PREDICTION OF STRONGLY DEPENDENT PROCESSES IN THE FREQUENCY DOMAIN WITH APPLICATION TO SIGNAL EXTRACTION

J. Hidalgo[*,†]
London School of Economics

Y. Yajima
University of Tokyo

October 19, 1999

Abstract

It is frequent to observe that one of the aims of time series analysts is to predict future values of the data. For weakly dependent data, when the model is known up to a finite set of parameters, its statistical properties are well documented and exhaustively examined. However, if the model was misspecified, the predictors would no longer be correct. Motivated by this observation and due to the interest to obtaining accurate and reliable predictors, Bhansali (1974) examined the properties of a nonparametric predictor based on the canonical factorization of the spectral density function given in Whittle (1963) and known as $FLES$.

However, the above work does not cover the so-called strongly dependent data. Due to the interest in this type of processes, one of our objectives in this paper is to examine the properties of the $FLES$ for these processes. In addition, we illustrate how the $FLES$ can be adopted to recover the signal of a strongly dependent process, showing its consistency. The proposed method is semiparametric, in the sense that, in contrast to other methods, we do not need to assume any particular model for the noise except that it is weakly dependent.

[*]The first author gratefully acknowledge the research support of the Economic and Social Research Council (ESRC) reference number: R000238212.

[†]Mailing address: J. Hidalgo. Econ. Department. London School of Economics. Houghton Street. London WC2A 2AE. U.K..

1. INTRODUCTION

In empirical studies, it is very common to observe that one of the aims to model time series data is for prediction purposes. In the context of linear models, the Wiener-Kolmogorov theory, see for instance Hannan $(1970, Ch.3)$, has been described as a great achievement in that direction. For weakly (linear) dependent data, the statistical properties of predicted future values are very well documented. Existing procedures lie in two main categories, namely the parametric and nonparametric approach. In the former, the parameters of the model are estimated, via either time or frequency domain approaches, and plugged into the Wiener-Kolgomorov formula. However, since always there is some degree of uncertainty about the correct specification of the model, Bhansali $(1974, 1977)$ described a nonparametric predictor, based on a factorization of a "windowed" estimate of the spectral density function, which computes the constants which enter in the Wiener-Kolmogorov formula. The algorithm, denoted $FLES$ (factorized logarithm of the estimated spectrum), guarantees that the predictor will always be consistent with no need for the practitioner to decide any specific model for the data.

On the other hand, it has been observed that in many areas, such as hydrology or economics, the data exhibits strong dependence, characterized by having a non-summable autocovariance function. Statistical properties of predictors in parametric models exhibiting this type of dependence has not been studied as deeply as with weakly dependent data, although some research has been done in that direction. Among them, we can mention Peiris and Perera (1988), Beran (1994), Crato and Ray (1996). Some empirical examples of prediction with strongly dependent data can be found in Porter-Hudak (1990) and Ray (1993). The former showed the superiority of predictors based on a parametric fractional autoregressive integrated moving average ($FARIMA\,(p, \alpha/2, q)$) for USA data of monetary aggregates model compared to predictors based on more traditional autoregressive integrated moving average ($ARIMA$) models.

Because, as was mentioned above, there is always a degree of uncertainty about the correct specification of the data, the first objective of this paper is to study and examine the properties of the predictor based on the $FLES$ algorithm of a covariance stationary linear possibly, strongly dependent process. The second objective of the paper is to illustrate how the $FLES$ algorithm can be adopted for the purpose of signal recovering of a covariance stationary linear process which exhibits strong dependence. More specifically, assuming that the process can be decomposed in such a way that the noise is weakly dependent and the signal strongly dependent, we describe how we can extract the signal without assuming any parametric model of the former.

The organization of the paper is as follows. In the next section, we present the $FLES$ algorithm and its estimator. In Section 3, we delimit our framework and examine the properties of the estimators given in Section 2. In Section 4, we describe how the $FLES$ algorithm can be adopted to extract the signal of a strongly dependent process, showing its statistical properties. In Section 5, we provide the proofs of our results which apply some technical lemmas given in Section 6. Finally, in the last section, we present conclusions and possible extensions.

2. THE FLES ALGORITHM AND ITS ESTIMATE

Let $\{x_t\}$ be a covariance stationary linear process which is observed at times $t = 1, ..., n$, having mean that is zero and with absolute continuous spectral distribution, so that its spectral density function, denoted $f_x(\lambda)$, is defined as

$$\gamma_x(j) \overset{def}{=} E(x_0 x_j) = \int_{-\pi}^{\pi} f_x(\lambda) \cos(j\lambda)\, d\lambda, \;\; j = 0, 1, 2, ...$$

We will assume that the process x_t admits the following representations

$$x_t = \sum_{j=0}^{\infty} b_j \varepsilon_{t-j}, \; \sum_{j=0}^{\infty} b_j^2 < \infty, \, b_0 = 1 \tag{2.1}$$

and

$$\sum_{j=0}^{\infty} a_j x_{t-j} = \varepsilon_t, \, a_0 = 1, \tag{2.2}$$

where ε_t is a process of uncorrelated random variables with mean zero and variance σ_ε^2, and a_j and b_j are constants. Following (2.1) or (2.2), the spectral density function can be written as

$$f_x(\lambda) = \frac{\sigma_\varepsilon^2}{2\pi} |A(\lambda)|^{-2} = \frac{\sigma_\varepsilon^2}{2\pi} |B(\lambda)|^2$$

where $A(\lambda) = \sum_{j=0}^{\infty} a_j e^{ij\lambda}$ is the spectral transfer function of the coefficients a_j. Likewise, $B(\lambda)$ is the transfer function of the coefficients b_j. All throughout our basic assumption on $f_x(\lambda)$ is that

$$f_x(\lambda) \sim C\lambda^{-\alpha} \text{ as } \lambda \to 0+ \tag{2.3}$$

where $C \in (0, \infty)$, $\alpha \in [0, 1)$ and " $\sim$ " means that the ratio of the left- and right-hand sides tends to one and differentiable in any open set outside the origin. When $\alpha = 0$, we say that the data is weakly dependent, whereas for $\alpha \in (0, 1)$, we say that the data exhibits the property of strong dependence.

Examples of processes with $\alpha = 0$ are the familiar autoregressive moving average ($ARMA(p, q)$) and Bloomfield's (1973) Exponential processes, whereas examples with $\alpha \in (0, 1)$, we can mention the $FARIMA(p, \alpha/2, q)$ and the Bloomfield's fractional integrated Exponential model, see Granger and Joyeux (1980) and Hosking (1981) and Robinson (1994) respectively. Thus, our framework simultaneously allows for both weak and strongly dependent processes. The latter two models have a spectral density function defined as

$$f_x(\lambda) = \frac{\sigma_\varepsilon^2}{2\pi} \left|1 - e^{i\lambda}\right|^{-\alpha} \left| \frac{\Theta(e^{i\lambda})}{\Phi(e^{i\lambda})} \right|^2, \; -\pi < \lambda \leq \pi,$$

where $\Phi(\cdot)$ and $\Theta(\cdot)$ are the AR and MA polynomials, respectively, having no zeroes in or on the unit circle, and

$$f_x(\lambda) = \left|1 - e^{i\lambda}\right|^{-\alpha} \exp\left[\sum_{k=1}^{p-1} \beta_j \cos\{(k-1)\lambda\}\right], \; -\pi < \lambda \leq \pi, \tag{2.4}$$

respectively.

As an earlier example of a process exhibiting the property of strong dependence is the Gaussian fractional noise model introduced by Mandelbrot and Van Ness (1968), and whose spectral density function, obtained by Sinai (1976), is

$$f_x(\lambda) = \frac{4\sigma_x^2 \Gamma(\alpha)}{(2\pi)^{3+\alpha}} \cos(\pi\alpha/2) \sin^2(\lambda/2) \sum_{j=-\infty}^{\infty} \left| j + \frac{\lambda}{2\pi} \right|^{-2-\alpha}, \quad -\pi < \lambda \leq \pi,$$

where $\sigma_x^2 = Var(x_t)$ and $\Gamma(\cdot)$ is the gamma function. A common feature of all the above models is that their spectral density function satisfies (2.3).

Given observations $\{x_{n-j}, j = 1, 2, ...\}$ on the infinite past of the series x_t, let the linear predictor of x_{n+h} $(h = 0, 1, ...)$ be denoted by $\widehat{x}_{n+h}$ and the mean-square prediction error by σ_{h+1}^2. Then

$$\widehat{x}_n = -\sum_{u=1}^{\infty} a_u x_{n-u} \quad \text{and} \quad \widehat{x}_{n+h} = -\sum_{u=1}^{h} a_u \widehat{x}_{n+h-u} - \sum_{u=1}^{\infty} a_{u+h} x_{n-u} \tag{2.5}$$

with

$$\sigma_{h+1}^2 = \sigma_\varepsilon^2 \sum_{u=0}^{h} b_u^2. \tag{2.6}$$

We notice that as $h \to \infty$, the mean-square prediction error approaches the variance of x_t. So, as $h \to \infty$, knowledge of the past does not help to predict future values.

