot apply. Rosenblatt (1961) showed that for a specific long-memory Gaussian process x_t,

$$\frac{1}{n^{2H-1}}\sum_{t=1}^{n} (x_t^2-1)$$

does not converge to a normal variate for H in between 3/4 and 1.

Taqqu (1975), in a fundamental paper, analyzed the **convergence** in distribution of properly normalized sums of functions of Gaussian long-memory processes as:

$$\sum_{t=1}^{n} G(x_t) \ . \tag{1.4.1}$$

We state a basic result: "let $\{ x_j, j\geq 1\}$ be a stationary Gaussian sequence satisfying $Ex_1=0$, $Ex_1^2=1$ and[11] $Ex_1x_{1+k}\sim L(k)k^{2H-2}$ as $k\rightarrow\infty$, G( ) satisfies $EG(x_1)$, $EG^2(x_1)<\infty$ and has Hermite rank[12] m, then:

---

[11] Where L( ) is a slowly varying function at infinity, that is a positive function so that

$$\frac{L(tk)}{L(k)}\rightarrow 1 \ \ as \ k \rightarrow \infty \ , for \ all \ t{>}0.$$

[12] The Hermite rank m is the minimum j for which c_j is different to zero in the Hermite expansion of G( ):

$$G(x)=\sum_{j=0}^{\infty}\frac{c_j}{j!}H_j(x) \ , \quad H_j(x)=(-1)^j e^{\frac{x^2}{2}}\frac{d^j}{dx^j}e^{-\frac{x^2}{2}} \ .$$

* if 1-1/2m<H<1:

$$d^2(n)=E(\sum_{j=1}^{n} G(x_j))^2 \sim \left(\frac{J(m)}{m!}\right)^2 \frac{n^{2(1-m+mH)}L^m(n)}{(1-2m+2mH)(1-m+mH)}$$

and

$$\frac{1}{d(n)}\sum_{j=1}^{[nt]} G(x_j) \to_d \overline{Z_m}(t)$$

where $\overline{Z_m}(t)$ has a complicated expression but for m=1 it is Fractional Brownian Motion and for m=2 it is called the Rosenblatt process, see p.41;

* if 1/2< H< 1-1/2m then:

$$\frac{1}{\sqrt{n}}\sum_{j=1}^{n} G(x_j) \to_d Z$$

where $Z \sim N(0,\sigma_z^2)$ and

$$\sigma_z^2=\lim_{n\to\infty}\sum_{j=1}^{n}\sum_{l=1}^{n} E[G(x_j)G(x_l)]."$$

There is a crucial point to notice. Convergence depends on m and H. They determine not only if there is convergence to a Gaussian or non-Gaussian process but also the rate of convergence that does not need to be $n^{1/2}$. In chapter 3 we will provide a result that even though does not follow directly from these results, it is also evidence of this distinctive behaviour.

Although most of the work done has assumed Gaussianity, this assumption has been relaxed and there have been studies analyzing convergence of linear processes and of their functionals, as (1.4.1) with x_t linear, see Robinson (1994c). Rates of convergence can be arbitrarily slow and in the functional case they will depend on H and on m, the rank in the Appell expansion of G( ), where:

$$G(x)=\sum_{j=0}^{\infty}\frac{e_j}{j!}A_j(x)\ ,$$

where A_j are called Appell polynomials and are defined by[13]:

$$\sum_{j=0}^{\infty}\frac{z^j}{j!}A_j(x)=\frac{e^{zx}}{E(e^{zx})}.$$

We can have normal or nonnormal convergence depending on m and H. A review of this results up to 1990 is Robinson (1994c).

In particular, as basis of statistical inference, a great deal of attention has been focused in the convergence of quadratic forms; the main references are Giriatis and Surgialis (1990) and Terrin and Taqqu (1991).

In the next chapter we will analyze the question of estimating H. Its importance is twofold: H is the parameter that characterizes the long-memory behaviour and H may appear in the rate of convergence of some statistics.

Adenstedt (1974) and Samarov and Taqqu (1987) analyze the **relative efficiency** of ordinary least squares (OLS), compared with generalized least squares (GLS), in estimating the mean of a long-memory process. They obtain that OLS is not asymptotically efficient, but if $H>1/2$ the loss of efficiency is quite small, while if $H<1/2$ the loss can be bigger.

Yajima (1988) extends this result to the case of having a trend as a regressor and a long-memory disturbance with $H>1/2$. In this situation the loss of efficiency is even greater.

Furthermore Yajima (1991) extends the result of Grenander (1954) and Grenander-Rosenblatt (1957) about the relative efficiency of OLS, compared with GLS, in a regression

---

[13]Appell polynomials are the extension to non-Gaussian cases of Hermite polynomials; if x is N(0,1) then the Appell are the Hermite Polynomials.

model where the disturbances exhibit long-memory. For regressors satisfying "Grenander's conditions"[14] he shows that OLS will be asymptotically efficient if the spectral density of the disturbance is constant on each element of the regression spectrum[15]. The additional restriction, with respect to Grenander's case[16], is that the zero frequency is excluded in the regression spectrum.

Because of the slow decay of the autocovariances, **simulation** of long-memory series has turned out to be a difficult task.

McLeod and Hipel (1978) proposed to decompose the theoretical covariance matrix, Γ, of the process that we want to simulate by Cholesky: Γ=MM', and then, filter a white noise, e, to obtain a long-memory series y=Me.

Granger and Joyeux (1980) modified that method in the sense that, only the first 100

---

[14]Suppose z_{jt} denotes the t-th observation of the j-th regressor, j=1,...,p, t=1,...,n; denote also

$$d_j^2=\sum_{t=1}^{n} z_{jt}^2,\quad D=diag(d_1,...,d_p),\quad Z=\{z_{jt}\},\quad \hat{R}=D^{-1}Z'ZD^{-1}\ ,$$

then, "Grenander's conditions"are:

$$a)\ d_j\to\infty,\ for\ all\ j,\quad as\ n\to\infty\ ,$$

$$b)\lim_{n\to\infty}\ \max_{1\le t\le n}\ \frac{|z_{jt}|}{d_j}=0\ ,\ \ for\ all\ j\ ,$$

$$c)\lim_{n\to\infty}\ \hat{R}=R>0\ .$$

[15] The regression spectrum is the set of points λ_i where $0<\lambda_i\le\pi$, i=1,...p, for which the spectral distribution function of the regressors jump.

[16]Grenander assumed that the spectral density of the disturbance was continuous and positive at the zero frequency, and this rules out condition 1.A.

values are obtained in that way and then, they use a truncated autoregression to obtain the further values of a fractional ARIMA (0,d,0) series:

$$\sum_{j=0}^{100} \pi_j x_{t-j} = e_t \,, \quad t=101,102,...$$

where π_j are given by (1.3.2).

Geweke and Porter-Hudak (1983) use the Levinson-Durbin-Whittle algorithm (that uses the Cholesky decomposition of a Toeplitz matrix) to simulate long-memory series.

All these methods imply a high cost in computational time; a much faster approach, that uses the fast Fourier transform (FFT), was proposed in the appendix of Davies-Harte (1987).

The models we have seen in the previous section are very simple. There have been generalizations worthy of mention. Taqqu (1987) examines a very general stochastic process model:

$$Z(t) = \sum_{\phi_1=-M}^{M} \cdots \sum_{\phi_m=-M}^{M} f_t(\phi_1,\phi_2,...\phi_m)\Delta B_\alpha(\phi_1)...\Delta B_\alpha(\phi_m) \,,$$

with

$$\phi_1 \neq \phi_2 \neq ... \neq \phi_m \,,$$

where ΔB_α are white noise, $\alpha \in (0,2)$ is the stability index of a stable distribution, f_t is the kernel that generates dependence by mixing the noises, and, f_t with m control the degree of nonlinearity. This model is very general and Taqqu (1987, p.9685) shows that fractional Gaussian noise can be derived as a particular case.

Gray et al. (1989) analyze a generalization of fractional ARIMA processes proposed by Hosking (1984):

$$\phi(L)(1-2uL+L^2)^\lambda \; x_t = \theta(L)\epsilon_t$$

using Gegenbauer's polynomials[17]. This model allows the singularity of the spectral density to be at frequencies other than zero.

Long-memory has also been used to characterize **second moments.** Robinson (1991a) extended autoregressive conditional heteroskedasticity models by the cases:

$$V(u_t|F_t)=\{1+\sum_{j=1}^{\infty}\phi_j^2(\theta)\}^{-1}\ \{\sigma+\sum_{j=1}^{\infty}\phi_j(\theta)u_{t-j}\}^2 \tag{1.4.2}$$

and

$$V(u_t|F_t)=\sigma^2+\sum_{j=1}^{\infty}\phi_j(\theta)(u_{t-j}^2-\sigma^2) \tag{1.4.3}$$

where u_t is the disturbance in a linear regression model and the coefficients $\phi_j(\theta)$ are uniquely defined functions of the vector θ, such that $\phi_j(\theta)=0$ for all $j\geq 1$ if and only if $\theta=0$, and $\phi_j(\theta)$ can decay slowly enough to allow for long memory and F_t is the σ-field of events generated by u_s, $s<t$. He uses Lagrange multiplier (LM) tests for the null hypothesis of no dynamic conditional heteroskedasticity ($\theta=0$), i.e., $V(u_t|F_t)=\sigma^2$ and obtains limiting χ^2 distributions. Baillie et al. (1993) and Harvey (1993) extend autoregressive conditionally heteroscedasticity and stochastic volatility models, respectively, to allow for long-memory. In particular Baillie et al. attempt to study the fractionally integrated generalized autoregressive conditionally heteroscedasticity process:

---

[17]That is, using:

$$(1-2uL+L^2)^{-\lambda}=\sum_{k=0}^{\infty}C_n^{(\lambda)}(u)L^n\ ,$$

where

$$C_n^{(\lambda)}(u)=\sum_{k=0}^{[\frac{n}{2}]}(-1)^k\frac{\Gamma(\lambda+n-k)}{\Gamma(\lambda)}\frac{(2u)^{n-2k}}{k!(n-2k)!}\ .$$

$$\phi(L)(1-L)^d y_t^2 = w + \theta(L)\nu_t \ , \tag{1.4.4}$$

where y_t is a zero mean serially uncorrelated process, $0<d<1$, the roots of $\phi(L)$ and $\theta(L)$ are outside the unit circle, w is a constant and

$$\nu_t = y_t^2 - \sigma_t^2, \quad E(y_t^2 | \Omega_{t-1}) = \sigma_t^2 \ .$$

Notice that when w=0 (1.4.4) is a special case of (1.4.2) or (1.4.3). Harvey's model is:

$$y_t = \sigma_t \epsilon_t \ , \qquad \epsilon_t \sim IID(0,1) \ , \qquad t=1,2...,$$

$$\sigma_t^2 = \sigma^2 \exp(h_t) \ , \ (1-L)^d h_t = \eta_t \ , \qquad \eta_t \sim NID(0,\sigma_\eta^2) \ .$$

These models may be particular relevant when analyzing financial data.

Robinson (1991b) compares a **nonparametric kernel estimate of the probability density function** under both weak and strong dependence, from a theoretical point of view, and also provides Monte Carlo evidence. He finds that in the long-memory case rates of convergence of the estimates are slower than in the weak dependent case and that in the strong dependent case, estimates evaluated at two different points are perfectly correlated. Monte Carlo evidence confirms this, showing that when the positive strong dependence is high the performance of the estimates is very poor. He also analyzes the optimal bandwidth question. He finds that, depending on the values of H, the properties that hold under weak dependence may continue to hold. But, if the dependence is strong enough, then the usual results are invalidated. Cheng and Robinson (1990) and Hall and Hart (1989) extend these results.

Testing with long-memory series is an area of research that is receiving a growing attention. Robinson (1994d) establishes a very general framework in which many long-memory as well as nonstationary models can be considered as null or alternative hypothesis.

The general model he uses is:

$$\rho(L,\theta)x_t=u_t \ ,$$

where x_t can be the residuals of a regression model and

$$\rho(L,\theta)=(1-L)^{\gamma_1+\theta_1}(1+L)^{\gamma_2+\theta_2}\prod_{j=3}^{h}(1-2\cos\omega_j L+L^2)^{\gamma_j+\theta_j} \ ,$$

for given $\gamma_1, \gamma_2,\ldots\gamma_h$; the null hypothesis is[18]:

$$H_0{:}\theta_1=\theta_2=\ldots=\theta_h=0 \ .$$

Using Lagrange multiplier tests in the frequency domain, he proves that asymptotic distributions are χ^2 , in contrast with much of the literature on unit root testing that end up with non-standard distributions. These tests are carried out in the frequency domain[19]. More popular among applied econometricians are tests in the time domain. Robinson (1991a) proposes a simple LM test for a null hypothesis of absence of any autocorrelation. The alternatives are of the class:

$$x_t=\sum_{j=1}^{\infty}\phi_j(\theta)x_{t-j}+u_t$$

and the $\phi_j(\theta)$ have been defined in p.23 and can decay slowly enough to allow for strongly autocorrelated alternatives. Agiakloglou and Newbold (1994) examine LM tests of ARMA(p,q) against fractional ARIMA(p,d,q) alternatives. They show that the tests will have low power when the orders (p,q) are over-specified.

Beran (1992) analyzes for long-memory series a **goodness-of-fit test,** proposed by

[18] Some of the null hypothesis that are include are: a unit root (ρ(L)=1-L), quarterly unit root ($\rho(L)=1-L^4$), "1/f" noise ($\rho(L)=(1-L)^{1/2}$), etc.

[19]Although they could be performed in the time domain.

Milhoj (1981) in the frequency domain. This test is an extension of the Box-Pierce (1970) statistic when one takes into account all the computable correlations. The asymptotic distribution under the null hypothesis is the same as in the weakly dependent case.

Hidalgo and Robinson (1993) analyze a Wald test for structural break at a known time τ in a linear regression model:

$$y_t = \beta_t' x_t + u_t \qquad \beta_t = \begin{cases} \beta_1 & t=1,2\ldots\tau \\ \beta_2 & t=\tau+1,\ldots T \end{cases},$$

with u_t being a Gaussian long-memory series. The usual structural break tests based on u_t being weakly dependent will not hold.

Analysis of prediction with long-memory series has focused on fractional ARIMA models. Compared with ARIMA forecasting, the additional problem is that $(1-L)^d$ has to be truncated for a finite series. Ray (1993) approximates a fractional ARIMA(0,d,0) with an AR(p). Pieris and Perera (1988) provides useful formulae while Porter-Hudak (1990) shows with a monetary series that a seasonal fractional model forecasts better than, the widely applied, ARIMA "airline" model.

Chapter 2

Estimation of H

2.1 Introduction

In this chapter we analyze the different procedures used to estimate H, the parameter that characterizes the presence of long-memory in a series. As we have seen previously, this is an important issue, not only because H reflects the degree of strong dependence in a series, but also because rates of convergence of some statistics that are relevant for statistical inference depend on H.

There have been basically two main approaches. The first one is a parametric approach, in which we specify and estimate a parametric model in which H is just one more parameter. The second one is a semiparametric approach, in which we focus on estimating H based on the definition of long-memory, i.e., based on either (1.2.1) or (1.2.2). This implies using estimates of the spectral density in a neighbourhood of the zero frequency, or, using estimates of the autocovariances of long lags only.

Before focusing in the semiparametric approach, we will review briefly the parametric approach.

2.2 Parametric estimates

H is estimated jointly with all the other parameters that specify the model. The analysis can be carried out in the frequency or in the time domain. In the former one it is assumed that the spectral density is known, up to a certain parameter vector θ ($H \in \theta$): $f(\lambda, \theta)$, with $\lambda \in (-\pi, \pi]$, $\theta \in \Theta$, θ_0 is supposed to be the true value, and the estimation procedure consists in estimating θ by some maximum likelihood method.

Fox and Taqqu (1986) assumed Gaussianity of the process, and minimized the Whittle function (an approximation to the exact likelihood function):

$$\frac{1}{4\pi}\int_{-\pi}^{\pi}\left(\log f(\lambda,\theta)+\frac{I(\lambda)}{f(\lambda,\theta)}\right)d\lambda \ ,$$

where $I(\lambda)$ is the periodogram of the process x_t evaluated at frequency λ:

$$I(\lambda)=\frac{1}{2\pi n}|\sum_{t=1}^{n}x_t e^{it\lambda}|^2 \ . \qquad (2.2.1)$$

Sowell (1992) analyses the maximum likelihood estimates of the parameters of a univariate fractional ARIMA. There is a limitation in his procedure: the roots of the AR polynomial cannot be multiple.

Dahlhaus (1989) also assumes Gaussianity but considers the exact likelihood function and minimizes:

$$\frac{1}{2n}\log|T_n(f(\theta))|+\frac{1}{2n}(x_n-\mu_n)'T_n(f(\theta))^{-1}(x_n-\mu_n) \ ,$$

where $T_n(f(\theta))$ is a nxn matrix with (r,s) element:

$$\{T_n(f(\theta))\}_{(r,s)}=\int_{-\pi}^{\pi}f(\lambda,\theta)e^{i(r-s)\lambda}d\lambda \quad for\ r,s=1,...n \ ,$$

μ_n estimates consistently the mean μ_0 and n denotes the sample size.

Dahlhaus (1989) proves the asymptotic efficiency of both MLE estimates, i.e., their asymptotic variance is the inverse of the information matrix ($\Gamma(\theta_0)$):

$$\sqrt{n}(\hat{\theta}-\theta_0)\to_d N(0,\Gamma(\theta_0)^{-1}) \ .$$

Giriatis and Surgailis (1990) relax the Gaussianity assumption and analyze the Whittle estimate for linear processes. Asymptotic normality is achieved, but the estimate is no longer asymptotically efficient.

It is worth to pointing out that these parametric estimates have the same asymptotic properties as in the weakly dependent case: the rate of convergence is $n^{1/2}$ and, in the Gaussian case, will achieve asymptotic efficiency.

2.3 Semiparametric estimates

If our main interest is the estimation of H then a parametric approach (that will produce consistent and efficient estimates if the parameterization is right, but inconsistent estimates if the parameterization is not correct) may not be desirable. We prefer to guarantee consistency of our estimates at the expense of losing efficiency. This is the justification of the "semiparametric"[1] approach. In this approach, we restrict our attention to the long-memory features of the series, i.e., the behaviour of the spectral density function (sdf, hereafter) close to the zero frequency, or the behaviour of the autocovariances for long lags only. In particular, if we work in the frequency domain, we will assume that the sdf (apart of being integrable due to covariance stationarity) behaves as (1.2.1), i.e., we only assume that we know the form of the sdf in a neighbourhood of zero. This is the definition of long-memory in the frequency domain and it is the only assumption on the spectral density function we want to make so far. It is important to stress the generality of this approach: we are allowing, practically, any behaviour of the process in the short run because we are not imposing any restriction (apart from integrability) on the sdf away from the zero frequency.

In order to implement the semiparametric approach in the frequency domain, we need

---

[1]Robinson(1994) introduced this terminology.

to define a bandwidth parameter, m, so that $\lambda_m = 2\pi m/n$ is going to be the highest frequency at which we are going to evaluate our estimates. We will need to assume that m tends to infinity, but slowly compared with n, so that, λ_m tends to 0 as n tends to infinity.

In the time domain we can carry out the semiparametric approach in the following way: we define a bandwidth number p and we assume that the relation: $\gamma_j = Cj^{2H-2}$ holds for $j=n-p, n-p+1, \ldots n-1$, i.e., it will be valid only for the p higher autocovariances. Similarly, to achieve consistency we will need to impose that p tends to infinity but slowly compared with n, so that p/n tends to zero.

As our estimates will be based in m(p) pieces of information, and $m/n \to 0$ ($p/n \to 0$), these estimates will be inefficient compared with the parametric estimates of the previous section. In fact, the asymptotic efficiency will be zero, but this is the cost we have to pay in order to ensure that our estimates of H will be consistent under any unknown short-run behaviour of the process.

2.3.1 Semiparametric estimates in the time domain

Before considering the semiparametric estimates we mention the first estimate proposed (Hurst, 1951) that is a nonparametric one and it is based on the R/S statistic:

$$R/S = \frac{\max\limits_{1 \le j \le n} \sum_{t=1}^{j} (x_t - \bar{x}_n) - \min\limits_{1 \le j \le n} \sum_{t=1}^{j} (x_t - \bar{x}_n)}{\left(\frac{1}{n} \sum_{t=1}^{n} (x_t - \bar{x}_n)^2 \right)^{1/2}} .$$

The specific estimate of H, Mandelbrot-Wallis(1969), is given by:

$$\frac{\log(R/S)}{\log n}$$

Its properties has been analyzed in Mandelbrot-Wallis (1969), Mandelbrot (1972, 1975) and Mandelbrot-Taqqu (1979). Beran (1994) provides a neat explanation of the way to implement the R/S procedure. Apart from its complication it is not clear the efficiency of this estimate. Lo (1991) modifies the R/S statistic to be robust to weak dependence.

Robinson (1994c) has proposed two semiparametric estimates:

- the first one can be motivated by assuming that the autocovariances will eventually be positive (as it happens in the fractional ARIMA with $d>0$) and taking logarithms in:

$$\gamma_j \sim C j^{2H-2} \rightarrow \log\gamma_j \sim \log C + (2H-2)\log j \ .$$

We can estimate it by OLS and this gives:

$$\hat{H}_1 = 1 + \frac{\sum_{j=n-p}^{n-1} \log\hat{\gamma}_j(\log j - \overline{\log j})}{2\sum_{j=n-p}^{n-1} (\log j - \overline{\log j})^2} \ , \qquad \overline{\log j} = \frac{1}{p}\sum_{j=n-p}^{n-1} \log j \ .$$

The main drawback of this estimate is that it will only be defined if:

$$\hat{\gamma}_j > 0 \quad for \quad j = n-p, n-p+1, \dots n-1;$$

- the second one is a minimum distance estimate; it is defined implicitly by:

$$(\hat{H}_2, \hat{C}) = arg\ \min_{\Theta} \sum_{j=n-p}^{n-1} (\hat{\gamma}_j - C j^{2H-2})^2 \ ,$$

where $\Theta = \{H \in (0,1)\ ;\ C \in R\}$.

The properties of these estimates are not known yet. Delgado and Robinson (1994) provide some evidence about the behaviour of these estimates with Spanish inflation data.

Most of the recent studies instead of analyzing estimates in the time domain have focused in the frequency domain.

2.3.2 Semiparametric estimates in the frequency domain

These estimates are based on the long-memory definition in the frequency domain (1.2.1) and the basic idea is to use any estimation procedure in the frequency domain, limited to frequencies close to zero, and replacing the sdf by its estimate (e.g. periodogram). We are going to review three:

I) The log-periodogram estimate (LPE) was proposed by Geweke-Porter-Hudak (1983), GPH hereafter, and modified by Künsch (1986) and Robinson (1992). We can motivate it by looking at the sdf of a fractional ARIMA(0,d,0):

$$(1-L)^d x_t = u_t \ ,$$

where u_t is stationary with sdf

$$\frac{\sigma^2}{2\pi} f_u(\lambda) \ ,$$

and $f_u(\lambda)$ is bounded and bounded away from zero at $\lambda=0$. Then the sdf of x_t is:

$$f(\lambda) = \frac{\sigma^2}{2\pi}\left(2\sin\frac{\lambda}{2}\right)^{-2d} f_u(\lambda) \ ,$$

i.e.,

$$\log f(\lambda) = \log\left(\frac{\sigma^2}{2\pi} f_u(0)\right) - 2d\log\left(2\sin\frac{\lambda}{2}\right) + \log\left(\frac{f_u(\lambda)}{f_u(0)}\right) ,$$

introducing the periodogram into this:

$$\log I(\lambda) = \log\left(\frac{\sigma^2}{2\pi} f_u(0)\right) - 2d\log\left(2\sin\frac{\lambda}{2}\right) + \log\left(\frac{f_u(\lambda)}{f_u(0)}\right) + \log\left(\frac{I(\lambda)}{f(\lambda)}\right)$$

and we focus on Fourier frequencies close to zero:

$$\lambda_j=\frac{2\pi j}{n}\ ,\quad j=1,...m \quad and \quad \frac{m}{n}\rightarrow 0\ ,$$

so that we can consider $\log(f_u(\lambda)/f_u(0))$ negligible[2]. Robinson (1992) also simplifies, using λ_j instead of $2\sin(\lambda_j/2)$ and we finish with a regression model like:

$$\log I(\lambda_j)=C-2d\log\lambda_j+\epsilon_j\ , \qquad (*)$$

where

$$C=\log[\frac{\sigma^2}{2\pi}f_u(0)]\ ,\quad \epsilon_j=\log[\frac{I(\lambda_j)}{f(\lambda_j)}]\ .$$

The estimate of d is just the OLS estimate of d in (*). Unfortunately it has not been proved that this estimate is consistent for d. The reason is the distinctive behaviour of the normalized periodogram for frequencies very close to zero. Künsch (1986), Robinson (1992) and Hurvich-Beltrao (1993) show that for j fixed the expectation and variance of $I(\lambda_j)/f(\lambda_j)$ depend on j while the covariance of $I(\lambda_j)/f(\lambda_j)$ and $I(\lambda_k)/f(\lambda_k)$ depends on j and k. Robinson (1992) modifies the former regression introducing 2 modifications:

- use a pooled periodogram instead of the raw periodogram,
- in order to avoid the inconsistency problem mentioned above: introduce a trimming number l, so that frequencies $\lambda_j=2\pi j/n$, $j=1,..l$, are excluded from the regression, where l tends to infinity slower than m, so that l/m tends to zero.

So, the final regression model is:

$$Y_k^{(J)}=C^{(J)}-2d\log\lambda_k+U_k^{(J)}\ ,$$

where

---

[2] Agiakloglou et al. (1993) warn about the bias (that can be quite severe in finite samples) of this procedure when the weak dependent component of the series has relatively high (or low) mass around zero.

$$Y_k^{(J)}=\log[\sum_{j=1}^{J} I(\lambda_{k+j-J})] \ , \quad k=l+J,l+2J,...m \ .$$

J controls the pooling (J=1,2..) and l controls the trimming (J=1, l=0 is the GPH case).

Assuming Gaussianity Robinson proves the consistency and asymptotic normality of this estimate in a multivariate framework. The asymptotic covariance matrix has a complicated expression for the off-diagonal elements while the diagonal elements satisfy:

$$\sqrt{m}(\hat{d}^{(J)}-d)\to_d N(0,\frac{1}{4}J\psi'(J)) \ ,$$

where Ψ' is the derivative of the digamma function:

$$\psi(x)=\frac{d}{dx}\log\Gamma(x) \ ,$$

$\Gamma(\ )$ is the gamma function that we have seen in chapter 1. The asymptotic variance takes values: 1.645/4, 1.289/4, 1.185/4 for J=1,2,3 respectively.

II) The averaged periodogram estimate (APE) was proposed by Robinson (1994a). The basic idea is to look at the cumulative spectral density evaluated at two points close to the origin, λ and $q\lambda$:

$$F(\lambda)=\int_0^{\lambda} f(\theta)d\theta \sim \int_0^{\lambda} C\ \theta^{1-2H}\ d\theta=\frac{C\lambda^{2-2H}}{2-2H} \ , \quad F(q\lambda)\sim\frac{Cq^{2-2H}\lambda^{2-2H}}{2-2H} \ ,$$

then

$$\frac{F(q\lambda)}{F(\lambda)}\sim q^{2-2H} \ , \quad i.e. \ , \quad H\sim 1-\frac{\log\left(\frac{F(q\lambda)}{F(\lambda)}\right)}{2\log q}$$

suggesting the estimate:

$$\hat{H}=1-\frac{\log\left(\hat{F}(q\lambda_m)/\hat{F}(\lambda_m)\right)}{2\log q} \ , \quad where \quad \lambda_m=\frac{2\pi m}{n} \ , \quad \frac{m}{n}\to 0 \ ,$$

with $q\in(0,1)$ and $\hat{F}(\lambda_m)$ is defined in section 3.3. Robinson proved the consistency of this estimate under very mild conditions. In chapter 3 we will analyze this estimate with more detail. In particular we will look under what conditions we can achieve asymptotic normality.

III) Quasi maximum likelihood estimate (QMLE). It is analyzed in the univariate case in Robinson (1993a). This estimate is basically a "Whittle estimate" in the frequency domain considering a band of frequencies that degenerates to zero. Instead of minimizing:

$$\sum_{i=1}^{n}\left(\log f(\lambda_j,\theta)+\frac{I(\lambda_j)}{f(\lambda_j,\theta)}\right),$$

the objective function[3] is:

$$Q(C,H)=\frac{1}{m}\sum_{j=1}^{m}\left(\log[C\lambda_j^{1-2H}]+\frac{I(\lambda_j)}{C\lambda_j^{1-2H}}\right) ;$$

we can concentrate C out and get:

$$\hat{H}=arg\ \min_{\Delta}R(H),\ R(H)=\log\hat{C}(H)-(2H-1)\frac{1}{m}\sum_{j=1}^{m}\log\lambda_j\ ,$$

where $\Delta=(0,1)$ and

$$\hat{C}(H)=\frac{1}{m}\sum_{j=1}^{m}I(\lambda_j)\lambda_j^{2H-1}\ .$$

Under finiteness of the fourth moment and other conditions Robinson (1993a) proves the asymptotic normality of this estimate:

$$\sqrt{m}(\hat{H}-H)\rightarrow_d N(0,\frac{1}{4})\ .$$

Note that the asymptotic variance is lower than that of the LPE. In chapter 5 we will analyze this estimate in a multivariate set up.

---

[3]Consider $f(\lambda,\theta)$ as $C\lambda^{1-2H}$.

To summarize, the comparison between the semiparametric and the parametric approach can be seen as a question of priorities. The parametric approach leads to $n^{1/2}$-consistent estimates while the semiparametric only achieve $m^{1/2}$-consistency. Asymptotically, the latter are inefficient compared with the former. The advantage of the semiparametric estimates is their robustness.

It is important to stress the generality of the semiparametric approach. The conditions on the spectral density we impose are mild and restricted to a neighbourhood of the zero frequency. We assume that the process has finite variance and so the spectral density belongs to L_1, i.e., it is integrable. In order to develop the theory we will need to assume some degree of smoothness or differentiability of the spectral density on a neighbourhood of the zero frequency. But away from zero frequency no assumptions whatsoever will be imposed. In most of our analysis we will not demand the spectral density to be in L_p for any $p > 1$ and we will not assume any degree of smoothness or parametric behaviour away from the zero frequency. This is what makes the semiparametric approach relevant.

In the parametric approach, if the short run behaviour of the process is misspecified, this will lead to inconsistency in the estimation of H; on the other hand the semiparametric approach will proportionate consistent estimation of H under any short-run behaviour of the process. As H is the parameter that characterizes the long-run behaviour of the process, its consistent estimation should be of main interest.

A semiparametric estimate, on the other hand, can be used, as a first step, in the estimation of a parametric (e.g. fractional ARIMA) model.

Chapter 3

Analysis of the Averaged Periodogram Estimate

In this chapter we analyze the Averaged Periodogram Estimate (APE). In particular in section 3.1 we state its consistency proved in Robinson (1994a) while in section 3.2 we examine its asymptotic distribution. In section 3.3 we analyze some issues concerning inference with this estimate. Finite sample behaviour is analyzed in section 3.4.

3.1 Consistency of the APE

The APE is proposed in Robinson (1994a). There he proves the consistency of this estimate under very mild conditions; in particular under:

Condition A:

$$f(\lambda) \sim L(\frac{1}{\lambda})\lambda^{1-2H}, \quad as \quad \lambda \to 0^+ , \quad for \ H \in (\frac{1}{2},1);$$

Condition B: as $n \to \infty$,

$$\frac{1}{m}+\frac{m}{n} \to 0;$$

and Condition C:

$$x_t=\mu+\sum_0^\infty \alpha_j e_{t-j} , \quad \sum_0^\infty \alpha_j^2<\infty ,$$

where

(i) $E(e_t e_u)=0, \ t>u;$

(ii) $E(e_r e_s e_t e_u)=\begin{cases} \sigma^4, \ if \ r=s>t=u; \\ 0, \ if \ r=s>t>u, \ or \ r>s=t>u, \ or \ r>s>t>u, \ or \ r>s>t=u \ ; \end{cases}$

(iii) there exists a non-negative random variable e such that for all $\eta>0$ and some $K<1$:

$$E(e^2)<\infty, \quad P(|e_t|>\eta)\leq KP(e>\eta);$$

(iv) $\frac{1}{n}\sum_{t=1}^{n} E(e_t^2|e_s^2,\ s<t)\rightarrow_p \sigma^2,\ as\ n\rightarrow\infty.$

Condition A is a very general condition , where L( ) is a slowly varying function at infinity. Note that H is restricted to be in between 1/2 and 1.

Condition B is the usual "semiparametric" condition: m are the number of ordinates of the periodogram that we are going to use in the estimate, n is the sample size.

Condition C is satisfied when e_t and $e_t^2 - \sigma^2$ are integrable martingale difference sequences, that is, $E(e_t|F_{t-1})=0$, $E(e_t^2|F_{t-1})=\sigma^2$, $E(e_t^2)<\infty$, see Robinson (1994a).

This estimate poses several difficulties. As it depends on two user-chosen numbers, q and m, it is important to give some criteria about these elections. Other issue that is crucial in order to make statistical inference is to derive the asymptotic distribution of this estimate. In particular to analyze under what circumstances asymptotic normality can be achieved. It is relevant also to consider how this estimate performs in finite samples. These questions will be addressed in the next three sections.

3.2 Asymptotic distribution

This issue is analyzed in Lobato and Robinson (1994). We have to distinguish between 2 cases: when $H\in(1/2,3/4)$ and when $H\in(3/4,1)$. We can only obtain asymptotic normality when $H\in(1/2,3/4)$. Heuristically we can invoke Hannan (1976) to motivate why. Hannan analyzed a central limit theorem for a finite set of serial covariances of a linear process with finite fourth moments. He proved that a necessary and sufficient condition to obtain asymptotic normality was that the spectral density of the process was square integrable. This may happen when $H \in (1/2, 3/4)$ but not for $H \in (3/4, 1)$[1].

In order to analyze the asymptotic distribution we introduce the following conditions:

C3.1: For some $E_\alpha \neq 0$ and $\alpha \in (0,2]$,

$$\frac{f(\lambda)}{g(\lambda)} = 1 + E_\alpha \lambda^\alpha + o(\lambda^\alpha), \quad as \quad \lambda \to 0^+ ,$$

where

$$g(\lambda) = C\lambda^{1-2H}.$$

C3.2: For any $\delta \in (0,1)$, $\Delta \in (1,\infty)$,

$$\sup_{-\Delta\lambda \leq \mu \leq \delta\lambda} \frac{|f(\lambda) - f(\lambda-\mu)|}{|\mu| g(|\mu|)} = O(\frac{1}{\lambda}), \quad as \ \lambda \to 0^+.$$

C3.3: As $n \to \infty$,

$$\frac{1}{m} + \frac{m^{2\alpha+1}}{n^{2\alpha}} \to 0.$$

C3.4: x_t is a Gaussian process.

C3.5: The autocovariances γ_j are quasi-monotonically convergent to zero, see expression (1.2.3).

---

[1]Under

$$f(\lambda) \sim C\lambda^{1-2H} , \quad as \quad \lambda \to 0^+ ,$$

and assuming

$$\int_\lambda^\pi f(v)^2 dv = C_1 < \infty ,$$

$$\int_0^\pi f(v)^2 dv = \int_0^\lambda f(v)^2 dv + \int_\lambda^\pi f(v)^2 dv = O(C_1 + C^2 \int_0^\lambda v^{2-4H} dv) = O(1 + C^2 \frac{\lambda^{3-4H}}{3-4H}) =$$

$$= O(1) \quad if \ H < \frac{3}{4}.$$

C3.1 introduces a rate in the bias of the spectral density.

C3.3 strengthens the condition on the bandwidth parameter m, implying that m tends to infinity slower than $n^{4/5}$.

C3.5 is stronger than C3.2 (see Robinson (1994b), lemma 8) and imposes a restriction on $f(\lambda)$ outside the neighbourhood of 0: $f(\lambda)$ satisfies a Lipschitz condition of order 2-2H. Long-memory at other frequencies apart from 0 is ruled out. Under C3.5: $\gamma_j \sim Dj^{2H-2}$ as $j \to \infty$, where $D=2C\Gamma(2-2H)\cos((1-H)\pi)$, see Yong (1974), p.71.

In theorem 3.1 we prove the asymptotic normality of the APE when $H \in (1/2,3/4)$. In theorem 3.2 we analyze the asymptotic distribution when $H \in (3/4,1)$.

Theorem 3.1 Under C3.1, C3.2, C3.3 and C3.4, for $H \in (1/2,3/4)$,

$$\sqrt{m}(\hat{H}-H) \to_d N(0, \frac{1+q^{-1}-2q^{1-2H}}{(\log q)^2} \frac{(1-H)^2}{(3-4H)}).$$

Proof: see appendix 3.1.

Theorem 3.2 Under C3.1, C3.3, C3.4 and C3.5, for $H \in (3/4,1)$,

$$m^{2-2H}(\hat{H}_q-H) \to_d \frac{(1-q^{2H-2})}{\log q} \frac{(1-H)\Gamma(2(1-H))\cos((1-H)\pi)}{(2\pi)^{2-2H}} T.$$

T is the random variable towards which T_n converges in distribution, where T_n is:

$$T_n = \frac{(S_n - ES_n) - n(\bar{x}-Ex_1)^2}{Dn^{2H-1}},$$

and S_n is:

$$S_n = \sum_{t=1}^{n} (x_t - Ex_1)^2 .$$

The proof of theorem 3.2 is in appendix 3.2 at the end of the chapter. We denote by T the random variable whose distribution is the limiting distribution of T_n because we cannot provide a neat expression for it. We can employ the results in Taqqu (1975) to analyze the limiting distributions of both components of T_n. Lemma 3.1 and Proposition 6.1 in Taqqu (1975) establish: " Let $\{x_i\}$ be a normalized stationary Gaussian sequence with zero mean with $r_j \equiv Ex_i x_{i+j}$, j=1,2... and $r_j \sim Kj^{-\alpha}$ as $j \to \infty$, with $\alpha \in (0,1/2)$, then:

Lemma 3.1: $\sum_{i=1}^{n}\sum_{j=1}^{n} r_{i-j} \sim \frac{2K}{(1-\alpha)(2-\alpha)} n^{2-\alpha}$;

Proposition 6.1: $\frac{1}{n^{\alpha-1}K} \sum_{i=1}^{n} (x_i^2-1) \to_d R$" ;

where R is an stochastic process called "Rosenblatt" by Taqqu:

$$R = \frac{1}{2\Gamma(2(1-H))\cos((1-H)\pi)} \int^{\prime\prime} \{\frac{e^{i(u_1+u_2)}-1}{i(u_1+u_2)}\} |u_1u_2|^{1/2-H} dW(u_1)dW(u_2),$$

where W is a complex-valued Gaussian white noise measure on R^1, and the integral is over R^2 except for the diagonals $u_1 = \pm u_2$.

In our case K=D, α=2-2H, so, we can deduce that:

1) the sample mean of x will have a normal distribution with mean Ex_1 and the variance will be:

$$\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} r_{i-j} \sim n^{2H-2} \frac{D}{H(2H-1)} ,$$

i.e.:

$$n^{1-H}(\bar{x}-Ex_1) \to_d N(0, \frac{D}{H(2H-1)}) ;$$

2) while:

$$\frac{(S_n - ES_n)}{Dn^{2H-1}} = \frac{\sum_{i=1}^{n} [(x_t - Ex_1)^2 - 1]}{Dn^{2H-1}} \rightarrow_d R;$$

so both terms of T_n converge in distribution:

$$\frac{S_n - ES_n}{Dn^{2H-1}} \rightarrow_d R,$$

and

$$\frac{n(\bar{x} - Ex_1)^2}{Dn^{2H-1}} \rightarrow_d \frac{1}{H(2H-1)} \chi^2 ,$$

but the variables are not asymptotically independent and so we cannot establish a neat limiting distribution.

Theorems 1 and 2 show a discontinuity in the asymptotic distribution of the APE around H=3/4. For $H \in (1/2, 3/4)$ the asymptotic distribution is normal and statistical inference can be performed (even though the variance depends on q and H, and so, in order to estimate the variance consistently, we need to substitute H by a consistent estimate). When $H \in (3/4, 1)$ the asymptotic distribution is not normal, in fact it is a functional of a Rosenblatt variate. As far as we know, this process is not tabulated and consequently, statistical inference can not be performed in this situation. Of course, the possibility of tabulating this process, and also the variate T, by Monte Carlo simulations could be considered. But it is not clear how to proceed in this way and the exercise looks computationally very intensive.

3.3 Optimal m and q

The problem of the optimal m is addressed in Robinson (1994b). The optimality criterion that he proposes is to minimize the scaled mean squared error:

$$M\hat{S}E_m = E\left[\frac{\hat{F}(\lambda_m)}{G(\lambda_m)} - 1\right]^2 ,$$

where $\lambda_m = 2\pi m/n$ and:

$$\hat{F}(\lambda_m) = \frac{2\pi}{n}\sum_{j=1}^{m} I(\lambda_j) \ , \quad G(\lambda_m) = \int_0^{\lambda_m} g(\theta)d\theta = \frac{C}{2-2H}\lambda_m^{2-2H} .$$

Under conditions similar to those we have seen in the previous section he obtains, see Corollary 1 and Corollary 2 in Robinson (1994b), that the optimal bandwidths are:

* for $H \in (1/2, 3/4)$:

$$m^* = K_1(\alpha, H, E_\alpha) n^{\frac{2\alpha}{2\alpha+1}} ,$$

* for $H \in (3/4, 1)$:

$$m^* = K_2(\alpha, H, E_\alpha) n^{\frac{\alpha}{2-2H+\alpha}} .$$

Some comments worth to make about these results are: a) first, notice that the rate depends on H when $H \in (3/4,1)$ but not in the other case; b) as

$$\frac{d\left(n^{\frac{2\alpha}{2\alpha+1}}\right)}{d\alpha} > 0 \ , \quad \frac{d\left(n^{\frac{\alpha}{2-2H+\alpha}}\right)}{d\alpha} > 0 \ ,$$

and as α controls the smoothness of the spectral density (the bigger α the smoother the spectrum), then the smoother the spectral density the bigger is the rate, i.e., we can increase the number of periodogram ordinates we use; c) notice also that when $H \in (3/4,1)$

$$\frac{d\left(n^{\frac{\alpha}{2-2H+\alpha}}\right)}{dH} > 0 \ ,$$

i.e., the bigger H (long-memory is stronger) the bigger the optimal rate, i.e., we need to increase the number of the ordinates of the periodogram we use[2].

The question about the optimal q is examined in Lobato and Robinson (1994). It can be addressed from two points of view. First we can use as a criterion function the mean squared error of $\hat{H}_{m,q}$, so that both components, bias and variance, are considered. There is a substantial problem here, these formulae, apart of their complexity, involve unknown quantities as α and E_α. The second approach is a simpler one. It consists in just looking at the variance of the limiting distribution when $H \in (1/2,3/4)$; the factor in which q appears is:

$$\frac{1+q^{-1}-2q^{1-2H}}{(\log q)^2}.$$

This expression has an unique minimum for every value of H and in table 3.1 these are tabulated. When $H \in (3/4,1)$ the factor that affects the limiting distribution is:

$$\frac{1-q^{2H-2}}{\log q}.$$

This function has a minimum as q approaches 1 and, therefore, nothing can be said about the optimal q.

3.4 Finite sample behaviour

This is also analyzed in Lobato and Robinson (1994). We examine by a Monte Carlo study the behaviour of the APE for series of size n=256. We have chosen this number

---

[2]When $H \in (1/2,3/4)$ we can appreciate a similar effect:

$$\frac{d\, K_1(\alpha,H,E_\alpha)}{dH}>0$$

and when H approaches 3/4, $K_1(\alpha,H,E_\alpha)$ tends to infinity.

because it is representative of the minimum size of a series in which these semiparametric methods can be applied. We have selected two bandwidth numbers, 32 and 64, and four values of H, 0.55, 0.7, 0.8 and 0.95. The series are generated from a Gaussian fractional noise with variance 1 and autocovariances given by:

$$\gamma_j=\frac{1}{2}\{|j+1|^{2H}-2|j|^{2H}+|j-1|^{2H}\}$$

using the algorithm provided in the appendix of Davies-Harte (1987). We base all our results in 10,000 replications.

In table 3.2 and figures 1 and 2 we examine the behaviour of this estimate for the different values of H given q=0.5. First, it is clear that m=64 is a better choice than m=32. The bias, the variance, the skewness and the kurtosis are smaller with m=64. Another point to notice is that all the estimates present a negative bias, negative skewness and some degree of kurtosis. All these features being more relevant as H increases. In fact, for H=0.95, the bias and the kurtosis are especially severe. In general the normal approximation for H=0.55 and H=0.7 does not look too bad, especially for H=0.55.

Table 3.3 and figures 3 to 10 examine the sensitivity to q. For H=0.55 and H=0.7 we analyze four values of q: 0.2,0.5,0.8 and the optimal value derived from table 3.1: 0.23 for H=0.55 and 0.4 for H=0.7. For H=0.8 and H=0.95 we report only the results for three values of q: 0.2, 0.5 and 0.8. In every case we appreciate the same features we have commented previously, negative bias and skewness and moderate kurtosis. It is interesting to notice the results for q=0.8. In this case the performance of the estimates is very poor, especially for m=32. The variance of the estimates is more than twice the variance of the estimates for the other values of q for H=0.55 and H=0.7. Even for the biggest values of

H, q=0.8 is the worst election in terms of mean squared error, skewness and kurtosis. It is also interesting to point that the optimal values of q for the cases H=0.55 and H=0.7, even displaying slightly more bias than q=0.5, turn out to be effectively the best options in terms of mean squared error. In general it can be said that the selection of q should not be crucial and setting q=0.5 should be a reasonable selection.

In the cases when H=0.55 and H=0.70 we can compare the limiting variance with the one derived for the optimal q. When H=0.55 the optimal q is 0.23 and the limiting variance for m=32 is 0.0028 while with m=64 it is 0.0014, both are substantially lower than the Monte Carlo result. If H=0.70 the optimal q is 0.40 and the limiting variance with m=32 is 0.0026 while for m=64 is 0.0013. There is a significative reduction in the divergence compared with the other case.

The APE has as main drawback its discontinuity in the asymptotic distribution. For this reason we will use it as a testing tool instead of employing it as an estimating tool. This will be done in chapter 6 in which we will apply the consistency of the cross-periodogram as a tool for proving the consistency of a LM test for I(0).

TABLE 3.1

Optimal q and limiting variance

H	Optimal q	Limiting Variance
0.51	0.21	0.0949
0.52	0.21	0.0934
0.53	0.22	0.0918
0.54	0.22	0.0903
0.55	0.23	0.0888
0.56	0.23	0.0873
0.57	0.24	0.0859
0.58	0.25	0.0845
0.59	0.26	0.0831
0.60	0.26	0.0818
0.61	0.27	0.0806
0.62	0.28	0.0795
0.63	0.29	0.0786
0.64	0.30	0.0778
0.65	0.31	0.0772
0.66	0.33	0.0770
0.67	0.34	0.0771
0.68	0.36	0.0778
0.69	0.38	0.0794
0.70	0.40	0.0824
0.71	0.43	0.0876
0.72	0.47	0.0973
0.73	0.53	0.1178
0.74	0.62	0.1795

TABLE 3.2

Summary statistics of $\hat{H}_q$ for various m and H for q=0.5

H	Mean m: 32	Mean 64	Variance 32	Variance 64	Skewness 32	Skewness 64	Kurtosis 32	Kurtosis 64
0.55	.536	.546	.015	.007	-.548	-.398	.388	.209
0.70	.666	.689	.010	.004	-.652	-.504	.578	.373
0.80	.742	.770	.007	.003	-.735	-.580	.748	.527
0.95	.838	.867	.004	.002	-.888	-.724	1.111	.875

TABLE 3.3

Summary statistics of $\hat{H}_q$ for various m, q and H.

H	q	Mean m: 32	Mean 64	Variance 32	Variance 64	Skewness 32	Skewness 64	Kurtosis 32	Kurtosis 64
	0.20	.502	.519	.014	.007	-.567	-.413	.485	.280
0.55	0.23	.508	.524	.013	.006	-.554	-.405	.397	.284
	0.50	.536	.546	.015	.007	-.548	-.398	.388	.209
	0.80	.490	.541	.038	.017	-.736	-.574	.736	.431
	0.20	.630	.659	.011	.005	-.680	-.505	.718	.415
0.70	0.40	.636	.676	.010	.004	-.631	-.444	.567	.178
	0.50	.666	.689	.010	.004	-.653	-.504	.578	.373
	0.80	.638	.691	.021	.009	-.830	-.663	1.001	.654
	0.20	.706	.740	.009	.004	-.777	-.574	.968	.539
0.80	0.50	.742	.770	.007	.003	-.735	-.580	.748	.528
	0.80	.725	.775	.014	.005	-.918	-.747	1.273	.890
	0.20	.802	.840	.006	.003	-.989	-.730	1.664	.886
0.95	0.50	.838	.867	.004	.002	-.888	-.724	1.111	.876
	0.80	.831	.875	.007	.002	-1.093	-.909	1.865	1.394

Figure 1

Histograms of $\hat{H}q$ for m=32, q=0.5 and various H.

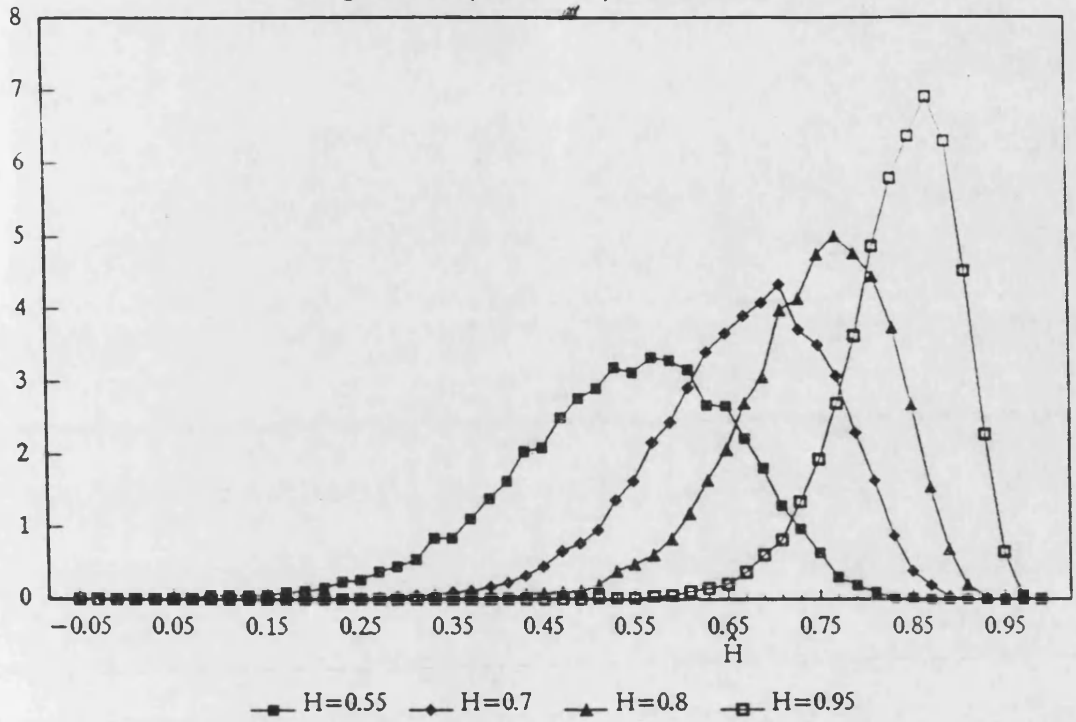


Figure 2

Histograms of $\hat{H}q$ for m=64, q=0.5 and various H.

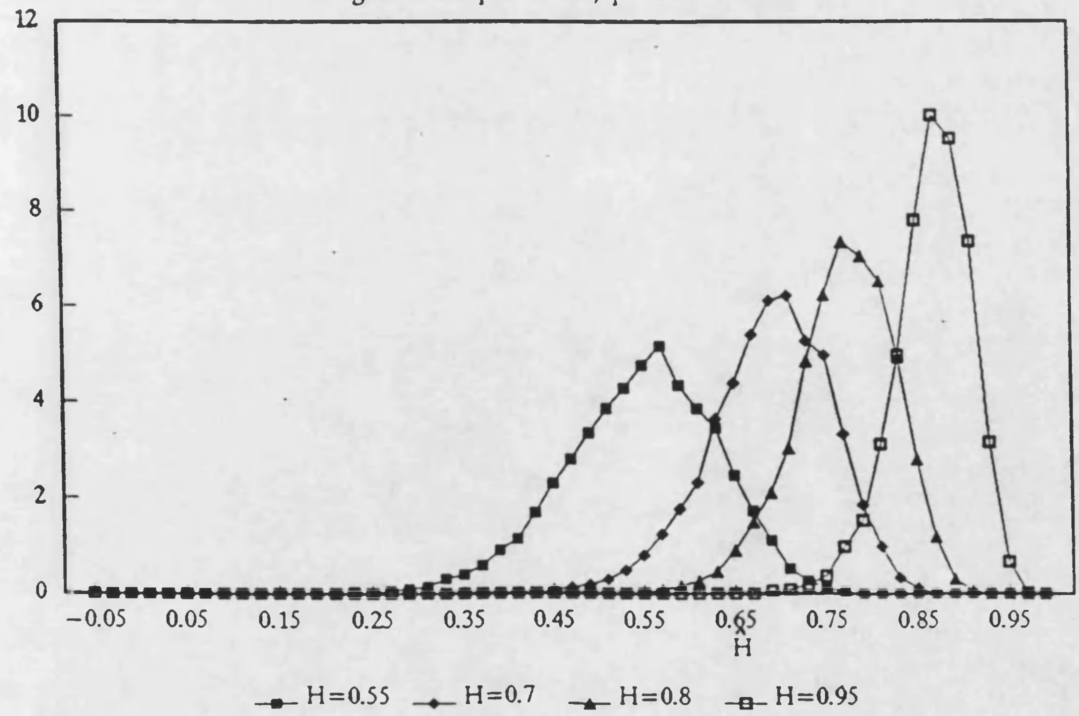


Figure 3

Histograms of $\hat{H}q$ for m=32, H=0.55 and various q.

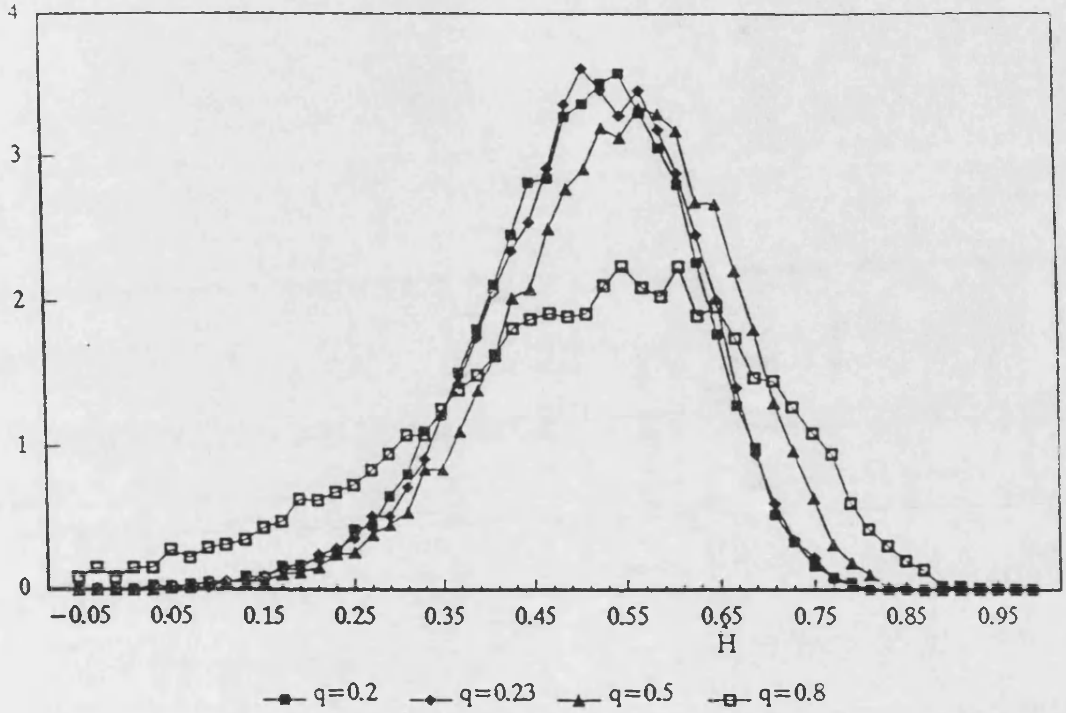


Figure 4

Histograms of $\hat{H}q$ for m=64, H=0.55 and various q.

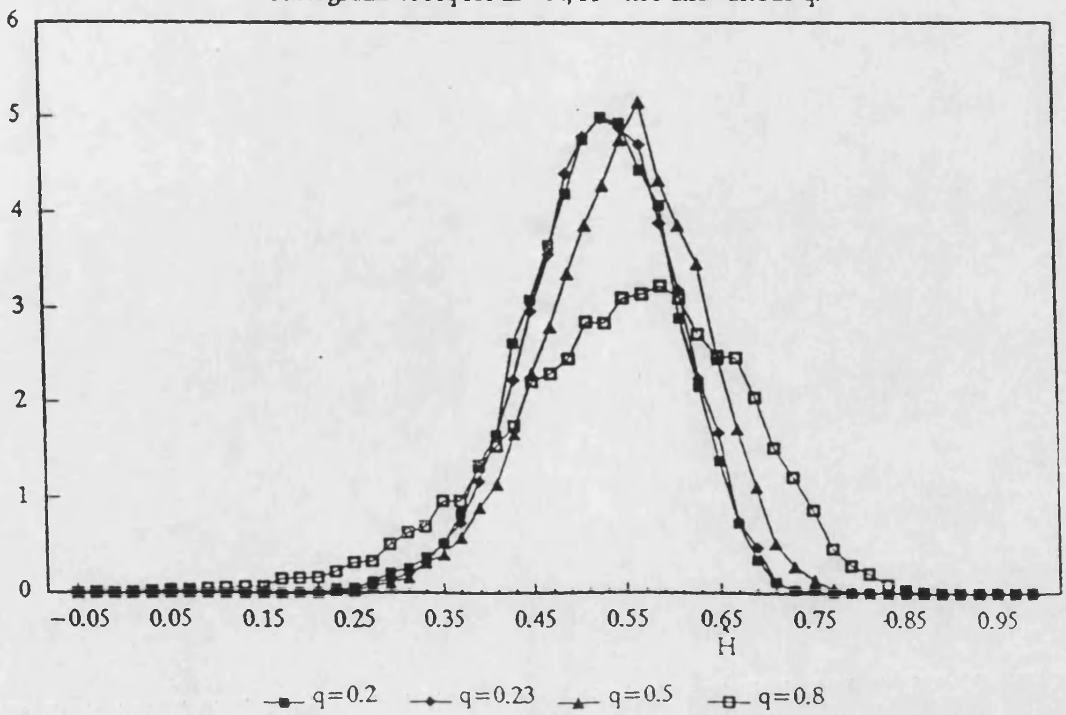


Figure 5

Histograms of $\hat{H}q$ for m=32, H=0.70 and various q.

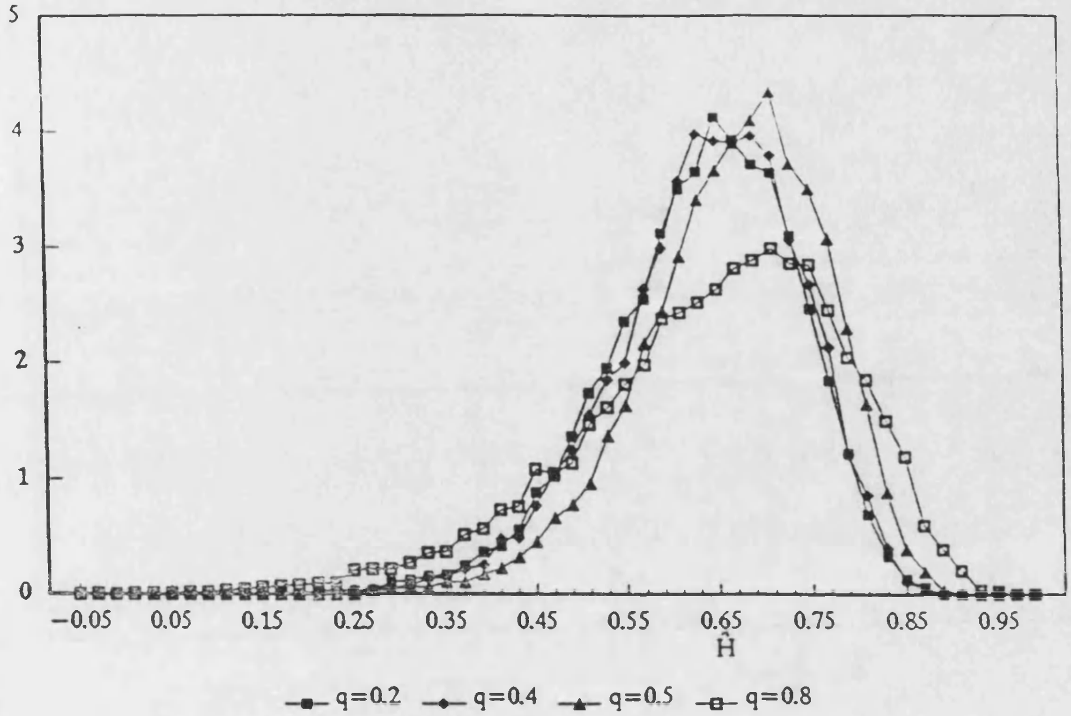


Figure 6

Histograms of $\hat{H}q$ for m=64, H=0.7 and various q.

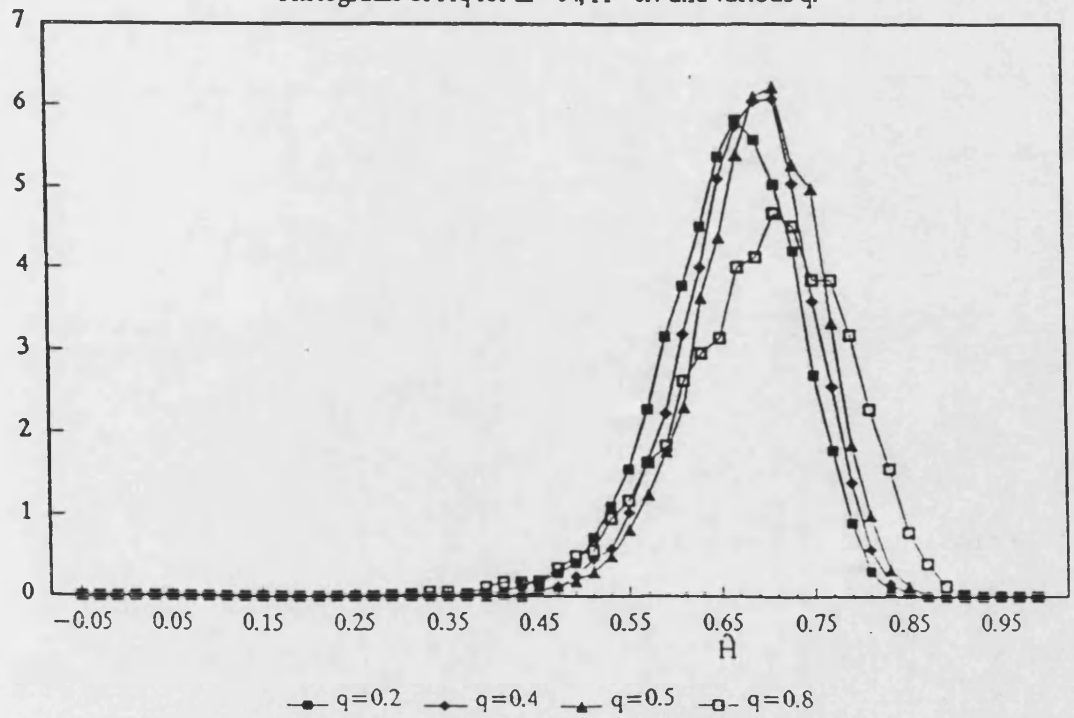


Figure 7

Histograms of $\hat{H}q$ for m=32, H=0.8 and various q.

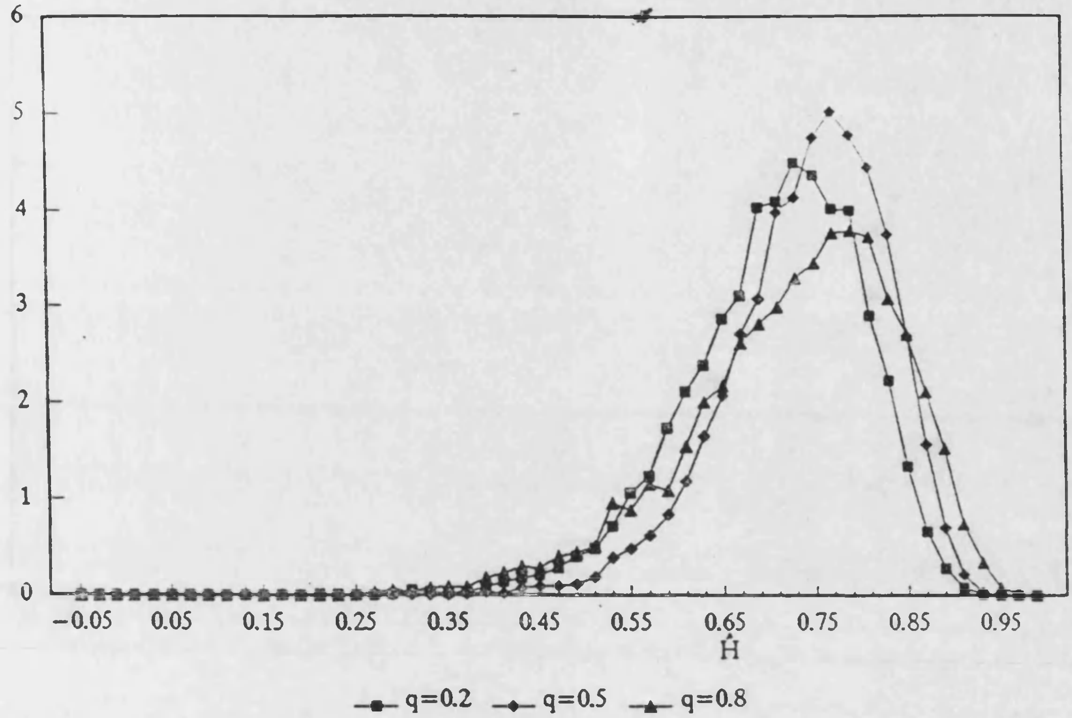


Figure 8

Histograms of $\hat{H}q$ for m=64, H=0.8 and various q.

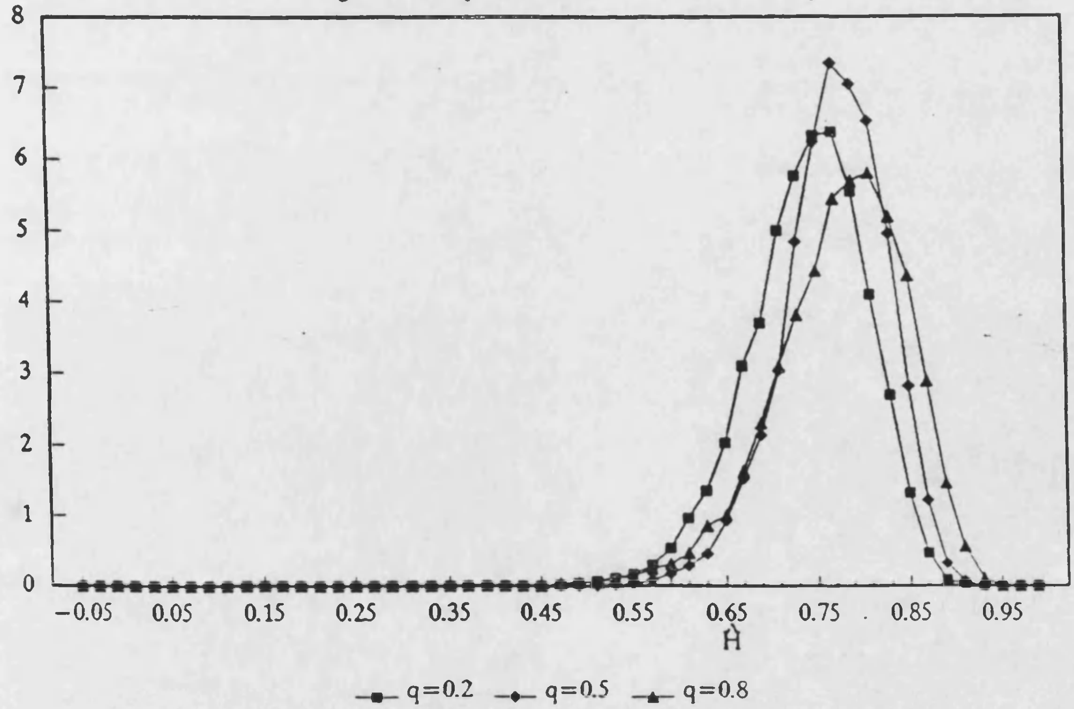


Figure 9

Histograms of $\hat{H}q$ for m=32, H=0.95 and various q.

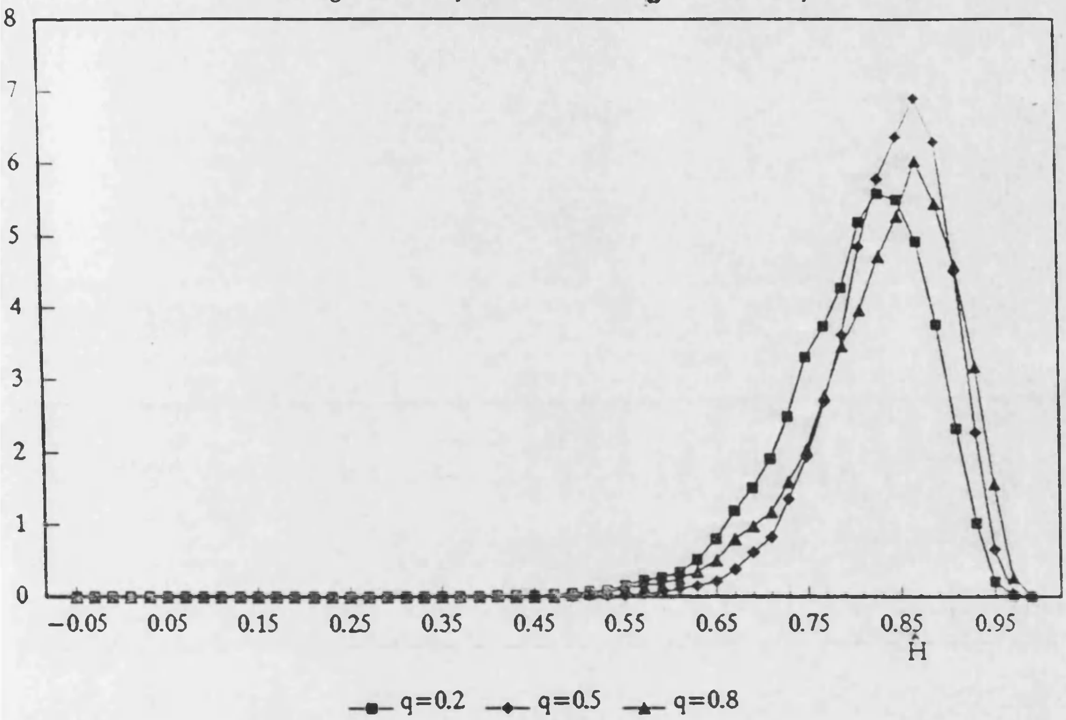
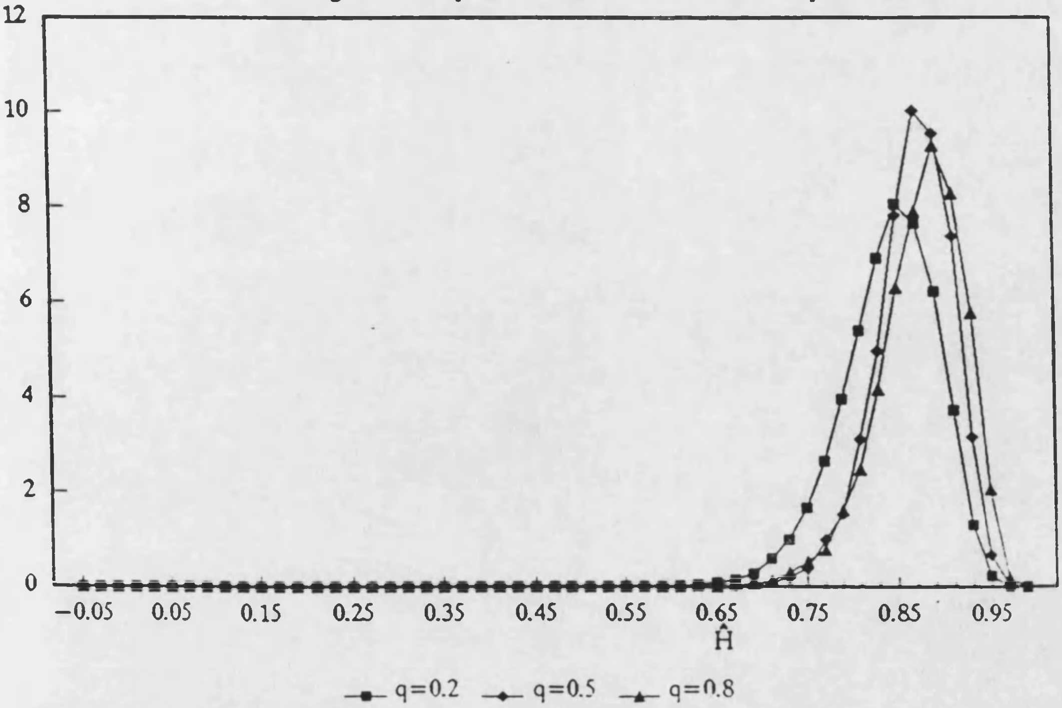


Figure 10

Histograms of $\hat{H}q$ for m=64, H=0.95 and various q



Appendix 3.1

As

$$G(\lambda_m)=\frac{C\lambda_m^{2-2H}}{2-2H} \quad and \quad G(q\lambda_m)=\frac{Cq^{2-2H}\lambda_m^{2-2H}}{2-2H},$$

$$H=1-\frac{1}{2\log q}\log\left(\frac{G(q\lambda_m)}{G(\lambda_m)}\right),$$

while

$$\hat{H}=1-\frac{1}{2\log q}\log\left(\frac{\hat{F}(q\lambda_m)}{\hat{F}(\lambda_m)}\right),$$

so

$$\hat{H}-H=\frac{1}{2\log q}\left(\log\left(\frac{\hat{F}(\lambda_m)}{G(\lambda_m)}\right)-\log\left(\frac{\hat{F}(q\lambda_m)}{G(q\lambda_m)}\right)\right)=$$

$$\frac{1}{2\log q}\left[\log(1+R(\lambda_m))-\log(1+R(q\lambda_m))\right] \qquad \text{(A3.1.1)}$$

where

$$R(\lambda)=\frac{\hat{F}(\lambda)}{G(\lambda)}-1.$$

Using $|\log(1+x)-x| \leq x^2/2$ for $x>0$, (A3.1.1) is equal to:

$$\frac{1}{2\log q}[R(\lambda_m)-R(q\lambda_m)]+O_p(E\{R^2(\lambda_m)\})+O_p(E\{R^2(q\lambda_m)\}).$$

Theorem 1 in Robinson (1994b) analyzes $E(R^2(\lambda_m))$ under the same assumptions we have. He obtains:

$$E(R^2(\lambda_m))\sim 4(1-H)^2\left[\frac{1}{(3-4H)m}+\left(\frac{2\pi m}{n}\right)^{2\alpha}\left(\frac{E_\alpha}{2-2H+\alpha}\right)^2\right],$$

now, as condition C3.3 implies $(m/n)^{2\alpha}=o(1/m)$, we have that:

$$E\{R^2(\lambda_m)\}=O(\frac{1}{m})\ ,\quad E\{R^2(q\lambda_m)\}=O(\frac{1}{m})\ .$$

In particular they are:

$$E\{R^2(\lambda_m)\}\to\frac{4(1-H)^2}{(3-4H)m}\ ,\ E\{R^2(q\lambda_m)\}\to\frac{4(1-H)^2}{(3-4H)qm}\ , \tag{A3.1.2}$$

and so, (A3.1.1) is:

$$\frac{1}{2\log q}[R(\lambda_m)-R(q\lambda_m)]+O_p(\frac{1}{m}).$$

Therefore:

$$\sqrt{m}(\hat{H}-H)=\frac{\sqrt{m}}{2\log q}[R(\lambda_m)-R(q\lambda_m)]+O_p(\frac{1}{\sqrt{m}})=$$

$$=\frac{\sqrt{m}}{2\log q}[R(\lambda_m)-R(q\lambda_m)]+o_p(1)\ .$$

So, in order to prove the theorem we need to prove:

a)

$$\frac{\sqrt{m}}{2\log q}R(\lambda_m)\to_d N\left(0,\frac{1}{(2\log q)^2}\frac{(2-2H)^2}{3-4H}\right),$$

and

b)

$$E\left(\frac{\sqrt{m}}{2\log q}R(\lambda_m)\frac{\sqrt{m}}{2\log q}R(q\lambda_m)\right)=\frac{(1-H)^2q^{1-2H}}{(3-4H)(\log q)^2}\ .$$

Proof of a). Call $g_j=C\lambda_j^{1-2H}$, then

$$m^{1/2}R(\lambda_m)=\sqrt{m}\left(\frac{\frac{2\pi}{n}\sum_1^m I(\lambda_j)}{G(\lambda_m)}-1\right)=\frac{(2-2H)}{m^{3/2-2H}}\sum_1^m j^{1-2H}\frac{I(\lambda_j)}{g_j}-\sqrt{m}=$$

$$\frac{(2-2H)}{m^{3/2-2H}}\sum_1^m j^{1-2H}\left(\frac{I(\lambda_j)}{g_j}-1\right)+\left(\frac{(2-2H)}{m^{3/2-2H}}\sum_1^m j^{1-2H}-\sqrt{m}\right). \qquad (A3.1.3)$$

As

$$\sqrt{m}(2-2H)\sum_{j=1}^m\left(\frac{j}{m}\right)^{1-2H}\frac{1}{m}\sim\sqrt{m}(2-2H)\int_0^1 x^{1-2H}dx=\sqrt{m}\ ,$$

the second component of (A3.1.3) tends to 0, as n tends to infinity.

So we need to show:

$$\frac{1}{m^{3/2-2H}}\sum_1^m j^{1-2H}\left(\frac{I(\lambda_j)}{g_j}-1\right)\rightarrow_d N(0,\frac{1}{3-4H})\ .$$

The method of proof is developed in Robinson (1993a). In order to prove this CLT we will employ the following formulae that are derived in Robinson (1993a) under the conditions we have assumed, as n $\rightarrow\infty$:

$$\sum_{j=1}^r(\frac{I(\lambda_j)}{g_j}-2\pi J_j)=O_p(r^{1/3}(\log r)^{2/3}+\frac{r^{\alpha+1}}{n^\alpha}+\frac{r^{1/2}}{n^{1/4}}); \qquad (A3.1.4)$$

$$\sum_{t=1}^{n-1}\sum_{s=1}^{n-t}\cos^2(s\lambda_j)=\frac{(n-1)^2}{4},\ for\ \lambda_j=\frac{2\pi j}{n},j=1,...m, \qquad (A3.1.5)$$

$$\sum_{t=1}^{n-1}\sum_{s=1}^{n-t}\left[\cos\{s(\lambda_j+\lambda_k)\}+\cos\{s(\lambda_j-\lambda_k)\}\right]=-n\ ,\ for\ \lambda_j\neq\lambda_k,\ \ j,k=1,...m. \qquad (A3.1.6)$$

Then:

$$m^{2H-\frac{3}{2}}\sum_1^m j^{1-2H}\left(\frac{I(\lambda_j)}{g_j}-1\right)=m^{2H-\frac{3}{2}}\sum_1^m j^{1-2H}\left(\frac{I(\lambda_j)}{g_j}-2\pi J_j\right)+$$

$$+m^{2H-\frac{3}{2}}\sum_1^m j^{1-2H}(2\pi J_j-1)=(X1)+(X2)\ ,$$

where J_j is the periodogram of iid standard Gaussian variates e_t.

(X1) by summation by parts and using (A3.1.4) is:

$$O_p(m^{-\frac{1}{6}}(\log m)^{2/3}+\frac{m^{\alpha+\frac{1}{2}}}{n^{\alpha}}+n^{-\frac{1}{4}})=o_p(1)\ ,$$

while (X2) is:

$$m^{2H-\frac{3}{2}}\sum_{1}^{m} j^{1-2H}\left(\frac{1}{n}\sum_{1}^{n} e_t^2-1\right)+$$

$$+m^{2H-\frac{3}{2}}\sum_{1}^{m} j^{1-2H}\left(\frac{2}{n}\sum_{t=2}^{n}\sum_{s=1}^{t-1} e_t e_s \cos(t-s)\lambda_j\right)$$

$$=\ (X21)+\ (X22).$$

As

$$\frac{1}{n}\sum_{1}^{n} e_t^2-1$$

has zero mean and variance 2/n, so

$$\frac{1}{n}\sum_{1}^{n} e_t^2-1=O_p(\frac{1}{\sqrt{n}})$$

so

$$(X21)=O_p(\frac{m^{2H-3/2}}{\sqrt{n}}\sum_{j=1}^{m} j^{1-2H})=O_p((\frac{m}{n})^{\frac{1}{2}}).$$

We can rewrite (X22) as:

$$(X22)=\sum_{t=1}^{n} z_t\ ,\ z_1=0,\ z_t=e_t\sum_{s=1}^{t-1} e_s c_{t-s}\quad t\geq 2\ ,$$

with

$$c_s=\frac{2m^{2H-\frac{3}{2}}}{n}\sum_{j=1}^{m} j^{1-2H}\cos s\lambda_j\ .$$

z_t is a zero mean martingale difference so we use a standard CLT (Brown, 1971) for martingale differences stated in Appendix 5.2. In order to prove:

$$\sum_{t=1}^{n} z_t \to_d N(0, \frac{1}{3-4H}),$$

we just need to prove:

$$(*)\ \sum_{t=1}^{n} E(z_t^2 | F_{t-1}) - \frac{1}{3-4H} \to_p 0\ , \tag{A3.1.7}$$

and

$$(**) \sum_{1}^{n} E(z_t^2 I(|z_t|>\delta)) \to 0 \ \textit{for all}\ \delta>0\ . \tag{A3.1.8}$$

where I() is the indicator function. First we proof (A3.1.7):

$$LHS(*) = \left(\sum_{t=2}^{n} \sum_{s=1}^{t-1} e_s^2 c_{t-s}^2 - \frac{1}{3-4H} \right) + \sum_{t=2}^{n} \sum_{r \neq} \sum_{s} e_r e_s c_{t-r} c_{t-s} = (1) + (2)\ ,$$

$$(1) = \left(\sum_{t=1}^{n-1} (e_t^2 - 1) \sum_{s=1}^{n-t} c_s^2 \right) + \left(\sum_{t=1}^{n-1} \sum_{s=1}^{n-t} c_s^2 - \frac{1}{3-4H} \right) = (1A) + (1B)\ ,$$

now consider

$$\sum_{t=1}^{n-1} \sum_{s=1}^{n-t} c_s^2 = \frac{4m^{4H-3}}{n^2} \sum_{j=1}^{m} j^{2-4H} \sum_{t=1}^{n-1} \sum_{s=1}^{n-t} \cos^2 s\lambda_j +$$

$$\frac{8m^{4H-3}}{n^2} \sum_{j\neq} \sum_{k} j^{1-2H} k^{1-2H} \sum_{t=1}^{n-1} \sum_{s=1}^{n-t} [\cos[s(\lambda_j+\lambda_k)] + \cos[s(\lambda_j-\lambda_k)]]$$

$$= (0) + (00)\ ;$$

using (A3.1.5) and (A3.1.6) we get:

$$(0) \sim \frac{4m^{4H-3}}{n^2} \frac{(n-1)^2}{4} \frac{m^{3-4H}}{3-4H} \sim \frac{1}{3-4H}\ ,$$

$$(00) = O(\frac{m^{4H-3}}{n^2} n \sum_{j\neq}^{m} \sum_{k}^{m} j^{1-2H} k^{1-2H}) = O(\frac{m}{n}) = o(1)\ ,$$

so,

$$(1B) \to 0.$$

While (1A) has zero mean and variance:

$$=O(\sum_{t=1}^{n-1}(\sum_{s=1}^{n-t}c_s^2)^2) \; .$$

As:

$$|c_s| \leq \frac{2m^{2H-\frac{3}{2}}}{n}\sum_{1}^{m} j^{1-2H} = O(\frac{\sqrt{m}}{n}) \; ,$$

and using Zygmund (1977), p.70:

$$\sum_{j=1}^{m} j^{1-2H} cos s\lambda_j = O(\lambda_s^{2H-2}) \; , \; as \quad m \to \infty, \quad for \; 0<\lambda_s \leq \pi \; ,$$

so:

$$|c_s| = O(\frac{s^{2H-2} m^{2H-\frac{3}{2}}}{n^{2H-1}}) \; , \; for \quad s<\frac{n}{2} \; .$$

Then

$$\sum_{s=1}^{n} c_s^2 = \sum_{s=1}^{[\frac{n}{m}]} c_s^2 + \sum_{[\frac{n}{m}]+1}^{n} c_s^2 = O(\frac{n}{m}\frac{m}{n^2} + \frac{m^{4H-3}}{n^{4H-2}}\sum_{[\frac{n}{m}]}^{n} s^{4H-4}) = O(\frac{1}{n}) \; ,$$

so

$$variance[(1A)] = O(n\frac{1}{n^2}) = O(\frac{1}{n}) \; .$$

So: (1) → 0. Now (2) has zero mean and variance:

$$=O(\sum_{t=2}^{n-1}\sum_{r\neq}\sum_{s} c_{t-r}^2 c_{t-s}^2 + 4\sum_{t=3}^{n-1}\sum_{u=2}^{t-1}\sum_{r\neq}^{u-1}\sum_{s}^{u-1} c_{t-r}c_{t-s}c_{u-r}c_{u-s}) =$$

$$O\left(n(\sum_{1}^{n} c_s^2)^2 + n(\sum_{1}^{n} c_s^2)\sum_{1}^{[n/2]} jc_j^2\right) = O\left(\frac{1}{n} + \sum_{1}^{[\frac{n}{m^{2/3}}]} jc_j^2 + \sum_{[\frac{n}{m^{2/3}}]+1}^{[n/2]} jc_j^2\right) =$$

$$O\left(\frac{1}{n}+\frac{m}{n^2}\left(\frac{n}{m^{2/3}}\right)^2+\frac{m^{4H-3}}{n^{4H-2}}\sum_{[\frac{n}{m^{2/3}}]+1}^{[n/2]} j^{4H-3}\right)=$$

$$=O\left(\frac{1}{n}+\frac{1}{m^{1/3}}+\frac{1}{m^{(5-4H)/3}}\right)=o(1)$$

so (A3.1.7) is proven.