It is clear that if the coefficients a_j were known, the prediction problem would be solved. Similarly, if $f_x(\lambda)$ was known, these coefficients a_j could be obtained by using the canonical factorization of the spectral density function, see Whittle $(1963, p.26)$ or Brillinger's (1981) Theorem 3.8.4. Specifically, we have that

$$a_j = \frac{1}{2\pi} \int_{-\pi}^{\pi} A(\lambda) e^{ij\lambda} d\lambda, \tag{2.7}$$

$$\sigma_\varepsilon^2 = 2\pi e^{c_0}, \tag{2.8}$$

where

$$A(\lambda) = \exp\left\{ -\sum_{u=1}^{\infty} c_u e^{-iu\lambda} \right\} \tag{2.9}$$

and

$$c_u = \frac{1}{\pi} \int_0^{\pi} \log(f_x(\lambda)) \cos(u\lambda) d\lambda. \tag{2.10}$$

The coefficients b_j can likewise be obtained from $B(\lambda) = A^{-1}(\lambda)$. In practice $f_x(\lambda)$ is unknown, so to apply (2.10) and therefore equations (2.7) – (2.10), $f_x(\lambda)$ needs to be estimated.

To that end, introduce the periodogram of x_t

$$I_x(\lambda) = (2\pi n)^{-1} \left| \sum_{t=1}^{n} x_t e^{-it\lambda} \right|^2.$$

We estimate $f_x(\lambda)$ by

$$\widehat{f}_x(\lambda) = \frac{|\lambda|^{-\widehat{\alpha}}}{2m+1} \sum_j |\lambda_j + \lambda|^{\widehat{\alpha}} I_x(\lambda + \lambda_j) \tag{2.11}$$

where $\lambda_j = (2\pi j)/n$, $j = 0, 1, ..., n-1$, $m = m(n)$ a number which increases slowly with n, $\sum_j = \sum_{j=-m}^{m}$, and $\widehat{\alpha}$ is a semiparametric estimator of α, for instance that obtained in Robinson (1995).

Thus $I_x(\lambda)$ is damped around zero frequency prior to the usual periodogram averaging (which is of the sort stressed by Brillinger, 1981), whereas $\widehat{f}_x$ will typically exhibit a pole at zero frequency. We can regard the estimator in (2.11) as a prewhitened estimator in the frequency domain, in contrast that in the time domain suggested in Press and Tukey (1956) when f is believed to have sharp peaks, as is our case. Moreover, see Lemma 1 in Section 6, that this estimator will have good bias properties compared to the usual average periodogram estimate.

Let $\widetilde{\lambda}_j = (\pi j)/M$, $j = 0, \pm 1, ..., \pm M$, where $M = [n/4m]$ and $[a]$ indicates the integer part of a. Abbreviating $\phi\left(\widetilde{\lambda}_\ell\right)$ by ϕ_ℓ, for a generic function $\phi(\lambda)$, (2.7)−(2.10) are then estimated by

$$\widehat{c}_u = \frac{1}{M}\sum_{\ell=1}^{M-1}\left(\log \widehat{f}_{x,\ell}\right)\cos\left(u\widetilde{\lambda}_\ell\right), \qquad u = 0, 1, ..., M-1, \tag{2.12}$$

$$\widehat{A}_j = \exp\left\{-\sum_{u=1}^{M-1}\widehat{c}_u e^{-iu\widetilde{\lambda}_j}\right\} = \overline{\widehat{A}_{-j}}, \quad j = 0, 1, ..., M-1, \tag{2.13}$$

$$\widehat{a}_u = \frac{1}{2M}\sum_{j=-M+1}^{M}\widehat{A}_j e^{iu\widetilde{\lambda}_j}, \qquad u = 1, ..., M-1, \tag{2.14}$$

$$\widehat{\sigma}_\varepsilon^2 = 2\pi e^{\widehat{c}_0}, \tag{2.15}$$

where $\overline{d}$ denotes the conjugate of the complex number d. Thus, (2.5) is estimated by

$$\widehat{x}_n = -\sum_{u=1}^{M-1}\widehat{a}_u\widetilde{x}_{n-u} \quad \text{and} \quad \widehat{x}_{n+h} = -\sum_{u=1}^{h}\widehat{a}_u\widehat{x}_{n+h-u} - \sum_{u=1}^{M-h}\widehat{a}_{u+h}\widetilde{x}_{n-u} \tag{2.16}$$

where $\{\widetilde{x}_{n-j}, j = 1, 2, ..., M-1\}$ are a new set of observations, with the same statistical properties as x_t, not used in the estimation of the spectral density $f_x(\lambda)$.

To finish this section, it should be noted two points. Firstly, since we only have a finite record of x_t, $\widehat{A}_j$, $\widehat{a}_u$ and $\widehat{\sigma}_\varepsilon^2$ do not actually estimate the true functions/parameters $A(\lambda)$, a_u and σ_ε^2 respectively, but rather their finite parameter versions $A_n(\lambda)$, $a_{u,n}$ and $\sigma_{\varepsilon,n}^2$. The latter can be obtained by replacing $\widehat{c}_u$ by $c_{u,n}$, $u = 0, 1, ..., M-1$ in (2.13) and (2.15), and $\widehat{A}_j$ by $A_n(\lambda)$ in (2.14). $c_{u,n}$ is obtained from (2.12) by replacing $\log\left(\widehat{f}_{x,\ell}\right)$ by $\log(f_{x,\ell})$. Following Bhansali (1974), we refer to the above method of prediction, c.f. (2.16), as the $FLES$ predictor of the data.

The second point is that, following the results of next section, the $FLES$, and in particular (2.12), provides $root-n$ consistent estimators for the parameters β_j in (2.4) and an initial $root-m$ consistent estimator for the parameter α. See also comments after Theorem 3.1 below.

3. ASYMPTOTIC PROPERTIES OF THE ESTIMATES

Before we establish the asymptotic properties of our estimates given in (2.12) – (2.15), and thus the predicted future value of x_t in (2.16), we introduce the following conditions:

C.1 $f_x(\lambda) = \lambda^{-\alpha} g_x(\lambda)$, $0 < \lambda \leq \pi$, where $0 \leq \alpha < 1$ and $g_x(\lambda)$ is a positive, symmetric around zero and twice continuously differentiable function.

C.2

$$x_t = \sum_{j=0}^{\infty} b_j \varepsilon_{t-j}, \quad \sum_{j=0}^{\infty} b_j^2 < \infty, \quad b_0 = 1,$$

where $E(\varepsilon_t | \mathcal{F}_{t-1}) = 0$, $E(\varepsilon_t^2 | \mathcal{F}_{t-1}) = \sigma_\varepsilon^2$, $E(\varepsilon_t^\ell | \mathcal{F}_{t-1}) = \mu_\ell < \infty$, for $\ell = 3$ and 4, and where $\mathcal{F}_t$ the $\sigma-$algebra of events generated by ε_s $s \leq t$, with joint fourth cumulant of $\varepsilon_{t_1}, \varepsilon_{t_2}, \varepsilon_{t_3}$ and ε_{t_4} satisfying

$$cum(\varepsilon_{t_1}, \varepsilon_{t_2}, \varepsilon_{t_3}, \varepsilon_{t_4}) = \begin{cases} \kappa_4, & \text{if } t_1 = t_2 = t_3 = t_4, \\ 0, & \text{otherwise.} \end{cases}$$

In addition, x_t^4 is uniformly integrable.

C.3 $B(\lambda)$ is twice continuously differentiable in any open set outside the origin, and satisfies

$$\frac{\partial}{\partial \lambda} |B(\lambda)| = O\left(\lambda^{-1} |B(\lambda)|\right) \text{ as } \lambda \to 0+.$$

C.4 As $n \to \infty$, $m^4/n^3 + n^2/m^3 \to 0$.

Some comments about Conditions $C.1 - C.4$ are in place. Condition $C.1$ and $C.3$ is common when analyzing processes which may exhibit strong dependence, see Robinson (1995), Hidalgo (1998) or Hidalgo and Yajima (1999), so their comments apply here. A sufficient condition for the first part of $C.2$ is $\sup_t E|x_t|^{4+c} < \infty$ for some $c > 0$. This last part of $C.2$ is needed in the proof of Theorem 3.4, and in particular to justify that, for example, $E(Y_n) \to E(Y)$ if $Y_n \xrightarrow{d} Y$ and Y_n is uniformly integrable, see Serfling (1980, p. 14). Finally, Condition $C.4$ gives upper and lower bounds on the rate of increase to infinity of the smoothing parameter m.

Theorem 3.1. *Define* $\widehat{\zeta}_j = \widehat{c}_j - c_{j,n}$. *Assuming C.1-C.4, for any finite collection* $j_1, ..., j_q$,

$$n^{1/2} \left(\widehat{\zeta}_{j_1}, ..., \widehat{\zeta}_{j_q}\right) \xrightarrow{d} N(0, \Omega_c)$$

where the $ij-th$ *element of* Ω_c *is* δ_{i-j} *and* $\delta_i = 1$ *if* $i = 0$ *and* $= 0$ *otherwise.*

Theorem 3.1 indicates that the results obtained in Bhansali (1974) for weakly dependent data, that is $\alpha = 0$, are also applicable for strongly dependent data. Thus, Theorem 3.1 generalizes Bhansali's results to any covariance stationary linear process.

Moreover, if it was known that x_t followed a Bloomfield's exponential model

$$f_x(\lambda) = \exp\left[\sum_{k=1}^{p-1} \beta_j \cos\{(k-1)\lambda\}\right], \quad -\pi < \lambda \leq \pi,$$

it is easily shown that $c_{j,n} - c_j = o\left(n^{-1/2}\right)$, so that we would obtain n-root consistent estimators of β_j. Additionally, if x_t followed the model (2.4), then the parameter α would be root-m consistency by some semiparametric estimator such as that provided in Robinson (1995), whereas the parameters β_j would still be n-root consistently estimated if

$$\widehat{f}_x(\lambda) = \frac{1}{2m+1}\sum_j \left|1 - e^{i(\lambda_j+\lambda)}\right|^{\widehat{\alpha}} I_x(\lambda+\lambda_j)$$

were used instead of (2.11).