Now instead of (A3.1.8) we proof a sufficient condition for (A3.1.8) that is[1]:

$$\sum_{t=1}^{n} E(z_t^4)\to 0 \ ,$$

using the previous results we get:

$$\sum_{1}^{n} E(z_t^4)=O\left(n\left(\sum_{1}^{n} c_t^2\right)^2\right)=o(1) \ .$$

Proof of b).

We are going to prove:

$$E(mR(\lambda_m)R(q\lambda_m))=\frac{4(1-H)^2 q^{1-2H}}{(3-4H)} \ , \tag{A3.1.9}$$

define:

$$\hat{F}_1=\hat{F}(q\lambda_m),\ \tilde{F}_1=\int_0^{q\lambda_m}\tilde{I}(\lambda)d\lambda,\quad F_1=\int_0^{q\lambda_m} f(\lambda)d\lambda,\quad G_1=G(q\lambda) \ ,$$

$$\hat{F}_2=\frac{2\pi}{n}\sum_{j=[qm]+1}^{m} I(\lambda_j),\ \tilde{F}_2=\int_{q\lambda_m}^{\lambda_m}\tilde{I}(\lambda)d\lambda \ ,$$

$$\tilde{I}(\lambda)=\frac{1}{2\pi n}\left|\sum_{t=1}^{n}(x_t-Ex_1)e^{it\lambda}\right|^2 \ ,$$

[1] See (A5.2.6) in appendix 5.2.

$$F_2=\int_{q\lambda_m}^{\lambda_m} f(\lambda)d\lambda,\; G_2=G(\lambda_m)-G_1=(1-q^{2-2H})G(\lambda_m)\ .$$

Now, using $G_1=q^{2-2H}G(\lambda_m)$:

$$R(\lambda_m)=q^{2-2H}R(q\lambda_m)+(1-q^{2-2H})S_{qm} \quad where \quad S_{qm}=\left(\frac{\hat{F}_2}{G_2}-1\right)\ .$$

So:

$$LHS((A3.1.9))=q^{2-2H}E\{mR^2(q\lambda_m)\}+(1-q^{2-2H})E(mR(q\lambda_m)S_{qm}\ ).$$

Now, from (A3.1.2), we have that the first term tends to:

$$\frac{4(1-H)^2}{(3-4H)}q^{1-2H}\ ;$$

so, it remains to prove that the second term asymptotically vanishes,i.e.:

$$E(mR(q\lambda_m)S_{mq})=o(1)\ .$$

Define:

$$R(q\lambda_m)=A_1+B_1+C_1+D_1\ ,\quad S_{mq}=A_2+B_2+C_2+D_2\ ,$$

where:

$$A_i=\frac{\hat{F}_i-E\hat{F}_i}{G_i},\ B_i=\frac{E\hat{F}_i-E\tilde{F}_i}{G_i},\ C_i=\frac{E\tilde{F}_i-F_i}{G_i},\ D_i=\frac{F_i}{G_i}-1\ ,$$

for i=1,2. In order to examine the order of magnitude of these terms, we need to state lemmas 1,3,4 and 7 of Robinson (1994b) that are established under the same conditions as ours:

<u>Lemma 1</u>

$$\frac{F(\lambda)}{G(\lambda)}-1\sim E_\alpha\frac{2(1-H)}{2(1-H)+\alpha}\lambda^\alpha,\ as\ \lambda\to0^+\ ;$$

Lemma 3

$$E\{\tilde{F}(\lambda)-F(\lambda)\}=O(\frac{g(\lambda)}{n})\ ,\ as\ \lambda\to 0^{+}\ ;$$

Lemma 4

$$E\{\tilde{F}(\lambda_m)-\hat{F}(\lambda_m)\}=O(m^{\eta}n^{2H-2}),\ as\ n\to\infty\ ,\ for\ any\ \eta>0\ ;$$

and Lemma 7

$$E[\hat{F}(\lambda_m)-E\hat{F}(\lambda_m)]^2\sim\frac{1}{m}\frac{\lambda_m^{4-4H}}{3-4H}\ ,\ as\ n\to\infty\ .$$

Then, using these lemmas for D_1, C_1, B_1 and A_1, respectively:

$D_1=O((m/n)^{\alpha})=o(m^{-1/2})$, using also condition C3,

$$C_1=O\left(\frac{g(\lambda_m)}{G(\lambda_m)n}\right)=O(\frac{1}{m})=o(m^{-1/2}),$$

$$B1=O\left(\frac{m^{\eta}n^{2H-2}}{G(\lambda_m)}\right)=O(m^{2H-2+\eta})=o(m^{-1/2})\ for\ 0<\eta<\frac{3}{2}-2H,\ \text{and}$$

$$A1=O_p\left(\sqrt{E\{R(\lambda_m)\}^2}\right)=O_p\left(\sqrt{\frac{1}{m}\frac{\lambda_m^{4-4H}}{\lambda_m^{4-4H}}}\right)=O_p(m^{-1/2}).$$

And similarly using that G_2 has the same order of magnitude as $G(\lambda_m)$, i.e., λ_m^{2-2H}:

$$D_2=\frac{F(\lambda_m)-F_1-(G(\lambda_m)-G_1)}{G_2}=\frac{F(\lambda_m)-G(\lambda_m)}{G_2}-\frac{F_1-G_1}{G_2}=o(m^{-1/2}),$$

$$C_2=\frac{E\tilde{F}(\lambda_m)-E\tilde{F}_1-(F-F_1)}{G_2}=\frac{E\tilde{F}(\lambda_m)-F}{G_2}-\frac{E\tilde{F}_1-F_1}{G_2}=o(m^{-1/2}),$$

$$B_2=\frac{E\hat{F}(\lambda_m)-E\hat{F}_1-(E\tilde{F}(\lambda_m)-E\tilde{F}_1)}{G_2}=\frac{E\hat{F}(\lambda_m)-E\tilde{F}(\lambda_m)}{G_2}-\frac{E\hat{F}_1-E\tilde{F}_1}{G_2}=o(m^{-1/2}),\ \text{and}$$

$$A_2=\frac{\hat{F}(\lambda_m)-\hat{F}_1-(E\hat{F}(\lambda_m)-E\hat{F}_1)}{G_2}=\frac{\hat{F}(\lambda_m)-E\hat{F}(\lambda_m)}{G_2}-\frac{\hat{F}_1-E\hat{F}_1}{G_2}=O_p(m^{-1/2}).$$

So all cross products between $R(q\lambda_m)$ and S_{qm} will be inmediately $o(m^{-1})$ or $o_p(m^{-1})$ except the one between A_1 and A_2.

So, in order to complete the proof we just need to prove:

$$E(A_1A_2)=o(m^{-1}),\ as\ n\to\infty\ .$$

As

$$E(A_1A_2)=G_1^{-1}G_2^{-1}E[(\hat{F}_1-E\hat{F}_1)(\hat{F}_2-E\hat{F}_2)]\ ,$$

first

$$G_1^{-1}G_2^{-1}=O\left(\left(\frac{m}{n}\right)^{4H-4}\right)\ ,$$

and

$$E[(\hat{F}_1-E\hat{F}_1)(\hat{F}_2-E\hat{F}_2)]=\left(\frac{2\pi}{n}\right)^2\sum_{j=1}^{[qm]}\sum_{k=[qm]+1}^{m}cov\{I(\lambda_j),I(\lambda_k)\},\tag{A3.1.10}$$

now, for zero-mean Gaussian variates, x, y, z and u:

$$E(xyzu)=E(xy)E(zu)+E(xz)E(yu)+E(xu)E(yz)\ ,$$

so we have for Fourier frequencies λ_j,λ_k:

$$cov(I(\lambda_j),I(\lambda_k))=$$

$$=E(w(\lambda_j)\overline{w(\lambda_j)}w(\lambda_k)\overline{w(\lambda_k)})-E(w(\lambda_j)\overline{w(\lambda_j)})E((w(\lambda_k)\overline{w(\lambda_k)})=$$

$$=E(w(\lambda_j)w(\lambda_k))E(w(-\lambda_j)w(-\lambda_k))+E(w(\lambda_j)w(-\lambda_k))E(w(\lambda_k)w(-\lambda_j))\ ,$$

now using that

$$E(w(\lambda_j))=0\ ,\quad for\ j=1,...m\ ,$$

and

$$E(w(\lambda_j)w(\lambda_k))=\frac{1}{2\pi n}\int_{-\pi}^{\pi}D(u)D(\lambda_j+\lambda_k-u)f(\lambda_j-u)du\ ,$$

where $D(\lambda)$ is the Dirichlet kernel:

$$D(\lambda)=\sum_{t=1}^{n} e^{it\lambda} ,$$

and following Robinson (1994b) notation:

$$Q(\lambda,\theta)=\int_{-\pi}^{\pi} D(u)D(\lambda+\theta-u)f(\lambda-u)du ,$$

we get that (A3.1.10) =

$$\frac{1}{n^4}\sum_{j=1}^{[qm]}\sum_{k=[qm]+1}^{m}(Q(\lambda_j,-\lambda_k)Q(\lambda_k,-\lambda_j)+Q(\lambda_k,\lambda_j)Q(-\lambda_j,-\lambda_k)). \qquad \text{(A3.1.11)}$$

Consequently define:

$$R(\lambda,\theta)=\int_{-\pi}^{\pi} D(u)D(\lambda+\theta-u)\{f(\lambda-u)-f(\lambda)\}du ,$$

then using:

$$\int_{-\pi}^{\pi} D(u)D(\lambda-u)du=2\pi D(\lambda) ,$$

and

$$D(\lambda_j-\lambda_k)=0 \ , for\ j\neq k \ , \ j,k=1...m ,$$

we obtain for $j \neq k$:

$$Q(\lambda_j,-\lambda_k)=R(\lambda_j,-\lambda_k) ,$$

$$Q(\lambda_j,\lambda_k)=R(\lambda_j,\lambda_k) ,$$

$$Q(-\lambda_j,-\lambda_k)=R(-\lambda_j,-\lambda_k) ,$$

as we can check for instance:

$$R(\lambda_j,-\lambda_k)=\int_{-\pi}^{\pi} D(u)D(\lambda_j-\lambda_k-u)f(\lambda_j-u)du-\int_{-\pi}^{\pi} D(u)D(\lambda_j-\lambda_k-u)f(\lambda_j)du=$$
$$Q(\lambda_j,-\lambda_k)-2\pi D(\lambda_j-\lambda_k)f(\lambda_j)=Q(\lambda_j,-\lambda_k) .$$

In order to evaluate the summands in (A3.1.11) we need to use lemma 6 in Robinson

(1994b) that establishes that:

$$\max_{j<k\leq m}\ \{|R(\lambda_j,\lambda_k)|+|R(\lambda_k,\lambda_j)|\}=O((logm)\frac{g(\lambda_j)}{\lambda_k})\ , \tag{L6.1}$$

$$\max_{j<k\leq m}\ |R(\lambda_k,-\lambda_j)|=O((logm)\frac{g(\lambda_j)}{\lambda_k-\lambda_j})\ , \tag{L6.2}$$

$$\max_{j<k\leq 2j\leq m}\ |R(\lambda_k,-\lambda_j)|=O((logm)\frac{g(\lambda_k-\lambda_j)}{\lambda_j})\ , \tag{L6.3}$$

$$\max_{j<k\leq m}\ |R(\lambda_j,-\lambda_k)|=O((logm)\frac{g(\lambda_j)}{\lambda_k-\lambda_j})\ , \tag{L6.4}$$

$$\max_{j<k\leq 2j\leq m}\ |R(\lambda_j,-\lambda_k)|=O((logm)\frac{g(\lambda_k-\lambda_j)}{\lambda_j})\ . \tag{L6.5}$$

Then:

$$(A3.1.11)=\frac{1}{n^4}\sum_{j=1}^{[qm]}\sum_{k=[qm]+1}^{m}(R(\lambda_j,-\lambda_k)R(\lambda_k,-\lambda_j))+ \qquad (x0)$$

$$+\frac{1}{n^4}\sum_{j=1}^{[qm]}\sum_{k=[qm]+1}^{m}(R(\lambda_k,\lambda_j)R(-\lambda_j,-\lambda_k))\ , \qquad (x01)$$

(x01) is straightforwardly using (L6.1) and $R(-\lambda_j,-\lambda_k)=\overline{R(\lambda_j,\lambda_k)}$:

$$O(\frac{(logm)^2}{n^4}\sum_{j=1}^{[qm]}\sum_{k=[qm]+1}^{m}\frac{g(\lambda_j)^2}{\lambda_k^2})=O(\frac{(logm)^2}{n^4}n^{4H}\sum_{j=1}^{[qm]}j^{2-4H}\sum_{k=[qm]+1}^{m}k^{-2})=$$

$$=O(\frac{(logm)^2}{n^{4-4H}}\frac{m^{3-4H}}{m})=O(\frac{(logm)^2}{n^{4-4H}}m^{2-4H})\ ,$$

while

$$(x0)=\frac{1}{n^4}\sum_{j=1}^{[\frac{1}{2}qm]}\sum_{k=[qm]+1}^{m}(R(\lambda_j,-\lambda_k)R(\lambda_k,-\lambda_j))+ \qquad (a1)$$

$$+\frac{1}{n^4}\sum_{j=[\frac{1}{2}qm]+1}^{[qm]}\sum_{k=[qm]+1}^{\min(2j,m)}(R(\lambda_j,-\lambda_k)R(\lambda_k,-\lambda_j))+ \qquad (a2)$$

$$+\frac{1}{n^4}\sum_{j=[\frac{1}{2}qm]+1}^{[qm]}\sum_{k=2j+1}^{m}(R(\lambda_{j},-\lambda_k)R(\lambda_{k},-\lambda_j)) \ , \qquad (a3)$$

where (a3) is zero when $j \geq m/2$.

Then (a1), using (L6.2) and (L6.4), is:

$$O(\frac{(logm)^2}{n^4}\sum_{j=1}^{[\frac{1}{2}qm]}\sum_{k=[qm]+1}^{m}\frac{(g(\lambda_j)^2}{(\lambda_k-\lambda_j)^2})=O(\frac{(logm)^2}{n^{4-4H}}\sum_{j=1}^{[\frac{1}{2}qm]}j^{2-4H}\sum_{k=[qm]+1}^{m}(k-j)^{-2})=$$

$$=O(\frac{(logm)^2}{n^{4-4H}}\sum_{j=1}^{m}j^{2-4H}\frac{1}{m})=O(\frac{(logm)^2}{n^{4-4H}}m^{2-4H}) \ ,$$

(a2), using (L6.3) and (L6.5), is:

$$O(\frac{(logm)^2}{n^4}\sum_{j=[\frac{1}{2}qm]+1}^{[qm]}\sum_{k=[qm]+1}^{\min(2j,m)}\frac{g(\lambda_k-\lambda_j)^2}{\lambda_j^2})=$$

$$=O(\frac{(logm)^2}{n^{4-4H}}\sum_{j=[\frac{1}{2}qm]+1}^{[qm]}j^{-2}\sum_{k=[qm]+1}^{\min(2j,m)}(k-j)^{2-4H})=$$

$$=O(\frac{(logm)^2}{n^{4-4H}}\sum_{j=[\frac{1}{2}qm]+1}^{m}j^{-2}m^{3-4H})=O(\frac{(logm)^2}{n^{4-4H}}m^{2-4H}) \ ,$$

and, finally, (a3) using (L6.4) is:

$$O(\frac{(logm)^2}{n^{4-4H}}\sum_{j=[\frac{1}{2}qm]+1}^{[qm]}j^{2-4H}\sum_{k=2j+1}^{m}\frac{1}{(k-j)^2})=O(\frac{(logm)^2}{n^{4-4H}}\sum_{j=[\frac{1}{2}qm]+1}^{[qm]}j^{2-4H}\frac{1}{j})=$$

$$=O(\frac{(logm)^2}{n^{4-4H}}\sum_{j=[\frac{1}{2}qm]+1}^{\infty}j^{1-4H})=O(\frac{(logm)^2}{n^{4-4H}}m^{2-4H}) \ ,$$

i.e.:

$$(x0)=O(\frac{(logm)^2}{n^{4-4H}}m^{2-4H}) \ ,$$

the same rate as (x01), and, so:

$$(A3.1.10)=O(\frac{(logm)^2}{n^{4-4H}}m^{2-4H}) \ ;$$

so,

$$E(A_1A_2)=O(\left(\frac{m}{n}\right)^{4H-4}\frac{(logm)^2}{n^{4-4H}}m^{2-4H})=O(\frac{(logm)^2}{m^2})=o(\frac{1}{m}) \ ,$$

as we wanted to prove.

Appendix 3.2

From appendix 3.1 we have:

$$\hat{H}_q-H=\frac{1}{2\log q}[R(\lambda_m)-R(q\lambda_m)]+O_p(E\{R^2(\lambda_m)\})+O_p(E\{R^2(q\lambda_m)\}) \ .$$

As theorem 2 in Robinson (1994b) establishes:

$$E\{R^2(\lambda_m)\} \sim A_1m^{4H-4}+A_2\frac{m^{2H-2+\alpha}}{n^\alpha}+A_3\left(\frac{m}{n}\right)^{2\alpha} \ ,$$

and condition C3.3 implies:

$$n^{-\alpha}=o(m^{-\alpha-1/2}) \ , \ so \ , \ \frac{m^{2H-2+\alpha}}{n^\alpha}=o(m^{2H-5/2}) \quad (A)$$

$$and \quad \left(\frac{m}{n}\right)^{2\alpha}=o(m^{-1}) \ , \quad (B)$$

both, (A) and (B), are o(m^{4H-4}).

So:

$$E\{R^2(\lambda_m)\}=O(m^{4H-4}), \quad E\{R^2(q\lambda_m)\}=O(m^{4H-4}) \ ,$$

therefore:

$$m^{2-2H}(\hat{H}_q-H)=\frac{1}{2\log q}m^{2-2H}[R(\lambda_m)-R(q\lambda_m)]+o_p(1)\ .$$

Now, if we are able to prove that

$$m^{2-2H}R(\lambda_m)=\frac{D(1-H)}{C(2\pi)^{2-2H}}T_n+o(1)\ , \qquad \text{(A3.2.1)}$$

where D has been defined in p.40, then this implies that

$$m^{2-2H}R(q\lambda_m)=\frac{D(1-H)q^{2H-2}}{C(2\pi)^{2-2H}}T_n+o(1)\ ,$$

and so,

$$\frac{m^{2-2H}}{2\log q}[R(\lambda_m)-R(q\lambda_m)]=\frac{(1-H)\Gamma(2(1-H))\cos((1-H)\pi)(1-q^{2H-2})}{(\log q)(2\pi)^{2-2H}}T_n+o(1)\ ,$$

and this converges in distribution to:

$$\frac{(1-q^{2H-2})}{\log q}\frac{(1-H)\Gamma(2(1-H))\cos((1-H)\pi)}{(2\pi)^{2-2H}}(1-q^{2H-2})T\ .$$

So, it just remains to prove (A3.2.1), i.e.:

$$m^{2-2H}R(\lambda_m)=\frac{D(1-H)}{C(2\pi)^{2-2H}}\frac{(S_n-ES_n)-n(\bar{x}-Ex_1)^2}{Dn^{2H-1}}+o_p(1)\ .$$

First we analyze $R(\lambda_m)$:

$$R(\lambda_m)=\frac{\hat{F}(\lambda_m)}{G(\lambda_m)}-1=\frac{\hat{F}(\lambda_m)-E\hat{F}(\lambda_m)}{G(\lambda_m)}+\frac{E\hat{F}(\lambda_m)-G(\lambda_m)}{G(\lambda_m)}=(AA)+(BB)\ ;$$

we analyze (BB) first:

$$\begin{aligned}E\hat{F}(\lambda_m)-G(\lambda_m)&=E\{\hat{F}(\lambda_m)-\tilde{F}(\lambda_m)\}+E\{\tilde{F}(\lambda_m)-F(\lambda_m)\}+\{F(\lambda_m)-G(\lambda_m)\}\\&=(b1)+(b2)+(b3)\ .\end{aligned}$$

In order to deal with (b1) we apply lemma 9 in Robinson (1994b) that states that, under the assumptions we have, as $n\to\infty$:

$$E\{\tilde{F}(\lambda_m)-\hat{F}(\lambda_m)\} \sim \frac{1}{2}V(\bar{x})+o(n^{2H-2}) \ ,$$

so,

$$(b1) \sim -\frac{1}{2}V(\bar{x})+o(n^{2H-2}) \ ;$$

now, using C3.3, we have:

$$(b3)=O(\left(\frac{m}{n}\right)^{2-2H+\alpha})=o(n^{2H-2}) \ ,$$

and, to analyze (b2) we use lemma 3 in Robinson (1994b) that states that:

$$E\{\tilde{F}(\lambda_m)-F(\lambda_m)\}=O(\frac{g(\lambda_m)}{n}) \ ,$$

so

$$(b2)=O(\frac{\lambda_m^{1-2H}}{n})=O(m^{1-2H}n^{2H-2})=o(n^{2H-2}) \ ;$$

therefore:

$$(BB)=-\frac{1}{2}\frac{V(\bar{x})}{G(\lambda_m)}+o(m^{2H-2}) \ .$$

In order to analyze (AA) we need to evaluate $\hat{F}(\lambda_m)$ *and* $E\hat{F}(\lambda_m)$. In order to do that, first we recall Parseval's identity:

$$S_n=\sum_{t=1}^{n}(x_t-\bar{x})^2=2\pi\sum_{j=1}^{n}\tilde{I}(\lambda_j) \ .$$

Now, for n odd, using that $\tilde{I}(\lambda_j)=\tilde{I}(2\pi-\lambda_j)$ *and* $\tilde{I}(\lambda_n)=\tilde{I}(0)=\frac{n}{2\pi}(\bar{x}-Ex_1)^2$:

$$\sum_{j=1}^{n} \tilde{I}(\lambda_j) = 2\sum_{j=1}^{\frac{n-1}{2}} \tilde{I}(\lambda_j) + I(0) = 2\sum_{j=1}^{\frac{n-1}{2}} \tilde{I}(\lambda_j) + \frac{n}{2\pi}(\bar{x} - Ex_1)^2 \ ,$$

so

$$S_n = 2\pi\left(2\sum_{j=1}^{m} \tilde{I}(\lambda_j) + 2\sum_{j=m+1}^{\frac{n-1}{2}} \tilde{I}(\lambda_j) + \frac{n}{2\pi}(\bar{x} - Ex_1)^2\right) .$$

then, using that: $\tilde{I}(\lambda_j) = I(\lambda_j)$ *for* $j = 1,...m < n$,

$$\frac{S_n}{2n} = \hat{F}(\lambda_m) + \frac{2\pi}{n}\sum_{j=m+1}^{\frac{n-1}{2}} \tilde{I}(\lambda_j) + \frac{(\bar{x} - Ex_1)^2}{2} \ ; \tag{A3.2.2}$$

and, for n even, using the same results, we get:

$$\sum_{j=1}^{n} \tilde{I}(\lambda_j) = 2\sum_{j=1}^{\frac{n}{2}} \tilde{I}(\lambda_j) - \tilde{I}(\pi) + \frac{n}{2\pi}(\bar{x} - Ex_1)^2 \ ,$$

$$S_n = 2\pi\left(2\sum_{j=1}^{m} \tilde{I}(\lambda_j) + 2\sum_{j=m+1}^{\frac{n}{2}} \tilde{I}(\lambda_j) - \tilde{I}(\pi) + \frac{n}{2\pi}(\bar{x} - Ex_1)^2\right) ,$$

$$\frac{S_n}{2n} = \hat{F}(\lambda_m) + \frac{2\pi}{n}\sum_{j=m+1}^{\frac{n}{2}} \tilde{I}(\lambda_j) - \frac{\pi}{n}\tilde{I}(\pi) + \frac{(\bar{x} - Ex_1)^2}{2} \ . \tag{A3.2.3}$$

Now, consider:

$$\frac{2\pi}{n}\sum_{j=m+1}^{\frac{n}{2}} \{\tilde{I}(\lambda_j) - E\tilde{I}(\lambda_j)\} \ ,$$

using that $X - EX = O_p(\sqrt{V(X)})$, and also that in lemma 10 in Robinson (1994b) it is proven

that:

$$V\left(\frac{2\pi}{n}\sum_{j=m+1}^{n/2}\tilde{I}(\lambda_j)\right)=o(n^{4H-4}) \ , \tag{A3.2.4}$$

and also that:

$$\frac{\pi}{n}\{\tilde{I}(\pi)-E\tilde{I}(\pi)\}=o_p(n^{2H-2}) \ , \tag{A3.2.5}$$

then, using (A3.2.4):

$$\frac{2\pi}{n}\sum_{j=m+1}^{n/2}\{\tilde{I}(\lambda_j)-E\tilde{I}(\lambda_j)\}=o_p(n^{2H-2}) \ ,$$

so, using this and (A3.2.5), we can rewrite (A3.2.3):

$$\frac{S_n}{2n}=\hat{F}(\lambda_m)+\frac{2\pi}{n}\sum_{j=m+1}^{\frac{n}{2}}E\tilde{I}(\lambda_j)-\frac{\pi}{n}\tilde{I}(\pi)+\frac{(\bar{x}-Ex_1)^2}{2}+o_p(n^{2H-2}) \ . \tag{A3.2.6}$$

Now, we can write, using (A3.2.2), for n odd:

$$\hat{F}(\lambda_m)=\frac{S_n}{2n}-\frac{2\pi}{n}\sum_{j=m+1}^{\frac{n-1}{2}}E\tilde{I}(\lambda_j)-\frac{(\bar{x}-Ex_1)^2}{2}+o_p(n^{2H-2}) \ ,$$

$$E\hat{F}(\lambda_m)=\frac{ES_n}{2n}-\frac{2\pi}{n}\sum_{j=m+1}^{\frac{n-1}{2}}E\tilde{I}(\lambda_j)-\frac{V(\bar{x})}{2} \ ,$$

so

$$\hat{F}(\lambda_m)-E\hat{F}(\lambda_m)=\frac{S_n-ES_n}{2n}-\frac{(\bar{x}-Ex_1)^2-V(\bar{x})}{2}+o_p(n^{2H-2}) \ ;$$

and using (A3.2.3) and (A3.2.6), for n even:

$$\hat{F}(\lambda_m)=\frac{S_n}{2n}-\frac{2\pi}{n}\sum_{j=m+1}^{\frac{n}{2}}E\tilde{I}(\lambda_j)-\frac{(\bar{x}-Ex_1)^2}{2}+\frac{\pi}{n}\tilde{I}(\pi)+o_p(n^{2H-2}) \ ,$$

$$E\hat{F}(\lambda_m)=\frac{ES_n}{2n}-\frac{2\pi}{n}\sum_{j=m+1}^{\frac{n}{2}}E\tilde{I}(\lambda_j)-\frac{V(\bar{x})}{2}+\frac{\pi}{n}E\tilde{I}(\pi) \ ,$$

so:

$$\hat{F}(\lambda_m)-E\hat{F}(\lambda_m)=\frac{S_n-ES_n}{2n}-\frac{(\bar{x}-Ex_1)^2-V(\bar{x})}{2}+o_p(n^{2H-2}) \ .$$

Then

$$\frac{\hat{F}(\lambda_m)-E\hat{F}(\lambda_m)}{G(\lambda_m)}=\frac{S_n-ES_n-n(\bar{x}-Ex_1)^2+nV(\bar{x})}{2nG(\lambda_m)}+o_p(m^{2H-2}) \ ,$$

so

$$R(\lambda_m)=(AA)+(BB)=\frac{S_n-ES_n-n(\bar{x}-Ex_1)^2}{2nG(\lambda_m)}+o_p(m^{2H-2}) \ ,$$

i.e.:

$$m^{2-2H}R(\lambda_m)=\frac{(1-H)[S_n-ES_n-n(\bar{x}-Ex_1)^2]}{n^{2H-1}C(2\pi)^{2-2H}}+o_p(1)=\frac{(1-H)DT_n}{C(2\pi)^{2-2H}}+o_p(1) \ ,$$

i.e., we have proven (A3.2.1) and the proof is complete.

Chapter 4

Consistency of the Cross-averaged Periodogram

In this chapter we establish a multivariate framework to analyze long-memory series. We provide some notation and then examine general conditions under which we can achieve consistency of the averaged cross-periodogram. These conditions are more general than those under which we will prove consistency for a quasi-maximum likelihood estimate in chapter 5. We restrict ourselves to prove consistency of our estimates and do not attempt to examine the limiting distribution because as we have seen in the previous chapter it has a discontinuity around H=3/4 in the univariate case. On the other hand, consistency of the averaged cross-periodogram will be relevant when analyzing in chapter 6 the asymptotic behaviour of a Lagrange multiplier test for I(0).

4.1 Notation

We consider r covariance stationary series that are observed in n moments of time. We use the index t to denote time (t=1,...n) and a to denote the series (a=1,...r), x_t is a rx1 vector with a-th element x_t^a. We define the normalized discrete Fourier Transform(DFT) of series a to be:

$$w_a(\lambda)=\frac{1}{\sqrt{2\pi n}}\sum_{t=1}^{n} x_t^a e^{it\lambda} \;, \quad a=1,...r, \quad \lambda\in(-\pi,\pi],$$

and for the vector x_t:

$$w(\lambda)=\begin{bmatrix} w_1(\lambda) \\ \cdot \\ \cdot \\ w_r(\lambda) \end{bmatrix},$$

and the cross-periodogram of series a and b:

$$I_{ab}(\lambda)=w_a(\lambda)\overline{w_b(\lambda)} \ ,$$

where the line over $w_b(\lambda)$ denotes complex conjugate, and the periodogram matrix (rxr) for the vector x_t is:

$$I(\lambda)=w(\lambda)w^*(\lambda), \tag{4.1.1}$$

where * denotes complex conjugation combined with transposition.

When representing our series x_t as a linear process we use the notation:

$$x_t=\begin{pmatrix} x_t^1 \\ \cdot \\ x_t^r \end{pmatrix}=Ex_0+\sum_{j=0}^{\infty} A_j e_{t-j} \ , \quad where\ A_j=\begin{pmatrix} \alpha_j^1 \\ \cdot \\ \alpha_j^r \end{pmatrix}, \quad e_t=\begin{pmatrix} e_t^1 \\ \cdot \\ e_t^r \end{pmatrix} \quad and \quad \alpha_j^a=(\alpha_j^{a_1},...\alpha_j^{a_r})$$

with e_t, a martingale difference sequence, $E(e_t \mid F_{t-1})=0$, $E|e_t|<\infty$ with $E(e_te_t' \mid F_{t-1})=R$ where F_{t-1} is the σ-field of events generated by e_s, $s\leq t$; and the typical element, series a, as:

$$x_t^a=\mu_a+\sum_{j=0}^{\infty} \alpha_j^a e_{t-j} \ ,$$

A_j is rxr, α_j^a is 1xr and e_t is rx1.

We define also the normalized DFT of e_t as:

$$v(\lambda)=\begin{pmatrix} v_1(\lambda) \\ \cdot \\ \cdot \\ v_r(\lambda) \end{pmatrix}=\frac{1}{\sqrt{2\pi n}}\sum_{t=1}^{n} e_t e^{it\lambda} \ ,$$

and the periodogram matrix for e_t:

$$J(\lambda)=v(\lambda)v(\lambda)^{*}=\begin{bmatrix} v_1(\lambda)\overline{v_1(\lambda)} & \dots & v_1(\lambda)\overline{v_r(\lambda)} \\ \dots & \dots & \dots \\ v_r(\lambda)\overline{v_1(\lambda)} & \dots & v_r(\lambda)\overline{v_r(\lambda)} \end{bmatrix},$$

where $v(\lambda)$ is rx1 and $J(\lambda)$ is rxr.

Let $A(\lambda)$ be the DFT of the weights A_j:

$$A(\lambda)=\sum_{j=0}^{\infty} A_j e^{ij\lambda}=\begin{pmatrix} \sum_{j=0}^{\infty} \alpha_j^1 e^{ij\lambda} \\ \cdot \\ \cdot \\ \sum_{j=0}^{\infty} \alpha_j^r e^{ij\lambda} \end{pmatrix}=\begin{pmatrix} A_1(\lambda) \\ \cdot \\ \cdot \\ A_r(\lambda) \end{pmatrix},$$

where

$$A_a(\lambda)=(A_{a_1}(\lambda),\dots,A_{a_r}(\lambda)),$$

$A_a(\lambda)$ is 1xr and $A(\lambda)$ is rxr. Consequently:

$$f(\lambda)=A(\lambda)f_e(\lambda)A(\lambda)^*.$$

Specifically the spectrum of e_t is[1]:

$$f_e(\lambda)=\frac{1}{2\pi}\begin{bmatrix}1 & r_{12} & .. & r_{1r}\\ ... & ... & ... & ...\\ r_{1r} & ... & ... & 1\end{bmatrix}=\frac{1}{2\pi}R,$$

and the spectrum of x_t:

$$f(\lambda)=\begin{bmatrix}f_{11}(\lambda) & .. & f_{1r}(\lambda)\\ ... & ... & ...\\ f_{r1}(\lambda) & .. & f_{rr}(\lambda)\end{bmatrix}=\begin{pmatrix}A_1(\lambda)\\ ..\\ A_r(\lambda)\end{pmatrix}\frac{R}{2\pi}(A_1^*(\lambda)\ ..\ A_r^*(\lambda))=$$

$$=\frac{1}{2\pi}\begin{bmatrix}A_1(\lambda)RA_1^*(\lambda) & .. & A_1(\lambda)RA_r^*(\lambda)\\ ... & ... & ...\\ A_r(\lambda)RA_1^*(\lambda) & .. & A_r(\lambda)RA_r^*(\lambda)\end{bmatrix},$$

and so the typical element, the cross spectral density function of series a and b is:

$$f_{ab}(\lambda)=A_a(\lambda)\frac{R}{2\pi}A_b^*(\lambda).$$

Let the cumulative cross spectral density function between series a and b be:

$$F_{ab}(\lambda)=\int_0^{\lambda}f_{ab}(\theta)d\theta\ ,$$

with

[1]Notice that for identification we set the variance of the innovations equal to 1 instead of the vector of coefficients A_0.

$$F_{ab}(\lambda)=ReF_{ab}(\lambda)+iImF_{ab}(\lambda)\ ,\ f_{ab}(\lambda)=Ref_{ab}(\lambda)+iImf_{ab}(\lambda)\ ,$$

so,

$$ReF_{ab}(\lambda)=\int_0^\lambda Ref_{ab}(\theta)d\theta\ ,\quad ImF_{ab}(\lambda)=\int_0^\lambda Imf_{ab}(\theta)d\theta.$$

The basic statistic we will consider is the averaged cross-periodogram:

$$\hat{F}_{ab}(\lambda)=\frac{2\pi}{n}\sum_{j=1}^{\left[\frac{n\lambda}{2\pi}\right]} I_{ab}(\lambda_j).$$

This statistic is just the (a,b) component of:

$$\hat{F}(\lambda)=\frac{2\pi}{n}\sum_{j=1}^{\left[\frac{n\lambda}{2\pi}\right]} I(\lambda_j)$$

where $I(\lambda_j)$ is the periodogram matrix. This is a fundamental statistic in multiple time series as an estimate of the spectral measure:

$$F(\lambda)=\int_0^\lambda f(\theta)d\theta\ .$$

The statistic has the advantage of being invariant to the mean of the process.

4.2 Consistency of the averaged periodogram

Before establishing the theorem, we are going to introduce general conditions under which consistency can be proved.

Condition C4.1: $f(\lambda) \sim \Lambda^0 G_0 \Lambda^0$ where $\Lambda^0=diag\{\lambda^{1/2-H_a}\}$ and G_0 is a Hermitian positive definite matrix, then for any two typical series, $a,b=1,\ldots r$, for H_a, $H_b \in (1/2,1)$,

$$f(\lambda)=\begin{bmatrix} f_{aa}(\lambda) & f_{ab}(\lambda) \\ f_{ba}(\lambda) & f_{bb}(\lambda) \end{bmatrix} \sim \begin{bmatrix} g_{aa}\lambda^{1-2H_a} & g_{ab}\lambda^{1-H_a-H_b} \\ \overline{g_{ab}}\lambda^{1-H_a-H_b} & g_{bb}\lambda^{1-2H_b} \end{bmatrix}, \quad as \quad \lambda \to 0^+ ,$$

with $0<g_{aa}<\infty, 0<|g_{ab}|<\infty$.

This is a <u>general</u> specification[2] that includes the behaviour of any fractional model as a particular case as we are going to show next.

For instance, the simplest bivariate fractional case:

$$(1-L)^{H_a-1/2}x_t^a=e_t^a \ \ and \ \ (1-L)^{H_b-1/2}x_t^b=e_t^b \ ,$$

where e_t^a, e_t^b are white noise with:

$$E((e_t^a)^2)=1, \ E((e_t^b)^2)=1, \ E(e_t^a e_t^b)=r_{ab} \ ,$$

then,

$$(1-e^{-i\lambda})^{H_a-1/2}(1-e^{i\lambda})^{H_b-1/2}f_{ab}(\lambda)=\frac{r_{ab}}{2\pi}.$$

Now 1-exp{-iλ} is a complex number with modulus 2sin(λ/2) and argument β=arctg(sinλ/1-cosλ). As $\lambda\to 0^+$, sinλ/1-cosλ→∞ , so $\beta\to\pi/2$. Also as $\lambda\to 0$, sin λ ~ λ. Then, we can represent 1-exp{-iλ} as λ exp{i $\pi/2$}. So,

---

[2] We could think that we could achieve more generality by allowing the cross-spectral density function to behave as:

$$f_{ab}(\lambda)\sim g_{ab}\lambda^{1-2H_{ab}} \ \ as \ \ \lambda\to 0^+, \ \ with \ \ H_{ab}<\frac{H_a+H_b}{2}$$

but consider in this case the coherency as $\lambda\to 0^+$:

$$coh(\lambda)\sim\frac{|g_{ab}|^2}{g_{aa}g_{bb}}\lambda^{H_b+H_a-2H_{ab}}\to 0 \ , \ \ as \ \ \lambda\to 0^+$$

and we can get this with our specification by letting $g_{ab}=0$.

$$(1-e^{-i\lambda})^{H_a-1/2}(1-e^{i\lambda})^{H_b-1/2} \sim \lambda^{H_a+H_b-1} e^{\frac{i\pi}{2}(H_a-H_b)} ,$$

then

$$f_{ab}(\lambda) \sim \frac{r_{ab}}{2\pi}\lambda^{1-(H_a+H_b)} e^{\frac{i\pi}{2}(H_b-H_a)} \quad for\ a,b=1,...r.$$

So this is a particular case of our specification with $g_{ab}=\frac{r_{ab}}{2\pi}\exp\{i\frac{\pi}{2}(H_b-H_a)\}$.

Now, for more general fractional models as:

* Model I:

$$A(L)diag\{(1-L)^{H_a-1/2}\}X_t=B(L)e_t$$

where A(L) has all its roots outside the unit circle; when λ approaches 0:

$$A(1)diag\{\lambda^{H_a-1/2}e^{i\frac{\pi}{2}(H_a-1/2)}\}f_x(\lambda)diag\{\lambda^{H_a-1/2}e^{-i\frac{\pi}{2}(H_a-1/2)}\}A(1)'=$$

$$B(1)f_\epsilon(\lambda)B(1)' ;$$

as $|A(1)|\neq 0$, we have:

$$f_{ab}(\lambda)\sim g_{ab}\lambda^{1-(H_a+H_b)}e^{i\frac{\pi}{2}(H_b-H_a)} , \quad as \quad \lambda\to 0^+ ,$$

where $\{g_{ab}\}$ are constants:

$$\{g_{ab}\}=A(1)^{-1}B(1)f_\epsilon(\lambda)B(1)'(A(1)')^{-1} ,$$

* Model II:

$$diag\{(1-L)^{H_a-1/2}\}A(L)X_t=B(L)e_t ,$$

so as λ approaches 0:

$$f_x(\lambda)=A(1)^{-1}\Big[diag\{\lambda^{1/2-H_a}e^{-i\frac{\pi}{2}(H_a-1/2)}\}B(1)f_\epsilon(\lambda)B(1)'\cdot$$
$$\cdot diag\{\lambda^{1/2-H_a}e^{i\frac{\pi}{2}(H_a-1/2)}\}\Big](A(1)')^{-1} ,$$

so, as $\lambda \to 0^+$,

$$f_{ab}(\lambda) \sim \sum_{s=1}^{r} \alpha_{as} \sum_{h=1}^{r} c_{sh} \lambda^{1-(H_s+H_h)} e^{i\frac{\pi}{2}(H_h-H_s)} \alpha_{bh} \,, \quad (**)$$

where

α_{ab} *is the* (a,b) *element of* $A(1)^{-1}$ $a,b=1,\ldots,r$; *and*

c_{sh} *is the* (s,h) *element of* $B(1)f_\epsilon(\lambda)B(1)'$ $s,h=1,\ldots r$,

so that, in general, expression (**) will be dominated by the term with largest $|1-(H_s+H_h)|$ and nonzero coefficient. So we have shown that condition C4.1 includes any fractional model as a particular case.

Condition C4.2. The minimum condition on the bandwidth m: as $n \to \infty$,

$$\frac{1}{m} + \frac{m}{n} \to 0.$$

Condition C4.3:

e_t and e_te_t'-R are uniformly integrable martingale difference sequences.

We note the implications of this condition:

(i) $E(e_te_u')=0$, $t \neq u$.

Proof:

$$\textit{for } t>u \quad E[e_te_u'] = E[E(e_t|F_{t-1})e_u'] = 0 \;.$$

(ii) $E(e_we_s' \otimes e_te_u') = R \otimes R$ for $w=s \neq t=u$;

$= 0$ for: $w>s,t,u$; $s>w,t,u$; $t>w,s,u$; $u>w,s,t$;

$w=s>t,u$; $w=t>s,u$; $w=u>s,t$;

$s=t>w,u$; $s=u>w,t$; $t=u>w,s$.

Proof:

for $w=s>t=u$, $E[e_w e_w^{/} \otimes e_t e_t^{/}] = E[E(e_w e_w^{/}/F_{w-1}) \otimes e_t e_t^{/}] = E[R \otimes e_t e_t^{/}] = R \otimes R$;

for $w>s$, $w>t$, $w>u$, $E[e_w e_s^{/} \otimes e_t e_u^{/}] = E[E(e_w/F_{w-1}) e_s \otimes e_t e_u^{/}] = 0$;

for $w=s>t,u$, $E(e_w e_w^{/} \otimes e_t e_u) = E[E(e_w e_w^{/}/F_{w-1}) \otimes e_t e_u) = E(R \otimes e_t e_u^{/}) = R \otimes E(e_t e_u^{/}) = 0$;

for $w=u>s,t$, $E[e_w e_s^{/} \otimes e_t e_w^{/}]$ is a $r^2 x r^2$ matrix with typical element:

$$E(e_w^l e_w^m e_s^n e_t^o) = E[E(e_w^l e_w^m / F_{w-1}) e_s^n e_t^o) = \sigma_{lm} E(e_s^n e_t^o) = 0.$$

(iii) by the weak law of large numbers for uniformly integrable martingale difference sequences[3]:

$$\frac{1}{n}\sum_{t=1}^{n} e_t e_t^{/} \to_p R.$$

Theorem 4.1

Under conditions C4.1, C4.2 and C4.3:

$$\hat{F}_{ab}(\lambda_m) - F_{ab}(\lambda_m) = o_p(F_{aa}(\lambda_m)^{1/2} F_{bb}(\lambda_m)^{1/2})$$

Proof: see appendix 4.1.

Estimates for the parameters of the cross-spectral density function, H_a, H_b and g_{ab} based on the averaged periodogram could be straightforwardly proposed and their consistency proved but we will not pursue further with these estimates. In chapter 5 we will analyze a procedure that will provide us with more straightforward tools for statistical inference. Theorem 4.1 is relevant because we will employ it in chapter 6 to study the asymptotic behaviour of a Lagrange multiplier test for I(0).

---

[3] See Heyde and Seneta (1972),theorem 1.

Appendix 4.1

Before proving the theorem we need to prove 4 propositions:

Proposition 1a:

Under conditions C4.1 and C4.2; as $n\to\infty$:

$$\frac{2\pi}{n}\sum_{j=1}^{m} Ref_{ab}(\lambda_j)-ReF_{ab}(\lambda_m)=o(|F_{ab}(\lambda_m)|) .$$

Proof: As $n\to\infty$,

$$LHS=\sum_{j=1}^{m}\int_{\lambda_{j-1}}^{\lambda_j}(Ref_{ab}(\lambda_j)-Ref_{ab}(\lambda))d\lambda=$$

$$\sum_{j=1}^{m}\int_{\lambda_{j-1}}^{\lambda_j}Re(g_{ab})[\lambda_j^{(1-H_a-H_b)}-\lambda^{(1-H_a-H_b)}]d\lambda+r= \qquad (A4.1.1)$$

$$\sum_{j=1}^{m}\int_{\lambda_{j-1}}^{\lambda_j}Re(g_{ab})[\lambda_j^{(-H_a-H_b)}][\lambda_j-\left(\frac{\lambda}{\lambda_j}\right)^{(-H_a-H_b)}\lambda]d\lambda+r=$$

$$Re(g_{ab})\sum_{j=1}^{m}[\lambda_j^{(-H_a-H_b)}]\int_{\lambda_{j-1}}^{\lambda_j}[\lambda_j-\left(\frac{\lambda}{\lambda_j}\right)^{(-H_a-H_b)}\lambda]d\lambda+r , \qquad (+)$$

where

$$r=o(\frac{1}{n}\sum_{j=1}^{m}\lambda_j^{1-H_a-H_b}+\lambda_m^{2-H_a-H_b})=o(\lambda_m^{2-H_a-H_b})=o(|F_{ab}(\lambda_m)|) .$$

Now, under condition C4.2, for n sufficiently large the leading term in (+) has order of magnitude:

$$O\left(\sum_{j=1}^{m}\lambda_j^{(-H_a-H_b)}\int_{\lambda_{j-1}}^{\lambda_j}\left|(\lambda_j-\left(\frac{\lambda}{\lambda_j}\right)^{(-H_a-H_b)}\lambda)\right|d\lambda\right)\ , \qquad (*)$$

now, because:

$$|\lambda_j-\left(\frac{\lambda}{\lambda_j}\right)^{(-H_a-H_b)}\lambda|<|\lambda_j-\lambda|\ ,\ \ for\ \ \lambda\in(\lambda_{j-1},\lambda_j)\ ,$$

then, (*) is

$$O(|\sum_{j=1}^{m}\lambda_j^{(-H_a-H_b)}\int_{\lambda_{j-1}}^{\lambda_j}(\lambda_j-\lambda)d\lambda|)\ , \qquad (**)$$

as

$$\int_{\lambda_{j-1}}^{\lambda_j}(\lambda_j-\lambda)d\lambda=\frac{\pi^2}{2n^2}\ ,$$

$$(**)=O(\frac{1}{n^2}\sum_{j=1}^{m}\lambda_j^{(-H_a-H_b)})=O(n^{(H_a+H_b-2)}\sum_{j=1}^{m}j^{(-H_a-H_b)})=$$

$$O(m^{H_a+H_b-2}\lambda_m^{2-H_a-H_b})=o(|F_{ab}(\lambda_m)|)\ . \qquad \text{(A4.1.2)}$$

<u>Proposition 1b</u>: Under C4.1, C4.2; as $n\to\infty$:

$$\frac{2\pi}{n}\sum_{j=1}^{m}Imf_{ab}(\lambda_j)-ImF_{ab}(\lambda_m)=o(|F_{ab}(\lambda_m)|)\ .$$

<u>Proof</u>: as the LHS is equal to

$$\sum_{j=1}^{m}\int_{\lambda_{j-1}}^{\lambda_j}(Imf_{ab}(\lambda_j)-Imf_{ab}(\lambda))d\lambda=\sum_{j=1}^{m}\int_{\lambda_{j-1}}^{\lambda_j}Im(g_{ab})[\lambda_j^{(1-H_a-H_b)}-\lambda^{(1-H_a-H_b)}]d\lambda+r$$

and the proof follows the one above.

Proposition 2:

Under conditions C4.1 and C4.2, as $n \to \infty$:

$$E\{\frac{2\pi}{n}\sum_{j=1}^{m} I_{aa}(\lambda_j)\} = O(F_{aa}(\lambda_m)) \text{ for } a=1,...r.$$

Proof: Our conditions C4.1 and C4.2 are more restrictive on x_t^a, for every $a=1...r$, than those under which Robinson (1994a), proposition 2, establishes that:

$$E\{\frac{2\pi}{n}\sum_{j=1}^{m} I(\lambda_j)\} = O(F(\lambda_m)) \ , \ as\ n \to \infty \ .$$

Proposition 3:

Under conditions C4.1, C4.2 and C4.3:

$$Re[\frac{2\pi}{n}\sum_{j=1}^{m} \{A_a(\lambda_j)J(\lambda_j)A_b^*(\lambda_j) - f_{ab}(\lambda_j)\}] =$$

$$o_p(F_{aa}(\lambda_m)^{0.5} F_{bb}(\lambda_m)^{0.5}) \ ,$$

and

$$Im[\frac{2\pi}{n}\sum_{j=1}^{m} \{A_a(\lambda_j)J(\lambda_j)A_b^*(\lambda_j) - f_{ab}(\lambda_j)\}] =$$

$$o_p(F_{aa}(\lambda_m)^{0.5} F_{bb}(\lambda_m)^{0.5}) \ .$$

Proof (we drop the arguments λ_j for simplicity):

$$\frac{2\pi}{n}\sum_{j=1}^{m} \{A_a(\lambda_j)J(\lambda_j)A_b^*(\lambda_j) - f_{ab}(\lambda_j)\} =$$

$$\frac{2\pi}{n}\sum_{j}\{A_a J A_b^* - A_a \frac{R}{2\pi} A_b^*\} = (*)$$

as

$$J=\frac{1}{2\pi n}\left(\begin{bmatrix}\sum_{t=1}^{n} e_t^{1^2} & .. & \sum_t e_t^1 e_t^r \\ \dots & .. & \dots \\ \dots & .. & \sum_t e_t^{r^2}\end{bmatrix}+\begin{bmatrix}\sum_{s\neq}\sum_t e_t^1 e_s^1 & .. & \sum_{s\neq}\sum_t e_t^1 e_s^r \\ \dots & .. & \dots \\ \dots & .. & \sum_{s\neq}\sum_t e_t^r e_s^r\end{bmatrix} e^{i(s-t)\lambda_j}\right),$$

so

$$(*)=\frac{2\pi}{n}\sum_j\left(\frac{1}{2\pi}A_a\begin{bmatrix}\frac{1}{n}\sum_t e_t^{1^2}-1 & .. & \frac{1}{n}\sum_t e_t^1 e_t^r - r_{1r} \\ \dots & .. & \dots \\ .. & .. & \sum_t e_t^{r^2}-1\end{bmatrix}A_b^*\right)+ \qquad (1)$$

$$+\frac{2\pi}{n}\sum_j\left(\frac{1}{2\pi n}A_a\begin{bmatrix}\sum_{s\neq}\sum_t e_t^1 e_s^1 & .. & \sum_{s\neq}\sum_t e_t^1 e_s^r \\ .. & .. & .. \\ \dots & .. & \sum_{s\neq}\sum_t e_t^r e_s^r\end{bmatrix} e^{i(s-t)\lambda_j} A_b^*\right), \qquad (2)$$

now, call

$$D=\begin{bmatrix}\frac{1}{n}\sum_t e_t^{1^2}-1 & .. & \frac{1}{n}\sum_t e_t^1 e_t^r - r_{1r} \\ \dots & .. & \dots \\ .. & .. & \sum_t e_t^{r^2}-1\end{bmatrix},$$

then

$$|(1)|=\left|\frac{1}{n}\sum_{j=1}^{m}A_a(\lambda_j)DA_b^*(\lambda_j)\right|\leq\frac{1}{n}\sum_{j=1}^{m}\|A_a(\lambda_j)\|\|D\|\|A_b^*(\lambda_j)\|\leq$$

$$\leq\|D\|\frac{1}{n}\left(\sum_{j=1}^{m}\|A_a(\lambda_j)\|^2\|A_b^*(\lambda_j)\|^2\right)^{\frac{1}{2}}=(*) \qquad \text{(A4.3.1)}$$

as

$$f_{aa}(\lambda)=\sum_{k=1}^{r}|A_a^k(\lambda)|^2+\sum_{k\neq}^{r}\sum_{l}^{r}A_a^k(\lambda)\overline{A_a^l(\lambda)}r_{kl}=$$

$$C\|A_a(\lambda)\|^2\ +o(\|A_a(\lambda)\|^2),$$

$$\textit{where}\ \|A_a(\lambda)\|^2=\max_i|A_a^i(\lambda)|^2\ \textit{and}\ C>0,$$

$$\|A_a(\lambda)\|=O(f_{aa}(\lambda)^{1/2})\ ,$$

and similarly:

$$\|A_b(\lambda)\|=O(f_{bb}(\lambda)^{1/2})\ ,$$

so,

$$(*)=O_p\left(\|D\|\left(\frac{1}{n}\sum_{j=1}^{m}f_{aa}(\lambda_j)\right)^{\frac{1}{2}}\left(\frac{1}{n}\sum_{j=1}^{m}f_{bb}(\lambda_j)\right)^{\frac{1}{2}}\right)=$$

$$o_p\left(F_{aa}(\lambda_m)^{0.5}F_{bb}(\lambda_m)^{0.5}\right)\ ,$$

because by proposition 1:

$$\frac{1}{n}\sum_{j=1}^{m} f_{aa}(\lambda_j)=O(F_{aa}(\lambda_m)) \ and \ \frac{1}{n}\sum_{j=1}^{m} f_{bb}(\lambda_j)=O(F_{bb}(\lambda_m)) \ ,$$

and by condition C4.3:

$$\|D\|=o_p(1).$$

Now,

$$(2)=\frac{1}{n^2}\sum_{j=1}^{m} A_a(\lambda_j)\left(\sum_{s\neq}\sum_{t} e_t e_s^{\prime}\right)A_b^{*}(\lambda_j)e^{i(s-t)\lambda_j}=$$

$$\frac{1}{n}\sum_{s\neq}\sum_{t} e_t^{\prime}\Gamma_{s-t,m}e_s \ ,$$

where

$$\Gamma_{s-t,m}=\frac{1}{n}\sum_{j=1}^{m} A_a^{\prime}(\lambda_j)\overline{A_b(\lambda_j)}e^{i(s-t)\lambda_j} \ ,$$

$$E|(2)|^2=E(2)(2)^*=E\left[\frac{1}{n}\sum_{s\neq}\sum_{t} e_t^{\prime}\Gamma_{s-t,m}e_s\frac{1}{n}\sum_{u\neq}\sum_{v} e_v^{\prime}\overline{\Gamma_{u-v,m}}e_u\right]$$

$$=E\left[\left(\frac{1}{n}\sum_{s\neq}\sum_{t} (e_s^{\prime}\otimes e_t^{\prime})vec\Gamma_{s-t,m}\right)\left(\frac{1}{n}\sum_{u\neq}\sum_{v} (e_u^{\prime}\otimes e_v^{\prime})vec\overline{\Gamma_{u-v,m}}\right)\right]$$

$$=\frac{1}{n^2}\sum_{s\neq}\sum_{t}\sum_{u\neq}\sum_{v} vec^{\prime}\Gamma_{s-t,m}E(e_s\otimes e_t)(e_u^{\prime}\otimes e_v^{\prime})vec\overline{\Gamma_{u-v,m}} \ , \quad (*)$$

now, by condition C4.3 (ii),

$$E(e_s\otimes e_t)(e_u^{\prime}\otimes e_v^{\prime})=E(e_s e_u^{\prime}\otimes e_t e_v^{\prime})=(R\otimes R)I(s=u,t=v) \ ,$$

where I( ) is the indicator function, i.e., I(A)=1 if A is true and I(A)=0 if A is false,

$$(*)=\frac{1}{n^2}\sum_{s*}\sum_{t} vec'\Gamma_{s-t,m}(R\otimes R)vec\overline{\Gamma_{s-t,m}}$$

$$=\frac{1}{n}\sum_{t=1}^{n-1}(1-\frac{t}{n})vec'(\Gamma_{t,m}+\Gamma_{-t,m})(R\otimes R)vec(\overline{\Gamma_{t,m}}+\overline{\Gamma_{-t,m}})$$

$$=\frac{1}{n}\sum_{t=1}^{n-1}(1-\frac{t}{n})vec'(\frac{2}{n}\sum_{j=1}^{m}A_a'(\lambda_j)\overline{A_b(\lambda_j}\cos t\lambda_j))(R\otimes R)\cdot$$
$$\cdot\ vec(\frac{2}{n}\sum_{k=1}^{m}A_a^*(\lambda_k)A_b(\lambda_k)\cos t\lambda_k)$$

$$=\frac{1}{n}\sum_{t=1}^{n-1}(1-\frac{t}{n})(\frac{2}{n}\sum_{j=1}^{m}vec'(A_a'(\lambda_j)\overline{A_b(\lambda_j}\cos t\lambda_j)))(R\otimes R)\cdot$$
$$\cdot\ (\frac{2}{n}\sum_{k=1}^{m}vec(A_a^*(\lambda_k)A_b(\lambda_k)\cos t\lambda_k))\ ,$$

$$=\frac{1}{n}\sum_{t=1}^{n-1}(1-\frac{t}{n})(\frac{2}{n})^2\sum_{j=1}^{m}[A_b^*(\lambda_j)\cos t\lambda_j\otimes A_a'(\lambda_j)]'(R\otimes R)\cdot$$
$$\cdot\ \sum_{k=1}^{m}[(A_b'(\lambda_k)\otimes A_a^*(\lambda_k)\cos t\lambda_k)])=(*)$$

call

$$B_a(\lambda)=A_a(\lambda)R^{\frac{1}{2}}\ and\ B_b(\lambda)=A_b(\lambda)R^{\frac{1}{2}}\ ,$$

then

$$(*)=\frac{1}{n}\sum_{t=1}^{n-1}(1-\frac{t}{n})(\frac{2}{n})^2\left(\sum_{j=1}^{m}B_b^*(\lambda_j)\otimes B_a'(\lambda_j)cost\lambda_j\right)'\cdot$$
$$\cdot\left(\sum_{k=1}^{m}B_b'(\lambda_k)\otimes B_a^*(\lambda_k)\cos t\lambda_k\right)$$

$$\leq\frac{C}{n}\sum_{t=1}^{n-1}\|\frac{1}{n}\sum_{j=1}^{m}B_b^*(\lambda_j)\otimes B_a'(\lambda_j)cost\lambda_j\|^2=(*)\ ,$$

call the r^2x1 vector

$$c_j=B_b^*(\lambda_j)\otimes B_a'(\lambda_j)\ ,$$

from condition C4.1

$$f_{aa}(\lambda)=\frac{1}{2\pi}B_a(\lambda)B_a^*(\lambda)\sim g_{aa}\lambda^{1-2H_a}\ and\ f_{bb}(\lambda)=\frac{1}{2\pi}B_b(\lambda)B_b^*(\lambda)\sim g_{bb}\lambda^{1-2H_b}$$

then, for at least one l: $|B_{al}(\lambda)|^2\sim C_{al}\lambda^{1-2H_a}$, and for the rest, k=1,...r, k≠l: $|B_{ak}(\lambda)|^2=o(\lambda^{1-2H_a})$, also, for at least one q: $|B_{bq}(\lambda)|^2\sim C_{bq}\lambda^{1-2H_b}$, and for the rest, p=1,...r, p≠q: $|B_{bp}(\lambda)|^2=o(\lambda^{1-2H_b})$, so, for at least one v: $c_{jv}=O(\lambda_j^{1-H_a-H_b})$, and for the rest, h=1,...r,h≠v: $c_{jh}=o(\lambda_j^{1-H_a-H_b})$, so,

$$(*)=\frac{C}{n}\sum_{t=1}^{n-1}\|\frac{1}{n}\sum_{j=1}^{m}c_jcost\lambda_j\|^2=\frac{C}{n}\sum_{t=1}^{n-1}\sum_{h=1}^{k^2}\left(\frac{1}{n}\sum_{j=1}^{m}c_{jh}cost\lambda_j\right)^2\leq$$
$$\frac{C}{n}\sum_{t=1}^{n-1}\left[\left(\frac{1}{n}\sum_{j=1}^{m}\lambda_j^{1-H_a-H_b}cost\lambda_j\right)^2+a_t^2\right] \qquad \text{(A4.3.2)}$$

where

$$a_t = o\left(\frac{1}{n}\sum_{j=1}^{m}\lambda_j^{1-H_a-H_b}\right) = o(F_{aa}(\lambda_m)^{0.5}F_{bb}(\lambda_m)^{0.5}) \ ,$$

now

$$\left|\frac{1}{n}\sum_{j=1}^{m}\lambda_j^{1-H_a-H_b}cost\lambda_j\right| \leq \frac{1}{n}\sum_{j=1}^{m}\lambda_j^{1-H_a-H_b} = O(F_{aa}(\lambda_m)^{0.5}F_{bb}(\lambda_m)^{0.5}) \ ,$$

so (*) is

$$O\left(\frac{r}{n}F_{aa}(\lambda_m)F_{bb}(\lambda_m) + \max_{r<t<n}\left|\frac{1}{n}\sum_{j=1}^{m}\lambda_j^{1-H_a-H_b}cost\lambda_j\right|^2\right) =$$

$$= o(F_{aa}(\lambda_m)F_{bb}(\lambda_m))$$

then by lemma 7 (Robinson 1994a):

$$\max_{r<t<n}\left|\frac{1}{n}\sum_{j=1}^{m}\lambda_j^{1-H_a-H_b}\cos t\lambda_j\right| = O(r^{H_a+H_b-2}) \ ,$$

then choose

$$r = nm^{\frac{H_a+H_b-2}{3-H_a-H_b}} \ ,$$

then

$$\max_{r<t<n}\left|\frac{1}{n}\sum_{j=1}^{m}\lambda_j^{1-H_a-H_b}\cos t\lambda_j\right|^2 = O(r^{2H_a+2H_b-4}) =$$

$$O\left(\left(\frac{m}{n}\right)^{4-2H_a-2H_b} m^{\frac{2(H_a+H_b-2)}{(3-H_a-H_b)}}\right) = O(F_{aa}(\lambda_m)F_{bb}(\lambda_m)m^{-\delta}) = \qquad \text{(A4.3.3)}$$

$$o(F_{aa}(\lambda_m)F_{bb}(\lambda_m)) \ ,$$

so

$$E|(2)|^2 = o(F_{aa}(\lambda_m)F_{bb}(\lambda_m)) \ ,$$

$$E|(2)| = o(F_{aa}(\lambda_m)^{0.5}F_{bb}(\lambda_m)^{0.5}) \ ,$$

therefore

$$|(1)|=o(F_{aa}(\lambda_m)^{0.5}F_{bb}(\lambda_m)^{0.5}) \text{ and } |(2)|=o(F_{aa}(\lambda_m)^{0.5}F_{bb}(\lambda_m)^{0.5}) ,$$

so that

$$Re(1)=o_p(F_{aa}(\lambda_m)^{0.5}F_{bb}(\lambda_m)^{0.5}) , Im(1)=o_p(F_{aa}(\lambda_m)^{0.5}F_{bb}(\lambda_m)^{0.5}),$$

$$Re(2)=o_p(F_{aa}(\lambda_m)^{0.5}F_{bb}(\lambda_m)^{0.5})$$

and

$$Im(2)=o_p(F_{aa}(\lambda_m)^{0.5}F_{bb}(\lambda_m)^{0.5}) .$$

Proposition 4:

Under C4.1, C4.2 and C4.3, as $n \to \infty$:

$$\frac{1}{n}\sum_{j=1}^{m} |w_a(\lambda_j)-A_a(\lambda_j)v(\lambda_j)|^2=o_p(F_{aa}(\lambda_m)) . \qquad (*)$$

Proof: As LHS(*) is nonnegative, we have to show its expectation has order of RHS(*),

$$E(LHS)=E\left[\frac{1}{n}\sum_{j=1}^{m}\frac{1}{2\pi n}\left(\sum_{t=1}^{n}x_{at}e^{it\lambda_j}-\sum_{h=0}^{\infty}\alpha_h^a e^{ih\lambda_j}\sum_{t=1}^{n}e_t e^{it\lambda_j}\right)\right.$$
$$\left.\cdot\left(\sum_{s=1}^{n}x_{as}e^{-is\lambda_j}-\sum_{h=0}^{\infty}\alpha_h^a e^{-ih\lambda_j}\sum_{s=1}^{n}e_s e^{-is\lambda_j}\right)\right]$$
$$=\frac{1}{2\pi n^2}\sum_{j=1}^{m}\left[\sum_{t=1}^{n}\sum_{s=1}^{n}E(x_{at}-\mu_a)(x_{as}-\mu_a)e^{it\lambda_j}e^{-is\lambda_j}-\right.$$
$$-\sum_{t=1}^{n}\sum_{h=0}^{\infty}\sum_{s=1}^{n}E\left[x_{at}e^{it\lambda_j}\alpha_h^a e^{-ih\lambda_j}e_s e^{-is\lambda_j}\right]-$$
$$-\sum_{t=1}^{n}\sum_{h=0}^{\infty}\sum_{s=1}^{n}E\left[\alpha_h^a e^{ih\lambda_j}e_t e^{it\lambda_j}x_{as}e^{-is\lambda_j}\right]+$$
$$\left.+\sum_{t=1}^{n}\sum_{h=0}^{\infty}\sum_{h'=0}^{\infty}\sum_{s=1}^{n}E\left[\alpha_h^a e^{ih\lambda_j}e_t e^{it\lambda_j}\alpha_h^a e^{-ih'\lambda_j}e_s e^{-is\lambda_j}\right]\right]$$

$$=\frac{1}{2\pi n^2}\sum_{j=1}^{m}\sum_{t=1}^{n}\sum_{s=1}^{n} e^{i(t-s)\lambda_j}\Big[cov(x_{at},x_{as})-E[x_{at}\overline{A_a}e_s]$$

$$-E[A_a e_t x_{as}]+E[A_a e_t e_s' A_a^*]\Big]$$

now,

$$E[x_{at}\overline{A_a}e_s]=E\left[\left(\mu_a+\sum_{f=0}^{\infty}\alpha_f^a e_{t-f}\right)\left(\sum_{k=0}^{\infty}\alpha_k^a e^{-ik\lambda_j}\right)e_s\right]$$

$$=\sum_{f=0}^{\infty}\sum_{k=0}^{\infty}\alpha_f^a RI(s=t-f)\alpha_k^{a'} e^{-ik\lambda_j}=\sum_{k=0}^{\infty}\alpha_{t-s}^a R\alpha_k^{a'} e^{-ik\lambda_j}=\alpha_{t-s}^a RA_a^*$$

because $f_{aa}(\lambda), A_a^*(\lambda)$ and $e^{ij\lambda}$ have period 2π, and

$$cov(x_{at},x_{as})=\int_{-\pi}^{\pi} f_{aa}(\lambda)e^{i(s-t)\lambda}d\lambda=$$

$$\int_{-\pi}^{\pi} A_a(\lambda)f_e(\lambda)A_a^*(\lambda)e^{i(s-t)\lambda}d\lambda\ ,$$

and

$$\alpha_{t-s}^a=\frac{1}{2\pi}\int_{-\pi}^{\pi} A_a(u)e^{i(s-t)u}du\ ,$$

so, for any λ_j:

$$cov(x_{at},x_{as})=\frac{1}{2\pi}\int_{-\pi}^{\pi} A_a(u+\lambda_j)RA_a^*(u+\lambda_j)e^{i(s-t)(u+\lambda_j)}du,$$

$$E[x_{at}\overline{A_a}e_s]=\alpha_{t-s}^a RA_a^*=\frac{1}{2\pi}\int_{-\pi}^{\pi} A_a(u+\lambda_j)RA_a^*(\lambda_j)e^{i(s-t)(u+\lambda_j)}du,$$

$$E[A_a e_t x_{as}]=\frac{1}{2\pi}\int_{-\pi}^{\pi} A_a(\lambda_j)RA_a^*(u+\lambda_j)e^{i(s-t)(u+\lambda_j)}du,$$

$$E[A_a e_t e_s A_a^*]=\frac{1}{2\pi}\int_{-\pi}^{\pi} A_a(\lambda_j)RA_a^*(\lambda_j)e^{i(s-t)(u+\lambda_j)}du;$$

where we have used $E[A_a e_t e_s A_a^*] = A_a(\lambda_j) R A_a^*(\lambda_j) I(s=t)$. Then E[LHS]=

$$=\frac{1}{(2\pi)^2 n^2}\sum_{j=1}^{m}\sum_{s=1}^{n}\sum_{t=1}^{n} e^{i(t-s)\lambda_j} \cdot$$

$$\cdot\left[\int_{-\pi}^{\pi}\left([A_a(u+\lambda_j)-A_a(\lambda_j)]R[A_a(u+\lambda_j)-A_a(\lambda_j)]^*\right)e^{i(s-t)(u+\lambda_j)}du\right]$$

$$=\frac{1}{(2\pi)^2 n}\sum_{j=1}^{m}\int_{-\pi}^{\pi}\frac{1}{n}\sum_{s=1}^{n}\sum_{t=1}^{n} e^{i(t-s)u}\left([A_a(u+\lambda_j)-A_a(\lambda_j)]R[A_a(u+\lambda_j)-A_a(\lambda_j)]^*\right)du$$

$$=\frac{1}{(2\pi)^2 n}\sum_{j=1}^{m}\int_{-\pi}^{\pi}K(u)\left([A_a(u+\lambda_j)-A_a(\lambda_j)]R[A_a(u+\lambda_j)-A_a(\lambda_j)]^*\right)du$$

$$=S_m[(-\pi,\pi)] \ ,$$

where

$$S_r(A)=\frac{1}{(2\pi)^2 n}\sum_{j=1}^{r}\int_A\left([A_a(u+\lambda_j)-A_a(\lambda_j)]R[A_a(u+\lambda_j)-A_a(\lambda_j)]^*\right)du \ ,$$

and

$$K(u)=\frac{1}{n}\left|\sum_{t=1}^{n} e^{itu}\right|^2 .$$

Now

$$[A_a(u+\lambda_j)-A_a(\lambda_j)]R[A_a(u+\lambda_j)-A_a(\lambda_j)]^* \le$$

$$\le K_1 A_a(u+\lambda_j) R A_a(u+\lambda_j)^* + K_2 A_a(\lambda_j) R A_a(\lambda_j)^*$$

$$=2\pi K_1 f_{aa}(u+\lambda_j)+2\pi K_2 f_{aa}(\lambda_j) \ ,$$

because

$$for\ R\ge 0 \quad (a-b)R(a-b)^* \le K_1 aRa^* + K_2 bRb^* \ .$$

So,

$$S_r(A)\le\frac{C}{n}\sum_{j=1}^{r} f_{aa}(\lambda_j)\int_A K(u)du+\frac{C}{n}\sum_{j=1}^{r}\int_A K(u) f_{aa}(u+\lambda_j)du \ ,$$

and the rest of the proof follows from proposition 4 in Robinson (1994a), who establishes under conditions weaker than ours that:

$$S_m([-\pi,\pi])=o_p(F(\lambda_m)) \ . \tag{A4.4.1}$$

Theorem 4.1 a

Under C4.1, C4.2 and C4.3:

$$Re\hat{F}_{ab}(\lambda_m)-ReF_{ab}(\lambda_m) = o_p(F_{aa}(\lambda_m)^{1/2}F_{bb}(\lambda_m)^{1/2}).$$

Proof:

$$\hat{F}_{ab}(\lambda_m)-F_{ab}(\lambda_m)=$$

$$=\frac{2\pi}{n}\sum_{j=1}^{m}\left(I_{ab}(\lambda_j)-A_a(\lambda_j)J(\lambda_j)A_b^*(\lambda_j)\right) \quad (1)$$

$$+\frac{2\pi}{n}\sum_{j=1}^{m}\left(A_a(\lambda_j)J(\lambda_j)A_b^*(\lambda_j)-f_{ab}(\lambda_j)\right) \quad (2)$$

$$+\frac{2\pi}{n}\sum_{j=1}^{m}f_{ab}(\lambda_j)-F_{ab}(\lambda_m) \ . \quad (3)$$

First, proposition 3 implies: $Re(2)=o_p(F_{aa}(\lambda_m)^{1/2}\ F_{bb}(\lambda_m)^{1/2})$, while Proposition 1a implies: $Re(3)=o(|F_{ab}(\lambda_m)|)$. Also let for convenience drop the argument λ_j in w_a, w_b, A_a, A_b, f_{aa}, f_{bb}, v and J. Then

$$(1)=\frac{2\pi}{n}\sum_{j=1}^{m}\{w_a\overline{w_b}-A_a vv^*A_b^*\}$$

$$=\frac{2\pi}{n}\sum\frac{1}{2}\{(w_a-A_a v)(\overline{w_b}+v^*A_b^*)+(w_a+A_a v)(\overline{w_b}-v^*A_b^*)\}$$

$$=\frac{\pi}{n}\sum Re\{(w_a-A_a v)(\overline{w_b}+v^*A_b^*)+(w_a+A_a v)(\overline{w_b}-v^*A_b^*)\}+$$

$$+i\frac{\pi}{n}\sum Im\{(w_a-A_a v)(\overline{w_b}+v^*A_b^*)+(w_a+A_a v)(\overline{w_b}-v^*A_b^*)\} \ ,$$

so,

$$|Re(1)| = |\frac{\pi}{n} Re\Big(\sum (w_a - A_a v)(\overline{w_b} + v^* A_b^*) + \sum (w_a + A_a v)(\overline{w_b} - v^* A_b^*)\Big)|$$

$$\leq \frac{\pi}{n} \{ |\sum (w_a - A_a v)(\overline{w_b} + v^* A_b^*)| + |\sum (w_a + A_a v)(\overline{w_b} - v^* A_b^*)| \} = \frac{\pi}{n}(E1 + E2) \ ,$$

then,

$$E1 \leq \sqrt{\sum |w_a - A_a v|^2 \sum |\overline{w_b} + v^* A_b^*|^2} \leq \sqrt{\sum |w_a - A_a v|^2 2 \sum (I_{bb} + A_b J A_b^*)} \ ,$$

and

$$E2 \leq \sqrt{\sum |w_b - A_b v|^2 2 \sum (I_{aa} + A_a J A_a^*)} \ ;$$

now, as $EJ = R/2\pi$,

$$E\Big[\sum A_a J A_a^*\Big] = \sum f_{aa} \quad and \quad E\Big[\sum A_b J A_b^*\Big] = \sum f_{bb} \ ,$$

because they are positive random variables their stochastic order of magnitude are those of their expectations, then proposition 1 implies:

$$\frac{1}{n}\sum_{j=1}^{m} A_a J A_a^* = O_p(F_{aa}(\lambda_m)) \quad and \quad \frac{1}{n}\sum_{j=1}^{m} A_b J A_b^* = O_p(F_{bb}(\lambda_m)) \ ,$$

and proposition 2 implies:

$$E\left(\frac{1}{n}\sum_{j=1}^{m} I_{aa}(\lambda_j)\right) = O(F_{aa}(\lambda_m)) \quad and \quad E\left(\frac{1}{n}\sum_{j=1}^{m} I_{bb}(\lambda_j)\right) = O(F_{bb}(\lambda_m)) \ ,$$

and proposition 4 implies:

$$\frac{1}{n}\sum |w_a - A_a v|^2 = o_p(F_{aa}(\lambda_m)) \quad and \quad \frac{1}{n}\sum |w_b - A_b v|^2 = o_p(F_{bb}(\lambda_m)) \ ,$$

then

$$\frac{\pi}{n}(E1+E2)\leq\pi\sqrt{\frac{1}{n}\sum|w_a-A_a v|^2\ 2\ (\frac{1}{n}\sum I_{bb}+\frac{1}{n}\sum A_b J A_b^*)}+$$

$$+\pi\sqrt{\frac{1}{n}\sum|w_b-A_b v|^2\ 2\ (\frac{1}{n}\sum I_{aa}+\frac{1}{n}\sum A_a J A_a^*)}$$

$$=\pi\sqrt{o_p(F_{aa}(\lambda_m)[O(F_{bb}(\lambda)+O_p(F_{bb}(\lambda_m))]}+$$

$$\pi\sqrt{o_p(F_{bb}(\lambda_m)[O(F_{aa}(\lambda)+O_p(F_{aa}(\lambda_m))]}=$$

$$=\sqrt{o_p(F_{aa}(\lambda_m)F_{bb}(\lambda_m))}=o_p(\sqrt{F_{aa}(\lambda_m)F_{bb}(\lambda_m)})\ .$$

So,

$$Re\hat{F}_{ab}(\lambda_m)-ReF_{ab}(\lambda_m)=o_p(F_{aa}(\lambda_m)^{0.5}F_{bb}(\lambda_m)^{0.5})$$

Theorem 4.1 b

Under C4.1, C4.2 and C4.3:

$$Im\hat{F}_{ab}(\lambda_m)-ImF_{ab}(\lambda_m)=o_p(F_{aa}(\lambda_m)^{1/2}F_b(\lambda_m)^{1/2})\ .$$

$$Im\hat{F}_{ab}(\lambda_m)-ImF_{ab}(\lambda_m)=o_p(F_{aa}(\lambda_m)^{1/2}F_b(\lambda_m)^{1/2})\ .$$

Proof: the proof is analogous to that of theorem 4.1 a.

Chapter 5

Analysis of a pseudo maximum likelihood estimate

In this chapter we analyze in a multivariate set-up pseudo-maximum likelihood estimates. By this we mean that they are based on a discrete version in the neighbourhood of zero frequency of an approximation to the Gaussian likelihood function in the frequency domain. These estimates have not an explicit form, which makes them harder to obtain. On the other hand their consistency and asymptotic normality can be proved under fairly general conditions (especially, it is not necessary to impose Gaussianity). Robinson (1993a) analyzes the univariate case. In section 1 we motivate the objective function. Section 2 presents the asymptotic properties of these estimates and some results on inference for the bivariate case. Section 3 provides some finite sample analysis and in section 4 we offer an empirical application. In chapter 6 we will analyze a Lagrange multiplier test for weak dependence, I(0), that is based on results obtained in this chapter.

5.1 Introduction

To motivate our objective function we start with the Whittle approximation of the Gaussian likelihood function in the frequency domain:

$$\mathcal{L}(\theta)=\int_{-\pi}^{\pi}\{\log|f(\lambda,\theta)|+tr[f(\lambda,\theta)^{-1}I(\lambda)]\}d\lambda \ .$$

Where $I(\lambda)$ is the rxr periodogram matrix defined in (4.1.1) and $f(\lambda,\theta)$ is the rxr spectral density matrix. If instead of considering all the frequencies $(-\pi,\pi]$ we focus on a neighbourhood of the zero frequency and in particular for Fourier frequencies $\lambda_j=2\pi j/n$ with j=1,...m, we get:

$$\mathcal{L}(\theta)=\sum_{j=1}^{m}\{\log f(\lambda_j,\theta)+tr[f(\lambda_j,\theta)^{-1}I(\lambda_j)]\}. \qquad (5.1.1)$$

Another main aspect concerns the form of the spectral density we are going to assume. In the previous chapter we considered condition C4.1:

$$f(\lambda)\sim\Lambda^0 G_0\Lambda^0 \quad as\ \lambda\to 0^+, \qquad \Lambda^0=diag\{\lambda^{1/2-H_a^0}\}, \qquad (5.1.2)$$

where G_0 is a positive definite Hermitian matrix and we proved that this was a very general specification. In this chapter we denote by H^0 and G_0 the true values and by H and G any admissible values. Under (5.1.2) we could rewrite (5.1.1) as our objective function:

$$\mathcal{L}(G,H)=\sum_{j=1}^{m}\{\log|\Lambda_j G\Lambda_j|+tr[\Lambda_j^{-1}G^{-1}\Lambda_j^{-1}I(\lambda_j)]\}, \qquad \Lambda_j=diag\{\lambda_j^{1/2-H_a}\} \qquad (5.1.3)$$

but notice that we can concentrate G out using standard matrix differential calculus (see, for instance chapter 10 in Graybill (1983)):

$$\mathcal{L}(G,H)=\sum_{j=1}^{m}\{2\log|\Lambda_j|+\log|G|+tr[G^{-1}\Lambda_j^{-1}I(\lambda_j)\Lambda_{j_1}]\},$$

$$\frac{\partial\mathcal{L}}{\partial G}=\sum_{j=1}^{m}\{(G^{-1})'-(G^{-1}\Lambda_j^{-1}I(\lambda_j)\Lambda_j^{-1}G^{-1})'\},$$

$$\frac{\partial\mathcal{L}}{\partial G}=0 \Rightarrow \hat{G}(H)=\frac{1}{m}\sum_{j=1}^{m}\{\Lambda_j^{-1}I(\lambda_j)\Lambda_j^{-1}\}, \qquad (5.1.4)$$

then define:

$$R_m(H)=\mathcal{L}(\hat{G},H)=\sum_{j=1}^{m}\{2\log|\Lambda_j|+\log|\hat{G}(H)|+tr(\hat{G}^{-1}\Lambda_j^{-1}I(\lambda_j)\Lambda_j^{-1}\}=$$

$$=\sum_{j=1}^{m} 2\log\left(\prod_{a=1}^{r} \lambda_j^{\frac{1-2H_a}{2}}\right)+m\log|\hat{G}(H)|+$$

$$+tr\{\left(\frac{1}{m}\sum_{j=1}^{m} \Lambda_j^{-1}I(\lambda_j)\Lambda_j^{-1}\right)^{-1}\sum_{j=1}^{m} \Lambda_j^{-1}I(\lambda_j)\Lambda_j^{-1}\}$$

$$=\sum_{j=1}^{m} 2\sum_{a=1}^{r} \frac{(1-2H_a)}{2}\log\lambda_j+m\log|\hat{G}(H)|+mr.$$

We can define our objective function to be:

$$R(H)=\frac{R_m(H)}{m}-r=\sum_{a=1}^{r} (1-2H_a)\frac{1}{m}\sum_{j=1}^{m} \log\lambda_j+\log|\hat{G}(H)|, \tag{5.1.5}$$

so we can carry out the estimation procedure in two steps:

1.- Estimate H minimizing R(H):

$$\hat{H}=arg\ \min_{\Theta} R(H)$$

where R(H) is defined in (5.1.5) and

$$\Theta=\{H\in R^r;\ H_i\in[\Delta_i^1,\Delta_i^2],\ i=1,...r\} \tag{5.1.6}$$

and Δ_i^1,Δ_i^2 are user-chosen and, in principle, $0<\Delta_i^1<\Delta_i^2<1$.

2.- Estimate G by plugging the estimate of H obtained in the previous step in (5.1.4):

$$\hat{G}=\hat{G}(\hat{H})=\frac{1}{m}\sum_{j=1}^{m} \hat{\Lambda}_j^{-1}I(\lambda_j)\hat{\Lambda}_j^{-1}$$

$$\hat{\Lambda}_j^{-1}=diag\{\lambda_j^{\hat{H}_a-1/2}\}.$$

These estimates are asymptotically globally identified under the assumptions we introduce in the next section as we are going to show next. We only need to check that H^0 is identified: first write

$$R(H)-R(H^0)=U(H)-T(H)$$

where

$$U(H)=2\sum_{i=1}^{r}(H_i-H_i^0)-\sum_{i=1}^{r}\log\{1+2(H_i-H_1^0)\} \tag{5.1.7}$$

and

$$T(H)=2\sum_{i=1}^{r}(H_i-H_i^0)[\frac{1}{m}\sum_{j=1}^{m}\log j-\log m+1]+\log|\Gamma^{-1}\hat{G}(H^0)|-\\ -\log|Z\Gamma^{-1}M\hat{G}(H)| \tag{5.1.8}$$

with

$$\Gamma=diag\{g_{ii}\}\ ,\quad M=diag\{\left(\frac{2\pi m}{n}\right)^{2(H_i^0-H_i)}\}\ ,\quad Z=diag\{1+2(H_i-H_i^0)\}\ . \tag{5.1.9}$$

Under those assumptions $T(H)\rightarrow_p 0$ as is proved in appendix 5.1, then: $R(H)-R(H^0)\rightarrow_p U(H)$.