>From the above theorem and a simple application of delta methods, we obtain

Corollary 1. *Let $\sigma^2_{\varepsilon,n} = 2\pi e^{c_{0,n}}$. Assuming C.1-C.4,*

$$n^{1/2}\left(\widehat{\sigma}^2_{\varepsilon} - \sigma^2_{\varepsilon,n}\right) \xrightarrow{d} N\left(0, \left(2\sigma^4_{\varepsilon} + \kappa_4\right)\right).$$

Theorem 3.2. *Define $\widehat{\zeta}_j = \widehat{A}_j - A_{j,n}$. Assuming C.1-C.4, for any finite collection $j_1, ..., j_q$,*

$$n^{1/2} M^{-1/2}\left(\widehat{\zeta}_{j_1}, ..., \widehat{\zeta}_{j_q}\right) \xrightarrow{d} N(0, \Omega_A)$$

where $\Omega_{A,i,j} = \left(\phi_{|j-i|} + K^2 \phi_j \phi_i\right) A_i A_j$ is the $ij-th$ element of Ω_A with $\delta_j + (1-\cos(j\pi))/j\pi = \phi_j$, δ_j as defined in Theorem 3.1 and $K = \lim_{M\to\infty} \sum_{\ell=1}^{M-1} g_{x,\ell}\left(\int_{-1}^{1} \log\left(1 + \frac{v}{2\ell}\right) dv\right)$.

Theorem 3.3. *Define $\widehat{\zeta}_j = \widehat{a}_j - a_{j,n}$. Assuming C.1-C.4, for any finite collection $j_1, ..., j_q$,*

$$n^{1/2}\left(\widehat{\zeta}_{j_1}, ..., \widehat{\zeta}_{j_q}\right) \xrightarrow{d} N(0, \Omega_a)$$

where the $\ell j - th$ element of Ω_a is $2^{-1}\int_{-\pi}^{\pi} |A(\lambda)|^2 e^{i\lambda(\ell-j)} d\lambda$.

>From the results of Theorem 3.3, we observe that the asymptotic covariance matrix is the same as that obtained for different, although related problems. More specifically, in the estimation of the parameters of a distributed lag regression model, see for instance Hannan (1967) or Brillinger (1981) for weakly dependent data, when the order of the polynomial is unknown or infinite, and Hidalgo (1998) for its extension to strongly dependent data.

Once we have obtained the asymptotic properties of the estimators of a_j, $j = 1, ..., M-1$, we are in the position to study the asymptotic properties of the predictor of the data.

Theorem 3.4. *Assuming C.1-C.4,*

$$\begin{aligned} (a)\quad AE\left(\widehat{x}_n - \widetilde{x}_n\right)^2 &\to \sigma^2_{\varepsilon} \\ (b)\quad AE\left(\widehat{x}_{n+h} - \widetilde{x}_{n+h}\right)^2 &\to \sigma^2_{h+1}, \end{aligned}$$

for $h = 1, ..., V$, with $V > 1$ and where AE denotes the asymptotic expectation.

Theorem 3.4 illustrates that once again the results obtained by Bhansali (1974) for weakly dependent data extrapolates to data which may exhibit strong dependence.

4. SIGNAL EXTRACTION

4.1. Statement of the problem and parametric estimation of the signal

The problem that we are interested in is as follows. Given a covariance stationary linear process y_t, which is observed at times $t = 1, ..., n$, it is decomposed as

$$y_t = x_t + z_t, \quad t = 0, \pm 1, \pm 2, ... \tag{4.1}$$

where x_t and z_t denote the signal and noise respectively. The purpose of this section is, given the observed data y_t, $t = 1, .., n$, to estimate the signal x_t. Denote the estimate by

$$\widehat{x}_{t|\infty} = E\left[x_t \,\widehat{|y_1, ..., y_n}\right].$$

We assume that x_t and z_t satisfy

$$(1 - L)^{\alpha/2} x_t = \varepsilon_t^x \tag{4.2}$$

where $\alpha \in (0, 1)$ and

$$z_t = \sum_{j=0}^{\infty} a_j^z \varepsilon_{t-j}^z, \ \sum_{j=0}^{\infty} \left|a_j^z\right| < \infty, \ \ a_0^z = 1, \tag{4.3}$$

respectively, where ε_t^x and ε_t^z are mutually independent processes satisfying the same conditions given in $C.2$ for ε_t. So, the spectral density function of y_t is

$$f_y(\lambda) = f_x(\lambda) + f_z(\lambda) \tag{4.4}$$

where $f_w(\lambda)$ denotes the spectral density function of a generic covariance stationary linear process w_t. As an example, suppose that (4.2) follows an $ARMA(p, q)$ process, then

$$f_y(\lambda) = \frac{1}{2\pi}\left(\sigma_{\varepsilon^x}^2 \left|1 - e^{i\lambda}\right|^{-\alpha} + \sigma_{\varepsilon^z}^2 \left|\frac{\Theta\left(e^{i\lambda}\right)}{\Phi\left(e^{i\lambda}\right)}\right|^2\right),$$

where $\Phi(L)$ and $\Theta(L)$ are the AR and MA polynomials, respectively, with non common roots and having no zeroes in or on the unit circle.

It is known that if all the past and future values of y_t were observed, then the Kolmogorov-Wiener formula would provide the best linear predictor (BLP) of x_t given y_s, $s = 0, \pm 1, \pm 2, ...$, that is, $x_{t|\infty} = E\left[x_t \left|y_s, s = 0, \pm 1, \pm 2, ...\right.\right]$, and defined by

$$x_{t|\infty} = \frac{\sigma_{\varepsilon^x}^2 (1 - L)^{-\alpha/2}\left(1 - L^{-1}\right)^{-\alpha/2}}{\sigma_{\varepsilon^x}^2 (1 - L)^{-\alpha/2}\left(1 - L^{-1}\right)^{-\alpha/2} + \sigma_{\varepsilon^z}^2 \frac{\Theta(L)\Theta(L^{-1})}{\Phi(L)\Phi(L^{-1})}} y_t \tag{4.5}$$

where L and L^{-1} are the backward and forward operators, respectively. However, in empirical studies, because not all values of y_t are observed, some truncation will be needed when (4.5) is implemented, so as to obtain $\widehat{x}_{t|\infty}$.

When a full parameterization of the process generating z_t is known, for example z_t follows an $ARMA(p, q)$, where $\Theta(L) = \Theta(L; \vartheta)$ and $\Phi(L) = \Phi(L; \vartheta)$, then $f_y(\lambda) = f_y(\lambda; \tau)$, for all $\lambda \in (0, \pi]$, where $\tau = \left(\alpha, \sigma_{\varepsilon^x}^2, \sigma_{\varepsilon^z}^2, \vartheta'\right)'$. If the parameters τ were known, the signal extraction problem would be solved by plugging those values of τ into the right side of (4.5) to obtain

$$x_{t|\infty} = \sum_{j=-\infty}^{\infty} \psi_j(\tau)\, y_{t+j} \tag{4.6}$$

where $\psi_j(\tau)$ denotes the jth coefficient in the expansion of

$$\frac{\sigma^2_{\varepsilon^x}(1-L)^{-\alpha/2}\left(1-L^{-1}\right)^{-\alpha/2}}{\sigma^2_{\varepsilon^x}(1-L)^{-\alpha/2}\left(1-L^{-1}\right)^{-\alpha/2}+\sigma^2_{\varepsilon^z}\frac{\Theta(L;\vartheta)\Theta(L^{-1};\vartheta)}{\Phi(L;\vartheta)\Phi(L^{-1};\vartheta)}}. \tag{4.7}$$

However, in practice τ is unknown, and thus to implement (4.6), τ is replaced by the Whittle estimate $\widetilde{\tau}$ which, under suitable conditions, is known to be $n^{1/2}$-consistent and asymptotically normal, see Fox and Taqqu (1986), Dahlhaus (1989), Giraitis and Surgailis (1990) or Hosoya (1997) among others. Given $\widetilde{\tau}$, $x_{t|\infty}$ is then estimated by plugging $\widetilde{\tau}$ into the right side of (4.6), obtaining

$$\widetilde{x}_{t|\infty} = \sum_{j=-\infty}^{\infty} \psi_j(\widetilde{\tau})\, y_{t+j}.$$

This parametric approach suffers from two possible drawbacks. First, its implementation can be difficult since to obtain the "reduced" form parameters in the denominator of (4.7) can be quite complex, even in simpler situations than those considered here. The second, and possibly more important, drawback is that the procedure described is very sensitive to a correct specification of the spectral density function of y_t, that is $f_y(\lambda)$. In particular, on a correct specification of the order of the polynomials $\Phi(L)$ and $\Theta(L)$ if indeed the noise z_t followed an $ARMA$ process, which might not even be the case. For instance, z_t may follow a Bloomfield's exponential model instead of an $ARMA$ one. If that was the case it would lead to inconsistent estimates of $f_y(\lambda)$ and so, the estimates of $x_{t|\infty}$, $\widehat{x}_{t|\infty}$, would be inaccurate and "inconsistent".