Now if we prove that $U(H)>0$ for all $H\in\Theta-\{H^0\}$ then H^0 is globally identified (because the existence of a H_1 observationally equivalent to H^0 would imply $U(H_1)=0$). Now, calling $H_i-H_i^0=x_i$ we have that $U(H)=U(x)=\sum_{i=1}^{r}(2x_i-\log(1+2x_i))$ and as $2x-\log(1+2x)\geq\frac{2}{3}x^2>0$ we
$H_i-H_i^0=x_i$ we have that $U(H)=U(x)=\sum(2x_i-\log(1+2x_i))$ and as $2x-\log(1+2x)\geq\frac{2}{3}x^2>0$ we

The QMLE has not an explicit form what makes it more difficult to get than the APE and the LPE. On the other hand, with very simple iterative procedures, such as a golden search, we have found in simulations and in real data that, for moderate r, the estimates converge quickly so that, in our experience, estimation has not been a difficult problem.

5.2. Consistency and Asymptotic Normality of the QMLE

We introduce first the conditions under which we can obtain consistency of the QMLE. These conditions are very general. We just restrict the behaviour of the spectral density matrix close to the zero frequency and assume that the process has a linear representation in terms of a square integrable martingale difference sequence. Robinson

(1993a) under similar conditions proved the consistency of the QMLE in the univariate case. We have to state one limitation of our procedure compared with the univariate case. Robinson's admissible estimates lay in the interval (0,1), i.e., including both cases, the "long-memory" case ($1/2<H<1$) and the "antipersistent" case ($0<H<1/2$). Here, due to a technical problem, we will not be able to prove consistency of our estimate for H_a when the admissible interval is (0,1) but we will suppose that in the case of "long-memory" we will restrict our possible set of estimates to (1/2,1). This is not so restrictive as it may appear. In any practical situation it should be clear by a simple inspection of the periodogram if the series belong to the "long-memory" or to the "antipersistent" case.

Conditions:

C5.1: as $\lambda \to 0^+$, $f(\lambda) \sim \Lambda^0 G_0 \Lambda^0$ where G_0 is Hermitian positive definite with typical element (a,b), g_{ab}, and $\Lambda^0 = diag\{\lambda^{0.5-H_i^0}\}$ where $H^0_i \in [\Delta^1_i, \Delta^2_i]$ is the interval of admissible estimates. In principle we can choose Δ_i^1 and Δ_i^2 so that $0<\Delta^1_i<\Delta^2_i<1$, but we will assume that if $H_i^0>0.5$ then we will pick $\Delta_i^1 \geq 0.5$.

C5.2: In a neighbourhood $(0,\epsilon)$ of the origin, $f_{aa}(\lambda)$ is differentiable with

$$\frac{d}{d\lambda}\log f_{aa}(\lambda) = O(\lambda^{-1}), \quad as \ \lambda \to 0^+, \ for \ \ a=1,...r.$$

C5.3: Condition C4.3 in chapter 4.

C5.4: Condition C4.2 in chapter 4.

Theorem 5.1

Under C5.1, C5.2, C5.3 and C5.4:

$$\hat{H} \to_p H.$$

Proof: see appendix 5.1.

To prove asymptotic normality we need to restrict our assumptions. We introduce:

C5.1': for $\beta \in (0,2]$: $f_{ab}(\lambda) \sim g_{ab}\lambda^{1-H_a^0-H_b^0}(1+O(\lambda^{\beta}))$ *as* $\lambda \to 0^+$.

C5.2': in a neighbourhood of 0

$$\frac{d}{d\lambda}A_a^k(\lambda)=O\left(\frac{|A_a^k(\lambda)|}{\lambda}\right) \text{ as } \lambda \to 0^+ \text{ for all } a,k=1,...r,$$

where $A_a^k(\lambda)$ has been defined in the previous chapter.

C5.3': similar to C5.3 with:

$$E(e_a(t)e_b(t)e_c(t)|F_{t-1})=\mu_{abc}^{(3)}, \quad |\mu_{abc}^{(3)}|<\infty$$

$$E(e_a(t)e_b(t)e_c(t)e_d(t)|F_{t-1})=3+\kappa_{abcd}, \quad |\kappa_{abcd}|<\infty \qquad a,b,c,d=1,...r$$

C5.4':

$$\frac{1}{m}+\frac{m^{1+2\beta}(\log m)^2}{n^{2\beta}}\to 0.$$

Theorem 5.2 :

Under C5.1', C5.2', C5.3' and C5.4':

$$\sqrt{m}(\hat{H}-H^0)\to_d N_r(0,E^{-1}), \tag{5.2.1}$$

where E is rxr matrix: $E=2I_r+2Re(G_0*(G_0^{-1})')$ and * denotes the Hadamard product, so the typical (a,b) element is:

$$\{E_{ab}\}=\begin{cases}2+2g_{aa}g^{aa} & \text{if } a=b\\ 2Re\,g_{ab}g^{ba} & \text{if } a\neq b.\end{cases}$$

Proof: see appendix 5.2.

This result on asymptotic normality of the QMLE provides us with a tool on which we can base our inference about H. It is interesting to notice that the elements of the covariance matrix E^{-1} depend on elements of G_0 and G_0^{-1} only. As we can estimate G_0

consistently using $\hat{G}_0=\hat{G}(\hat{H})$, we can estimate G_0^{-1} and so E consistently by:

$$\{\hat{E}_{ab}\}=\begin{cases}2+2\hat{g}_{aa}\hat{g}^{aa} & if \quad a=b\\ 2Re\hat{g}_{ab}\hat{g}^{ba} & if \quad a\neq b.\end{cases}$$

It is very interesting to notice also that in the univariate case the asymptotic variance does not depend on any unknown parameter, $E_{aa}=2+2g_{aa}g^{aa}=4$, so that the asymptotic distribution, first analyzed in Robinson (1993a), is just:

$$\sqrt{m}(\hat{H}-H_0)\to_d N(0,\frac{1}{4}),$$

so, statistical inference is particularly immediate in this case.

If we analyze the asymptotic distribution for the bivariate case we obtain that:

$$E=\begin{pmatrix}2+\dfrac{2}{1-c^2} & \dfrac{-2c^2}{1-c^2}\\ \dfrac{-2c^2}{1-c^2} & 2+\dfrac{2}{1-c^2}\end{pmatrix}$$

i.e.,

$$E^{-1}=\frac{1}{8}\begin{pmatrix}2-c^2 & c^2\\ c^2 & 2-c^2\end{pmatrix},$$

where

$$c^2=\frac{|g|^2}{g_1g_2} \quad and \quad G_0=\begin{pmatrix}g_1 & g\\ \bar{g} & g_2\end{pmatrix}$$

i.e., c^2 is the squared coherency as $\lambda\to 0^+$. The asymptotic variance of $\hat{H}_1$ or $\hat{H}_2$ will be $\frac{1}{4}-\frac{c^2}{8}$. As $0\leq c^2<1$ (due to G_0 being positive definite) then, the greater the coherency the more efficiently we will estimate the coefficient H. Note also that c^2 can be estimated consistently using that G_0 can be estimated by:

$$\hat{G}_0=\frac{1}{m}\sum_{j=1}^{m}\begin{pmatrix}\lambda_j^{\hat{H}_1-1/2} & 0\\ 0 & \lambda_j^{\hat{H}_2-1/2}\end{pmatrix}\begin{pmatrix}I_{11}(\lambda_j) & I_{12}(\lambda_j)\\ \overline{I_{12}(\lambda_j)} & I_{22}(\lambda_j)\end{pmatrix}\begin{pmatrix}\lambda_j^{\hat{H}_1-1/2} & 0\\ 0 & \lambda_j^{\hat{H}_2-1/2}\end{pmatrix}=$$

$$=\begin{pmatrix}\frac{1}{m}\sum_{j=1}^{m}I_{11}(\lambda_j)\lambda_j^{2\hat{H}_1-1} & \frac{1}{m}\sum_{j=1}^{m}I_{12}(\lambda_j)\lambda_j^{\hat{H}_1+\hat{H}_2-1}\\ \frac{1}{m}\sum_{j=1}^{m}\overline{I_{12}(\lambda_j)}\lambda_j^{\hat{H}_1+\hat{H}_2-1} & \frac{1}{m}\sum_{j=1}^{m}I_{22}(\lambda_j)\lambda_j^{2\hat{H}_2-1}\end{pmatrix}$$

then,

$$\hat{c}^2=\frac{\left|\frac{1}{m}\sum_{j=1}^{m}I_{12}(\lambda_j)\lambda_j^{\hat{H}_1+\hat{H}_2-1}\right|^2}{\left(\frac{1}{m}\sum_{j=1}^{m}I_{11}(\lambda_j)\lambda_j^{2\hat{H}_1-1}\right)\left(\frac{1}{m}\sum_{j=1}^{m}I_{22}(\lambda_j)\lambda_j^{2\hat{H}_2-1}\right)}$$

and this estimate is consistent for c^2 (by Slutsky): because $\hat{c}^2=\frac{|\hat{g}|^2}{\hat{g}_{11}\hat{g}_{22}}$ and $|\hat{g}|\to_p |g|$, $\hat{g}_1\to_p g_1$ *and* $\hat{g}_2\to_p g_2$.

We can use (5.2.1) for statistical inference. Consider as null hypothesis a linear set of q ($\leq$r) independent restrictions on H^0:

$$H_0\text{: } RH^0=v$$

where R is qxr and v is qx1; then, asymptotically, under H_0:

$$m(R\hat{H}-v)'(R\hat{E}^{-1}R')^{-1}(R\hat{H}-v)\sim\chi_q^2.$$

Two interesting cases are:

a) Equality of H^0 across the series. In this case v is a vector of r-1 zeros and R is

$$R=(I_{r-1}\,\vdots\,0)-(0\,\vdots\,I_{r-1})$$

with dimension (r-1)xr, where 0 is a rx1 vector of zeros and I_{r-1} is the identity matrix of order r-1.In section 4 we will apply this test to some data.

b) The vector process is I(0), that is, $H^0=1/2$ for every series. In this case R is the identity matrix of order r and v is a vector with all its components equal to 1/2.

5.3. Optimal m

In this section we analyze the finite sample performance of the QMLE and, in particular, we focus on the selection of m, the bandwidth parameter. As we have seen in chapter 2 the APE and the LPE both depend on two user-chosen numbers: in the APE, q and m, in the LPE, l and m. The QMLE depends only on one user-chosen number, m, and this is a clear advantage of this estimate.

Optimality of m for the case $H \in (1/2,1)$ has been analyzed in Robinson (1994b) and in chapter 3 we have already commented his results. For $H \in (0,1)$ Hurvich and Beltrao (1994) have heuristically analyzed two cross-validated criteria for selecting m. In this section we just consider some Monte Carlo results for the univariate case. We consider two sample sizes: n=128 and n=256, and eight possible values for m: 4, 8, 12, 16, 24, 32, 40 and 48 for the first sample size and 8, 16, 24, 32, 48, 64, 80 and 96 for the second. These sets of values for m should be enough for our purposes. We generate fractional Gaussian noise using the procedure of Davies and Harte (1987). The QMLE is obtained by a golden search procedure. We analyze nine possible values for H: 0.1(0.1)0.9. We perform 10000 replications in each case. Table 5.1 provides the results for n=128 and table 5.2 for n=256.

First, for sample size of 128, we appreciate that for H small the bias tends to be small for relatively small values of m (8-12) while for large H then the bias is small for bigger values for m (32). For instance, with H=0.1 the bias is less with m=12 and for

H=0.2 and H=0.3 the bias is less for m=8, but for values of H bigger than 0.5, the long-memory case, then m=32 is the selection that proportionates less bias. When n=256 we get similar results: for H small we get that small m (like 16) will render less bias, while for big H we will need bigger m (like 48).

The asymptotic variance is 1/4 and as m increases we get that the finite sample variance approximates more to this benchmark as expected.

If we base our optimality criterion for m in the minimum mean squared error we will get that only for H=0.1 or H=0.2 we can get a reasonable m, in the other cases this criterion would imply a fairly big value for m. In particular, for n=128, the optimal m will be around 24 and 32 for H=0.1 and H=0.2 and for n=256 it would be 32-48 and 48-64 for H=0.1 and H=0.2, respectively.

An interesting aspect is the lack of skewness and kurtosis in these distributions except for the case H=0.1 (especially for big values of m). This confirms the normal approximation.

We can compare these Monte Carlo results with those reported in chapter 3 for the APE. Columns for m=32 and m=64 in table 5.2 are the counterpart of table 3.3 in chapter 3. Also notice that only values for H>0.5 can be compared. The first striking feature is the difference in the bias, especially for m=32. The APE shows a fairly big negative bias while the QMLE has approximately zero bias. Also the degree of skewness and kurtosis of the APE is very severe compared with the QMLE in which they are almost nonexistent (we have to remind here that only for H<0.75 we got a normal distribution for the APE).

The only positive feature of the APE with respect to the QMLE is the slight less variance of that estimate, especially for H=0.8 in which case for q=0.5 the variance of the APE is substantially inferior to that of the QMLE.

5.4. Empirical application

In this section we apply the QMLE to some exchange rate data. We use daily and weekly exchange rate data from January 1989 to July 1994. There are four series: BP/$, BP/DM, BP/JYn and BP/SwFr. The sample sizes are 292 and 1459 for the weekly and daily data respectively. The series we analyze are the squares of the first differences of the logarithm of the spot exchange rates. These series can be interpreted as a measure of the volatility. We estimate the different H's using QMLE, employing the downhill simplex method that is a very robust way of finding the minimum of a multivariate function. The subroutine we use is due to Press et al. (1990). Convergence is achieved very quickly. In tables 5.3 and 5.4 we present the results for weekly and daily data respectively. We have chosen as representatives two values for m. For weekly series, n=292, the chosen m are 20 and 40. We present the estimates of H and also the 95% asymptotic confidence interval. The estimates are greater than 0.5 and show evidence of long-memory, especially the DM/BP series (but notice that for m=40 the asymptotic confidence intervals include H=1/2 in three cases). We also present the estimate of the matrix of square coherencies at zero frequency. The most clear feature is the high coherency between the DM/BP and the SwFr/BP.

Furthermore, we perform a test of equality of H for the four series. This test is immediate based on (5.2.1). Under the null hypothesis of equality of all H's:

$$m(R\hat{H})'(R\hat{E}^{-1}R')^{-1}R\hat{H} \sim \chi_3^2 , \quad R=\begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}.$$

When we don't reject the null hypothesis based on the asymptotic value we also provide the common estimate of H.

For daily data, n=1459, we have chosen m=40 and 80. The evidence of long-memory here is more clear.

In chapter 6 we will apply a Lagrange multiplier test for I(0) to this data set.

TABLE 5.1

mean

H\m	4	8	12	16	24	32	40	48
0.1	0.238	0.142	0.101	0.074	0.044	0.028	0.019	0.014
0.2	0.283	0.213	0.185	0.167	0.144	0.128	0.118	0.112
0.3	0.335	0.289	0.276	0.268	0.258	0.252	0.247	0.244
0.4	0.394	0.373	0.371	0.370	0.371	0.370	0.370	0.369
0.5	0.457	0.462	0.470	0.473	0.480	0.483	0.486	0.487
0.6	0.523	0.554	0.570	0.577	0.587	0.593	0.598	0.601
0.7	0.589	0.645	0.670	0.679	0.692	0.701	0.707	0.712
0.8	0.654	0.731	0.763	0.777	0.795	0.806	0.815	0.822
0.9	0.714	0.807	0.845	0.864	0.887	0.901	0.913	0.922

variance

H\m	4	8	12	16	24	32	40	48
0.1	0.092	0.034	0.018	0.011	0.005	0.003	0.002	0.001
0.2	0.105	0.047	0.029	0.020	0.013	0.009	0.007	0.006
0.3	0.116	0.058	0.037	0.026	0.016	0.012	0.009	0.007
0.4	0.126	0.066	0.042	0.029	0.017	0.012	0.009	0.007
0.5	0.132	0.072	0.044	0.031	0.018	0.012	0.009	0.007
0.6	0.134	0.073	0.045	0.031	0.018	0.012	0.009	0.007
0.7	0.131	0.070	0.043	0.030	0.017	0.012	0.009	0.007
0.8	0.124	0.062	0.037	0.026	0.016	0.012	0.009	0.007
0.9	0.114	0.050	0.029	0.020	0.012	0.009	0.006	0.005

mse

H\m	4	8	12	16	24	32	40	48
0.1	0.111	0.036	0.018	0.012	0.008	0.008	0.008	0.008
0.2	0.111	0.047	0.029	0.021	0.016	0.014	0.014	0.014
0.3	0.117	0.058	0.037	0.027	0.018	0.014	0.012	0.010
0.4	0.126	0.067	0.042	0.030	0.018	0.013	0.010	0.008
0.5	0.134	0.073	0.045	0.031	0.018	0.012	0.009	0.007
0.6	0.140	0.075	0.045	0.031	0.018	0.012	0.009	0.007
0.7	0.144	0.073	0.044	0.030	0.017	0.012	0.009	0.007
0.8	0.146	0.067	0.039	0.027	0.016	0.012	0.009	0.007
0.9	0.149	0.059	0.032	0.021	0.012	0.009	0.007	0.006

skewness

H\m	4	8	12	16	24	32	40	48
0.1	1.130	1.398	1.474	1.597	1.865	2.196	2.579	2.927
0.2	0.889	0.929	0.759	0.669	0.496	0.436	0.395	0.319
0.3	0.642	0.582	0.360	0.216	0.005	-0.078	-0.126	-0.164
0.4	0.395	0.291	0.089	-0.037	-0.170	-0.200	-0.198	-0.197
0.5	0.147	0.027	-0.094	-0.174	-0.217	-0.213	-0.191	-0.187
0.6	-0.105	-0.238	-0.251	-0.269	-0.228	-0.213	-0.182	-0.180
0.7	-0.363	-0.519	-0.453	-0.386	-0.270	-0.226	-0.179	-0.179
0.8	-0.635	-0.847	-0.753	-0.640	-0.466	-0.368	-0.304	-0.269
0.9	-0.924	-1.247	-1.206	-1.107	-0.988	-0.923	-0.886	-0.864

kurtosis

H\m	4	8	12	16	24	32	40	48
0.1	0.108	1.554	2.074	2.442	3.460	5.046	7.004	9.780
0.2	-0.475	0.271	0.096	-0.023	-0.376	-0.444	-0.427	-0.471
0.3	-0.917	-0.351	-0.366	-0.331	-0.314	-0.194	-0.012	0.028
0.4	-1.217	-0.650	-0.396	-0.153	0.039	0.151	0.263	0.198
0.5	-1.381	-0.749	-0.290	0.051	0.228	0.244	0.290	0.203
0.6	-1.406	-0.693	-0.195	0.128	0.282	0.272	0.287	0.204
0.7	-1.284	-0.455	-0.076	0.079	0.219	0.214	0.256	0.184
0.8	-0.989	0.062	0.248	0.153	0.104	0.040	0.035	-0.002
0.9	-0.495	1.066	1.244	0.948	0.757	0.524	0.411	0.282

TABLE 5.2

mean

H\m	8	16	24	32	48	64	80	96
0.1	0.164	0.102	0.073	0.054	0.030	0.017	0.010	0.007
0.2	0.224	0.186	0.171	0.161	0.143	0.129	0.121	0.115
0.3	0.296	0.277	0.275	0.273	0.266	0.259	0.255	0.251
0.4	0.376	0.374	0.378	0.380	0.380	0.378	0.377	0.375
0.5	0.464	0.473	0.481	0.485	0.489	0.490	0.492	0.492
0.6	0.555	0.574	0.583	0.588	0.595	0.599	0.603	0.605
0.7	0.645	0.674	0.685	0.691	0.699	0.705	0.712	0.716
0.8	0.730	0.772	0.787	0.794	0.804	0.811	0.819	0.825
0.9	0.806	0.859	0.879	0.889	0.903	0.913	0.922	0.929

variance

H\m	8	16	24	32	48	64	80	96
0.1	0.038	0.014	0.008	0.005	0.002	0.001	0.001	0.000
0.2	0.049	0.022	0.014	0.010	0.007	0.005	0.004	0.003
0.3	0.059	0.027	0.017	0.012	0.007	0.005	0.004	0.003
0.4	0.067	0.029	0.018	0.012	0.007	0.005	0.004	0.003
0.5	0.073	0.030	0.018	0.012	0.007	0.005	0.004	0.003
0.6	0.074	0.031	0.018	0.012	0.007	0.005	0.004	0.003
0.7	0.070	0.030	0.018	0.012	0.007	0.005	0.004	0.003
0.8	0.063	0.027	0.017	0.012	0.007	0.005	0.004	0.003
0.9	0.051	0.021	0.013	0.009	0.006	0.004	0.003	0.003

mse

H\m	8	16	24	32	48	64	80	96
0.1	0.042	0.014	0.009	0.007	0.007	0.008	0.009	0.009
0.2	0.050	0.022	0.015	0.012	0.010	0.010	0.010	0.010
0.3	0.059	0.027	0.018	0.013	0.009	0.007	0.006	0.006
0.4	0.068	0.030	0.018	0.013	0.008	0.006	0.005	0.004
0.5	0.074	0.031	0.018	0.012	0.008	0.005	0.004	0.003
0.6	0.076	0.031	0.018	0.012	0.007	0.005	0.004	0.003
0.7	0.073	0.030	0.018	0.012	0.007	0.005	0.004	0.004
0.8	0.067	0.027	0.017	0.012	0.007	0.005	0.004	0.004
0.9	0.060	0.022	0.013	0.009	0.006	0.005	0.004	0.004

skewness

H\m	8	16	24	32	48	64	80	96
0.1	1.226	1.171	1.261	1.390	1.803	2.405	2.908	3.368
0.2	0.866	0.547	0.365	0.235	0.134	0.096	0.045	0.035
0.3	0.561	0.196	0.004	-0.099	-0.165	-0.186	-0.187	-0.155
0.4	0.288	-0.006	-0.132	-0.185	-0.184	-0.192	-0.181	-0.146
0.5	0.034	-0.115	-0.182	-0.201	-0.182	-0.191	-0.174	-0.137
0.6	-0.222	-0.202	-0.206	-0.203	-0.182	-0.191	-0.169	-0.131
0.7	-0.507	-0.343	-0.258	-0.218	-0.182	-0.191	-0.165	-0.126
0.8	-0.849	-0.615	-0.442	-0.327	-0.221	-0.208	-0.164	-0.128
0.9	-1.264	-1.103	-0.938	-0.809	-0.677	-0.635	-0.581	-0.551

kurtosis

H\m	8	16	24	32	48	64	80	96
0.1	1.015	0.895	1.090	1.500	2.962	6.117	9.237	12.799
0.2	0.088	-0.272	-0.438	-0.504	-0.493	-0.434	-0.422	-0.374
0.3	-0.415	-0.376	-0.265	-0.163	-0.011	0.113	0.077	0.070
0.4	-0.678	-0.199	0.017	0.094	0.075	0.156	0.078	0.074
0.5	-0.769	-0.018	0.178	0.179	0.078	0.171	0.073	0.071
0.6	-0.715	0.082	0.224	0.195	0.085	0.185	0.069	0.067
0.7	-0.478	0.090	0.183	0.160	0.094	0.195	0.065	0.060
0.8	0.076	0.251	0.105	0.035	0.020	0.153	0.044	0.026
0.9	1.139	1.236	0.707	0.378	0.084	0.138	-0.066	-0.150

TABLE 5.3
(weekly data)

USD/BP, DM/BP, JYn/BP, SFr/BP

	m=20		m=40	
H1	0.78	(0.66-0.91)	0.60	(0.50-0.70)
H2	0.72	(0.60-0.84)	0.66	(0.57-0.75)
H3	0.76	(0.61-0.91)	0.62	(0.49-0.75)
H4	0.68	(0.56-0.81)	0.57	(0.48-0.68)

Normalized |G|

1	0.79	0.74	0.78	1	0.73	0.70	0.72
	1	0.73	0.93		1	0.62	0.88
		1	0.67			1	0.52
			1				1

	m=20	m=40
Chi-test	2.28	5.09
Common H	0.66	0.67

TABLE 5.4
(daily data)

	m=40		m=80	
H1	0.85	(0.72-0.97)	0.75	(0.66-0.85)
H2	0.78	(0.67-0.90)	0.72	(0.63-0.81)
H3	0.74	(0.59-0.89)	0.71	(0.60-0.83)
H4	0.86	(0.74-0.99)	0.75	(0.65-0.85)

Normalized |G|

1	0.45	0.47	0.40	1	0.40	0.38	0.33
	1	0.47	0.93		1	0.42	0.85
		1	0.40			1	0.33
			1				1

	m=40	m=80
Chi-test	8.30	1.47
Common H	---	0.74

Appendix 5.1

This theorem is an application of theorem 29 of P. Robinson's "Quantitative Techniques" lecture notes that states that:

" Let $Q_n(\theta)$ be a scalar function of a rx1 vector θ and of random variables, with sample size n; let $\Theta \subset R^r$, let:

$$\hat{\theta}=\hat{\theta}_n=arg\ \min_{\theta\in\Theta} Q_n(\theta),$$

if:

(i): $\theta_0 \in \Theta$ and

(ii): $Q_n(\theta)-Q_n(\theta_0)=S(\theta)-T_n(\theta)$,

such that $S(\theta)$ is nonstochastic and constant over n and for all $\epsilon>0$ there exists $\eta>0$ such that: $inf_{|\theta-\theta_0|\geq\epsilon}\ S(\theta)\geq\eta$

and $T_n(\theta)$ satisfies: $sup_{\theta\in\Theta}\ |T_n(\theta)|\to_p 0$

then: $\hat{\theta}\to_p \theta_0$."

In our case: (i) is by assumption and in order to prove (ii) we need to show:

1). $R(H)-R(H^0)=U(H)-T(H)$,

2a). $inf_{\bar{N}_\delta\cap\Theta}\ U(H)\geq\eta$,

2b). $sup_\Theta |T(H)|\to_p 0$,

where Θ is given in (5.1.6) and for $0<\delta<0.5$:

$$N_\delta=\{H:\|H-H^0\|<\delta\} \quad and \quad \bar{N}_\delta=\mathbb{R}^r-N_\delta \ .$$

where $\| \ \|$ denotes the supremum norm in R^r : $\|A\|=\max_i(|A_i|)$.

Proof of 1). Define S(H)=R(H)-R(H_0), then

$$S(H)=\frac{1}{m}\sum_{j=1}^{m}\log\lambda_j\cdot 2\sum_{i=1}^{r}(H_i^0-H_i)+\log\frac{|\hat{G}(H)|}{|\hat{G}(H^0)|}$$

also recalling the definitions of Γ, M and Z of chapter 5, equation (5.1.9), then:

$$S(H)=2\sum_{i=1}^{r}(H_i^0-H_i)\frac{1}{m}\sum_{j=1}^{m}\log j+2\sum_{i=1}^{r}(H_i^0-H_i)\log\left(\frac{2\pi}{n}\right)+\log|Z\Gamma^{-1}M\hat{G}(H)|$$

$$-\log|Z|-\log|M|-\log|\Gamma^{-1}\hat{G}(H^0)| \ .$$

And:

$$S(H)=-T(H)+U(H) \ ,$$

where U(H) has been defined in (5.1.7) and T(H) in (5.1.8).

Proof of 2a). As

$$U(H)=\sum_{i=1}^{r}U_i(H_i) \ where \ U_i(H_i)=2(H_i-H_i^0)-\log(1+2(H_i-H_i^0))$$

and so

$$inf_{\bar{N}_\delta\cap\Theta}U(H)=\sum_{i=1}^{r}inf_{\bar{N}_\delta\cap\Theta^{(i)}}\ U_i(H_i) \qquad (*)$$

where:

$$\Theta=\mathop{X}_{i=1}^{r}\Theta^{(i)} \ and \ \Theta^{(i)}=[\Delta_i^1,\Delta_i^2]$$

so, as $x-\log(1+x)\geq\frac{1}{6}x^2$ and $-x-\log(1-x)\geq\frac{1}{2}x^2$,

$$inf_{\bar{N}_\delta\cap\Theta^{(i)}}\ U_i(H_i)\geq\min(2\delta-\log(1+2\delta),-2\delta-\log(1-2\delta))\geq\frac{\delta^2}{2}$$

so,

$$(*)\geq\frac{r\delta^2}{2}=\eta>0.$$

Proof of 2b): as $H^0\in\Theta$,

$$sup_\Theta|T(H)|\leq 2r|\frac{1}{m}\sum_{j=1}^{m}\log j-\log m+1|+2sup_\Theta|\log|Z\Gamma^{-1}M\hat{G}(H)||\ ,$$

then, using $\log|A|\leq tr(A-I)$

$$\log|Z\Gamma^{-1}M\hat{G}(H)|\leq tr(Z\Gamma^{-1}M\hat{G}(H)-I_r)=$$

$$=tr(Z\Gamma^{-1}M\frac{1}{m}\sum_{j=1}^{m}\Lambda_j^{-1}I(\lambda_j)\Lambda_j^{-1}-I_r)=tr(\frac{1}{m}\sum_{j=1}^{m}ZM\Lambda_j^{-2}\Psi_j^{-1}\Psi_j\Gamma^{-1}I(\lambda_j)-I_r)=$$

$$=tr(\frac{1}{m}\sum_{j=1}^{m}ZM\Lambda_j^{-2}\Psi_j^{-1}\Phi_j^{-1}\Phi_j\Psi_j\Gamma^{-1}I(\lambda_j)-I_r)=tr(\frac{1}{m}\sum_{j=1}^{m}Z\Phi_j\Psi_j\Gamma^{-1}I(\lambda_j)-I_r)=$$

$$=tr(Z\frac{1}{m}\sum_{j=1}^{m}\Phi_j[\Psi_j\Gamma^{-1}I(\lambda_j)-I_r])+tr(\frac{1}{m}\sum_{j=1}^{m}(Z\Phi_j-I_r))= \qquad (A5.1.1)$$

where

$$\Psi_j=diag\{\lambda_j^{2H_i^0-1}\}\ ,\ \Phi_j=diag\{\left(\frac{j}{m}\right)^{2(H_i-H_i^0)}\}\ ,$$

and we have used that

$$M\Lambda_j^{-2}\Psi_j^{-1}\Phi_j^{-1}=I_r\ .$$

Then

$$(A5.1.1)=\sum_{i=1}^{r}\left[(1+2(H_i-H_i^0))\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2(H_i-H_i^0)}-1\right]+$$

$$+\sum_{i=1}^{r}(1+2(H_i-H_i^0))\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2(H_i-H_i^0)}\left(g_{ii}^{-1}\lambda_j^{2H_i^0-1}I_{ii}(\lambda_j)-1\right)\ ,$$

then

$$sup_\Theta|T(H)|\leq 2r|\frac{1}{m}\sum_{j=1}^{m}\log j-\log m+1|+ \qquad (*)$$

$$2sup_\Theta\left|\sum_{i=1}^{r}\left[(1+2(H_i-H_i^0))\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2(H_i-H_i^0)}-1\right]\right|+ \qquad (**)$$

$$2sup_\Theta\left|\sum_{i=1}^{r}(1+2(H_i-H_i^0))\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2(H_i-H_i^0)}\left(g_{ii}^{-1}\lambda_j^{2H_i^0-1}I_{ii}(\lambda_j)-1\right)\right| \qquad (***)$$

(*)=O(logm/m) using lemma 2 of Robinson (1993a);

(**)= $O\left(m^{-1-2\min_i(\Delta_i^1-H_i^0)}\right)=O(m^{-\epsilon})$ *for* $\epsilon>0$ *because* $(\Delta_i^1-H_i^0)\in(-0.5,0)$ using lemma 1 in Robinson (1993a) and C5.1; and:

(***)= $2sup_\Theta\left|\sum_{i=1}^{r}A_i(H_i)\right|$, where

$$A_i(H_i)=(1+2(H_i-H_i^0))\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2(H_i-H_i^0)}\left(g_{ii}^{-1}\lambda_j^{2H_i^0-1}I_{ii}(\lambda_j)-1\right)$$

$$(***)\le 2\sum_{i=1}^{r}sup_{\Theta^{(i)}}|A_i(H_i)|\rightarrow_p 0$$

$$as \quad sup_{\Theta^{(i)}}|A_i(H_i)|\rightarrow_p 0 \quad for \quad all \quad i,$$

as is proven in Robinson (1993a), to finish the proof of 2b).

Appendix 5.2

This theorem is an application of theorem 32 of P.Robinson's "Quantitative Techniques" lecture notes on asymptotic normality of extremum estimates. This theorem, with the same framework than the consistency theorem we have seen in Appendix 5.1 states that: " if:

(i) θ_0 is an interior point of the compact set Θ,

(ii) $Q_n(\theta)$ is twice differentiable and

$$\sqrt{n}\frac{\partial Q_n(\theta_0)}{\partial\theta}\rightarrow_d N_r(0,D),$$

and for any $\tilde{\theta}\rightarrow_p \theta_0 \Rightarrow \dfrac{\partial^2 Q_n(\tilde{\theta})}{\partial\theta\partial\theta'}\rightarrow_p E>0,$

(iii) $\hat{\theta} \to_p \theta_0$,
then $\sqrt{n}(\hat{\theta}-\theta_0) \to_d N_r(0, E^{-1}DE^{-1})$".

In our case we need to examine 5 points:

(1) H_0 is an interior point of Θ;

(2) R(H) is twice continuously differentiable and for every λ:

$$\sqrt{m}\ \lambda' \frac{dR(H)}{dH}\Big|_{H^0} \to_d N(0, \lambda' E \lambda)$$

(where we have used $x \sim N_r(\mu, \Sigma) \leftrightarrow \lambda' x \sim N_1(\lambda'\mu, \lambda'\Sigma\lambda)$ *for any* λ);

(3) *for any* $\tilde{H} \to_p H^0 \Rightarrow \frac{d^2R(\tilde{H})}{dH_a dH_b} \to_p E_{ab}$;

(4) E>0;

(5) $\hat{H} \to_p H^0$.

Now we analyze them:

(1) by assumption;

(5) by the theorem of consistency, proved in appendix 5.1.

(3) by lemma 1, proved in appendix 5.4.

(4) This is immediate considering that:

$$E = 2I_r + 2Re\Omega \quad where \quad \Omega = G_0 * (G_0^{-1})'$$

and as the Hadamard product of two positive definite matrices is also positive definite (see Styan (1973) p.221), then Ω is positive definite too.

Let analyze (2). We have to prove:

$$\sqrt{m}\ \lambda' \frac{dR(H)}{dH}\Big|_{H^0} \to_d N(0, \lambda' E \lambda). \qquad (A5.2.1)$$

where the variance can be written as:

$$\lambda' E\lambda = \sum_{a=1}^{r}\sum_{b=1}^{r}\lambda_a\lambda_b E_{ab}.$$

$$LHS(A5.2.1) = \sqrt{m}\sum_{a=1}^{r}\lambda_a \frac{dR(H)}{dH_a}|_{H^0} \qquad \text{(A5.2.2)}$$

As

$$\frac{dR(H)}{dH_a} = -2K_1 + tr\left[\hat{E}(H)^{-1}\frac{d\hat{E}(H)}{dH_a}\right] =$$

$$= -2K_1 + tr\left[\hat{G}(H)^{-1}\left(\frac{1}{m}\sum_{j=1}^{m}(\log j)\Lambda_j^{-1}\{i_a I(\lambda_j) + I(\lambda_j)i_a\}\Lambda_j^{-1}\right)\right]$$

where we define the matrix i_a as the rxr matrix with every element equal to zero except the a-th diagonal that is one and

$$K_1 = \frac{1}{m}\sum_{j=1}^{m}\log j\ ,\ \hat{E}(H) = \frac{1}{m}\sum_{j=1}^{m}J_j^{-1}I_j J_j^{-1}, \qquad \text{(A5.2.2a)}$$

where

$$J_j = diag\{j^{H_a - 0.5}\}.$$

Then we can write (A5.2.2) as:

$$-\sqrt{m}K_1\sum_{a=1}^{r}\lambda_a + \sqrt{m}\sum_{a=1}^{r}\lambda_a tr\left[\hat{G}(H)^{-1}\left(\frac{1}{m}\sum_{j=1}^{m}(\log j)\Lambda_j^{-1}\{i_a I(\lambda_j) + I(\lambda_j)i_a\}\Lambda_j^{-1}\right)\right]|_{H_0} \quad (\&)$$

Now, as

$$\hat{G}(H^0)^{-1} = G_0^{-1} + op(1),$$

(this is due to:

$$\|\hat{G}(H^0) - G_0\| = O_p[\max_{a,b}|\frac{1}{m}\sum_{j=1}^{m}(\lambda_j^{H_a^0 + H_b^0 - 1}I_{ab}(\lambda_j) - g_{ab})|] =$$

$$= O_p(\left(\frac{m}{n}\right)^{\beta} + m^{-2/3}(\log m)^{2/3} + \frac{\log m}{m} + m^{-1/2}n^{-1/4} + \frac{1}{\sqrt{m}}) = o_p(1)$$

using (A5.3.1), (A5.3.2) and (A5.3.3) proved in appendix 5.3), so (&) will be asymptotically equivalent to

$$-2K_1\sqrt{m}\sum_{a=1}^{r}\lambda_a+\sqrt{m}\sum_{a=1}^{r}\lambda_a tr(G_0^{-1}\hat{G}^{(a)})$$

where $\hat{G}^{(a)}=\hat{G}(H^0)i_a+i_a\hat{G}(H^0)$

Then we have:

$$tr[G_0^{-1}\hat{G}^{(a)}]=2Re\left(\sum_{k=1}^{r}g^{ak}\frac{1}{m}\sum_{j=1}^{m}(\log j)\lambda_j^{H_k^0+H_a^0-1}I_{ka}(\lambda_j)\right)$$

and so

$$\sqrt{m}\lambda'\frac{dR(H)}{dH}\Big|_{H_0}=\frac{2}{\sqrt{m}}\sum_{a=1}^{r}\lambda_a\sum_{j=1}^{m}\nu_j[\psi_a(\lambda_j)-1](1+op(1)) \qquad \text{(A5.2.4)}$$

where

$$\psi_a(\lambda_j)=Re\sum_{k=1}^{r}g^{ak}\lambda_j^{H_k^0+H_a^0-1}I_{ka}(\lambda_j)=Re(g^a\Lambda_j^{0^{-1}}I_{a}(\lambda_j)\lambda_j^{H_a^0-\frac{1}{2}})$$

and

$$\nu_j=logj-\frac{1}{m}\sum_{j=1}^{m}logj \qquad \text{(A5.2.5)}$$

where g^a is the a-th row of the inverse of G_0 and $I_{.a}(\lambda_j)$ is the a-th column of $I(\lambda_j)$; then (A5.2.4) is asymptotically equivalent to:

$$\sum_{a=1}^{r}\lambda_a\frac{2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j[Re[g^a\Lambda_j^{0-1}I_a(\lambda_j)\lambda_j^{H_a^0-\frac{1}{2}}]-Re[g^a\Lambda_j^{0^{-1}}A(\lambda_j)JA_a^*(\lambda_j)\lambda_j^{H_a^0-\frac{1}{2}}]] \qquad (1)$$

$$+\sum_{a=1}^{r}\lambda_a\frac{2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j[Re[g^a\Lambda_j^{0^{-1}}A(\lambda_j)JA_a^*(\lambda_j)\lambda_j^{H_a^0-\frac{1}{2}}]-1] \qquad (2)$$

now (1) is negligible because:

$$(1)=\sum_{a=1}^{r}\lambda_a\frac{2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j[Re[g^{a}\Lambda_j^{0^{-1}}(I_a(\lambda_j)-A(\lambda_j)JA_a^*(\lambda_j))\lambda_j^{H_a^0-\frac{1}{2}}]]=$$

$$=\sum_{a=1}^{r}\lambda_a\frac{2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j[\sum_{b=1}^{r}Re[g^{ba}\lambda_j^{H_b^0-\frac{1}{2}}(I_{ba}(\lambda_j)-A_b(\lambda_j)JA_a^*(\lambda_j))\lambda_j^{H_a^0-\frac{1}{2}}]]=o(1)$$

using (A5.3.2) proved in appendix 5.3 and summation by parts; and we have (2)=

$$\sum_{a=1}^{r}\lambda_a\frac{2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\left[Re\left[g^{a}\Lambda_j^{0^{-1}}A(\lambda_j)\left[\frac{1}{2\pi n}\sum_{t=1}^{n}\sum_{s=1}^{n}e_te_s'e^{i(t-s)}\lambda_j\right]A_a^*(\lambda_j)\lambda_j^{H_a^0-\frac{1}{2}}\right]-1\right]=$$

$$=\sum_{a=1}^{r}\lambda_a\frac{2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\left[Re\left[g^{a}\Lambda_j^{0^{-1}}A(\lambda_j)\left(\frac{1}{2\pi n}\sum_{t=1}^{n}e_te_t'\right)A_a^*(\lambda_j)\lambda_j^{H_a^0-\frac{1}{2}}\right]-1\right]+$$

$$+\sum_{a=1}^{r}\lambda_a\frac{2}{\sqrt{m}}\sum_{j=1}^{m}\nu_jRe\left[g^{a}\Lambda_j^{0^{-1}}A(\lambda_j)\left(\frac{1}{2\pi n}\sum_{t\neq}\sum_{s}e_te_s'e^{i(t-s)}\lambda_j\right)A_a^*(\lambda_j)\lambda_j^{H_a^0-\frac{1}{2}}\right]$$

$$(w1) \quad + \quad (w2)$$

(w1) is negligible by lemma 2 in appendix 5.5, and (w2) is

$$\sum_{a=1}^{r}\lambda_a\frac{2}{\sqrt{m}}\sum_{j=1}^{m}\nu_jRe\left[g^{a}\Lambda_j^{0^{-1}}A(\lambda_j)\left(\frac{1}{2\pi n}\sum_{t\neq}\sum_{s}e_te_s'e^{i(t-s)\lambda_j}\right)A_a^*(\lambda_j)\lambda_j^{H_a^0-\frac{1}{2}}\right]=$$

$$\sum_{t=1}^{n}e_t'\sum_{s=1}^{t-1}\Gamma_{t-s}^{m}e_s\ ,$$

where

$$\Gamma_{t-s}^{m}=\frac{1}{\pi\sqrt{mn}}\sum_{j=1}^{m}\nu_j\Omega_j\cos(t-s)\lambda_j\quad ,$$

$$\Omega_j=\sum_{a=1}^{r}\lambda_aRe\left[A(\lambda_j)'\Lambda_j^{0^{-1}}g^{a\prime}\overline{A_a(\lambda_j)}+A_a'(\lambda_j)\overline{g^{a}\Lambda_j^{0^{-1}}A(\lambda_j)}\right]\lambda_j^{H_a^0-\frac{1}{2}}\ ,$$

so,

$$(*)=\sum_{t=1}^{n} e_t'\sum_{s=1}^{t-1}\Gamma_{t-s}^{m}e_s=\sum_{t=1}^{n} z_t \quad where \quad z_t=e_t'\sum_{s=1}^{t-1}\Gamma_{t-s}^{m}e_s$$

(for simplicity let's omit the superindex m of Γ from now on).