Looking at equations (4.1), (4.2) and (4.3), the model can be regarded as semiparametric, in that only the term which we are interested in is parameterized. That is, we have a parametric model for the underlying structure of the signal x_t, while the model that the noise z_t follows is left unspecified. So, in the terminology employed in semiparametric statistics, we can consider the parameters $\left(\sigma^2_{\varepsilon^z}, \vartheta'\right)'$ as nuisance parameters. Thus, the question of interest is whether we can "estimate" the parametric part of the model, that is, to extract the signal x_t, in the presence of those nuisance parameters represented by the noise process z_t. This is answered in the next subsection.

4.2. Semiparametric estimation of the signal x_t

Our main concern lies in the estimation of $x_{t|\infty}$, that is $\widehat{x}_{t|\infty}$. Assuming that y_t follows (4.1), (4.2) and (4.3), and that y_t is a covariance stationary linear process which admits a representation as an infinite autoregressive model as in (2.2), then

$$x_{t|\infty} = \frac{\sigma^2_{\varepsilon^x}(1-L)^{-\alpha/2}\left(1-L^{-1}\right)^{-\alpha/2}}{\sigma^2_{\varepsilon^y}\left(1-\sum_{j=1}^{\infty}\beta_j L^j\right)^{-1}\left(1-\sum_{j=1}^{\infty}\beta_j L^{-j}\right)^{-1}} y_t,$$

where $\sigma^2_{\varepsilon^y}$ is the variance of the innovations ε^y_t in the Wold decomposition of y_t given by

$$y_t = \sum_{j=0}^{\infty} \upsilon_j \varepsilon^y_{t-j}, \quad \sum_{j=0}^{\infty} \upsilon_j^2 < \infty, \quad \upsilon_0 = 1,$$

where the coefficients υ_j and β_j satisfy the relation, with $\beta_0 = 1$,

$$\left(\sum_{j=0}^{\infty} \upsilon_j L^j \right) \left(\sum_{j=0}^{\infty} \beta_j L^j \right) = 1.$$

Observe that such a Wold decomposition is possible since $\int_{-\pi}^{\pi} (\log f_x(\lambda))\, d\lambda$ and $\int_{-\pi}^{\pi} (\log f_z(\lambda))\, d\lambda > -\infty$ imply that $\int_{-\pi}^{\pi} (\log f_y(\lambda))\, d\lambda > -\infty$ with the innovation ε^y_t satisfying the same conditions as ε_t in $C.2$ and

Thus, the results of Section 3 suggest to estimate $x_{t|\infty}$ by

$$\begin{aligned} \widehat{x}_{t|\infty} &= \frac{\widehat{\sigma}^2_{\varepsilon^x} (1-L)^{-\widehat{\alpha}/2} \left(1-L^{-1}\right)^{-\widehat{\alpha}/2}}{\widehat{\sigma}^2_{\varepsilon^y} \left(1-\sum_{j=1}^{N} \widehat{\beta}_j L^j\right)^{-1} \left(1-\sum_{j=1}^{N} \widehat{\beta}_j L^{-j}\right)^{-1}} y_t \qquad (4.8) \\ &= \frac{\widehat{\sigma}^2_{\varepsilon^x}}{\widehat{\sigma}^2_{\varepsilon^y}} \left(\left(1-\sum_{j=1}^{N} \widehat{\beta}_j L^j\right) (1-L)^{-\widehat{\alpha}/2} \left(1-\sum_{j=1}^{N} \widehat{\beta}_j L^{-j}\right) \left(1-L^{-1}\right)^{-\widehat{\alpha}/2} \right) y_t, \end{aligned}$$

where $\widehat{\sigma}^2_{\varepsilon^y}$, $\widehat{\beta}_j$, $j = 1, ..., N$, are estimated using the $FLES$ algorithm described in $(2.12)-(2.15)$, and where $\widehat{\sigma}^2_{\varepsilon^x}$ and $\widehat{\alpha}$ are estimates of the parameters $\sigma^2_{\varepsilon^x}$ and α. Thus, the problem to obtain $\widehat{x}_{t|\infty}$ in (4.8) is reduced to obtaining estimates of the latter two parameters.

Given our model (4.2), $f_x(\lambda) = \left(\sigma^2_{\varepsilon^x}/2\pi\right) \left|1-e^{i\lambda}\right|^{-\alpha}$ which together with (4.1) and (4.3), implies that $f_y(\lambda) \sim C\lambda^{-\alpha}$ as $\lambda \to 0+$, c.f. (2.3). To observe the latter claim, assume that z_t follows an $ARMA$ process for expositional simplicity. From (4.4), we have that

$$\begin{aligned} f_y(\lambda) &= \frac{1}{2\pi} \left(\sigma^2_{\varepsilon^x} \left|1-e^{i\lambda}\right|^{-\alpha} + \sigma^2_{\varepsilon^z} \left| \frac{\Theta\left(e^{i\lambda}\right)}{\Phi\left(e^{i\lambda}\right)} \right|^2 \right) \\ &\sim \frac{\sigma^2_{\varepsilon^x}}{2\pi} \lambda^{-\alpha} + O(1) \qquad \text{as } \lambda \to 0+ \end{aligned}$$

because $\left|\Theta\left(e^{i\lambda}\right)/\Phi\left(e^{i\lambda}\right)\right|^2$ is continuous for all $\lambda \in [0, \pi]$, and

$$\begin{aligned} \left|1-e^{i\lambda}\right|^{-\alpha} &= 2^{-\alpha} (\sin|\lambda/2|)^{-\alpha} \\ &\sim \lambda^{-\alpha}\left(1+C_1\lambda^2\right) \qquad \text{as } \lambda \to 0+, \end{aligned}$$

where C_1 is a finite positive constant. Thus, we conclude that

$$f_y(\lambda) \sim \frac{\sigma^2_{\varepsilon^x}}{2\pi} \lambda^{-\alpha} \qquad \text{as } \lambda \to 0+, \qquad (4.9)$$

so that $\sigma^2_{\varepsilon^x}/2\pi$ and α can be estimated as in Section 2, i.e. employing Robinson's (1995) estimator, noting that $\sigma^2_{\varepsilon^x}/2\pi$ is identified as C in (2.3).

Let us introduce the following condition,

C.5 $N^{-1} + N^{3+\alpha} m^{-2} \to 0$ and $NM^{-1} \leq 1$.

Observe that from C.4, we can choose N to be equal to M in $C.5$. However, we leave $C.5$ in its present form to give somehow more generality to the result in Theorem 4.1 below.

Theorem 4.1. *Assuming C.1-C.5, as* $n \to \infty$, $\widehat{x}_{t|\infty} - x_{t|\infty} \overset{P}{\to} 0$.

Theorem 4.1 indicates that the signal extraction algorithm, whose implementation is simple, is consistent. However, and more importantly, in the process of performing the signal extraction of x_t, there has been no need to specify any particular structure for the noise z_t. Thus, we have avoided the problem that a bad specification for the noise may induce on the extraction of the signal. The approach can thus be considered as semiparametric, where z_t, following the semiparametric terminology, is the nuisance "parameter" or function.

5. PROOFS

Proof of Theorem 3.1 By Wold device, it suffices to show that for finite constants φ_j,

$$n^{1/2} \sum_{j=p}^{q} \varphi_j \left(\widehat{c}_j - c_{j,n} \right) \overset{d}{\to} N \left(0, 2 \sum_{j=p}^{q} \varphi_j^2 \right). \tag{5.1}$$

>From the definitions of $c_{j,n}$ and $\widehat{c}_j$, a typical component on the left of (5.1) is

$$\left(\widehat{c}_j - c_{j,n} \right) = \frac{1}{M} \sum_{\ell=1}^{M-1} \left(\log \widehat{f}_{x,\ell} - \log \widetilde{f}_{x,\ell} \right) \cos \left(j \widetilde{\lambda}_\ell \right) + \left(\widetilde{c}_{j,n} - c_{j,n} \right), \tag{5.2}$$

where, for $j = 0, 1, ..., M-1$, $\widetilde{c}_{j,n} = M^{-1} \sum_{\ell=1}^{M-1} \left(\log \widetilde{f}_{x,\ell} \right) \cos \left(j \widetilde{\lambda}_\ell \right)$ with $\widetilde{f}_{x,\ell} = \lambda_{2m\ell}^{-\alpha} (2m+1)^{-1} \sum_p g_x \left(\lambda_{p+2m\ell} \right)$.

Now, the second term on the right of (5.2) is $o\left(n^{-1/2}\right)$ by Lemma 1 and $C.4$.

Next, since $\sup_\ell a_\ell^2 \leq \sum_\ell a_\ell^2$, $\sup_{\ell=1,..M} \left| \left(\widehat{f}_{x,\ell} - \widetilde{f}_{x,\ell} \right) / \widetilde{f}_{x,\ell} \right|^2$ is bounded by

$$\sum_{\ell=1}^{M} \left| \left(\widehat{f}_{x,\ell} - \widetilde{f}_{x,\ell} \right) / \widetilde{f}_{x,\ell} \right|^2 \leq 2 \sum_{\ell=1}^{M} \left| \left(\widehat{f}_{x,\ell} - \check{f}_{x,\ell} \right) / \widetilde{f}_{x,\ell} \right|^2 + 2 \sum_{\ell=1}^{M} \left| \left(\check{f}_{x,\ell} - \widetilde{f}_{x,\ell} \right) / \widetilde{f}_{x,\ell} \right|^2$$

where $\check{f}_{x,\ell} = \lambda_{2m\ell}^{-\alpha} (2m+1)^{-1} \sum_p \lambda_{j+2m\ell}^{\alpha} I_x \left(\lambda_{p+2m\ell} \right)$. Because by Lemmas 2 and 3 respectively and $C.4$, the right side of the last displayed inequality is $o_p(1)$, we have by Taylor expansion of $\log(x)$ that the first term on the right of (5.2) is

$$\frac{1}{M} \sum_{\ell=1}^{M-1} \left(\frac{\check{f}_{x,\ell} - \widetilde{f}_{x,\ell}}{\widetilde{f}_{x,\ell}} \right) \cos \left(j \widetilde{\lambda}_\ell \right) + \frac{1}{2M} \sum_{\ell=1}^{M-1} \left(\frac{\check{f}_{x,\ell} - \widetilde{f}_{x,\ell}}{\widetilde{f}_{x,\ell}} \right)^2 \cos \left(j \widetilde{\lambda}_\ell \right) (1 + o_p(1))$$

$$+ \frac{1}{M} \sum_{\ell=1}^{M-1} \left(\frac{\check{f}_{x,\ell} - \widehat{f}_{x,\ell}}{\widetilde{f}_{x,\ell}} \right) \cos \left(j \widetilde{\lambda}_\ell \right) + \frac{1}{2M} \sum_{\ell=1}^{M-1} \left(\frac{\check{f}_{x,\ell} - \widehat{f}_{x,\ell}}{\widetilde{f}_{x,\ell}} \right)^2 \cos \left(j \widetilde{\lambda}_\ell \right) (1 + o_p(1)), \tag{5.3}$$

where the $o_p(1)$ is uniformly in ℓ. The second term of (5.3) is $O_p\left(m^{-1}\right)$ by Lemma 2, whereas the third and fourth terms of (5.3) are $O_p\left(m^{-1/2}M^{-1}+m^{-1}\right)$ by $C.4$, Lemma 3 and $\left|\sum_{\ell=1}^{M} g_{x,\ell}\int_{-1}^{1}\log\left(1+\frac{v}{2\ell}\right)dv\right|<\infty$ after straightforward calculations. Thus, we conclude that the first term on the right of (5.2), i.e. $\widehat{c}_j-\widetilde{c}_{j,n}$, is

$$\frac{1}{M}\sum_{\ell=1}^{M-1}\left(\frac{\check{f}_{x,\ell}-\widetilde{f}_{x,\ell}}{\widetilde{f}_{x,\ell}}\right)\cos\left(j\widetilde{\lambda}_\ell\right)+O_p\left(m^{-1/2}M^{-1}+m^{-1}\right),$$

whose first term is

$$\frac{1}{M}\sum_{\ell=1}^{M-1}\left(\frac{\check{f}_{x,\ell}-\widetilde{f}_{x,\ell}}{f_{x,\ell}}\right)\cos\left(j\widetilde{\lambda}_\ell\right)+\frac{1}{M}\sum_{\ell=1}^{M-1}\left(\frac{\check{f}_{x,\ell}-\widetilde{f}_{x,\ell}}{f_{x,\ell}}\right)\left(\frac{f_{x,\ell}-\widetilde{f}_{x,\ell}}{\widetilde{f}_{x,\ell}}\right)\cos\left(j\widetilde{\lambda}_\ell\right). \tag{5.4}$$

The first absolute moment of the second term of (5.4) is bounded by

$$\frac{1}{M}\sum_{\ell=1}^{M-1}\left|\frac{f_{x,\ell}-\widetilde{f}_{x,\ell}}{f_{x,\ell}}\right|E\left|\frac{\check{f}_{x,\ell}-\widetilde{f}_{x,\ell}}{\widetilde{f}_{x,\ell}}\right|=O\left(\frac{1}{M^2m^{1/2}}\right)=o\left(n^{-1/2}\right)$$

by Lemma 2 and the proof of Lemma 1.

Thus, except negligible ending effects, by definition of $\widehat{f}_{x,\ell}$, $\widetilde{f}_{x,\ell}$ and $f_{x,\ell}$ and $C.4$

$$n^{1/2}\left(\widehat{c}_j-c_{j,n}\right)=\frac{n^{1/2}}{[n/2]}\sum_{s=1}^{[n/2]}\left(\frac{I_x\left(\lambda_s\right)-f_x\left(\lambda_s\right)}{f_x\left(\lambda_s\right)}\right)h_{j,n}(s)+o_p(1), \tag{5.5}$$

where $h_{j,n}(s)=g_{x,\ell}^{-1}g_x\left(\lambda_s\right)\cos\left(j\widetilde{\lambda}_\ell\right)$ if $(2\ell-1)m<s\leq(2\ell+1)m$ and $\ell=1,...,M-1$. That is, $h_{j,n}(s)$ is a step function in $[0,\pi]$. Now, by Lemmas 3.1 and 3.2 of Giraitis et al. (1999), which are a simple extension of the proof of (4.9) in Robinson (1995), after observing that $\left|h_{j,n}(s)\right|$ is a bounded function,

$$\sum_{s=1}^{[n/2]}\left(\frac{I_x\left(\lambda_s\right)}{f_x\left(\lambda_s\right)}-\frac{(2\pi)I_\varepsilon\left(\lambda_s\right)}{\sigma_\varepsilon^2}\right)h_{j,n}(s)=O_p\left(n^{1/3}\right) \tag{5.6}$$

where $I_\varepsilon\left(\lambda_s\right)$ denotes the periodogram of the innovations ε_t in (2.1). So, the right side of (5.5) is

$$\frac{n^{1/2}}{[n/2]}\sum_{s=1}^{[n/2]}\left(\frac{(2\pi)I_\varepsilon\left(\lambda_s\right)}{\sigma_\varepsilon^2}-1\right)h_{j,n}(s)+o_p(1),$$

and we conclude that (5.1) is

$$n^{1/2}\sum_{j=p}^{q}\varphi_j\left(\widehat{c}_j-c_{j,n}\right)=\frac{n^{1/2}}{[n/2]}\sum_{s=1}^{[n/2]}\left(\frac{(2\pi)I_\varepsilon\left(\lambda_s\right)}{\sigma_\varepsilon^2}-1\right)\sum_{j=p}^{q}\varphi_j h_{j,n}(s)+o_p(1).$$

But by an extension of Robinson's (1995) Theorem 3.2, see Giraitis et al. (1999), the right side converges in distribution to $N(0,V)$ where V is, by the definition of $h_{j,n}(s)$, $s=1,...,[n/2]$,

$$\sum_{j_1,j_2=p}^{q}\varphi_{j_1}\varphi_{j_2}\Omega_{c,j_1,j_2} = \lim_{n\to\infty}\frac{2}{M}\sum_{\ell=1}^{M-1}\sum_{j_1,j_2=p}^{q}\varphi_{j_1}\varphi_{j_2}\cos\left(j_1\widetilde{\lambda}_\ell\right)\cos\left(-j_2\widetilde{\lambda}_\ell\right)$$

$$
\begin{aligned}
& +\kappa_4 \left(\sum_{j=p}^{q} \varphi_{j_1} \lim_{n\to\infty} \frac{1}{M} \sum_{\ell=1}^{M-1} \cos\left(j\widetilde{\lambda}_\ell \right) \right)^2 \\
= \; & 2 \sum_{j_1,j_2=p}^{q} \varphi_{j_1}\varphi_{j_2} \int_0^1 \cos(\pi j_1 \lambda)\cos(-\pi j_2 \lambda)\, d\lambda \\
= \; & \sum_{j=p}^{q} \varphi_j^2,
\end{aligned}
$$

by elementary algebra and $g_{x,\ell}^{-1} g_x(\lambda_s) - 1 \to_{n\to\infty} 0$ by $C.1$ and $(2\ell-1)\,m < s \leq (2\ell+1)\,m$, for the second equality that $\cos(z)\cos(w)$ is a differentiable function, see Brillinger (1981, p.15) and that $M^{-1}\sum_{\ell=1}^{M-1}\cos\left(j\widetilde{\lambda}_\ell\right) \to \pi^{-1}\int_0^\pi \cos(ju)\,du = 0$ except for $j=0$ in which case is 1, whereas for the third equality we have used that the integral is $2^{-1}\delta_{j_1-j_2}$. ■

Proof of Theorem 3.2 By Wold device, it suffices to examine the behaviour of

$$
m^{1/2} \sum_{j=q}^{p} \varphi_j \left(\widehat{A}_j - A_{j,n} \right),
$$

for any set of constants φ_j such that $\sum_{j=q}^{p}\varphi_j^2 = 1$. Let $\widehat{d}_j = \log\left(\widehat{A}_j\right)$ and $\widetilde{d}_{j,n} = \log\left(\widetilde{A}_{j,n}\right)$, where $\widetilde{A}_{j,n} = \exp\left\{-\sum_{u=1}^{M-1}\widetilde{c}_{u,n}e^{-iu\widetilde{\lambda}_j}\right\}$. We start examining

$$
\widehat{d}_j - d_{j,n} = -\sum_{u=1}^{M-1}\left(\widehat{c}_u - \widetilde{c}_{u,n}\right)e^{-iu\widetilde{\lambda}_j} - \sum_{u=1}^{M-1}\left(\widetilde{c}_{u,n} - c_{u,n}\right)e^{-iu\widetilde{\lambda}_j}. \tag{5.7}
$$