Then z_t is a martingale difference and we are going to apply a central limit theorem for a martingale difference (chapter 11, Solo (1986) based on Brown (1971)), that states that:

" If z_t is a zero-mean martingale difference array and:

(i) $\sum_{t=1}^{n} E(z_t^2|F_{t-1})-1\to_p 0,$

(ii) $\sum_{t=1}^{n} E(z_t^2 I(|z_t|>\delta))\to 0$ for all $\delta>0$,

then: $\sum_{t=1}^{n} z_t\to_d N(0,1)$".

First we prove (i): i.e., $\sum_{t=1}^{n} E(z_t^2|F_{t-1})-\sum_{a=1}^{r}\sum_{b=1}^{r}\lambda_a\lambda_b E_{ab}\to_p 0$

$$\sum_{t=1}^{n} E(z_t^2|F_{t-1})=\sum_{t=1}^{n} E(\sum_{s=1}^{t-1} e_s'\Gamma_{t-s}^{*}e_t e_t'\sum_{s'=1}^{t-1}\Gamma_{t-s'}e_{s'}|F_{t-1})=$$

$$\sum_{t=1}^{n}\sum_{s=1}^{t-1} e_s'\Gamma_{t-s}^{*}R\sum_{s'=1}^{t-1}\Gamma_{t-s'}e_{s'}=$$

$$\sum_{t=1}^{n}\sum_{s=1}^{t-1} e_s'\Gamma_{t-s}^{*}R\Gamma_{t-s}e_s+\sum_{t=1}^{n}\sum_{s\neq}\sum_{s'} e_s'\Gamma_{t-s}^{*}R\Gamma_{t-s'}e_{s'}=(l1)+(l2)$$

(l2) is negligible by lemma 5 in appendix 5.5 and we have:

$$\sum_{t=1}^{n}\sum_{s=1}^{t-1} e_s'\Gamma_{t-s}^{*}R\Gamma_{t-s}e_s=\sum_{t=1}^{n}\sum_{s=1}^{t-1} tr(\Gamma_{t-s}^{*}R\Gamma_{t-s}e_s e_s')=$$

and by lemma 4 in appendix 5.5 this is equivalent to:

$$\sum_{t=1}^{n}\sum_{s=1}^{t-1} tr(\Gamma_{t-s}^{*}R\Gamma_{t-s}R)=$$

$$\sum_{t=1}^{n}\sum_{s=1}^{t-1} tr\left[\left(\frac{1}{\pi\sqrt{mn}}\right)^2\sum_{j=1}^{m}\nu_j\Omega_j^{*}\cos(t-s)\lambda_j R\sum_{j'=1}^{m}\nu_{j'}\Omega_{j'}\cos(t-s)\lambda_{j'}R\right]=$$

$$\sum_{t=1}^{n}\sum_{s=1}^{t-1}\left(\frac{2}{\sqrt{mn}}\right)^2\sum_{j=1}^{m} v_j^2 tr\left[\Omega_j^* \cos(t-s)\lambda_j \frac{R}{2\pi}\Omega_j \cos(t-s)\lambda_j \frac{R}{2\pi}\right]+ \qquad (v1)$$

$$\sum_{t=1}^{n}\sum_{s=1}^{t-1} tr\left[\left(\frac{1}{\pi\sqrt{mn}}\right)^2\sum_{j\neq}^{m}\sum_{j'}^{m} v_j\Omega_j^* \cos(t-s)\lambda_j R v_{j'}\Omega_{j'}\cos(t-s)\lambda_{j'} R\right] \qquad (v2)$$

and by lemma 3 in appendix 5.5, (v2) is negligible.
As

$$tr\left[\Omega_j^* \frac{R}{2\pi}\Omega_j \frac{R}{2\pi}\right]=$$

$$tr\left[\sum_{a=1}^{r}\lambda_a Re\left[A_a'(\lambda_j)\lambda_j^{H_a^0-\frac{1}{2}} g^a \Lambda_j^{0^{-1}}\overline{A(\lambda_j)}+A'(\lambda_j)\Lambda_j^{0^{-1}} g^{a'}\lambda_j^{H_a^0-\frac{1}{2}}\overline{A_a(\lambda_j)}\right]\frac{R}{2\pi}\cdot\right.$$
$$\left.\sum_{a=1}^{r}\lambda_a Re\left[A'(\lambda_j)\Lambda_j^{0^{-1}} g^{a'}\lambda_j^{H_a^0-\frac{1}{2}}\overline{A_a(\lambda_j)}+A_a'(\lambda_j)\lambda_j^{H_a^0-\frac{1}{2}} g^a\Lambda_j^{0^{-1}}\overline{A(\lambda_j)}\right]\frac{R}{2\pi}\right]=$$

$$tr\left[Re\left[\sum_{a=1}^{r}\lambda_a A_a'(\lambda_j)\lambda_j^{H_a^0-\frac{1}{2}} g^a\Lambda_j^{0^{-1}}\overline{A(\lambda_j)}\right]\frac{R}{2\pi} Re\left[\sum_{a=1}^{r}\lambda_a A'(\lambda_j)\Lambda_j^{0^{-1}} g^{a'}\lambda_j^{H_a^0-\frac{1}{2}}\overline{A_a(\lambda_j)}\frac{R}{2\pi}\right]\right]+$$

$$tr\left[Re\left[\sum_{a=1}^{r}\lambda_a A_a'(\lambda_j)\lambda_j^{H_a^0-\frac{1}{2}} g^a\Lambda_j^{0^{-1}}\overline{A(\lambda_j)}\right]\frac{R}{2\pi} Re\left[\sum_{a=1}^{r}\lambda_a A_a'(\lambda_j)\lambda_j^{H_a^0-\frac{1}{2}} g^a\Lambda_j^{0^{-1}}\overline{A(\lambda_j)}\right]\frac{R}{2\pi}\right]+$$

$$tr\left[Re\left[\sum_{a=1}^{r}\lambda_a A'(\lambda_j)\Lambda_j^{0^{-1}} g^{a'}\lambda_j^{H_a^0-\frac{1}{2}}\overline{A_a(\lambda_j)}\right]\frac{R}{2\pi} Re\left[\sum_{a=1}^{r}\lambda_a A'(\lambda_j)\Lambda_j^{0^{-1}} g^{a'}\lambda_j^{H_a^0-\frac{1}{2}}\overline{A_a(\lambda_j)}\right]\frac{R}{2\pi}\right]+$$

$$tr\left[Re\left[\sum_{a=1}^{r}\lambda_a A'(\lambda_j)\Lambda_j^{0^{-1}} g^{a'}\lambda_j^{H_a^0-\frac{1}{2}}\overline{A_a(\lambda_j)}\right]\frac{R}{2\pi} Re\left[\sum_{a=1}^{r}\lambda_a A_a'(\lambda_j)\lambda_j^{H_a^0-\frac{1}{2}} g^a\Lambda_j^{0^{-1}}\overline{A(\lambda_j)}\right]\frac{R}{2\pi}\right]$$

$$=(s1)+(s2)+(s3)+(s4)$$

now, using C5.1' and the definition of f(λ), we get that (s1) and (s4) are asymptotically equivalent to $Re(\sum_{a=1}^{r}\sum_{b=1}^{r}\lambda_a\lambda_b g_{ab} g^{ba})$ and (s2) and (s3) to $\sum_{a=1}^{r}\lambda_a^2$, then using, see Robinson (1993a):

$$\sum_{t=1}^{n}\sum_{s=1}^{n-t}\cos^2(s\lambda_j)\sim\frac{(n-1)^2}{4}\quad and\quad \frac{1}{m}\sum_{j=1}^{m}v_j^2=1+O(\frac{(logm)^2}{m}),$$

we get that (v1) is asymptotically equivalent to: $\sum_{a=1}^{r}\sum_{b=1}^{r}\lambda_a\lambda_b E_{ab}$.

So we have proven (i), now we prove (ii):

$$\sum_{t=1}^{n}E(z_t^2 I(|z_t|>\delta)\to 0\quad for\quad any\quad \delta>0,$$

as

$$\sum_{t=1}^{n}E(z_t^2 I(|z_t|>\delta)\leq\sum_{t=1}^{n}\left\{\frac{Ez_t^4}{\delta^2}I(|z_t|>\delta)\right\}\leq\sum_{t=1}^{n}\left\{\frac{Ez_t^4}{\delta^2}\right\}=\frac{1}{\delta^2}\sum_{t=1}^{n}Ez_t^4 \tag{A5.2.6}$$

as δ is fixed then we just need to check the sufficient condition:

$$\sum_{t=1}^{n}E(z_t^4)\to 0.$$

As

$$\sum_{t=1}^{n}E(z_t^4)=\sum_{t=1}^{n}E(e_t'\sum_{s=1}^{t-1}\Gamma_{t-s}e_s)^4=$$

$$\sum_{t=1}^{n}E(\sum_{s=1}^{t-1}e_s'\Gamma_{t-s}'e_te_t'\sum_{r=1}^{t-1}\Gamma_{t-r}e_r\sum_{p=1}^{t-1}e_p'\Gamma_{t-p}'e_te_t'\sum_{q=1}^{t-1}\Gamma_{t-q}e_q)=$$

$$\sum_{t=1}^{n}E(\sum_{s=1}^{t-1}e_s'\Gamma_{t-s}'R\sum_{r=1}^{t-1}\Gamma_{t-r}e_r\sum_{p=1}^{t-1}e_p'\Gamma_{t-p}'R\sum_{q=1}^{t-1}\Gamma_{t-q}e_q)=(*)$$

now using

$$tr(ABCD)=vec'(C)(D\otimes B')vec(A'),$$

$$(*)=\sum_{t=1}^{n}tr(\sum_{s=1}^{t-1}\Gamma_{t-s}'R\Gamma_{t-s}R\Gamma_{t-s}'R\Gamma_{t-s})+\sum_{t=1}^{n}tr(\sum_{s=1}^{t-1}\Gamma_{t-s}'R\sum_{r=1}^{t-1}\Gamma_{t-r}R\Gamma_{t-r}'R\Gamma_{t-s})=$$

$$(1)+(2).$$

Now

$$(1)=O(\sum_{t=1}^{n}(\sum_{s=1}^{t-1}\|\Gamma_{t-s}\|^4))=O(n(\sum_{t=1}^{n}\|\Gamma_t\|^2)^2)=O(\frac{(logm)^4}{n})$$

$$(2)=\sum_{t=1}^{n}tr(\sum_{s=1}^{t-1}\sum_{r=1}^{t-1}\Gamma_{t-s}'R\Gamma_{t-r}R\Gamma_{t-r}'R\Gamma_{t-s})=\sum_{t=1}^{n}tr(\sum_{s=1}^{t-1}\Gamma_{t-s}'R\Gamma_{t-s}R\sum_{r=1}^{t-1}\Gamma_{t-r}'R\Gamma_{t-r}R)$$

$$=O(\sum_{t=1}^{n}(\sum_{s=1}^{t-1}\|\Gamma_{t-s}|^2)^2)=O(\frac{(logm)^4}{n})$$

using the bounds derived in lemma 4 in appendix 5.5.

Appendix 5.3

In this appendix we analyze the stochastic order of magnitude of:

$$\sum_{j=1}^{s}\left(I_{ab}(\lambda_j)\lambda_j^{H_a^0+H_b^0-1}-g_{ab}\right)=$$

$$\sum_{j=1}^{s}\left(I_{ab}(\lambda_j)-f_{ab}(\lambda_j)\right)\lambda_j^{H_a^0+H_b^0-1}+ \tag{A}$$

$$+\sum_{j=1}^{s}\left(\lambda_j^{H_a^0+H_b^0-1}f_{ab}(\lambda_j)-g_{ab}\right). \tag{B}$$

First, by assumption C5.1, immediately:

$$(B)=\sum_{j=1}^{s}\left(\lambda_j^{H_a^0+H_b^0-1}f_{ab}(\lambda_j)-g_{ab}\right)=O\left(\frac{s^{\beta+1}}{n^{\beta}}\right) \tag{A5.3.1}$$

while,

$$(A)=\sum_{j=1}^{s}\left(I_{ab}(\lambda_j)-A_a(\lambda_j)JA_b^*(\lambda_j)\right)\lambda_j^{H_a^0+H_b^0-1} \tag{x0}$$

$$+\sum_{j=1}^{s}\left(A_a(\lambda_j)JA_b^*(\lambda_j)-f_{ab}(\lambda_j)\right)\lambda_j^{H_a^0+H_b^0-1}\ , \tag{x1}$$

Then:

$$(x0)=O_p(s^{1/3}(logs)^{2/3}+logs+s^{1/2}n^{-1/4})\ , \tag{A5.3.2}$$

$$(x1)=O_p(s^{1/2}). \tag{A5.3.3}$$

Proof. First, we analyze (x0):

$$(x0)=\sum_{j=1}^{s}\left(I_{ab}(\lambda_j)-A_a(\lambda_j)JA_b^*(\lambda_j)\right)\lambda_j^{H_a^0+H_b^0-1}=$$

$$\sum_{j=1}^{l}\left(I_{ab}(\lambda_j)-A_a(\lambda_j)JA_b^*(\lambda_j)\right)\lambda_j^{H_a^0+H_b^0-1}+ \qquad (x00)$$

$$+\sum_{j=l+1}^{s}\left(I_{ab}(\lambda_j)-A_a(\lambda_j)JA_b^*(\lambda_j)\right)\lambda_j^{H_a^0+H_b^0-1} \qquad (x01)$$

In order to see the order of magnitude of (x0) we are going to employ repeatedly theorem 2 of Robinson (1992) that state that: "under conditions:

a) $f_{aa}(\lambda)=K_a\lambda^{1-2H_a}+O(\lambda^{1-2H_a+\alpha})$, $as\ \lambda\to 0^+$, $K_a\in(0,\infty)$, $H_a\in(0,1)$ $\alpha\in(0,2]$ for a=1,...r,

b) $|\frac{d}{d\lambda}\log f_{ab}(\lambda)|=O(\lambda^{-H_a-H_b})$, $as\ \lambda\to 0^+$, $a,b=1,...r$, and

c) for some $\beta\in[0,2]$ $|R_{ab}(\lambda)-R_{ab}(0)|=O(\lambda^{\beta})$ $as\ \lambda\to 0^+$, $a<b=2,...r$, where R_{ab} is the coherency between x_a and x_b; then: for any sequence of positive integers j=j(n) such that j/n $\neq 0$ as n→ ∞, for a,b=1,...r,

$$E\{v_a(\lambda_j)\overline{v_b(\lambda_j)}\}=R_{ab}(0)+O[\frac{\log j}{j}+(\frac{j}{n})^{min(\alpha,\beta)}]\ as\ n\to\infty, \qquad (A5.T5)$$

$$E\{v_a(\lambda_j)v_b(\lambda_j)\}=O[\frac{\log j}{j}]\ as\ n\to\infty, \qquad (A5.T6)$$

and for any sequences j=j(n), k=k(n) such that j>k and j/n →0 as n→ ∞,

$$E\{v_a(\lambda_j)\overline{v_b(\lambda_k)}\}=O[\frac{\log j}{k}]\ as\ n\to\infty, \qquad (A5.T7)$$

$$E\{v_a(\lambda_j)v_b(\lambda_k)\}=O[\frac{\log j}{k}]\ as\ n\to\infty, \qquad (A5.T8)$$

for a,b=1,...r."

To analyze the order of magnitude of (x00) we employ repeatedly (A5.T5), so:

$$(x00)=O_p(\sum_{j=1}^{l}E|I_{ab}(\lambda_j)\lambda_j^{H_a^0+H_b^0-1}|+\sum_{j=1}^{l}E|A_a(\lambda_j)JA_b^*(\lambda_j)\lambda_j^{H_a^0+H_b^0-1}|),$$

first,

$$E|I_{ab}(\lambda_j)\lambda_j^{H_a^0+H_b^0-1}|=O(E|w_a(\lambda_j)\lambda_j^{H_a^0-0.5}\overline{w}_b(\lambda_j)\lambda_j^{H_b^0-0.5}|)=O(1)$$

also, as

$$f_{aa}(\lambda)=\sum_{k=1}^{r}|A_a^k(\lambda)|^2 r_{kk}+\sum_{k\neq}^{r}\sum_{l}^{r}A_a^k(\lambda)\overline{A_a^l(\lambda)}r_{kl}=C\|A_a(\lambda)\|^2+o(\|A_a(\lambda)\|^2),$$

$$where\ \|A_a(\lambda)\|^2=\max_i|A_a^i(\lambda)|^2\ and\ C>0,$$

then $\|A_a(\lambda)\|=O(f_{aa}(\lambda)^{1/2})=O(\lambda^{1/2-H_a})$, (A5.3.4),

then

$$E|A_a(\lambda_j)vv^*A_b^*(\lambda_j)\lambda_j^{H_a^0+H_b^0-1}|=O(1),$$

so (x00) = O(l).

Now we analyze (x01):

$$(x01)=\sum_{j=l+1}^{s}\left(I_{ab}(\lambda_j)-A_a(\lambda_j)JA_b^*(\lambda_j)\right)\lambda_j^{H_a^0+H_b^0-1}=$$

$$=\sum_{j=l+1}^{s}\left(\tilde{w}_a(\lambda_j)\tilde{w}_b^*(\lambda_j)-\tilde{A}_a(\lambda_j)v(\lambda_j)\tilde{v}(\lambda_j)\tilde{A}_b^*(\lambda_j)\right)$$

$$where\quad \tilde{w}_a(\lambda_j)=w_a(\lambda_j)\lambda_j^{H_a^0-1/2}\quad \tilde{A}_a(\lambda_j)=A_a(\lambda_j)\lambda_j^{H_a^0-1/2}$$

$$\tilde{w}_b(\lambda_j)=w_b(\lambda_j)\lambda_j^{H_b^0-1/2}\ and\ \tilde{A}_b(\lambda_j)=A_b(\lambda_j)\lambda_j^{H_b^0-1/2}$$

then, using that $x=O_p(\sqrt{Ex^2})$

$$(x01)=O_p\left(\left(E\left[\sum_{j=l+1}^{s}\left(\tilde{w}_a(\lambda_j)\tilde{w}_b^*(\lambda_j)-\tilde{A}_a(\lambda_j)v(\lambda_j)\tilde{v}(\lambda_j)\tilde{A}_b^*(\lambda_j)\right)\right]\cdot\right.\right.$$

$$\left.\left.\left[\sum_{j=l+1}^{s}\left(\tilde{w}_a(\lambda_j)\tilde{w}_b^*(\lambda_j)-\tilde{A}_a(\lambda_j)v(\lambda_j)\tilde{v}(\lambda_j)\tilde{A}_b^*(\lambda_j)\right)\right]^*\right)^{1/2}\right)$$

for simplicity lets drop the argument λ_j

$$(x01)=O_p\left(\left(E\left[\sum_{j=l+1}^{s}\left(\tilde{w}_a\tilde{w}_b^*-\tilde{A}_a vv^*\tilde{A}_b^*\right)\right]\left[\sum_{j=l+1}^{s}\left(\tilde{w}_a\tilde{w}_b^*-\tilde{A}_a vv^*\tilde{A}_b^*\right)\right]^*\right)^{1/2}\right)=O_p((XX)^{1/2}),$$

where

$$XX=\sum_{j=l+1}^{s}E\left(\left(\tilde{w}_a\tilde{w}_b^*-\tilde{A}_a vv^*\tilde{A}_b^*\right)\left(\tilde{w}_a\tilde{w}_b^*-\tilde{A}_a vv^*\tilde{A}_b^*\right)^*\right)+\qquad\text{(XXA)}$$

$$+\sum_{j\neq}^{s}\sum_{k}^{s} E\left(\left(\tilde{w}_a\tilde{w}_b^* - \tilde{A}_a vv^*\tilde{A}_b^*\right)\left(\tilde{w}_a\tilde{w}_b^* - \tilde{A}_a vv^*\tilde{A}_b^*\right)^*\right) \;; \qquad \text{(XXB)}$$

first we analyze (XXA):

$$XXA = \sum_{j=l+1}^{s} E\left(\tilde{w}_a\tilde{w}_b^*\tilde{w}_b\tilde{w}_a^*\right) - \sum_{j=l+1}^{s} E\left(\tilde{w}_a\tilde{w}_b^*\tilde{A}_b vv^*\tilde{A}_a^*\right) -$$
$$-\sum_{j=l+1}^{s} E\left(\tilde{A}_a vv^*\tilde{A}_b^*\tilde{w}_b\tilde{w}_a^*\right) + \sum_{j=l+1}^{s} E\left(\tilde{A}_a vv^*\tilde{A}_b^*\tilde{A}_b vv^*\tilde{A}_a^*\right)$$

now we use that for zero mean variates:

$$E(wxyz) = E(wx)E(yz) + E(wy)E(xz) + E(wz)E(xy) + cum(w,x,y,z),$$

then we can decompose:

$$XXA = (*PAR1) + (*CUM1)$$

where

$$(*PAR1) = \sum_{j=l+1}^{s}\Big(\Big[E(\tilde{w}_a\tilde{w}_b)E(\tilde{w}_b^*\tilde{w}_a^*) +$$
$$[E\tilde{w}_a\tilde{w}_b^* - \tilde{A}_a R\tilde{A}_b^*][E\tilde{w}_b\tilde{w}_a^* - \tilde{A}_b R\tilde{A}_a^*]) + [E\tilde{w}_a\tilde{w}_a^* - \tilde{A}_a R\tilde{A}_a^*][E\tilde{w}_b\tilde{w}_b^* - \tilde{A}_b\tilde{A}_b^*] -$$
$$-E(\tilde{w}_a v'\tilde{A}_b')E(\tilde{w}_b^* v^*\tilde{A}_a^*) - (Ew_a\tilde{w}_b^* - \tilde{A}_a R\tilde{A}_b^*)(E\tilde{A}_b vv^*\tilde{A}_a^* - A_b R\tilde{A}_a^*)$$
$$-(E\tilde{w}_a v^*\tilde{A}_a^* - \tilde{A}_a \Delta\tilde{A}_a^*)(E\tilde{A}_b v\tilde{w}_b^* - \tilde{A}_b R\tilde{A}_b^*) - (E\tilde{A}_a vv^*\tilde{A}_b^* - \tilde{A}_a R\tilde{A}_b^*)(E\tilde{w}_b w_a^* - \tilde{A}_b R\tilde{A}_a^*)$$
$$+(E\tilde{A}_a vv^*\tilde{A}_b^* - \tilde{A}_a R\tilde{A}_b^*)(E\tilde{A}_b vv^*\tilde{A}_a^* - \tilde{A}_b R\tilde{A}_a^*) + (E\tilde{A}_a vv^*\tilde{A}_a^* - \tilde{A}_a R\tilde{A}_a^*)(E\tilde{A}_b vv^*\tilde{A}_b^* - \tilde{A}_b R\tilde{A}_b^*)$$
$$-[E(\tilde{A}_a v\tilde{w}_b)E(\tilde{w}_a^* v^*\tilde{A}_b)] - [E\tilde{w}_b v^*\tilde{A}_b^* - \tilde{A}_b R\tilde{A}_b^*)(E\tilde{A}_a v\tilde{w}_a^* - \tilde{A}_a R\tilde{A}_a^*)]$$
$$+[E(\tilde{A}_a v\tilde{A}_b v)E(v^*\tilde{A}_b^* v^*\tilde{A}_a^*] \quad \Big) \;;$$

and

$$(*CUM1) = \sum_{j=l+1}^{s}\Big[cum(\tilde{w}_a,\tilde{w}_b^*,\tilde{w}_b,\tilde{w}_a^*) - cum(\tilde{w}_a,\tilde{w}_b^*,\tilde{A}_b v, v^*\tilde{A}_a^*)$$
$$-cum(\tilde{A}_a v, v^*\tilde{A}_b^*,\tilde{w}_b,w_a^*) + cum(\tilde{A}_a v, v^*\tilde{A}_b^*,\tilde{A}_b v, v^*\tilde{A}_a^*)\Big]$$

(*CUM1) will be analyzed later, using (A5.T5) and (A5.T6) for I,J=a,b:

$$Evv^* = \frac{R}{2\pi},\quad E(\tilde{w}_I\tilde{w}_J^*) - \tilde{A}_I\frac{R}{2\pi}\tilde{A}_J^* = O\left(\frac{\log j}{j}\right),\quad E(\tilde{w}_I\tilde{w}_J) = O\left(\frac{\log j}{j}\right)$$

and

$$E(\tilde{w}_I v^* \tilde{A}_J^*) - \tilde{A}_I \frac{R}{2\pi} \tilde{A}_J^* = O(\frac{logj}{j}),$$

$$E(\tilde{A}_I v \tilde{w}_J) = O(\frac{logj}{j}), \quad E(\tilde{A}_I v v^* \tilde{A}_J) - \tilde{A}_I \frac{R}{2\pi} \tilde{A}_J^* = O(\frac{logj}{j}),$$

$$E(\tilde{A}_I v \tilde{A}_J v) = O(\frac{logj}{j});$$

so:

$$(*PAR1) = O(\sum_{j=l+1}^{s} (\frac{logj}{j})^2) = O(\frac{(logs)^2}{l}).$$

Analogously:

$$XXB = (*PAR2) + (*CUM2)$$

where

$$(*PAR2) = \sum_{j*}^{s} \sum_{k} \Big([E(\tilde{w}_a^j \tilde{w}_b^k) E(\tilde{w}_b^{*j} \tilde{w}_a^{*k}) +$$

$$[E\tilde{w}_a^j \tilde{w}_b^{*j} - \tilde{A}_a^j R \tilde{A}_b^{*j}][E\tilde{w}_b^k \tilde{w}_a^{*k} - \tilde{A}_b^k R \tilde{A}_a^{*k}]) + ([E\tilde{w}_a^j \tilde{w}_a^{*k}][E\tilde{w}_b^k \tilde{w}_b^{*j}]) -$$

$$-E(\tilde{w}_a^j v^{k'} \tilde{A}_b^{k'} E\tilde{w}_b^{*j} v^{*k} \tilde{A}_a^k) - (Ew_a^j \tilde{w}_b^{*j} - \tilde{A}_a^j \tilde{A}_b^{*j})(E\tilde{A}_b^k v^k v^{*k} \tilde{A}_a^{*k} - A_b^k R \tilde{A}_a^{*k})$$

$$-(E\tilde{w}_a^j v^{*k} \tilde{A}_a^{*k})(E\tilde{A}_b^k v^k \tilde{w}_b^{*j}) - (E\tilde{A}_a^j v^j v^{*j} \tilde{A}_b^{*j} - \tilde{A}_a^j R \tilde{A}_b^{*j})(E\tilde{w}_b^k w_a^{*k} - \tilde{A}_b^k R \tilde{A}_a^{*k})$$

$$+(E\tilde{A}_a^j v^j v^{*j} \tilde{A}_b^{*j} - \tilde{A}_a^j R \tilde{A}_b^{*j})(E\tilde{A}_b^k v^k v^{*k} \tilde{A}_a^{*k} - \tilde{A}_b^k R \tilde{A}_a^{*k}) + (E\tilde{A}_a^j v^j v^{*k} \tilde{A}_a^{*k})(E\tilde{A}_b^k v^k v^{*j} \tilde{A}_b^{*j})$$

$$-[E(\tilde{A}_a^j v^j \tilde{w}_b^k) E(w_a^{*k} v^{*j} \tilde{A}_b^{*j})] - [E\tilde{w}_b^k v^{*j} \tilde{A}_b^{*j})(E\tilde{A}_a^j v^j \tilde{w}_a^{*k})]$$

$$+[E(\tilde{A}_a^j v^j \tilde{A}_b^k v^k) E(v^{*j} \tilde{A}_b^{*j} v^{*k} \tilde{A}_a^{*k}] + [E(\tilde{A}_a^j v^j \tilde{A}_b^k v^k) E(v^{*j} \tilde{A}_b^{*j} v^{*k} \tilde{A}_a^{*k}]\Big) \; ;$$

and

$$(*CUM2) = \sum_{j*}^{s} \sum_{k} \Big(cum(\tilde{w}_a^j, \tilde{w}_b^{*j}, \tilde{w}_b^k, \tilde{w}_a^{*k}) - cum(\tilde{w}_a^j, \tilde{w}_b^{*j}, \tilde{A}_b^k v^k, v^{*k} \tilde{A}_a^{*k}) -$$

$$-cum(\tilde{A}_a^j v^j, v^{*j} \tilde{A}_b^{*j}, \tilde{w}_b^k, \tilde{w}_a^{*k}) + cum(\tilde{A}_a^j v^j, v^{*j} \tilde{A}_b^{*j}, \tilde{A}_b^k v^k, v^{*k} \tilde{A}_a^{*k}) \Big).$$

We analyze first (*PAR2) using (A5.T5), (A5.T6), (A5.T7) and (A5.T8):

$$E(\tilde{w}_I^l \tilde{w}_J^{*l}) - \tilde{A}_I^l \frac{R}{2\pi} \tilde{A}_J^l = O(\frac{logl}{l}) \text{ and}$$

$$E(\tilde{A}_I^l v v^* A_J^{*l} - \tilde{A}_I^l \frac{R}{2\pi} \tilde{A}_J^{*l}) = O(\frac{logl}{l}) \quad for \quad I,J=a,b \quad and \quad l=j,k \; ;$$

$$also\ E(\tilde{w}_I^l \tilde{w}_J^{*m}) = O(\frac{logl}{m})\ ,\ E(\tilde{w}_I^l \tilde{w}_J^m) = O(\frac{logl}{m}),$$

$$E(\tilde{A}_I^l v \tilde{w}_J^m) = O(\frac{logl}{m})\ ,\ E(\tilde{w}_I^l v^* \tilde{A}_J^{*m}) = O(\frac{logl}{m})\ ,\ E(\tilde{A}_I^l v v^* \tilde{A}_J^{*m}) = O(\frac{logl}{m}),$$

$$E(\tilde{A}_I^l v \tilde{A}_J^m v) = O(\frac{logl}{m})\ \ for\ I,J=a,b,\ \ and\ \ l,m=j,k,\ \ l>m.$$

so (*PAR2) is:

$$O\left(\sum_{j<} \sum_{k} \frac{(logk)^2}{j^2}\right) = O\left(\frac{s(logs)^2}{l}\right)$$

now choosing $l \sim s^{1/3}(logs)^{2/3}$ we get:

(*PAR2) $= O(s^{2/3}(logs)^{4/3})$ and

(*PAR1) $= O((logs)^{4/3}s^{-1/3})$.

Now we analyze the part with the cumulants, we examine first (*CUM2) =

$$cum(\tilde{w}_a^j, \tilde{w}_b^{*j}, \tilde{w}_b^k, \tilde{w}_a^{*k}) - cum(\tilde{w}_a^j, \tilde{w}_b^{*j}, \tilde{A}_b^k v^k, v^{*k} \tilde{A}_a^{*k}) -$$

$$- cum(\tilde{A}_a^j v^j, v^{*j} \tilde{A}_b^{*j}, \tilde{w}_b^k, \tilde{w}_a^{*k}) + cum(\tilde{A}_a^j v^j, v^{*j} \tilde{A}_b^{*j}, \tilde{A}_b^k v^k, v^{*k} \tilde{A}_a^{*k}) =$$

$$\lambda_j^{H_a^0 - H_b^0 - 1}\Big(cum(w_a^j, w_b^{*j}, w_b^k, w_a^{*k}) - cum(w_a^j, w_b^{*j}, A_b^k v^k, v^{*k} A_a^{*k}) -$$

$$- cum(A_a^j v^j, v^{*j} A_b^{*j}, w_b^k, w_a^{*k}) + cum(A_a^j v^j, v^{*j} A_b^{*j}, A_b^k v^k, v^{*k} A_a^{*k})\Big), \qquad (*)$$

using well known properties of cumulants (see for instance Brillinger , p19.,p.26,p.39) we get:

$$(*) = \sum_{\substack{j \neq \\ l+1}}^{s} \sum_{k}^{s} \lambda_j^{H_a^0 + H_b^0 - 1} \lambda_k^{H_a^0 + H_b^0 - 1} \frac{1}{(2\pi)^3} \sum_{k_1}^{r} \sum_{k_2}^{r} \sum_{k_3}^{r} \sum_{k_4}^{r} \frac{\kappa_{k_1 k_2 k_3 k_4}}{(2\pi n)^2} \cdot$$

$$\iiint_{-\pi}^{\pi} \left[A_a^{k_1}(\lambda_j + \lambda + \mu + \zeta) \overline{A_b^{k_2}(\lambda_j - \mu)} - A_a^{k_1}(\lambda_j) \overline{A_b^{k_2}(\lambda_j)}\right] \cdot$$

$$\left[A_b^{k_3}(\lambda_k - \lambda) \overline{A_a^{k_4}(\lambda_k - \zeta)} - A_b^{k_3}(\lambda_k) \overline{A_a^{k_4}(\lambda_k)}\right] E_{jk}(\lambda, \mu, \zeta) d\lambda d\mu d\zeta\ , \qquad (**)$$

where

$$A_a^{k_1}(\lambda) \textit{ is the } k_1 \textit{ element of } A_a(\lambda),$$

and

$$E_{jk}(\lambda,\mu,\zeta)=D(\lambda_j-\lambda-\mu-\zeta)D(\lambda_k+\lambda)D(\mu-\lambda_j)D(\zeta-\lambda_k)\ , \qquad D(\lambda)=\sum_{t=1}^{n} e^{it\lambda}.$$

As r is finite and

$$\kappa_{k_1k_2k_3k_4}<\infty \quad for \quad all \quad k_1,k_2,k_3,k_4=1,\ldots r,$$

and using:

$$(x_1x_2-y_1y_2)(x_3x_4-y_3y_4)=\prod_{i=1}^{4}(x_i-y_i)+\sum_{i=1}^{4}\prod_{\substack{j=1\\ j\neq i}}^{4}(x_j-y_j)y_i+$$
$$\sum_{i=1}^{2}\sum_{j=1}^{2}(x_i-y_i)(x_{j+2}-y_{j+2})y_{3-i}y_{5-j}\ ;$$

so, basically we have 3 types of summands in (**); the first is typified by:

$$\frac{1}{(2\pi)^3}\frac{\kappa_{k_1k_2k_3k_4}}{(2\pi n)^2}\int\int_{-\pi}^{\pi}\int\left[A_a^{k_1}(\lambda_j+\lambda+\mu+\zeta)-A_a^{k_1}(\lambda_j)\right]\left[\overline{A_b^{k_2}(\lambda_j-\mu)}-\overline{A_b^{k_2}(\lambda_j)}\right]\cdot \qquad (*)$$
$$\left[A_b^{k_3}(\lambda_k-\lambda)-A_b^{k_3}(\lambda_k)\right]\left[\overline{A_a^{k_4}(\lambda_k-\zeta)}-\overline{A_a^{k_4}(\lambda_k)}\right]E_{jk}(\lambda,\mu,\zeta)d\lambda d\mu d\zeta$$

using Cauchy-Schwartz inequality and periodicity:

$$(*)=O(P_j^{\frac{1}{2}}(a,k_1)\ P_j^{\frac{1}{2}}(b,k_2)\ P_k^{\frac{1}{2}}(a,k_3)\ P_k^{\frac{1}{2}}(b,k_4))$$

where

$$P_i(I,k_n)=\int_{-\pi}^{\pi}|A_I^{k_n}(\lambda_i+\lambda)-A_I^{k_n}(\lambda_i)|^2K(\lambda-\lambda_i)d\lambda\ ,$$

$$for \quad I=a,b\ ,\ n=1,2,3,4\ ,\ i=j,k \ where\ K(\lambda)=\frac{|D(\lambda)|^2}{2\pi n}\ ;$$

now as:

$$A_I(\lambda)=O(\lambda^{\frac{1}{2}-H_I^0}),$$

this implies, for at least one k_n:

$$A_I^{k_n}(\lambda)=O(\lambda^{\frac{1}{2}-H_I^0}),$$

and for the others:

$$A_I^{k_p}(\lambda)=o(\lambda^{\frac{1}{2}-H_I^0}),$$

then using lemma 3 of Robinson (1993a):

$$\int_{-\pi}^{\pi}|A_I^{k_n}(\lambda)-A_I^{k_n}(\lambda_i)|^2K(\lambda-\lambda_i)d\lambda=O(\frac{1}{i}\lambda_i^{1-2H_I^0})\ so$$

$$P_i(I,k_n)=O(\frac{1}{i}\lambda_i^{1-2H_I^0})\ for\ at\ least\ one\ k_n$$

$$and\ for\ the\ others:\ P_i(I,k_p)=o(\frac{1}{i}\lambda_i^{1-2H_I^0});$$

the second type of component is typified by:

$$\frac{1}{(2\pi)^3}\frac{\kappa_{k_1k_2k_3k_4}}{(2\pi n)^2}\int\int\int_{-\pi}^{\pi}\left[A_a^{k_1}(\lambda_j+\lambda+\mu+\zeta)-A_a^{k_1}(\lambda_j)\right]\overline{\left[A_b^{k_2}(\lambda_j-\mu)-A_b^{k_2}(\lambda_j)\right]}\cdot$$

$$\left[A_b^{k_3}(\lambda_k-\zeta)-A_b^{k_3}(\lambda_k)\right]\overline{A_a^{k_4}(\lambda_k)}E_{jk}(\lambda,\mu,\zeta)d\lambda d\mu d\zeta=(***)$$

as

$$A_a^{k_4}(\lambda_k)=O(\lambda_k^{\frac{1}{2}-H_a^0})$$

for at least one k_4 and for the others $=o(\lambda_k^{\frac{1}{2}-H_a^0})$, so (***) is:

$$=O(P_j^{\frac{1}{2}}(a,k_1)\ P_j^{\frac{1}{2}}(b,k_2)\ P_k^{\frac{1}{2}}(b,k_3)\ \lambda_k^{\frac{1}{2}-H_a^0}),$$

the third component is typified by:

$$\frac{1}{(2\pi)^3}\frac{\kappa_{k_1k_2k_3k_4}}{(2\pi n)^2}\int\int\int_{-\pi}^{\pi}\left[A_a^{k_1}(\lambda_j+\lambda+\mu+\zeta)-A_a^{k_1}(\lambda_j)\right]\left[\overline{A_b^{k_2}(\lambda_j-\lambda)}-\overline{A_b^{k_2}(\lambda_j)}\right]\cdot$$

$$A_b^{k_3}(\lambda_k)\overline{A_a^{k_4}(\lambda_k)}E_{jk}(\lambda,\mu,\zeta)d\lambda d\mu d\zeta=$$

$$\frac{1}{(2\pi)^3}\frac{\kappa_{k_1k_2k_3k_4}}{(2\pi n)^2}\int\int\int_{-\pi}^{\pi}\left[A_a^{k_1}(\lambda_j+\theta)-A_a^{k_1}(\lambda_j)\right]\left[\overline{A_b^{k_2}(\lambda_j-\lambda)}-\overline{A_b^{k_2}(\lambda_j)}\right]\cdot$$

$$A_b^{k_3}(\lambda_k)\overline{A_a^{k_4}(\lambda_k)}E_{jk}(\lambda,\theta-\lambda-\zeta,\zeta)d\lambda d\theta d\zeta=$$

$$\frac{1}{(2\pi)^2}\frac{\kappa_{k_1k_2k_3k_4}}{(2\pi n)^2}\int\int_{-\pi}^{\pi}\left[A_a^{k_1}(\lambda_j+\theta)-A_a^{k_1}(\lambda_j)\right]\left[\overline{A_b^{k_2}(\lambda_j)}-\overline{A_b^{k_2}(\lambda_j)}\right]\cdot$$

$$A_b^{k_3}(\lambda_k)\overline{A_a^{k_4}(\lambda_k)}D(\lambda_j-\theta)D(\lambda_k+\lambda)D(\theta-\lambda-\lambda_j-\lambda_k)d\lambda d\theta$$

that is :

$$=O(\frac{1}{\sqrt{n}}P_j^{\frac{1}{2}}(a,k_1)\ P_j^{\frac{1}{2}}(b,k_2)\ \lambda_k^{\frac{1}{2}-H_a^0}\ \lambda_k^{\frac{1}{2}-H_b^0}),$$

so

$$(*CUM2)=O(\sum_{\substack{j<\\ l+1}}^{s}\sum_{k}^{s}\lambda_k^{H_a^0+H_b^0-1}\lambda_j^{H_a^0+H_b^0-1}(j^{-1}k^{-1}\lambda_j^{1-H_a^0-H_b^0}\lambda_k^{1-H_a^0-H_b^0}+$$

$$j^{-\frac{1}{2}}k^{-1}\lambda_j^{1-H_a^0-H_b^0}\lambda_k^{1-H_a^0-H_b^0}+j^{-1}k^{-\frac{1}{2}}\lambda_j^{1-H_a^0-H_b^0}\lambda_k^{1-H_a^0-H_b^0}+$$

$$n^{-\frac{1}{2}}j^{-\frac{1}{2}}k^{-\frac{1}{2}}\lambda_j^{1-H_a^0-H_b^0}\lambda_k^{1-H_a^0-H_b^0}))=$$

$$O((logs)^2+s^{\frac{1}{2}}logs+sn^{-\frac{1}{2}})$$

and similarly,

$$(*CUM1)=O(\sum_{j=l+1}^{s}(j^{-2}+j^{-\frac{3}{2}}+n^{-\frac{1}{2}}j^{-1}))=O(1),$$

so we get that:

$$(x01)=O(s^{\frac{1}{3}}(logs)^{\frac{2}{3}}+s^{\frac{1}{2}}n^{-\frac{1}{4}}+(logs)).$$