The second term on the right of (5.7) is $o\left(m^{-1/2}\right)$, because from the definition of $\widetilde{c}_{u,n}$ and $c_{u,n}$, $\left(\widetilde{d}_{j,n} - d_{j,n}\right)$ is, by Taylor expansion of $\log\left(\widetilde{f}_{x,\ell}/f_{x,\ell}\right)$,

$$
\begin{aligned}
& \sum_{u=1}^{M-1}\frac{1}{M}\sum_{\ell=1}^{M}\left\{\left(\frac{\widetilde{f}_{x,\ell}-f_{x,\ell}}{f_{x,\ell}}\right) + \left(\frac{\widetilde{f}_{x,\ell}-f_{x,\ell}}{f_{x,\ell}}\right)^2 (1+o(1))\right\}\cos\left(u\widetilde{\lambda}_\ell\right)e^{-iu\widetilde{\lambda}_j} \\
= \; & \sum_{u=1}^{M-1}\frac{1}{M}\sum_{\ell=1}^{M}\left(\frac{\widetilde{f}_{x,\ell}-f_{x,\ell}}{f_{x,\ell}}\right)\cos\left(u\widetilde{\lambda}_\ell\right)e^{-iu\widetilde{\lambda}_j} + O\left(\frac{1}{M^3}\right) \\
= \; & \frac{1}{2M}\sum_{\ell=1}^{M}\left(\frac{\widetilde{f}_{x,\ell}-f_{x,\ell}}{f_{x,\ell}}\right)\sum_{u=1}^{M-1}\left(e^{iu\widetilde{\lambda}_{\ell-j}} + e^{iu\widetilde{\lambda}_{\ell+j}}\right) + O\left(\frac{1}{M^3}\right) = O\left(\frac{\log M}{M^2}\right)
\end{aligned}
$$

from the proof of Lemma 1 for the first equality and from the proof of Lemma 4 for the second equality. So, by $C.4$ we conclude that it is $o\left(m^{-1/2}\right)$.

Next, the first term on the right of (5.7). From the definition of $\widehat{c}_u - \widetilde{c}_{u,n}$ in (5.2) and its properties in (5.3), this term is

$$
-\sum_{u=1}^{M-1}\left(\frac{1}{M}\sum_{\ell=1}^{M-1}\left(\frac{\check{f}_{x,\ell}-\widetilde{f}_{x,\ell}}{\widetilde{f}_{x,\ell}}\right)\cos\left(u\widetilde{\lambda}_\ell\right) - \frac{1}{M}\sum_{\ell=1}^{M-1}\left(\frac{\widehat{f}_{x,\ell}-\check{f}_{x,\ell}}{\widetilde{f}_{x,\ell}}\right)\cos\left(u\widetilde{\lambda}_\ell\right)\right)e^{-iu\widetilde{\lambda}_j}
$$

$$+o_p\left(m^{-1/2}\right)$$

$$= -\sum_{u=1}^{M-1}\frac{1}{[n/2]}\sum_{\ell=1}^{[n/2]}\rho_\ell h_{u,n}(\ell)\, e^{-iu\widetilde{\lambda}_j}$$

$$-(\widehat{\alpha}-\alpha)\sum_{u=1}^{M-1}\frac{1}{M}\sum_{\ell=1}^{M-1} g_{x,\ell}\left(\int_{-1}^{1}\log\left(1+\frac{v}{2\ell}\right)dv\right)e^{-iu\widetilde{\lambda}_j}+o_p\left(m^{-1/2}\right),$$

by $C.4$ and Lemma 3, where $\rho_\ell = (2\pi)\,\sigma_\varepsilon^{-2} I_\varepsilon(\lambda_\ell) - 1$ and $h_{u,n}(\ell)$ as defined in Theorem 3.1. Thus,

$$m^{1/2}\left(\widehat{d}_j-\widetilde{d}_{j,n}\right) = -m^{1/2}\sum_{u=1}^{M-1}\left(\frac{1}{[n/2]}\sum_{\ell=1}^{[n/2]}\rho_\ell h_{u,n}(\ell)\right)e^{-iu\lambda_{2jm}} - \left(K\frac{1}{M}\sum_{u=1}^{M-1}e^{-iu\lambda_{2jm}}\right)m^{1/2}(\widehat{\alpha}-\alpha)+o_p(1). \tag{5.8}$$

since by standard calculations $\sum_{\ell=1}^{M-1} g_{x,\ell}\left(\int_{-1}^{1}\log\left(1+\frac{v}{2\ell}\right)dv\right) \to K$ as $M\to\infty$, where henceforth, K denotes a finite constant. But, the first term on the right of (5.8) is, because $m=[n/(4M)]$,

$$\frac{1}{[n/2]^{1/2}}\sum_{\ell=1}^{[n/2]}\rho_\ell M^{-1/2}\sum_{u=1}^{M-1}h_{u,n}(\ell)\,e^{-iu\lambda_{2jm}} = \frac{1}{[n/2]^{1/2}}\sum_{\ell=1}^{[n/2]}\rho_\ell\psi_{\ell,M(n)}(j)$$

where $\psi_{\ell,M(n)}(j) = M^{-1/2}\sum_{u=1}^{M-1}h_{u,n}(\ell)\,e^{-iu\lambda_{2jm}}$. Thus, $m^{1/2}\sum_{j=q}^{p}\varphi_j\left(\widehat{d}_j-\widetilde{d}_{j,n}\right)$ is

$$\frac{1}{[n/2]^{1/2}}\sum_{\ell=1}^{[n/2]}\rho_\ell\sum_{j=q}^{p}\varphi_j\psi_{\ell,M(n)}(j)+K\sum_{j=q}^{p}\varphi_j\phi_j m^{1/2}(\widehat{\alpha}-\alpha)+o_p(1). \tag{5.9}$$

Proceeding as in the proof of Theorem 3.1, the first term on the right of (5.8) converges in distribution to a complex normal random variable with variance

$$V = 2\lim_{n\to\infty}\sum_{j_1,j_2=q}^{p}\varphi_{j_1}\varphi_{j_2}\frac{1}{[n/2]}\sum_{\ell=1}^{[n/2]}\psi_{\ell,M(n)}(j_1)\,\psi_{\ell,M(n)}(j_2).$$

But, by definition of $h_{u,n}(\ell)$, the right side of the above equation is

$$2\lim_{n\to\infty}\sum_{j_1,j_2=q}^{p}\varphi_{j_1}\varphi_{j_2}\frac{1}{M^2}\sum_{\ell=1}^{M-1}\sum_{u_1=1}^{M-1}\sum_{u_2=1}^{M-1}\left\{\cos\left(u_1\widetilde{\lambda}_\ell\right)\cos\left(-u_2\widetilde{\lambda}_\ell\right)e^{-iu_1\widetilde{\lambda}_{j_1}+iu_2\widetilde{\lambda}_{j_2}}\right\}$$

$$= \sum_{j_1,j_2=q}^{p}\varphi_{j_1}\varphi_{j_2}\phi_{j_1-j_2},$$

by Lemma 4. Next, the second term on the right of (5.8) $\overset{d}{\to} N\left(0, K^2\left(\sum_{j=q}^{p}\varphi_j\phi_j\right)^2\right)$ by Robinson (1995), whereas the first and second terms of (5.8) are asymptotically independent since by Robinson (1995),

$$\left(\frac{1}{[n/2]^{1/2}}\sum_{\ell=1}^{[n/2]}\rho_\ell\right)m^{1/2}\left(\widehat{\alpha}-\alpha\right)=\frac{1}{[n/2]^{1/2}}\sum_{\ell=1}^{[n/2]}\rho_\ell\frac{1}{m^{1/2}}\sum_{\ell=1}^{m}v_\ell\rho_\ell+o_p\left(1\right),$$

where $v_\ell=\log\ell+m^{-1}\sum_{p=1}^{m}\log p$. But by Brillinger (1981), $Cov\left(I_\varepsilon(\lambda_{\ell_1}),I_\varepsilon(\lambda_{\ell_2})\right)=O\left(n^{-1}\log n\right)$ so that the expectation of the first term on the right side is $o(1)$. So, from the behaviour of (5.8) and the second term on the right of (5.7),

$$m^{1/2}\sum_{j=q}^{p}\varphi_j\left(\widehat{d}_j-d_{j,n}\right)\overset{d}{\to}N\left(0,\sum_{j_1,j_2=q}^{p}\varphi_{j_1}\varphi_{j_2}\left(\phi_{j_1-j_2}+K^2\phi_{j_1}\phi_{j_2}\right)\right).$$

But $\widehat{A}_j-A_{j,n}=\left(\exp\left(\widehat{d}_j-d_{j,n}\right)-1\right)A_{j,n}$ and $|A_{j,n}-A_j|\to 0$ from the proof of (6.6) in Lemma 6 which implies, by a simple application of delta methods

$$m^{1/2}\sum_{j=q}^{p}\varphi_j\left(\widehat{A}_j-A_{j,n}\right)\overset{d}{\to}N\left(0,\sum_{j_1,j_2=q}^{p}\varphi_{j_1}\Omega_{A,j_1,j_2}\varphi_{j_2}\right),$$

and the proof is completed. ■

Proof of Theorem 3.3 Write $\widetilde{a}_{v,n}=(2M)^{-1}\sum_{j=-M+1}^{M}\widetilde{A}_{j,n}e^{iv\widetilde{\lambda}_j}$. The proof is completed if

$$(a)\;\; n^{1/2}\sum_{v=p}^{q}\varphi_v\left(\widehat{a}_v-\widetilde{a}_{v,n}\right)\overset{d}{\to}N\left(0,\sum_{v_1,v_2=p}^{q}\varphi_{v_1}\varphi_{v_2}\Omega_{a,v_1v_2}\right) \tag{5.10}$$

$$(b)\;\; n^{1/2}\sum_{v=p}^{q}\varphi_v\left(\widetilde{a}_{v,n}-a_{n,v}\right)\to 0.$$

We begin with (a). From the definition of $\widehat{a}_v-\widetilde{a}_{v,n}$ and Taylor expansion of $\widehat{A}_j-\widetilde{A}_{j,n}$, a typical element on the left of (5.10) is

$$\frac{n^{1/2}}{2M}\sum_{j=-M}^{M-1}\left(\widehat{d}_j-\widetilde{d}_{j,n}\right)\widetilde{A}_{n,j}e^{iv\widetilde{\lambda}_j}+\frac{n^{1/2}}{4M}\sum_{j=-M}^{M-1}\left|\widehat{d}_j-\widetilde{d}_{j,n}\right|^2\left|\widetilde{A}_{n,j}\right|\left(1+o_p\left(1\right)\right). \tag{5.11}$$

First, since by Theorem 3.2 $m^{1/2}\left(\widehat{d}_j-\widetilde{d}_{j,n}\right)\overset{d}{\to}N\left(0,V\right)$ and x_t^4 is uniformly integrable by $C.2$, by Theorem A in Serfling (1980, p.14), $E\left|\widehat{d}_j-\widetilde{d}_{j,n}\right|^2=O\left(m^{-1}\right)$. So, from the definition of $\widehat{d}_j-\widetilde{d}_{j,n}$, (5.11) is

$$\begin{aligned}
&-\frac{n^{1/2}}{2M}\sum_{j=-M}^{M-1}\left(\sum_{u=1}^{M-1}\left(\widehat{c}_u-\widetilde{c}_{u,n}\right)e^{-iu\widetilde{\lambda}_j}\right)\widetilde{A}_{n,j}e^{iv\widetilde{\lambda}_j}+O_p\left(\frac{n^{1/2}}{m}\right)\\
=&-\frac{n^{1/2}}{2M}\sum_{j=-M}^{M-1}\left(\sum_{u=1}^{M-1}\left(\frac{1}{M}\sum_{\ell=1}^{M-1}\left(\frac{\widehat{f}_{x,\ell}-\widetilde{f}_{x,\ell}}{\widetilde{f}_{x,\ell}}\right)\cos\left(u\widetilde{\lambda}_\ell\right)\right)e^{-iu\widetilde{\lambda}_j}\right)\widetilde{A}_{n,j}e^{iv\widetilde{\lambda}_j}+o_p\left(1\right)\\
=&-\frac{n^{1/2}}{M}\sum_{\ell=1}^{M-1}\left(\frac{\widehat{f}_{x,\ell}-\widetilde{f}_{x,\ell}}{\widetilde{f}_{x,\ell}}\right)\sum_{u=1}^{M-1}\cos\left(u\widetilde{\lambda}_\ell\right)\left(\frac{1}{2M}\sum_{j=-M}^{M-1}\widetilde{A}_{n,j}e^{i(v-u)\widetilde{\lambda}_j}\right)+o_p\left(1\right)
\end{aligned}$$

$$= -\frac{n^{1/2}}{M}\sum_{\ell=1}^{M-1}\left(\frac{\widehat{f}_{x,\ell}-\widetilde{f}_{x,\ell}}{\widetilde{f}_{x,\ell}}\right)\sum_{u=1}^{M-1}\cos\left(u\widetilde{\lambda}_{\ell}\right)\left(\frac{1}{2M}\sum_{j=-M}^{M-1}A_j e^{i(v-u)\widetilde{\lambda}_j}\right)+o_p\left(1\right), \tag{5.12}$$

where for the first equality we have employed Taylor expansion and that by Proposition 3 of Hidalgo and Yajima (1999), $\widetilde{f}_{x,\ell}^{-1}\left(\widehat{f}_{x,\ell}-\widetilde{f}_{x,\ell}\right)=o_p\left(1\right)$ uniformly in ℓ and for the last equality that, by Lemma 6, $\left|\sum_{j=-M}^{M-1}\left(\widetilde{A}_{n,j}-A_j\right)\right|=O\left(1\right)$, $M^{-1}\sum_{u=1}^{M-1}\left|\cos\left(u\widetilde{\lambda}_{\ell}\right)\right|=O\left(\ell^{-1}\right)$, and Lemmas 2 and 3 and $C.4$.

Next, by the Euler-McLauring approximation of sums by integrals, see Brillinger (1981, p.15),

$$\left|\frac{1}{2M}\sum_{j=-M}^{M-1}A_j e^{i(v-u)\widetilde{\lambda}_j}-\frac{1}{2\pi}\int_{-\pi}^{\pi}A\left(\lambda\right)e^{i(v-u)\lambda}d\lambda-\frac{2}{M}\left(A\left(0\right)+A\left(\pi\right)\right)\right|\leq KM^{-1}.$$

So, proceeding as with the last term of (5.12), that expression is

$$\begin{aligned}
&-\frac{n^{1/2}}{M}\sum_{\ell=1}^{M-1}\left(\frac{\widehat{f}_{x,\ell}-\widetilde{f}_{x,\ell}}{\widetilde{f}_{x,\ell}}\right)\sum_{u=1}^{M-1}\cos\left(u\widetilde{\lambda}_{\ell}\right)\left(a_{v-u}+\frac{2}{M}\left(A\left(0\right)+A\left(\pi\right)\right)\right)+o_p\left(1\right)\\
=\;&-\frac{n^{1/2}}{M}\sum_{\ell=1}^{M-1}\left(\frac{\widehat{f}_{x,\ell}-\widetilde{f}_{x,\ell}}{\widetilde{f}_{x,\ell}}\right)\left(A_{\ell}e^{iv\widetilde{\lambda}_{\ell}}\left(1+o_p\left(1\right)\right)+\frac{2}{M}\left(A\left(0\right)+A\left(\pi\right)\right)\right), \qquad (5.13)
\end{aligned}$$

because $\sum_{u=1}^{M-1}\cos\left(u\widetilde{\lambda}_{\ell}\right)a_{v-u}-A_{\ell}e^{iv\widetilde{\lambda}_{\ell}}$ is

$$\begin{aligned}
&\frac{1}{2}\sum_{u=1}^{M-1}a_{v-u}e^{iu\widetilde{\lambda}_{\ell}}-A_{\ell}e^{iv\widetilde{\lambda}_{\ell}}=\left(\frac{1}{2}\sum_{|u|=1}^{M-1}a_u e^{iu\widetilde{\lambda}_{\ell}}-A_{\ell}\right)e^{iv\widetilde{\lambda}_{\ell}}+o\left(1\right)\\
=\;&e^{iv\widetilde{\lambda}_{\ell}}\frac{1}{2}\sum_{|u|=M+1}^{\infty}a_u e^{iu\widetilde{\lambda}_{\ell}}+o\left(1\right)=o\left(1\right)
\end{aligned}$$

uniformly in ℓ, as $a_u=O\left(u^{-1-\alpha/2}\right)$. But the right hand side of (5.13) is dominated by the contribution of $A_{\ell}e^{iv\widetilde{\lambda}_{\ell}}$, so is

$$\left(-\frac{n^{1/2}}{2M}\sum_{\ell=1}^{M-1}\left(\frac{\widehat{f}_{x,\ell}-\check{f}_{x,\ell}}{\widetilde{f}_{x,\ell}}\right)A_{\ell}e^{iv\widetilde{\lambda}_{\ell}}-\frac{n^{1/2}}{2M}\sum_{\ell=1}^{M-1}\left(\frac{\check{f}_{x,\ell}-\widetilde{f}_{x,\ell}}{\widetilde{f}_{x,\ell}}\right)A_{\ell}e^{iv\widetilde{\lambda}_{\ell}}\right)\left(1+o\left(1\right)\right).$$

The second term proceeding as with the third term of (5.3) is $O_p\left(n^{1/2}m^{-1/2}M^{-1}\right)=O_p\left(M^{-1/2}\right)$, whereas the first term of the last displayed expression, proceeding as in the derivation of (5.6) from (5.4) and (5.5), is

$$-\frac{1}{2n^{1/2}}\sum_{\ell=1}^{n/2}\left(\frac{\left(2\pi\right)I_{\varepsilon}(\lambda_{\ell})}{\sigma_{\varepsilon}^2}-1\right)A\left(\lambda_{\ell}\right)e^{iv\lambda_{\ell}}\left(1+o_p\left(1\right)\right)+O_p\left(n^{-1/6}\right).$$

Thus,

$$n^{1/2}\sum_{v=p}^{q}\varphi_v\left(\widehat{a}_v-\widetilde{a}_{v,n}\right)\;=\;\sum_{v=p}^{q}\varphi_v\frac{1}{2n^{1/2}}\sum_{\ell=1}^{n/2}\left(\frac{\left(2\pi\right)I_{\varepsilon}(\lambda_{\ell})}{\sigma_{\varepsilon}^2}-1\right)A\left(\lambda_{\ell}\right)e^{iv\lambda_{\ell}}+o_p\left(1\right)$$

$$\xrightarrow{d} N\left(0, \sum_{v_1,v_2=p}^{q} \varphi_{v_1}\varphi_{v_2}\Omega_{a,v_1v_2}\right)$$

proceeding as in the proof of Theorem 3.1. To complete the proof, we need to show (b), which follows by identical arguments to those in (5.12) and Lemma 1. ■