Then:

$$(x0)=O(s^{1/3}(logs)^{2/3}+s^{1/2}n^{-1/4}+(logs))$$

and we have proven (A5.3.2). Now we analyze (X1):

$$(X1)=\sum_{j=1}^{s}\left(A_a(\lambda_j)JA_b^*(\lambda_j)-f_{ab}(\lambda_j)\right)\lambda_j^{H_a^0+H_b^0-1},$$

as

$$f_{ab}(\lambda_j)=A_a(\lambda_j)\left(\frac{R}{2\pi}\right)A_b^*(\lambda_j)\ ,$$

and

$$J=\frac{1}{2\pi n}\left(\begin{bmatrix}\sum_{t=1}^{n}e_t^{(1)2} & .. & \sum_t e_t^{(1)}e_t^{(r)}\\ ... & .. & ...\\ ... & .. & \sum_t e_t^{(r)2}\end{bmatrix}+\begin{bmatrix}\sum_{u\neq}\sum_t e_t^{(1)}e_u^{(1)} & .. & \sum_{u\neq}\sum_t e_t^{(1)}e_u^{(r)}\\ ... & .. & ...\\ ... & .. & \sum_{u\neq}\sum_t e_t^{(r)}e_u^{(r)}\end{bmatrix}e^{i(t-u)\lambda_j}\right)$$

we can rewrite it as:

$$J=\frac{1}{2\pi n}[S_1+S_2e^{i(t-u)\lambda_j}],$$

so

$$(X1)=\sum_{j=1}^{s}\left(\tilde{A}_a(\lambda_j)\frac{1}{2\pi}\left[(\frac{1}{n}S_1-R)+\frac{1}{n}S_2 e^{i(t-u)\lambda_j}\right]\tilde{A}_b^*(\lambda_j)\right)=$$

$$\frac{1}{2\pi}\sum_{j=1}^{s}\left(\tilde{A}_a(\lambda_j)\left[(\frac{1}{n}S_1-R)\right]\tilde{A}_b^*(\lambda_j)\right)+\frac{1}{2\pi}\sum_{j=1}^{s}\tilde{A}_a(\lambda_j)\frac{1}{n}S_2 e^{i(t-u)\lambda_j}\tilde{A}_b^*(\lambda_j)=$$

$$(A1)+(A2),$$

(A1) and (A2) have both zero mean while:

$$\|(A1)\|\le\sum_{j=1}^{s}\|\tilde{A}_a(\lambda_j)\|\|\frac{1}{n}S_1-R\|\|\tilde{A}_b^*(\lambda_j)\|,$$

as we have seen before:

$$\|\tilde{A}_a(\lambda_j)\|=O(1)\quad \|\tilde{A}_b^*(\lambda_j)\|=O(1)$$

and

$$\|\frac{1}{n}S_1-R\|=O_p\left(\sqrt{\max_{x,y}E(\frac{1}{n}\sum_{t=1}^{n}e_t^x e_t^y-r_{xy})^2}\right)$$

as

$$E(\frac{1}{n}\sum_{t=1}^{n}e_t^x e_t^y-r_{xy})^2=E(\frac{1}{n^2}\sum_{t=1}^{n}\sum_{s=1}^{n}e_t^x e_t^y e_t^x e_s^y+r_{xy}^2-2r_{xy}\frac{1}{n}\sum_{t=1}^{n}e_t^x e_t^y)$$

$$=\frac{1}{n^2}n(3+\kappa_{xxyy})+\frac{1}{n^2}(n^2-n)r_{xy}^2+r_{xy}^2-2r_{xy}^2=\frac{3+\kappa_{xxyy}-r_{xy}^2}{n}$$

so

$$\|\frac{1}{n}S_1-R\|=O_p\left(\sqrt{\max_{x,y}\frac{3+\kappa_{xxyy}-r_{xy}^2}{n}}\right)=O_p(\frac{1}{\sqrt{n}})\qquad\text{(A5.3.5)}$$

where we have used that $x=O_p(\sqrt{Ex^2})$, so

$$(A1)=O_p(\frac{s}{\sqrt{n}});$$

while

$$(A2)=\frac{1}{2\pi}\sum_{j=1}^{s}\tilde{A}_a(\lambda_j)\frac{1}{n}S_2 e^{i(t-u)\lambda_j}\tilde{A}_b^*(\lambda_j)=$$

$$\frac{1}{2\pi}\sum_{j=1}^{s}\tilde{A}_a(\lambda_j)\frac{1}{n}\sum_{u\neq}\sum_{t} e_t e_u' e^{i(t-u)\lambda_j}\tilde{A}_b^*(\lambda_j)$$

$$=\sum_{u\neq}\sum_{t} e_t'\Gamma_{t-u}e_u \ ,$$

where

$$\Gamma_{t-u}=\frac{1}{2\pi n}\sum_{j=1}^{s}\tilde{A}_a^*(\lambda_j)\tilde{A}_b(\lambda_j)e^{i(t-u)\lambda_j},$$

(A2) has zero mean and variance:

$$=E\left(\sum_{u\neq}\sum_{t}\sum_{u'\neq}\sum_{t'} e_t'\Gamma_{t-u}e_u e_u'\Gamma^*_{t'-u'}e_{t'}\right)=$$

$$=\sum_{u\neq}\sum_{t}\sum_{u'\neq}\sum_{t'} vec'\Gamma_{t-u}E\big((e_u\otimes e_t)(e_{t'}\otimes e_{u'})\big)vec\Gamma^*_{t'-u'} =$$

$$=\sum_{u\neq}\sum_{t}\sum_{u'\neq}\sum_{t'} vec'\Gamma_{t-u}E\big((e_u e_{t'}\otimes e_t e_{u'})\big)vec\Gamma^*_{t'-u'} =$$

$$=\sum_{u\neq}\sum_{t}\sum_{u'\neq}\sum_{t'} vec'\Gamma_{t-u}(R\otimes R)vec\Gamma^*_{t'-u'} \quad I(u=t',t=u')=$$

$$=\sum_{u\neq}\sum_{t} vec'\Gamma_{t-u}(R\otimes R)vec\Gamma^*_{t'-u'} \quad (*)$$

and calling

$$\Omega=R\otimes R \ ,$$

we get:

$$(*)=n\sum_{t=1}^{n-1}\left(1-\frac{t}{n}\right)(vec'(\Gamma_t+\Gamma_{-t}))\Omega(vec(\Gamma_t^*+\Gamma_{-t}^*))\leq n\sum_{t=1}^{n-1}(vec'(Z_t))\Omega(vec(Z_t^*))=(**)$$

where

$$Z_t=\Gamma_t+\Gamma_{-t} ,$$

now:

$$Z_t=\frac{2}{n}\sum_{j=1}^{s}\tilde{A}_a^*(\lambda_j)\tilde{A}_b(\lambda_j)\cos t\lambda_j ,$$

so

$$\|Z_t\|\le\frac{2}{n}\sum_{j=1}^{s}\|\tilde{A}_a^*(\lambda_j)\|\|\tilde{A}_b(\lambda_j)\|=O(\frac{s}{n}) ,$$

and, as we can write:

$$\tilde{A}_a^*(\lambda_j)=K_a(1+o(1)) \;\; \tilde{A}_b(\lambda_j)=K_b(1+o(1)) ,$$

$$Z_t=(1+o(1))\frac{2}{n}K_aK_b\sum_{j=1}^{s}\cos t\lambda_j ,$$

using, see Zygmund(1977), p.2,

$$\sum_{j=a}^{b}\cos j\lambda=O(\frac{1}{\lambda}) \quad for \quad 0<|\lambda|<\pi ,$$

we have

$$Z_t=O(\frac{1}{t}) ,$$

then, as $\Omega=O(1)$

$$(**)=O(n\sum_{t=1}^{n}\|Z_t\|^2)=O(n\sum_{t=1}^{[\frac{n}{s}]}\|Z_t\|^2+n\sum_{t=[\frac{n}{s}]}^{n}\|Z_t\|^2)=O(n\frac{n}{s}\frac{s^2}{n^2}+n\sum_{[\frac{n}{s}]}^{\infty}t^{-2})=O(s)$$

so that (X1) is $O_p(s^{1/2})$ and we have proven (A5.3.3).

Appendix 5.4

Lemma 1:

$$\frac{d^2R(\tilde{H})}{dH_a dH_b} \to_p E_{ab} \quad for \quad all \quad \tilde{H}:\|\tilde{H}-H_0\| \leq \|\hat{H}-H_0\| \ ,$$

Proof:

First define $G_{1b}{}^0 = G^0 i_b + i_b G^0$ and $G_{2ab}{}^0 = i_a i_b G^0 + i_a G^0 i_b + i_b G^0 i_a + G^0 i_a i_b$. Then,

$$R(H) = K_0(r-2i'H) + \log|\hat{G}(H)| \ ,$$

can be rewritten as:

$$R(H) = K_1(r-2i'H) + \log|\hat{E}(H)| \ ,$$

where K_1 and $\hat{E}(H)$ have been defined in Appendix 5.2, so,

$$\frac{d^2R(H)}{dH_a dH_b} = tr[-\hat{E}(H)^{-1}\frac{d\hat{E}(H)}{dH_b}\hat{E}(H)^{-1}\frac{d\hat{E}(H)}{dH_a} + \hat{E}(H)\frac{d^2\hat{E}(H)}{dH_a dH_b}]. \qquad (A5.4.1)$$

Now define:

$$N^{0^{-1}} = diag\{\left(\frac{2\pi}{n}\right)^{H_a^0 - 0.5}\}$$

and:

$$\hat{F}(H) = N^{0^{-1}}\hat{E}(H)N^{0^{-1}},$$

$$\hat{F}_{1b}(H) = N^{0^{-1}}\frac{d\hat{E}(H)}{dH_b}N^{0^{-1}},$$

$$\hat{F}_{2ab}(H) = N^{0^{-1}}\frac{d^2\hat{E}(H)}{dH_a dH_b}N^{0^{-1}}.$$

Then:

$$\frac{d^2R(H)}{dH_a dH_b} = tr[-\hat{F}(H)^{-1}\hat{F}_{1b}(H)\hat{F}(H)^{-1}\hat{F}_{1a}(H) + \hat{F}(H)^{-1}\hat{F}_{2ab}(H)].$$

As

$$\|\hat{F}_{1b}(H)-\hat{F}_{1b}(H^0))\|=O_p\left(\|\hat{F}_1(H)-\hat{F}_1(H^0)\|\right),$$

$$\|\hat{F}_{2ab}(H)-\hat{F}_{2ab}(H^0))\|=O_p\left(\|\hat{F}_2(H)-\hat{F}_2(H^0)\|\right),$$

$$\|\hat{F}_{1b}(H^0)-G^0_{1b}\|=O_p\left(\|\hat{F}_1(H^0)-G^0\frac{1}{m}\sum_{j=1}^{m}logj\|\right),$$

and

$$\|\hat{F}_{2ab}(H^0)-G^0_{2ab}\|=O_p\left(\|\hat{F}_2(H^0)-G^0\frac{1}{m}\sum_{j=1}^{m}(logj)^2\|\right),$$

where

$$\hat{F}_1(H)=\frac{1}{m}\sum_{j=1}^{m}logj\ N^{0^{-1}}J_j^{-1}I_jJ_j^{-1}N^{0^{-1}},$$

and

$$\hat{F}_2(H)=\frac{1}{m}\sum_{j=1}^{m}(logj)^2\ N^{0^{-1}}J_j^{-1}I_jJ_j^{-1}N^{0^{-1}}\ ,$$

and as

$$\frac{1}{m}\sum_{j=1}^{m}(logj)^2-(\frac{1}{m}\sum_{j=1}^{m}logj)^2\to 1,$$

and

$$tr[G_0^{-1}G_{1b}G_0^{-1}G_{1a}]=tr[G_0^{-1}G_{2ab}],$$

we just need to show for k=0,1,2:

P1) $\hat{F}_k(\tilde{H})=\hat{F}_k(H_0)+o_p(1)$, and

P2) $\hat{F}_k(H^0)=G_0(\frac{1}{m}\sum_{j=1}^{m}(\log j)^k)+o_p(1)$.

Proof of P1):

$$P(\|\hat{F}_k(\tilde{H})-\hat{F}_k(H_0)\|>\eta)=$$

$$P(\|\hat{F}_k(\tilde{H})-\hat{F}_k(H_0)\|>\eta,(logm)^3\|\tilde{H}-H^0\|\leq\epsilon))+ \qquad (1)$$

$$P(\|\hat{F}_k(\tilde{H})-\hat{F}_k(H_0)\|>\eta,(logm)^3\|\tilde{H}-H^0\|>\epsilon)) \;; \qquad (2)$$

first we analyze (1):

$$\|\hat{F}_k(\tilde{H})-\hat{F}_k(H_0)\|=$$

$$O_p(\max_{a,b}|\frac{1}{m}\sum_{j=1}^{m}(\log j)^k\left(\frac{2\pi}{n}\right)^{H_a^0+H_b^0-1}I_{ab}(\lambda_j)(j^{\tilde{H}_a+\tilde{H}_b-1}-j^{H_a^0+H_b^0-1})|=$$

$$O_p(\max_{a,b}|\frac{1}{m}\sum_{j=1}^{m}(logj)^k\ \lambda_j^{H_a^0+H_b^0-1}I_{ab}(\lambda_j)(j^{\tilde{H}_a+\tilde{H}_b-H_a^0-H_b^0}-1)|=$$

$$O_p(\max_{a,b}\frac{1}{m}\sum_{j=1}^{m}(logj)^k\ \lambda_j^{H_a^0+H_b^0-1}|I_{ab}(\lambda_j)|\,|(j^{\tilde{H}_a+\tilde{H}_b-H_a^0-H_b^0}-1)|=(*1),$$

as

$$|j^{\tilde{H}_a+\tilde{H}_b-H_a^0-H_b^0}-1|\leq|\tilde{H}_a+\tilde{H}_b-H_a^0-H_b^0|\ e\ logj\ ,$$

$$(*1)=O_p(\max_{a,b}|\tilde{H}_a+\tilde{H}_b-H_a^0-H_b^0|\ e\ (logm)^{k+1}\frac{1}{m}\sum_{j=1}^{m}\lambda_j^{H_a^0+H_b^0-1}|I_{ab}(\lambda_j)|)$$

and using appendix 5.3:

$$\frac{1}{m}\sum_{j=1}^{m}\ \lambda_j^{H_a^0+H_b^0-1}\ |I_{ab}(\lambda_j)|\ \leq$$

$$\leq\left(\frac{1}{m}\sum_{j=1}^{m}\ \lambda_j^{2H_a^0-1}\ I_{aa}(\lambda_j)\ \right)^{1/2}\left(\frac{1}{m}\sum_{j=1}^{m}\ \lambda_j^{2H_b^0-1}\ I_{bb}(\lambda_j)\right)^{1/2}\to_p(g_{aa}g_{bb})^{1/2}$$

so,

$$(*1)=O_p(\max_{a,b}|\tilde{H}_a+\tilde{H}_b-H_a^0-H_b^0|e(logm)^{k+1}(g_{aa}g_{bb})^{1/2}); \qquad (*2)$$

also as

$$(logm)^3\|\tilde{H}-H^0\|\leq\epsilon,$$

we have

$$(*2)=O_p(2e\epsilon(logm)^{k-2}(g_{aa}g_{bb})^{1/2}),$$

so

$$(1)\le P(2e\epsilon(\log m)^{k-2}(g_{aa}g_{bb})^{1/2}>\eta)\to 0 \ .$$

Now we analyze 2):

$$(2)\le P(\|\hat{H}-H^0\|>\frac{\epsilon}{(\log m)^3})=P(\hat{H}\in\bar{M}\cap\Theta) \qquad (*3)$$

where

$$\bar{M}=\{H:\|H-H^0\|>\frac{\epsilon}{(\log m)^3}\}$$

and recalling the definition of S(H) in appendix 5.1: $S(H)=R(H)-R(H_0)$,

$$(*3)=P(\inf_{\bar{M}\cap\Theta} S(H)\le 0)=P(\inf_{\bar{M}\cap N_\delta\cap\Theta} S(H)\le 0)+P(\inf_{\bar{M}\cap\bar{N}_\delta\cap\Theta} S(H)\le 0)=(a)+(b)$$

$$(b)\le P(\inf_{\bar{N}_\delta\cap\Theta} S(H)\le 0)\to 0$$

using 2a) and 2b) in appendix 5.1.Now we analyze (a); we recall $S(H)=-T(H)+U(H)$ defined in appendix 5.1, then:

$$P(\inf_{\bar{M}\cap N_\delta\cap\Theta} S(H)\le 0)\le P(\inf_{\bar{M}\cap N_\delta\cap\Theta} U(H)\le \sup_{\bar{N}_\delta\cap\Theta} |T(H)|),$$

analogously as in the proof of consistency (see appendix 5.1) we get here:

$$\inf_{\bar{M}\cap N_\delta\cap\Theta} U(H)\ge \inf_{\bar{M}\cap\Theta} U(H)=\sum_{k=1}^{r}\inf_{\bar{M}\cap\Theta^{(k)}} U_k(H_k)\ge\frac{r}{2}\frac{\epsilon^2}{(\log m)^6}$$

and

$$\sup_{\Theta\cap N_\delta}|T(H)|\le \sup_\Theta|T(H)|\le 2r\left|\frac{1}{m}\sum_{j=1}^{m}\log j-\log m+1\right|+ \qquad (*)$$

$$\sup_\Theta\left|\sum_{k=1}^{r}\left[(1+2(H_k-H_k^0))\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2(H_k-H_k^0)}-1\right]\right|+ \qquad (**)$$

$$\sup_\Theta\left|\sum_{k=1}^{r}(1+2(H_k-H_k^0))\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2(H_k-H_k^0)}\left(g_{kk}^{-1}\lambda_j^{2H_k^0-1}I_{kk}(\lambda_j)-1\right)\right| \qquad (***)$$

as in the proof of consistency in appendix 5.1:

$$(*)=O(\frac{\log m}{m}) \qquad and \qquad (**)=O(\frac{1}{m^\psi}) \ , \quad \psi>0 \ ,$$

$$(***)=\sup_{N_\delta\cap\Theta}|\sum_{i=k}^{r}(1+2(H_k-H_k^0))\frac{1}{m}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^{2(H_k-H_k^0)}\left(g_{kk}^{-1}\lambda_j^{2H_k^0-1}I_{kk}(\lambda_j)-1\right)|\leq$$

$$\leq\sum_{k=1}^{r}\sup_{N_\delta\cap\Theta^{(k)}}|A_k(H_k)|=O_P((\frac{\log m}{m})^{2/3})$$

then

$$(b)\leq P(\frac{r\epsilon^2}{(\log m)^6}<\frac{1}{m^{\psi}}(1+o(1))\to 0 \quad for \quad all\ \epsilon\ ,\ \psi>0\ ,$$

so we have proven P1). Proof of P2):

$$\|\hat{F}_k(H^0)-G^0\frac{1}{m}\sum_{j=1}^{m}(\log j)^k\|=\|\frac{1}{m}\sum_{j=1}^{m}(\log j)^k(\Lambda_j^{0^{-1}}I_j\Lambda_0^{0^{-1}}-G^0)\|=$$

$$=\|\frac{1}{m}\sum_{j=1}^{m-1}((\log j)^k-(\log(j+1))^k)\sum_{k=1}^{j}(\Lambda_j^{0^{-1}}I_j\Lambda_0^{0^{-1}}-G^0)+$$

$$(\log m)^k\frac{1}{m}\sum_{j=1}^{m}(\Lambda_j^{0^{-1}}I_j\Lambda_0^{0^{-1}}-G^0)\|\leq$$

$$=\frac{1}{m}\sum_{j=1}^{m}((\log j)^k-(\log(j+1))^k)\|\sum_{k=1}^{j}(\Lambda_j^{0^{-1}}I_j\Lambda_0^{0^{-1}}-G^0)\|+$$

$$(\log m)^k\frac{1}{m}\|\sum_{j=1}^{m}(\Lambda_j^{0^{-1}}I_j\Lambda_0^{0^{-1}}-G^0)\|=(*4)\ ,$$

then as

$$\|\frac{1}{m}\sum_{j=1}^{m}(\Lambda_j^{0^{-1}}I_j\Lambda_0^{0^{-1}}-G^0)\|=O_p(\max_{a,b}|\frac{1}{m}\sum_{j=1}^{m}I_{ab}(\lambda_j)\lambda_j^{H_a^0+H_b^0-1}-g_{ab}|)=$$

$$O_p(\frac{m^{\beta}}{n^{\beta}}+m^{-2/3}(\log m)^{2/3}+\frac{\log m}{m}+\frac{1}{\sqrt{m}\,n^{1/4}}+m^{-1/2})$$

using (A5.3.1), (A5.3.2) and (A5.3.3) in appendix 5.3, we get:

$$(*4)=O_p(|\frac{1}{m}\sum_{s=1}^{m}\frac{(\log s)^{k-1}}{s}(s^{1/3}(\log s)^{2/3}+s^{\beta+1}n^{-\beta}+s^{1/2}n^{-1/4}+s^{1/2})+$$

$$\frac{1}{m}(\log m)^k(m^{1/3}(\log m)^{2/3}+m^{\beta+1}n^{-\beta}+m^{1/2}n^{-1/4}))=$$

$$=O_p\left(\frac{(logm)^{k+2/3}}{m^{2/3}}+\frac{(\log m)^{k-1}}{\sqrt{m}}+\left(\frac{m}{n}\right)^{\beta}(logm)^{k}\right)=o_p(1)$$

using condition C5.4', and we get the desired result.

Appendix 5.5

In this appendix we state and prove some lemmas that we use along chapter 5 and its appendices.

Lemma 2:

$$\frac{2}{\sqrt{m}}\sum_{j=1}^{m}v_j\left[Re\left[g^a\Lambda_j^{0^{-1}}A(\lambda_j)\left(\frac{1}{2\pi n}\sum_{t=1}^{n}e_te_t'\right)A_a^*(\lambda_j)\lambda_j^{H_a^0-\frac{1}{2}}\right]-1\right]=o_p(1).$$

Proof. Rewrite the LHS as:

$$\frac{2}{\sqrt{m}}\sum_{j=1}^{m}v_j\left[Re\left[g^a\Lambda_j^{0^{-1}}A(\lambda_j)\frac{1}{2\pi}\left(\frac{1}{n}\sum_{t=1}^{n}e_te_t'-R\right)A_a^*(\lambda_j)\lambda_j^{H_a^0-\frac{1}{2}}\right]\right]+ \qquad (*)$$

$$\frac{2}{\sqrt{m}}\sum_{j=1}^{m}v_j\left[Re\left[g^a\Lambda_j^{0^{-1}}A(\lambda_j)\frac{R}{2\pi}A_a^*(\lambda_j)\lambda_j^{H_a^0-\frac{1}{2}}\right]-1\right], \qquad (**)$$

then

$$(**)=\frac{2}{\sqrt{m}}\sum_{j=1}^{m}v_j\left[Re\left[g^a\Lambda_j^{0^{-1}}f_a(\lambda_j)\lambda_j^{H_a^0-\frac{1}{2}}\right]-1\right]$$

where

$$f_a(\lambda_j)=A(\lambda_j)\frac{R}{2\pi}A_a^*(\lambda_j),$$

using C5.1':

$$f_a(\lambda_j)=\Lambda_j^0g_{a}\lambda_j^{\frac{1}{2}-H_a^0}(1+O(\lambda_j^{\beta}))$$

implies

$$(**)=O(\frac{2}{\sqrt{m}}\sum_{j=1}^{m}|\nu_j|\lambda_j^{\beta})=O(logm\frac{m^{\beta+\frac{1}{2}}}{n^{\beta}})=o(1)$$

by assumption C5.4', now

$$(*)=\frac{2}{\sqrt{m}}\sum_{j=1}^{m}\nu_j\left[Re\left[g^{a}\Lambda_j^{0^{-1}}A(\lambda_j)\frac{1}{2\pi}\left(\frac{1}{n}\sum_{t=1}^{n}e_te_t'-R\right)A_a^*(\lambda_j)\lambda_j^{H_a^0-\frac{1}{2}}\right]\right], \qquad (o)$$

now,using:

$$Re\left[g^{a}\Lambda_j^{0^{-1}}A(\lambda_j)\frac{1}{2\pi}\left(\frac{1}{n}\sum_{t=1}^{n}e_te_t'-R\right)A_a^*(\lambda_j)\lambda_j^{H_a^0-\frac{1}{2}}\right]$$

$$=\sum_{b=1}^{r}Re\left[g^{ba}\lambda_j^{H_b^0-\frac{1}{2}}A_b(\lambda_j)\frac{1}{2\pi}\left(\frac{1}{n}\sum_{t=1}^{n}e_te_t'-R\right)A_a^*(\lambda_j)\lambda_j^{H_a^0-\frac{1}{2}}\right]=O_p(\frac{1}{\sqrt{n}})$$

where we have used (A5.3.4) and (A5.3.5) in appendix 5.3, then:

$$(o)=O_p(\frac{1}{\sqrt{m}}\frac{1}{\sqrt{n}}\sum_{j=1}^{m}|\nu_j|)=O_p(\log m\left(\frac{m}{n}\right)^{\frac{1}{2}})=o_p(1).$$

<u>Lemma 3</u>:

$$\left(\frac{1}{\pi\sqrt{mn}}\right)^2\sum_{j\neq}^{m}\sum_{j'}^{m}\nu_j\nu_{j'}\sum_{t=1}^{n}\sum_{s=1}^{t-1}tr\left[\Omega_j^*(a)R\Omega_{j'}(b)R\cos(t-s)\lambda_j\cos(t-s)\lambda_{j'}\right]=o_p(1)$$

for a,b=1,...r.

<u>Proof</u>:

The proof is immediate considering:

1) $\sum_{t=1}^{n}\sum_{s=1}^{t-1}\cos(t-s)\lambda_j\cos(t-s)\lambda_{j'}=-n$, see Robinson (1993a),

2) $\Omega_j=\sum_{a=1}^{r}\lambda_a Re\left[A(\lambda_j)'\Lambda_j^{0^{-1}}g^{a'}\overline{A_a(\lambda_j)}+A_a^*(\lambda_j)g^{a}\Lambda_j^{0^{-1}}A(\lambda_j)\right]\lambda_j^{H_a^0-\frac{1}{2}}=O(1)$,

this is due to $A(\lambda_j)\Lambda_j^{0^{-1}}=O(1)$ *and* $A_a(\lambda_j)\lambda_j^{H_a-\frac{1}{2}}=O(1)$, see (A5.3.4) in appendix 5.3, and

3) $\frac{1}{mn^2}n\sum_{j*}\sum_{j'}\nu_j\nu_j'=O(\sum_{j=1}^{m}(\frac{\nu_j^2}{n}))=O(\frac{m}{n})=o(1).$

Lemma4:

$$\sum_{t=1}^{n}\sum_{s=1}^{t-1}tr[\Gamma_{t-s}^*(a)R\Gamma_{t-s}(b)[e_se_s'-R]]=o_p(1).$$

Proof:

The LHS has zero mean and variance that tends to 0 as we are going to prove. Using

$$tr(ABCD)=vec'(C)(D\otimes B')vec(A'),$$

we get

$$Var(LHS)=E\Bigg(\sum_{t=1}^{n}\sum_{s=1}^{t-1}vec'(R)[\Gamma_{t-s}(b)\otimes\overline{\Gamma_{t-s}(a)}]vec[e_se_s'-R]\cdot$$

$$\sum_{u=1}^{n}\sum_{v=1}^{u-1}vec'[e_ve_v'-R][\Gamma_{u-v}^*(b)\otimes\Gamma_{u-v}(a)']vec(R)\Bigg) \qquad (*X)$$

and calling:

$$K=E(vec[e_se_s'-R]vec'[e_se_s'-R]) \ ,$$

then:

$$(*X)=\sum_{t=1}^{n}\sum_{s=1}^{t-1}vec'(R)[\Gamma_{t-s}(b)\otimes\overline{\Gamma_{t-s}(a)}]K\sum_{v=1}^{t-1}[\Gamma_{t-v}^*(b)\otimes\Gamma_{t-v}(a)']vec[R]=$$

$$\sum_{t=1}^{n}d_t^*Kd_t=O_p(\sum_{t=1}^{n}\|d_t\|^2) \ ,$$

where

$$d_t=\sum_{s=1}^{n-t}[\Gamma^*_{t-s}(b)\otimes\Gamma_{t-s}(a)']vec[R] \ ,$$

i.e.,

$$d_t=\sum_{s=1}^{n-t}\gamma_s^2 \ ,$$

where

$$\gamma_s^2=[\Gamma^*_{t-s}(b)\otimes\Gamma_{t-s}(a)']vec[R] \ ,$$

then

$$|\gamma_s^2|\leq\|\Gamma^*_{t-s}(b)R^{\frac{1}{2}}\|\cdot\|\Gamma_{t-s}(a)'R^{\frac{1}{2}}\|.$$

Consider:

1)

$$\| \Gamma_{t-s}(I)'R^{\frac{1}{2}}\|=O(\frac{1}{\sqrt{mn}}\sum_{j=1}^{m}|\nu_j|)=O(\frac{\sqrt{m}\log m}{n}) \ , \quad I=a,b \ ,$$

2)

$$\|\Gamma_{t-s}(I)'R^{\frac{1}{2}}\|=O\left(\frac{1}{\sqrt{mn}}\left(|\sum_{j=1}^{m-1}(\nu_j-\nu_{j+1})\sum_{k=1}^{j}\cos(t-s)\lambda_k|+\nu_m|\sum_{j=1}^{m}\cos(t-s)\lambda_j|\right)\right)=$$

$$=O(\frac{1}{\sqrt{m}}\frac{1}{s}\log m) \quad for \quad 1\leq s\leq\frac{n}{2} \ ,$$

using $|\sum_{j=1}^{k}\cos(s\lambda_j)|=O(\frac{n}{s})$, see Zygmund (1977), p.2. Then

$$d_t=\sum_{s=1}^{n-t}\gamma_s^2=O\left(\frac{n}{m}\left(\frac{\sqrt{m}logm}{n}\right)^2+\left(\frac{logm}{\sqrt{m}}\right)^2\sum_{s>\frac{n}{m}}^{\infty}\frac{1}{s^2}\right)=O(\frac{(\log m)^2}{n}) ,$$

and

$$\sum_{t=1}^{n}\|d_t\|^2=O(\frac{logm)^4}{n}).$$

So that the variance of the LHS tends to 0.

Lemma 5:

$$\sum_{t=1}^{n}\sum_{s\neq}^{t-1}\sum_{s'}\epsilon_s'\Gamma_{t-s}^*(a)R\Gamma_{t-s'}(b)\epsilon_{s'}=o_p(1).$$

Proof:

LHS has zero mean and variance:

$$\sum_{t=1}^{n}\sum_{u=1}^{n}\sum_{s\neq}^{\min(t-1,u-1)}\sum_{s'}tr(\Gamma_{t-s}^*(a)R\Gamma_{t-s'}(b)R\Gamma_{u-s'}^*(b)R\Gamma_{u-s}(a)R)=$$

$$O(\sum_{t=1}^{n}\sum_{u=1}^{n}\sum_{s\neq}^{\min(t-1,u-1)}\sum_{s'}\|\Gamma_{t-s}^*(a)\|\|R\|\|\Gamma_{t-s'}(b)\|\|R\|\|\Gamma_{u-s'}^*(b)\|\|R\|\|\Gamma_{u-s}(a)\|\|R\|)=$$

$$O(2\sum_{t=1}^{n}\sum_{s\neq}^{t-1}\sum_{s'}\|\Gamma_{t-s}(a)\|^2\|\Gamma_{t-s'}(b)\|^2+$$

$$+4\sum_{t=3}^{n}\sum_{u=2}^{t-1}\sum_{s\neq}^{u-1}\sum_{s'}\|\Gamma_{t-s}^*(a)\|\|\Gamma_{t-s'}(b)\|\|\Gamma_{u-s'}^*(b)\|\|\Gamma_{u-s}(a)\|)=$$

$$O(n\left(\frac{(logm)^2}{n}\right)^2+\sum_{t=1}^{n}\|\Gamma_{t-s}^*(a)\|^2\sum_{t=3}^{n}\sum_{u=2}^{t-1}\sum_{s=t-u+1}^{t-1}\|\Gamma_{t-s}(a)\|^2)=$$

$$O(\frac{(logm)^4}{n}+\frac{(logm)^4}{m^{\frac{1}{3}}}).$$

This is due to:

$$\sum_{t=1}^{n}\|\Gamma^*_{t-s}(I)\|^2=O(\frac{(logm)^2}{n})\quad I=a,b\ ,$$

so the first part is:

$$O\left(n\left(\frac{(logm)^2}{n}\right)^2\right),$$

and the second part is:

$$O(\sum_{t=1}^{n}\|\Gamma^*_{t-s}(a)\|^2\sum_{t=3}^{n}\sum_{u=2}^{t-1}\sum_{s=t-u+1}^{t-1}\|\Gamma_{t-s}(a)\|^2)=O(\frac{(logm)^2}{n}\cdot n\sum_{j=1}^{[\frac{n}{2}]}j\|\Gamma^*_t(a)\|^2)=$$

$$O(\frac{(logm)^4}{m^{1/3}}).$$

Chapter 6

A Lagrange multiplier test for I(0)

In this chapter we will consider a Lagrange multiplier test for testing that a vector process[1] is I(0) and analyze its asymptotic properties and finite sample behavior. Then we will apply this test to a multivariate financial data set. This test is proposed and analyzed with more generality in Lobato and Robinson (1994) where both types of alternatives (the so called "long-memory" case and the "antipersistent" case) are analyzed jointly. Here we just consider the case when the alternative is "long-memory". In section 1 we present the test and state its asymptotic properties. Section 2 examines finite sample performance for the univariate case. In section 3 we apply this test to exchange rate data in a multivariate setup. Technical details are in the appendices of the chapter.

6.1 A Lagrange multiplier test for I(0)

Much econometric literature (applied as well as theoretical) in recent years has made great emphasis in the concept of I(0). For instance, most of the cointegration literature seems to have identified the concept of stationarity with I(0). Also the literature on the Generalized Method of Moments (GMM) has assumed I(0) and estimated the asymptotic variance of the estimates by some function of a smoothed estimate of the spectral density at zero frequency, see for instance in Newey-West (1987). In both cases I(0) is usually taken as granted and it

[1] A vector process is I(0) when every one of its components is I(0).

would be useful to test that assumption. On the other hand, the presence of long memory is generally considered as a nuisance because the analysis usually simplifies a great deal under weak dependence. Therefore, there is a clear interest in testing the I(0) assumption against long memory.

Consider the specification of the spectral density matrix that we have seen in the preceding chapter:

$$f(\lambda) \sim \Lambda G \Lambda \quad as \quad \lambda \to 0^+ \ , \quad \Lambda = diag\{\lambda^{1/2-H_a}\} \ ,$$

with G Hermitian positive definite, we formulate the null hypothesis as:

$$H_0 : H_a = 1/2 \quad for \quad a = 1, \ldots r,$$

and the alternative:

$$H_1 : H_a > 1/2 \quad for \ at \ least \ one \ a, \ a = 1, \ldots r.$$

Under H_0 the vector series we have is I(0) while under the alternative we have that at least one component of the vector has long-memory.

In order to motivate the test consider the objective function in (5.5) and recall from appendix 5.2 the distribution of the score evaluated at the true value H^0:

$$\sqrt{m}\frac{dR(H)}{dH}|_{H^0} \to_d N_r(0,E).$$

We can thus suggest the test

$$LM = m\left(\frac{dR(H_0)}{dH}\right)' \hat{E}^{-1}\left(\frac{dR(H_0)}{dH}\right)$$

where we evaluate the score at the null hypothesis and $\hat{E}$ is defined in (6.1.1). In (A5.2.3) we have derived an expression for ($dR(H)/dH_a$) that we can evaluate at the null hypothesis to get, in the notation of chapter 5:

$$\left(\frac{dR(H_0)}{dH_a}\right)=-\frac{2}{m}\sum_{j=1}^{m}\log\lambda_j+tr\left(\left(\frac{1}{m}\sum_{j=1}^{m}I_j\right)^{-1}\left(\frac{1}{m}\sum_{j=1}^{m}(\log\lambda_j)\{i_aI_j+I_ji_a\}\right)\right)$$

So that we can rewrite this as:

$$\frac{dR(H_0)}{dH}=v(\hat{C}_1\hat{C}^{-1}+\hat{C}^{-1}\hat{C}_1)=\hat{e}.$$

with

$$\hat{C}_k=\frac{1}{m}\sum_{j=1}^{m}v_j^kI(\lambda_j)\quad for\quad k=0,1,\ with\ \hat{C}=\hat{C}_0\ ,$$

υ_j has been defined in (A5.2.5) and v(A) denotes the vector whose components are the diagonal elements of the matrix A.

With respect to $\hat{E}$ consider

$$\hat{E}=2(I_r+\hat{C}*\hat{C}^{-1}) \tag{6.1.1}$$

we will show later that under the null hypothesis it is a consistent estimate for E, justifying that selection.

So that we can express the test concisely as:

$$LM=m\hat{e}'\hat{E}^{-1}\hat{e}.$$

Our test procedure is:

"reject H_0 if $LM>\chi^2_{r,\alpha}$ "

where $\chi^2_{r,\alpha}$ verifies:

$$P(\chi_r^2>\chi_{r,\alpha}^2)=\alpha$$

and χ^2_r follows a chi-square distribution with r degrees of freedom.

The asymptotic properties of this test are considered in the next two theorems.

Theorem 6.1: Under conditions C5.1', C5.2', C5.3' and C5.4' that we have stated in chapter 5 the LM test converges under H_0 to a random variable that follows a χ_r^2 distribution, that is:

$$P(LM>\chi^2_{r,\alpha}|H_0)\to\alpha,$$

Proof: This is immediate considering that in appendix 5.2 we have proven that:

$$\sqrt{m}\frac{dR(H_0)}{dH}\to_d N_r(0,E)$$

where H_0 could take the value H=1\2 and under the null:

$$\hat{C}=\frac{1}{m}\sum_{j=1}^{m}I(\lambda_j)\to_p G>0,$$

so, immediately:

$$\hat{E}=2(I_r+\hat{C}*\hat{C}^{-1})\to_p 2(I_r+G*G^{-1})=E.$$

In order to prove the consistency of the test we can relax the conditions we have seen above, in fact we can prove the consistency of the test under very general assumptions, we just need to streghten slightly the conditions in theorem 4.1. We introduce the conditions:

C6.1: $A_a(\lambda)\sim A_a\lambda^{1/2-H_a}+O(\lambda^{1/2-H_a+\tau})\ \ as\ \lambda\to 0^+,$ for some $\tau>0$ where A_a is a 1xr vector of complex constants. Notice that this assumption implies that $f_{ab}(\lambda)\sim g_{ab}\lambda^{1-H_a-H_b}+O(\lambda^{1-H_a-H_b+\tau})\ \ as\ \lambda\to 0^+.$

C6.2: $1/m+m/n\to 0$ as $n\to\infty$.

C6.3: e_t and $e_te_t'-R$ are uniform integrable martingale difference sequences and:

$$\max_{t,j}|e_t^j|^{2+v}<\infty$$

for some $v>0$, $j=1,...r$ and $t\geq 1$.

Notice that C6.2 is the same as condition C4.2 in chapter 4 and C6.1 and C6.3 are just stronger versions of C4.1 and C4.3.

Theorem 6.2: Under C6.1, C6.2 and C6.3: the test is consistent under H_1, that is,

$$P(LM>c|H_1)\rightarrow 1, \quad for \quad all \quad c>0.$$

The proof is in appendix 6.1.

6.2 Finite Sample Performance

In this section we examine with a Monte Carlo experiment the finite sample behaviour of the test. We have chosen two sample sizes, 128 and 256, that, as we discussed in chapter 3, are extremely modest sample sizes for these semiparametric procedures. We analyze the performance of the test for several values of m. For n=128 we have chosen m=4, 12, 20, 28 and 36; while for n=256 the grid is 20, 28, 36, 44 and 52.

In order to analyze the size we consider data generated from an AR(1):

$$x_t=\phi x_{t-1}+\epsilon_t$$

with ϵ_t being normal iid(0,1) and ϕ taking values: -0.9,-0.6,-0.3,0.0,0.3,0.6 and 0.9. In order to analyze the power of the test we consider series generated by fractional Gaussian noise using the same procedure as in chapter 3. We just analyze the univariate case because that is enough in order to grasp the main features of the finite sample behaviour. The number of replications is 1000 in all the experiments we have performed.

In tables 6.1 and 6.2 we examine size and power respectively. It is clear from these tables that in order to get moderate size we have to choose a fairly small value for m. It is also clear that as the sample size is bigger the distortion of the sizes is smaller (compare the sizes for m=20 or m=36 with n=128 and n=256). In any case we still notice as a main feature of these tables the unusual high size when $\phi=0.90$. With respect to power the results are what we could expect. First notice that although our theoretical analysis has been established only for H>0.5, we report also power figures when H<0.5. As we could expect the power is higher with higher m and when H differs more from 0.5.

We have performed another experiment in order to examine if we can improve the sizes. We have replaced the periodogram in the LM tests by a smoothed periodogram, that is:

$$LM=m\left(\frac{\sum_{j=1}^{m} v\hat{f}_j}{\sum_{j=1}^{m}\hat{f}_j}\right)^2$$

where

$$\hat{f}_j=\frac{1}{2k+1}\sum_{i=-k}^{k} I(\lambda_{j-i})$$

In tables 6.3 and 6.4 we examine this test for different degrees of smoothing. We just report results for n=256, the results for n=128 are qualitatively similar, with m= 12, 20, 28, 36 and 44, and k=2, 4 and 12. As k is lower the degree of smoothing is lower and the results are more similar to the ones obtained with the raw periodogram. As the degree of smoothing increases we obtain lower sizes as we could expect but we have to state also the effect that smoothing has on power. The loss of power when we use the smoothed periodogram is especially severe when the degree of smoothing is very high (k=12).

6.3 Analysis of exchange rate data

In this section we apply the LM test in its original form to two financial data sets. The first one is the one used in Whistler (1990) and in Robinson (1991). They are data for four exchange rates: BP/$, $/DM, $/JYn and $/SwFr. There are three different records: daily, weekly and monthly. The daily set goes from October, 1st 1981 to June, 28th 1985, i.e., the sample size is 946. The weekly and monthly data cover the period from January 1974 to June 1985. For weekly data the sample sizes are 600 while for monthly is just 138. It is important to emphasize, as it has been done in the unit root literature, that as important as the sample size is the sample span. This is what makes the monthly and especially the weekly data interesting because they cover more than eleven years. The second data set is the one used in chapter 5, section 4.

We apply the LM test for the first differences of the data. This is an interesting test to do because the efficiency market hypothesis establishes that:

$$E(p_t|\Omega_{T-1})=p_{t-1},$$

or what is equivalent that:

$$E(\epsilon_t|\Omega_{t-1})=0, \tag{6.4.1}$$

where p_t is the exchange rate and

$$\epsilon_t = p_t - p_{t-1}.$$

Equation (6.4.1) establishes that ϵ_t is a martingale difference. Some tests for (6.4.1) have focused on looking for some short-term correlation structure in ϵ_t, i.e. they have tested (6.4.1) as null hypothesis with the alternative being that ϵ_t has some sort of weak correlation or weak dependence. But we can fail to reject (6.4.1) using those tests procedures if ϵ_t possesses strong dependence. This is why it is important to test for (6.4.1) considering long-memory alternatives.

The result of the LM test for the first data set is in table 6.5. There are five columns: the first four are for the univariate LM test and the fifth is for the multivariate version of the test. Test values greater than the 5% asymptotic critical value are marked with "*". From table 6.5 we can deduce that there's no evidence of long-memory for daily and monthly data. For weekly data there's some evidence, especially for the BP/$.

The results of the LM test for the second data set for several values of m are in tables 6.6a and 6.6b for weekly and daily data respectively. The main conclusion is the lack of long-memory for these series. Only for the BP/JYn we appreciate a slight indication of long-memory with daily data.

Therefore, with the data we have used we do not reject the efficiency hypothesis. Another idea that has been analyzed in some empirical papers is to look for evidence of long memory in ϵ_t^2 as a measure of volatility, see for instance Ding et al. (1993) where they look at stock market data. In tables 6.7 and 6.8 we look for evidence of long memory in ϵ_t^2. The results are striking. Except for monthly data in the first data set and for BP/$ weekly in the second set in all the others series the evidence of long-memory is overwhelming using the univariate or multivariate LM test. Especially for daily data we reject the null hypothesis that ϵ_t^2 is I(0) for all the series using the univariate or the multivariate version of the LM test for any value of m.[2]

[2] The only exception is for small m in the first data set for $/SwFr.

Table 6.1

n=128 c.v.=5% and 1%

φ\m	4		12		20		28		36	
-0.90	0.000	0.000	0.004	0.000	0.004	0.001	0.025	0.000	0.120	0.012
-0.60	0.000	0.000	0.005	0.000	0.004	0.001	0.024	0.000	0.089	0.007
-0.30	0.000	0.000	0.005	0.000	0.005	0.001	0.018	0.000	0.037	0.002
0.00	0.000	0.000	0.006	0.000	0.008	0.002	0.014	0.003	0.014	0.004
0.30	0.000	0.000	0.008	0.000	0.025	0.003	0.093	0.019	0.212	0.083
0.60	0.000	0.000	0.028	0.007	0.260	0.096	0.666	0.417	0.907	0.790
0.90	0.000	0.000	0.661	0.390	0.975	0.935	1.000	0.998	1.000	1.000

n=256 c.v.=5% and 1%

φ\m	20		28		36		44		52	
-0.90	0.013	0.000	0.009	0.002	0.012	0.003	0.022	0.004	0.039	0.001
-0.60	0.010	0.001	0.009	0.002	0.014	0.002	0.020	0.004	0.034	0.001
-0.30	0.009	0.001	0.010	0.002	0.017	0.002	0.019	0.004	0.023	0.004
0.00	0.011	0.001	0.012	0.002	0.016	0.004	0.024	0.007	0.024	0.006
0.30	0.015	0.003	0.025	0.004	0.047	0.014	0.082	0.030	0.151	0.045
0.60	0.052	0.013	0.180	0.058	0.430	0.206	0.696	0.480	0.903	0.746
0.90	0.873	0.680	0.992	0.973	1.000	0.999	1.000	1.000	1.000	1.000

Table 6.2

n=128 c.v.=5% and 1%

H\m	4		12		20		28		36	
0.10	0.000	0.000	0.018	0.000	0.428	0.028	0.851	0.355	0.985	0.751
0.20	0.000	0.000	0.006	0.000	0.142	0.003	0.440	0.064	0.744	0.286
0.30	0.000	0.000	0.002	0.000	0.026	0.000	0.117	0.007	0.299	0.030
0.40	0.000	0.000	0.002	0.000	0.004	0.000	0.020	0.001	0.046	0.002
0.50	0.000	0.000	0.012	0.002	0.015	0.005	0.018	0.004	0.016	0.003
0.60	0.000	0.000	0.035	0.010	0.080	0.023	0.128	0.043	0.187	0.071
0.70	0.000	0.000	0.093	0.030	0.259	0.118	0.431	0.243	0.567	0.383
0.80	0.000	0.000	0.217	0.083	0.525	0.314	0.741	0.567	0.873	0.761
0.90	0.000	0.000	0.374	0.191	0.739	0.572	0.907	0.818	0.968	0.924

n=256 c.v.=5% and 1%

H\m	20		28		36		44		52	
0.10	0.284	0.024	0.713	0.233	0.947	0.622	0.993	0.896	0.999	0.976
0.20	0.095	0.003	0.350	0.042	0.616	0.184	0.826	0.403	0.939	0.661
0.30	0.030	0.000	0.108	0.006	0.238	0.026	0.380	0.070	0.535	0.158
0.40	0.005	0.000	0.021	0.001	0.041	0.001	0.062	0.004	0.105	0.013
0.50	0.013	0.001	0.014	0.001	0.018	0.001	0.018	0.001	0.018	0.003
0.60	0.061	0.019	0.103	0.038	0.159	0.054	0.220	0.095	0.263	0.134
0.70	0.240	0.102	0.377	0.222	0.509	0.326	0.612	0.454	0.701	0.536
0.80	0.474	0.307	0.691	0.508	0.798	0.676	0.885	0.791	0.948	0.879
0.90	0.705	0.532	0.866	0.780	0.948	0.885	0.982	0.949	0.996	0.983

Table 6.3 a

n=256 k=2 c.v.=5% and 1%

φ\m	12		20		28		36		44	
-0.90	0.000	0.000	0.005	0.000	0.004	0.001	0.008	0.002	0.017	0.003
-0.60	0.000	0.000	0.003	0.000	0.005	0.000	0.012	0.002	0.015	0.004
-0.30	0.000	0.000	0.004	0.000	0.005	0.000	0.015	0.002	0.014	0.003
0.0	0.000	0.000	0.004	0.000	0.005	0.000	0.015	0.002	0.018	0.004
0.30	0.000	0.000	0.008	0.000	0.016	0.003	0.040	0.012	0.080	0.026
0.60	0.000	0.000	0.034	0.004	0.171	0.049	0.430	0.202	0.706	0.494
0.90	0.181	0.018	0.873	0.682	0.994	0.975	1.000	0.999	1.000	1.000

Table 6.3 b

n=256 k=4 c.v.=5% and 1%

φ\m	12		20		28		36		44	
-0.90	0.000	0.000	0.001	0.000	0.002	0.000	0.005	0.001	0.014	0.000
-0.60	0.000	0.000	0.000	0.000	0.003	0.000	0.007	0.001	0.013	0.000
-0.30	0.000	0.000	0.000	0.000	0.003	0.000	0.008	0.001	0.010	0.001
0.0	0.000	0.000	0.000	0.000	0.003	0.000	0.011	0.002	0.012	0.004
0.30	0.000	0.000	0.001	0.000	0.009	0.000	0.026	0.007	0.065	0.021
0.60	0.000	0.000	0.015	0.000	0.142	0.031	0.421	0.180	0.716	0.477
0.90	0.010	0.000	0.857	0.591	0.995	0.972	1.000	0.999	1.000	1.000

Table 6.3 c

n=256 k=12 c.v.=5% and 1%

φ\m	12		20		28		36		44	
-0.90	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.000
-0.60	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000
-0.30	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.000
0.0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.000
0.30	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.000	0.022	0.002
0.60	0.000	0.000	0.000	0.000	0.006	0.000	0.235	0.023	0.654	0.317
0.90	0.000	0.000	0.000	0.000	0.962	0.650	1.000	0.995	1.000	1.000

Table 6.4 a

n=256 k=2 c.v.=5% and 1%

H\m	12		20		28		36		44	
0.1	0.003	0.000	0.245	0.010	0.697	0.190	0.944	0.600	0.995	0.890
0.2	0.000	0.000	0.070	0.000	0.326	0.035	0.608	0.162	0.809	0.402
0.3	0.000	0.000	0.019	0.000	0.091	0.005	0.213	0.021	0.372	0.066
0.4	0.000	0.000	0.004	0.000	0.016	0.000	0.035	0.001	0.059	0.002
0.5	0.001	0.000	0.006	0.000	0.009	0.000	0.014	0.000	0.016	0.001
0.6	0.004	0.000	0.038	0.010	0.078	0.030	0.144	0.048	0.206	0.080
0.7	0.026	0.001	0.183	0.058	0.333	0.186	0.493	0.298	0.602	0.432
0.8	0.098	0.012	0.426	0.228	0.676	0.474	0.796	0.663	0.887	0.791
0.9	0.211	0.053	0.683	0.476	0.867	0.761	0.945	0.885	0.983	0.949

Table 6.4 b

n=256 k=4 c.v.=5% and 1%

H\m	12		20		28		36		44	
0.1	0.000	0.000	0.171	0.001	0.676	0.138	0.940	0.576	0.996	0.887
0.2	0.000	0.000	0.045	0.000	0.274	0.018	0.578	0.124	0.795	0.369
0.3	0.000	0.000	0.009	0.000	0.069	0.003	0.199	0.014	0.343	0.049
0.4	0.000	0.000	0.001	0.000	0.010	0.000	0.027	0.000	0.051	0.003
0.5	0.000	0.000	0.001	0.000	0.005	0.000	0.009	0.000	0.014	0.001
0.6	0.000	0.000	0.022	0.001	0.052	0.016	0.114	0.033	0.185	0.058
0.7	0.000	0.000	0.091	0.017	0.286	0.101	0.456	0.250	0.577	0.381
0.8	0.003	0.000	0.308	0.101	0.623	0.401	0.784	0.626	0.883	0.759
0.9	0.020	0.000	0.595	0.321	0.848	0.713	0.944	0.872	0.985	0.945

Table 6.4 c

n=256 k=12 c.v.=5% and 1%

H\m	12		20		28		36		44	
0.1	0.000	0.000	0.004	0.000	0.290	0.005	0.813	0.251	0.983	0.729
0.2	0.000	0.000	0.000	0.000	0.035	0.000	0.310	0.013	0.652	0.151
0.3	0.000	0.000	0.000	0.000	0.004	0.000	0.042	0.001	0.169	0.012
0.4	0.000	0.000	0.000	0.000	0.000	0.000	0.004	0.000	0.017	0.000
0.5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.6	0.000	0.000	0.000	0.000	0.000	0.000	0.010	0.000	0.058	0.004
0.7	0.000	0.000	0.000	0.000	0.017	0.000	0.187	0.032	0.383	0.161
0.8	0.000	0.000	0.000	0.000	0.132	0.003	0.578	0.244	0.794	0.560
0.9	0.000	0.000	0.000	0.000	0.468	0.063	0.864	0.637	0.964	0.888

TABLE 6.5a

m	BP/$	$/DM	$/Jyn	$/SFr	Joint
5	0.064	0.234	0.174	0.246	1.328
10	0.547	0.006	0.054	0.020	1.095
15	0.614	0.123	0.009	0.135	0.963
20	1.825	0.245	0.220	0.556	2.161
25	1.048	0.805	0.553	1.248	1.554
30	0.769	0.454	0.927	0.498	1.776
35	1.700	1.024	0.237	1.471	2.280
40	3.169	1.067	0.025	1.944	2.883
45	3.027	1.345	0.008	3.249	3.634
50	5.146	2.294	0.006	4.694	5.573

TABLE 6.5b

m	BP/$	$/DM	$/JYn	$/SFr	Joint
10	0.589	0.009	0.002	0.024	1.225
20	2.039	0.164	0.330	0.634	3.271
40	3.283	0.622	0.042	2.074	5.316
60	4.372*	0.036	0.029	0.721	10.863*
80	9.987*	0.201	2.833	2.391	21.226*
100	8.185*	1.719	2.429	4.592*	13.411*
130	8.778*	3.511	7.975	5.776*	16.723*
160	4.281*	2.969	12.447	5.274*	18.440*
200	2.546	1.237	18.260	1.818	21.978*
250	4.258*	3.437	25.252	3.552	31.186*

TABLE 6.5c

m	BP/$	S/DM	$/JYn	$/SFr	Joint
20	0.028	1.127	0.080	0.310	2.499
40	0.023	0.107	0.037	0.033	1.564
60	0.573	0.030	1.383	0.044	2.705
80	0.001	0.103	5.286	0.034	8.489
100	0.044	0.306	6.134	0.466	6.664
130	0.028	0.666	6.218	1.145	6.837
160	0.302	0.120	2.957	0.024	4.445
200	0.010	0.065	3.519	0.055	3.704
250	0.740	1.299	4.487	1.953	3.904
300	0.753	1.032	2.364	2.604	2.800

TABLE 6.6a

m	BP/$	BP/DM	BP/JYn	BP/SFr	Joint
10	0.033	0.099	0.734	0.011	4.909
20	0.147	0.067	0.004	0.544	0.528
30	0.638	1.126	0.860	0.005	3.777
40	1.101	1.096	0.633	0.000	4.367
50	0.672	0.464	0.487	0.000	2.786
60	1.028	1.114	0.775	0.266	2.730
70	1.185	0.713	1.689	0.013	3.324
80	1.846	1.205	2.492	0.193	4.020
90	2.378	2.195	3.778	1.024	4.967
100	2.929	2.281	5.147	1.035	5.421

TABLE 6.6b

m	BP/$	BP/DM	BP/JYn	BP/SFr	Joint
20	0.062	0.105	0.002	0.581	0.449
40	0.860	1.465	0.647	0.001	5.235
70	0.662	0.679	0.738	0.032	3.465
100	1.411	1.383	5.460*	0.485	6.052
140	0.834	2.419	2.462	0.428	2.173
180	2.842	0.116	1.724	0.401	2.438
230	3.973*	0.806	2.072	0.131	4.094
280	5.682*	2.625	6.274*	0.211	4.861
350	2.576	1.267	3.864*	0.307	2.387
450	6.014*	2.721	5.203*	0.656	3.291

TABLE 6.7a

m	BP/$	S/DM	S/JYn	$/SFr	Joint
5	0.115	0.229	0.138	0.006	0.834
10	0.674	1.192	0.808	1.015	2.383
15	0.946	1.939	2.101	1.021	3.256
20	1.158	1.017	2.063	0.610	4.978
25	1.113	0.012	0.632	0.300	4.247
30	0.320	0.000	1.318	0.000	2.997
35	0.277	0.195	0.472	0.121	2.502
40	0.525	0.908	0.758	0.869	3.724
45	0.494	1.129	0.846	1.246	2.886
50	1.195	1.031	0.917	1.417	4.383

TABLE 6.7b

m	BP/$	$/DM	$/JYn	$/SFr	Joint
10	1.142	1.453	0.300	0.516	3.023
20	9.457*	3.322	3.135	1.444	16.829*
40	12.830*	10.005*	10.710*	2.970	40.395*
60	19.029*	32.333*	26.771*	13.278*	66.647*
80	24.045*	62.196*	24.153*	30.146*	89.038*
100	32.199*	100.356*	13.852*	53.524*	101.492*
130	34.113*	115.263*	13.716*	63.059*	115.997*
160	38.037*	128.070*	22.502*	89.265*	130.718*
200	30.152*	160.217*	20.590*	92.443*	145.696*
250	36.084*	212.339*	29.614*	120.379*	188.654*

TABLE 6.7c

m	BP/$	$/DM	$/JYn	$/SFr	Joint
20	13.427*	13.031*	30.025*	0.994	28.478*
40	49.021*	43.990*	48.504*	2.759	64.595*
60	57.648*	60.561*	101.841*	2.487	126.003*
80	69.364*	80.126*	106.683*	3.874*	148.818*
100	77.886*	96.976*	130.407*	7.007*	184.731*
130	100.664*	107.023*	122.626*	11.071*	211.050*
160	120.389*	132.900*	160.435*	20.687*	274.840*
200	111.882*	144.207*	153.778*	42.254*	298.056*
250	107.884*	149.281*	152.249*	88.409*	336.335*
300	92.717*	141.004 *	97.190*	135.428*	313.851*

TABLE 6.8a

m	BP/$	BP/DM	BP/JYn	BP/SFr	Joint
10	0.075	0.557	0.298	0.829	1.317
20	0.451	0.683	0.335	0.878	1.778
30	0.905	2.527	0.500	1.342	6.437
40	1.252	4.750*	1.808	1.937	8.546
50	1.049	10.424*	1.790	4.239*	12.265*
60	1.090	15.698*	1.249	4.926*	19.264*
70	1.354	27.505*	1.964	7.285*	34.225*
80	1.057	33.257*	1.990	5.851*	46.162*
90	1.510	39.560*	2.265	6.140*	56.483*
100	2.045	42.875*	4.052*	5.036*	66.951*

TABLE 6.8b

m	BP/$	BP/DM	BP/JYn	BP/SFr	Joint
20	3.357	0.479	3.435	2.048	16.780*
40	18.619*	7.102*	14.308*	13.610*	38.219*
70	52.701*	29.284*	44.216*	48.104*	85.963*
100	72.341*	43.893*	66.118*	59.714*	93.146*
140	79.030*	25.084*	97.535*	16.456*	120.820*
180	105.363*	40.319*	120.413*	36.402*	141.643*
230	134.572*	65.893*	154.056*	50.513*	157.791*
280	151.564*	86.483*	179.571*	58.537*	165.556*
350	167.957*	125.580*	231.736*	89.909*	191.486*
450	190.416*	173.214*	271.186*	127.573*	218.944*

Appendix 6.1

In order to prove consistency first we are going to recall that in appendix 4.1 we have proven that[1]:

$$\Lambda_m^{-1}\hat{C}\Lambda_m^{-1}=\frac{1}{m}\Lambda_m^{-1}\left(\sum_{j=1}^{m} I_j\right)\Lambda_m^{-1}\to_p \Gamma \qquad \text{(A6.1.0)}$$

where

$$\Lambda_m=diag\{\lambda_m^{1/2-H_a}\}$$

and Γ has as (a,b) element:

$$\frac{g_{ab}}{2-H_a-H_b}\ .$$

We are going to use (A6.1.0) and the following two results that will be proved later:

$$\Lambda_m^{-1}\hat{C}_1\Lambda_m^{-1}\to_p \Delta\ , \qquad \text{(A6.1.1)}$$

where Δ has as (a,b) element:

$$g_{ab}\,\frac{1-H_a-H_b}{(2-H_a-H_b)^2}$$

and

$$\Gamma>0. \qquad \text{(A6.1.2)}$$

Now using that Λ_m is a diagonal matrix, we can write:

$$\hat{e}=v\left(\Lambda_m\hat{C}^{-1}\Lambda_m\Lambda_m^{-1}\hat{C}_1\Lambda_m^{-1}+\Lambda_m^{-1}\hat{C}_1\Lambda_m^{-1}\Lambda_m\hat{C}^{-1}\Lambda_m\right)\ ,$$

then, using (A6.1.0), (A6.1.1) and (A6.1.2) we get

[1]We have proven this for $1/2<H_a,H_b<1$. This is the main limitation of this proof, if we were able to prove (6.1.3) for $0<H_a,H_b<1$ then we could get a two sided test, i.e., the alternative would be " H_1: $H_a\neq 1/2$ for at least one a".

$$\hat{e}\to_p v(\Gamma^{-1}\Delta+\Delta\Gamma^{-1})=\ \Xi=(\Xi_1,...,\Xi_r)'.$$

Now we have to show that under the alternative hypothesis at least one Ξ_p is different to 0; as " at least one $H_a>1/2 \Rightarrow$ at least one $\Xi_p\neq 0$" is equivalent to "$\Xi_p=0$, for all p $\Rightarrow H_a=1/2$, for all a", we need to prove the second statement. Unfortunately we haven't been able to prove this statement for a general r yet. To check it for r=1 and r=2 is immediate and we conjecture that the statement is valid for a general r.

Then we could write:

$$\hat{e}=\Xi+o_p(1)$$

with at least one $\Xi_p\neq 0$, now using (A6.1.0) and (A6.1.2):

$$\hat{E}=2(I_r+\hat{C}*\hat{C}^{-1})\to_p 2(I_r+(\Gamma*(\Gamma^{-1})'))=M>0\ .$$

As M is positive definite then:

$$LM=4m(\Xi+o_p(1))'(M^{-1}+o_p(1))(\Xi+o_p(1))=4m\Xi'M^{-1}\Xi+o_p(m),$$

then

$$P(LM>c\,|H_1)=P(4m\Xi'M^{-1}\Xi+o_p(m)>c)\to 1$$

$$as\ m\to\infty\ for\ any\ c>0.$$

So it only remains to prove (A6.1.1) and (A6.1.2). First we prove that Γ is positive definite. We can rewrite Γ as:

$$\Gamma=\frac{1}{\lambda_m}\Lambda_m^{-1}\int_0^{\lambda_m}\Lambda(\lambda)G\Lambda(\lambda)d\lambda\ \Lambda_m^{-1}\qquad where\ \Lambda(\lambda)=diag\{\lambda^{1/2-H_a}\},$$

now for all complex vectors $x\neq 0$:

$$x^*\Gamma x \geq \frac{\epsilon}{\lambda_m} x^* \Lambda_m^{-1} \int_o^{\lambda_m} \Lambda(\lambda)^2 d\lambda \ \Lambda_m^{-1} x =$$

$$= \epsilon \ x^* \ diag(\frac{1}{2-2H_1},\dots,\frac{1}{2-2H_r}) \ x > 0$$

where ϵ is a lower bound for the eigenvalues of G_0.

Now we prove (A6.1.1). The typical element (a,b) is:

$$\frac{1}{m}\lambda_m^{H_a+H_b-1} \sum_{j=1}^{m} \nu_j I_{ab}(\lambda_j) \to_p g_{ab}\frac{1-H_a-H_b}{(2-H_a-H_b)^2}$$

we can rewrite the LHS as:

$$\frac{1}{m}\lambda_m^{H_a+H_b-1} \sum_{j=1}^{m} \nu_j (I_{ab}(\lambda_j) - g_{ab}\lambda_j^{1-H_a-H_b}) \qquad (A)$$

$$+\frac{1}{m}\lambda_m^{H_a+H_b-1} \sum_{j=1}^{m} \nu_j g_{ab}\lambda_j^{1-H_a-H_b} \qquad (B)$$

So, we are going to prove that:

$$(A) \to_p 0 \quad , \quad (B) \to g_{ab}\frac{1-H_a-H_b}{(2-H_a-H_b)^2} \ .$$

$$(B) = \frac{1}{m}\lambda_m^{H_a+H_b-1} \sum_{j=1}^{m} \nu_j g_{ab}\lambda_j^{1-H_a-H_b} = g_{ab} \sum_{j=1}^{m} \nu_{j/m}\left(\frac{j}{m}\right)^{1-H_a-H_b}\frac{1}{m} \qquad (*)$$

where

$$\nu_{j/m} = \log(\frac{j}{m}) - \frac{1}{m}\sum_{j=1}^{m} \log(\frac{j}{m}) = \nu_j \ ,$$

then:

$$(*) \sim g_{ab}\int_0^1 \nu(x) x^{1-H_a-H_b} dx \ , \qquad (**)$$

where

$$\nu(x)=\log x-\int_0^1 \log x dx=\log x+1 \ .$$

As

$$\int_0^1 \nu(x)x^{1-H_a-H_b}dx=\frac{1-H_a-H_b}{(2-H_a-H_b)^2},$$

the result for (B) follows immediately.

In order to analyze (A) we need to use the following result that will be proved later:

$$\frac{2\pi}{n}\sum_{j=1}^{r}(I_{ab}(\lambda_j)-g_{ab}\lambda_j^{1-H_a-H_b})=O_p(\lambda_r^{2-H_a-H_b}\{n^{\frac{-\nu}{2+\nu}}+r^{-\delta}+\left(\frac{r}{n}\right)^{\tau}\}) \qquad \text{(A6.1.3)}$$

Then, by summation by parts:

$$(A)=\frac{1}{m}\lambda_m^{H_a+H_b-1}\{\sum_1^{m-1}(\nu_j-\nu_{j+1})\sum_{k=1}^{j}(I_{ab}(\lambda_k)-g_{ab}\lambda_k^{1-H_a-H_b})+$$
$$+\nu_m\sum_1^m(I_{ab}(\lambda_j)-g_{ab}\lambda_j^{1-H_a-H_b})\}=(A1)+(A2)$$

and using (A6.1.3):

$$(A2)=O_p\left((\log m)\frac{n}{m}\lambda_m^{H_a+H_b-1}\lambda_m^{2-H_a-H_b}\{n^{\frac{-\nu}{2+\nu}}+m^{-\delta}+\left(\frac{m}{n}\right)^{\tau}\}\right)=o_p(1)$$

while using the same result and the mean value theorem:

$$(A1)=O_p(\frac{n}{m}\lambda_m^{H_a+H_b-1}\sum_1^m\{\frac{j^{1-H_a-H_b}}{n^{2-H_a-H_b}}(n^{\frac{-\nu}{2+\nu}}+j^{-\delta}+(\frac{j}{n})^{\tau}\})=o_p(1)$$

So it remains to prove (A6.1.3). This result is just an extension of Theorem 4.1. We just indicate how we need to modify that proof in order to get the result. We need to streghten propositions 1, 3 and 4 so that for some $\delta>0$:

$$\frac{2\pi}{n}\sum_{j=1}^{m} Ref_{ab}(\lambda_j)-ReF_{ab}(\lambda_m)=o(\lambda_m^{2-H_a-H_b}\{m^{-\delta}+\left(\frac{m}{n}\right)^{\tau}\})\ , \tag{4}$$

$$Re(\frac{2\pi}{n}\sum_{j=1}^{m}(A_a(\lambda_j)J(\lambda_j)A_b^*(\lambda_j)-f_{ab}(\lambda_j))=O_p(\lambda_m^{2-H_a-H_b}\{n^{\frac{-\nu}{2+\nu}}+m^{-\delta}+\left(\frac{m}{n}\right)^{\tau}\}) \tag{5}$$

and

$$\frac{1}{n}\sum_{j=1}^{m}|w_a(\lambda_j)-A_a(\lambda_j)v(\lambda_j)|^2=o_p(\lambda_m^{2-2H_a}\{m^{-\delta}+(\frac{m}{n})^{\tau}\}) \tag{6}$$

and the same results hold for the imaginary parts in (4) and (5). In order to prove (4) recall (A4.1.1), now:

$$r=o(\frac{1}{n}\sum_{j=1}^{m}\lambda_j^{1-H_a-H_b+\tau}+\lambda_m^{2-H_a-H_b+\tau})=o((\frac{m}{n})^{\tau}\ \lambda_m^{2-H_a-H_b})$$

while (A4.1.2) is

$$O(m^{H_a+H_b-2}\lambda_m^{2-H_a-H_b})=o(m^{-\delta}|F_{ab}(\lambda_m)|)$$

for some $\delta>0$.

In order to prove (5) recall (A4.3.1) and using C6.3:

$$\|D\|=O_p(n^{\frac{-2}{2+\nu}})$$

for some $\nu>0$, so that

$$(1)=O_p(F_{aa}(\lambda_m)^{1/2}F_{bb}(\lambda_m)^{1/2}n^{\frac{-2}{2+\nu}})$$

while in (A4.3.2) using C6.1 we get that

$$a_t=O((\frac{m}{n})^{\tau}\lambda_m^{2-H_a-H_b})$$

and in (A4.3.3) we have:

$$\max_{r<t<n}\left|\frac{1}{n}\sum_{1}^{m}\lambda_j^{1-H_a-H_b}\cos t\lambda_j\right|^2=O(F_{aa}(\lambda_m)^{1/2}F_{bb}(\lambda_m)^{1/2}m^{-\delta})$$

for some $\delta>0$ so that

$$(2)=o(\lambda_m^{2-H_a-H_b}\{m^{-\delta}+(\frac{m}{n})^{\tau}\}).$$

Now (6) follows immediately considering that (A4.4.1) is

$$o_p(\lambda_m^{2-2H_a}\{m^{-\delta}+(\frac{m}{n})^{\tau}\})$$

as proved in Robinson (1994).

Once we have strengthened propositions 1, 3 and 4, (A6.1.3) follows straightforwardly following the same steps as in the proof of Theorem 4.1.

Chapter 7

Conclusions

In this thesis we have examined semiparametric estimation procedures in the frequency domain for long-memory series. We have justified the semiparametric approach based on its robustness. A parametric approach, if the model has been correctly specified, can produce efficient estimates for H, but if not, it will lead to inconsistent estimates for H. The semiparametric approach is more general. We assume covariance stationarity and so, the spectral density belongs to L_1, but the rest of the assumptions on the spectral density concern only a neighbourhood of zero; away from zero we don't impose any smoothness or parametric behaviour. So, our estimates of H that are going to be a function of the ordinates of the periodogram that are close to the zero frequency, will be consistent estimates for H irrespectively of the short run behaviour of the process.

We have analyzed mainly two estimates, the averaged periodogram estimate (APE) and a quasi maximum likelihood estimate (QMLE). The most used in the applied literature has been the log-periodogram estimate (LPE). The comparison between the different semiparametric estimates has to be done according to several criteria. These include not only theoretical considerations (under what degree of strictness in our assumptions we get asymptotic normality, which estimate has less variance...) but also implementation problems (how immediate it is to get the estimates). With respect to the first criterion, the QMLE looks clearly as the best estimate. While we need to assume Gaussianity in the LPE or in the APE in order to achieve asymptotic normality, we do not need to impose it in the QMLE.

In fact, finiteness of fourth moments is enough. Another drawback of the APE was the discontinuity of the asymptotic theory around H=3/4. We only got asymptotic normality when $H \in (1/2,3/4)$. A shortcoming of LPE is the necessity of choosing a trimming number l because about its optimality nothing is known so far. If we compare the variance of the asymptotic distribution in the univariate case we get that if we choose the optimal q for every H, with $H \in (1/2,3/4)$, then APE will provide the estimate with less variance (e.g., H=0.7, $q^*=0.4$, variance=0.08); while LPE will provide the largest (e.g. with J=1, variance=0.411). In general, QMLE provides the most straightforward way of making inference, the asymptotic variance of this estimate being 0.25. With respect to the second consideration all the estimates are not difficult to implement, although LPE is particularly straightforward as it just reduces to OLS.

In a multivariate framework we appreciate a clear superiority of the QMLE and LPE to the APE . The multivariate version of the QMLE has been analyzed in chapter 5 while Robinson (1992) analyzed the LPE in a multivariate setup. But we haven't obtained the asymptotic distribution of multivariate APE because the difficulty of the univariate case suggests that the multivariate analysis has to be almost intractable.

So far we have discussed the different semiparametric procedures to estimate H. In chapter 6 we have seen an extension of these procedures to derive a Lagrange multiplier test for I(0), i.e., weak dependence. This is an important issue for economic policy. Much applied econometric assumes that economic time series have basically two components: a secular or trend component that accounts for the main institutional or technological factors underlying the economy, this trend can be stochastic or deterministic, and, a random component that superimposes this trend. This random component is assumed to have negligible influence in the long run. By this we mean that, if the series, because of some external factors, is moved away from its trend, there exist some mechanisms in the economy that will push the series to the trend until it is reached again. The standard econometric literature assumes that this stationary random component is weak dependent. In economic terms, this means that this adjustment process we have discussed above is fast.

It is very important from an economic policy point of view to determine if this adjustment is fast or slow. Economic policy has been classified in two main groups: structural policy that attempts to modify the trends, and stabilization policy that has a short run impact and mainly tries to accelerate the adjustment processes commented above. Stabilization policy is more seriously justified if the random component is not weak dependent than in the other case. If a series is weak dependent it means that any disequilibrium situation will be transitory and in a short time the series will reach its trending value, so the role of economic policy is minor. On the other hand, if the random component is not weak dependent then, even though eventually the series will get back to its trending value, this adjustment will be very slow and policy intervention to try to speed up this adjustment process can be justified.

With the test we have developed we can attempt to answer that question. The problem with real series is that usually the number of observations available is very limited and it is not clear that semiparametric procedures will be really informative. On the other hand financial series available are longer and semiparametric procedures are more adequate. This is why we have applied the LM test and the QMLE to some exchange rate data.

We have tested the efficiency hypothesis and we didn't reject it. Then we looked for evidence of long memory in the square of the first differences of the exchange rate series and we found an overwhelming evidence. This can be interpreted as some sort of long memory in the volatility of exchange rates. This result has important implications for financial analysts. Being able to forecast the volatility better than the market can lead to substantial monetary profits by means of "straddles". These deals consist on selling simultaneously put options (conferring a right to sell) and call options (a right to buy) if you think that the volatility is going to be less than what the market predicts and buying simultaneously put and call options if you anticipate greater volatility than the market.

Bibliography

Adenstedt, R.K. (1974). On large sample estimation for the mean of a stationary random sequence. Annals of Statistics, 2, 1095-1107.

Agiakloglou, C. and Newbold, P. (1993). Lagrange multiplier tests for fractional difference. Journal of Time Series Analysis, 15, 3, 253-262.

Agiakloglou, C., Newbold, P. and Wohar, M. (1994). Bias in an estimator of the fractional difference parameter. Journal of the Time Series Analysis, 14, 3, 235-246.

Baillie, R.T., Bollerslev, T. and Mikkelsen, H.O.A. (1993). Fractionally integrated generalized autoregressive conditional heteroskedasticity. Preprint.

Beran, J. (1991). M-estimators of location for data with slowly decaying serial correlations. Journal of the American Statistical Association, 86, 704-708.

Beran, J. (1992). A goodness of fit tests for time series with long-range dependence. Journal of the Royal Statistical Society Series B, 54, 749-760.

Beran, J. (1994). Statistics for Long-Memory Processes. Chapman and Hall, New York.

Box, G.E.P. and Jenkins, G.M. (1970). Time Series Analysis: Forecasting and Control. Holden-Day, San Francisco.

Box, G.E.P. and Pierce, D.A. (1970). Distribution of residual autocorrelations in autoregressive integrated moving average time series models. Journal of the American Statistical Association, 65, 1509-1526.

Brillinger, D.R. (1981). Time Series: Data Analysis and Theory. Expanded edition. McGraw-Hill.

Brown, B.M. (1971). Martingale central limit theorems. The Annals of Mathematical Statistics, Vol.42, No. 1, 59-66.

Cheng, B. and Robinson, P.M. (1991). Semiparametric density derivative estimates under strong dependence. Preprint.

Dahlhaus, R. (1989). Efficient parameter estimation for self-similar processes. Annals of Statistics, 17, 1749-1766.

Davies, R.B. and Harte, D.S. (1987). Tests for Hurst effect. Biometrika, 74, 95-102.

Delgado, M. and P.M. Robinson (1994). New methods for the analysis of long memory time series: application to Spanish inflation. Journal of Forecasting, 13, 97-107.

Fox, R. and Taqqu, M.S. (1985). Non-central limit theorems for quadratic forms in random variables having long-range dependence. Annals of Probability, 13, 428-446.

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

Geweke, J. and Porter-Hudak, S. (1983). The estimation and application of long memory time series models. Journal of Time Series Analysis, 4, 221-238.

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

Granger, C.W.J. (1966). The typical spectral shape of an economic variable. Econometrica, 34, 150-161.

Granger, C.W.J. (1980). Long memory relationships and aggregation of dynamic models. Journal of Econometrics, 14, 227-238.

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

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

Graybill, F.A.(1983). Matrices with applications in Statistics. Second Edition. Wadsworth International Group. Belmont, California.

Grenander, U. (1954). On the estimation of regression coefficients in the case of autocorrelated disturbance. Annals of Mathematical Statistics, 25, 252-272.

Grenander, U. and Rosenblatt, M. (1957). Analysis of Stationary Time Series. Wiley, New York.

Hall, P. and Hart, J.D. (1989). Nonparametric regression with long-range dependence. Technical Report 88, Dept. Statistics, Texas A&M Univ.

Hannan, E.J. (1976). The identification and parameterisation of ARMAX and state space forms. Econometrica, 44, 713-723.

Harvey, A.C. (1993). Stochastic volatility models. LSE. Preprint.

Hidalgo, J. and Robinson, P.M. (1993). Testing for structural change in a long memory environment. Preprint.

Hipel, K.W. and McLeod, A.I. (1978). Preservation of the Rescaled Adjusted Range. Water Resources Research, 14, 491-518.

Hosking, J.R.M. (1981). Fractional differencing. Biometrika, 68, 1, 165-176.

Hosking, J.R.M. (1984). Modelling persistence in hydrological time series using fractional differencing. Water Resources Research, 20, 1898-1908.

Hurst, H.E. (1951). Long-term storage capacity of reservoirs. Transactions of American Society Civil Engineers, 116, 770-779.

Hurvich, C.M. and Beltrao, K.I. (1993). Asymptotics for the low-frequency ordinates of the periodogram of a long memory time series. Journal of Time Series Analysis, 14, 455-472.

Hurvich, C.M. and Beltrao, K.I. (1994). Automatic semiparametric estimation of the memory parameter of a long-memory time series. Journal of Time Series Analysis, 15,285-302.

Künsch, H.R. (1986). Discrimination between monotonic trends and long-range dependence. Journal of Applied Probability, 23,1025-1030.

Li, W.K. and McLeod, A.I. (1986). Fractional time series modelling. Biometrika, 73, 217-221.

Lo, A.W. (1991). Long-term memory in stock market prices. Econometrica, 59, 1279-1313.

Lobato, I. and Robinson, P.M. (1994). Averaged periodogram estimation of long memory. Preprint.

Lobato, I. and Robinson P.M. (1995). A general test for I(0). Preprint.

Magnus, J.R. and Neudecker, H. (1988). Matrix differential calculus with applications in statistics and econometrics. John Wiley & Sons.

Mandelbrot, B.B. (1960). The Pareto-Lévy law and the distribution of income. International Economics Review, 1, 79-106.

Mandelbrot, B.B. (1969). Long-run linearity, locally Gaussian processes, H-spectra and infinite variances. International Economic Review, 10, 82-111.

Mandelbrot, B.B. (1971). When can price be arbitraged efficiently? A limit to the validity of the random walk and martingale models. Review of Economic and Statistics, LIII, 225-236.

Mandelbrot, B.B. (1972). Statistical methodology for non-periodic cycles: from the covariance to R/S analysis. Annals of Economic and Social Measurement, 1, 259-290.

Mandelbrot, B.B. (1975). Limit theorems on the self-normalized range for weakly and strongly dependent processes. Z. Wahrscheinlichkeitstheorie verw. Geb., 31, 271-285.

Mandelbrot, B.B. and Taqqu, M.S. (1979). Robust R/S analysis of long-run serial correlation. Proceedings of the 42nd Session of the International Statistical Institute, Manila. Bulletin of the I.S.I, Vol. 48, Book 2, 69-104.

Mandelbrot, B.B. and Van Ness, J.W. (1968). Fractional Brownian motions, fractional noises and applications. SIAM Review, 10, 422-437.

Mandelbrot, B.B. and Wallis, J.R. (1968). Noah, Joseph and operational hydrology. Water Resources Research, 4, 909-918.

Mandelbrot, B.B. and Wallis, J.R. (1969). Some long-run properties of geophysical records. Water Resources Research, 5, 321-340.

Milhoj, A. (1981). A test of fit in time series models. Biometrika, 68, 177-188.

Newey, N.K. and K.D. West (1987). A simple positive definite heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica, 55, 703-708.

Peiris, M.S. and Perera, B.J.C. (1988). On prediction with fractionally differenced ARIMA models. Journal of Time Series Analysis, 9, 215-220.

Porter-Hudak, S. (1990). An application of the seasonal fractionally differenced model to the monetary aggregates. Journal of the American Statistical Association, 85, 338-344.

Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vetterling, W.T. (1990). Numerical Recipes. Cambridge University Press.

Ray, B.K. (1993). Modelling long-memory processes for optimal long-range prediction. Journal of Time Series Analysis, 14, 511-525.

Robinson, P.M. Quantitative Techniques lecture notes. LSE.

Robinson, P.M. (1978). Statistical inference for a random coefficient autoregressive model. Scandinavian Journal of Statistics, 5, 163-168.

Robinson, P.M. (1991a). Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression. Journal of Econometrics, 47, 67-84.

Robinson, P.M. (1991b). Nonparametric function estimation for long memory time series. In Nonparametric and semiparametric methods in econometrics and statistics. (W. Barnett et al., eds), 437-457. Cambridge University Press. New York.

Robinson, P.M. (1992). Log-periodogram regression of time series with long range dependence. Preprint.

Robinson, P.M. (1993a). Gaussian semiparametric estimation of long range dependence. Preprint.

Robinson, P.M. (1994a). Semiparametric analysis of long-memory time series. Annals of Statistics, 22, 515-539.

Robinson, P.M. (1994b). Rates of convergence and optimal spectral bandwidth for long range dependence. Probability Theory and Related Fields, 99, 443-473.

Robinson, P.M. (1994c). Time series with strong dependence. In C.A. Sims, ed., Advances in Econometrics: Sixth World Congress, Vol. 1, 47-95, Cambridge University Press, Cambridge.

Robinson, P.M. (1994d). Efficient tests of nonstationary hypotheses. Journal of the American Statistical Association, 89, 428, 1420-1437.

Rosenblatt, M. (1961). Independence and dependence. Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability, 411-443. Berkeley: Univ. Calif. Press.

Samarov, A. and Taqqu, M.S. (1987). On the efficiency of the sample mean in long memory noise. Journal of Time Series Analysis, 9, 191-200.

Sinai, Ya.G. (1976). Self-similar probability distributions. Theory Probability Applied, 21, 64-80.

Solo, V. (1986). Topics in Advanced Time Series Analysis. Springer Verlag lecture notes in mathematics vol. 215.

Sowell, F. (1992). Maximum likelihood estimation of stationary univariate fractionally integrated time series models. Journal of Econometrics, 53, 165-188.

Styan, G.P.H. (1973). Hadamard products and multivariate statistical analysis. Linear Algebra and its Applications, 6, 217-240.

Taqqu, M.S. (1975). Weak convergence to fractional Brownian motion and to the Rosenblatt process. Z. Wahrscheinlichkeitstheorie verw. Geb., 31, 287-302.

Taqqu, M.S. (1985). A bibliographic guide to self-similar processes and long-range dependence. In Dependence in Probability and Statistics (E. Eberlein and M. S. Taqqu, eds) 137-165. Birkhäuser, Boston.

Taqqu, M.S. (1987). Random processes with long-range dependence and high variability. Journal of Geophysical Research, 92, 9683-6.

Terrin, N. and Taqqu, M.S. (1991). Convergence in distribution of sums of bivariate Appell polynomials with long-range dependence. Probability Theory and Related Fields, 90, 57-81.

Verwaat, W. (1987). Properties of general self-similar processes. Bulletin of the Institute International of Statistics, 52, 199-216.

Whittle, P. (1956). On the variation of yield variance with plot size. Biometrika, 43, 337-343.

Whittle, P.(1962). Topographic correlation, power-law covariance functions and diffusion. Biometrika, 49, 304-314.

Yajima, Y. (1988). On estimation of a regression model with long-memory stationary errors. Annals of Statistics, 16, 791-807.

Yajima, Y. (1991). Asymptotic properties of the LSE in a regression model with long-memory stationary errors. Annals of Statistics, 19, 158-177.

Yong, C.H. (1974). Asymptotic behaviour of trigonometric series. Hong Kong, Chinese University of Hong Kong.

Zygmund, A. (1979). Trigonometric Series. Vols. 1 and 2. Cambridge: Cambridge Univ. Press.