Proof of Theorem 3.4 (a) From the definition of $\widehat{x}_n$ and $\widetilde{x}_n$, their difference is

$$\varepsilon_t - \sum_{u=M}^{\infty} a_u \widetilde{x}_{n-u} + \sum_{u=1}^{M-1} \left(\widehat{a}_u - a_u\right) \widetilde{x}_{n-u}. \tag{5.14}$$

The second moment of the second term of (5.14) is

$$\sigma^2_{\widetilde{x}} \sum_{u=M}^{\infty} a_u^2 + 2 \sum_{M=u_1<u_2}^{\infty} a_{u_1} a_{u_2} \gamma_{\widetilde{x}}\left(u_1 - u_2\right) = O\left(M^{-1}\right).$$

because $a_u = O(u^{-1-\alpha/2})$ and $\gamma_{\widetilde{x}}(u) \sim Cu^{-1+\alpha}$, $|C| < \infty$ by C.1 and C.2. The second moment of the third term of (5.14) is

$$\sigma^2_{\widetilde{x}} \sum_{u=1}^{M-1} E\left(\widehat{a}_u - a_u\right)^2 + 2 \sum_{1=u_1<u_2}^{M-1} E\left[\left(\widehat{a}_{u_1} - a_{u_1}\right)\left(\widehat{a}_{u_2} - a_{u_2}\right)\right] \gamma_{\widetilde{x}}\left(u_1 - u_2\right) = O\left(M^{\alpha-1}\right),$$

after elementary algebra by the Cauchy-Schwarz inequality and Theorem 3.3 and Lemma 6 and since by uniform integrability of $\widehat{a}_u - a_u$ then by Serfling (1980, p.14). From here, the conclusion is immediate since $E\left(\varepsilon_t^2\right) = \sigma^2_{\varepsilon}$.

(b) By definition

$$\widetilde{x}_{n+h} = \sum_{j=0}^{h} b_j \varepsilon_{t+h-j} + \sum_{u=h+1}^{\infty} \left(\sum_{\ell=0}^{h} b_\ell a_{u-\ell}\right) \widetilde{x}_{t+h-u},$$

where b_j, $j = 0, 1, ...,$ satisfies $\left(\sum_{j=0}^{\infty} b_j L^j\right)\left(\sum_{j=0}^{\infty} a_j L^j\right) = 1$. Thus,

$$\begin{aligned} \widetilde{x}_{n+h} - \widehat{x}_{n+h} &= \sum_{j=0}^{h} b_j \varepsilon_{t+h-j} + \sum_{u=h+1}^{M} \left[\left(\sum_{\ell=0}^{h} b_\ell a_{u-\ell}\right) - \left(\sum_{\ell=0}^{h} \widehat{b}_\ell \widehat{a}_{u-\ell}\right)\right] \widetilde{x}_{t+h-u} \\ &\quad + \sum_{u=M+1}^{\infty} \left(\sum_{\ell=0}^{h} b_\ell a_{u-\ell}\right) \widetilde{x}_{t+h-u}. \end{aligned}$$

Now, by similar arguments to those used for the last two terms of (5.14), the last two terms on the right of the last equation are $o_p(1)$. So,

$$\widetilde{x}_{n+h} - \widehat{x}_{n+h} = \sum_{j=0}^{h} b_j \varepsilon_{t+h-j} + o_p(1).$$

>From here the conclusion follows by C.2. ■

Proof of Theorem 4.1 Defining $\widehat{\phi}_{j,n}$ from the relation
$\widehat{\sigma}_{\varepsilon^x}\left(1-\sum_{j=1}^{N}\widehat{\beta}_j L^j\right)(1-L)^{-\widehat{\alpha}/2}=\widehat{\sigma}_{\varepsilon^y}\sum_{j=1}^{\infty}\widehat{\phi}_{j,n}L^j$, the right side of (4.8) is

$$\widehat{x}_{t|\infty}=\left(\left(\sum_{j=1}^{N}\widehat{\phi}_{j,n}L^j\right)\left(\sum_{j=1}^{N}\widehat{\phi}_{j,n}L^{-j}\right)\right)y_t.$$

>From Theorem 3.3 and Corollary 1,

$$\widehat{\beta}_j-\widetilde{\beta}_{j,n}=O_p\left(n^{-1/2}\right)\text{ and }\widehat{\sigma}^2_{\varepsilon^y}-\sigma^2_{\varepsilon^y,n}=O_p\left(n^{-1/2}\right)$$

and by definition of $c_{0,n}$ and c_0 is easy to prove by Brillinger (1981, p.15) that $\sigma^2_{\varepsilon^y,n}-\sigma^2_{\varepsilon^y}=O\left(M^{-1}\right)$. Next, by Robinson (1995),

$$\widehat{\sigma}^2_{\varepsilon^x}-\sigma^2_{\varepsilon^x}=O_p\left(m^{-1/2}\right)\text{ and }(\widehat{\alpha}-\alpha)=O_p\left(m^{-1/2}\right),$$

which implies by continuity of $\widehat{\phi}_{j,n}$ with respect to $\widehat{\beta}_j$, $\widehat{\alpha}$, $\widehat{\sigma}^2_{\varepsilon^x}$ and $\widehat{\sigma}^2_{\varepsilon^y}$ that

$$\widehat{\phi}_{j,n}-\widetilde{\phi}_{j,n}=O_p\left(m^{-1/2}\right), \tag{5.15}$$

where $\widetilde{\phi}_{j,n}$ is defined from the relation
$\sigma_{\varepsilon^x}\left(1-\sum_{j=1}^{N}\widetilde{\beta}_{j,n}L^j\right)(1-L)^{-\alpha/2}=\sigma_{\varepsilon^y}\sum_{j=1}^{\infty}\widetilde{\phi}_{j,n}L^j$. Thus, $\widehat{x}_{t|\infty}-\widetilde{x}_{t|\infty}$ is

$$\left(\sum_{j=1}^{N}\left(\widehat{\phi}_{j,n}-\widetilde{\phi}_{j,n}\right)L^j\right)\left(\sum_{j=1}^{N}\left(\widehat{\phi}_{j,n}-\widetilde{\phi}_{j,n}\right)L^{-j}\right)y_t=\sum_{j=-N+1}^{N-1}\left(\widehat{\psi}_{j,n}-\widetilde{\psi}_{j,n}\right)y_{t-j}, \tag{5.16}$$

where

$$\widetilde{x}_{t|\infty}=\left(\left(\sum_{j=1}^{N}\widetilde{\phi}_{j,n}L^j\right)\left(\sum_{j=1}^{N}\widetilde{\phi}_{j,n}L^{-j}\right)\right)y_t.$$

The second moment of the right of (5.16) is

$$\sum_{j_1,j_2=-N+1}^{N-1}E\left[\left(\widehat{\psi}_{j_1,n}-\widetilde{\psi}_{j_1,n}\right)\left(\widehat{\psi}_{j_2,n}-\widetilde{\psi}_{j_2,n}\right)\right]E\left(y_{t-j_1}y_{t-j_2}\right)=O\left(m^{-2}N^{3+\alpha}\right),$$

since $E\left|\widehat{\psi}_{j,n}-\widetilde{\psi}_{j,n}\right|^2=O\left(Nm^{-2}\right)$ by (5.15) because $\widehat{\psi}_{j,n}-\widetilde{\psi}_{j,n}$ involves the contribution of $N-j$ terms of the type $\left(\widehat{\phi}_{j_1,n}-\widetilde{\phi}_{j_1,n}\right)\left(\widehat{\phi}_{j_2,n}-\widetilde{\phi}_{j_2,n}\right)$, where $j_1-j_2=j$, by $C.4$ $E\left(y_{t-j_1}y_{t-j_2}\right)=O\left(|j_1-j_2|^{\alpha-1}\right)$ and Cauchy inequality. So, by $C.5$ and Markov's inequality, the right side of (5.16) is $o_p(1)$.

To finish, it remains to prove that $\widetilde{x}_{t|\infty}-x_{t|\infty}$ is $o_p(1)$. But

$$\widetilde{x}_{t|\infty}-x_{t|\infty}=\left(\sum_{j=-N+1}^{N-1}\left(\widetilde{\psi}_{j,n}-\psi_j\right)y_{t-j}\right)+\sum_{|j|=N}^{\infty}\psi_j y_{t-j} \tag{5.17}$$

where ψ_j is defined from the relation $\left(\sum_{j=1}^{\infty}\phi_j L^j\right)\left(\sum_{j=1}^{\infty}\phi_j L^{-j}\right)=\sum_{j=-\infty}^{\infty}\psi_j L^j$, with

$$\frac{\sigma_{\varepsilon^x}}{\sigma_{\varepsilon^y}}\left(1-\sum_{j=1}^{N}\beta_j L^j\right)(1-L)^{-\alpha/2}=\sum_{j=1}^{\infty}\phi_j L^j.$$

The second moment of the second term on the right of (5.17) is

$$\sigma_y^2\sum_{|j|=N+1}^{\infty}\psi_j^2+2\sum_{N+1=|j_1|<|j_2|}^{\infty}\psi_{j_1}\psi_{j_2}\gamma_{j_2-j_1}=O\left(N^{-1-\alpha}+N^{-1}\right)=o\left(1\right)$$

by $C.5$, because $\gamma_{j_2-j_1}=O\left(|j_2-j_1|^{\alpha-1}\right)$, $\psi_j=O\left(j^{-1-\alpha/2}\right)$ and $N\to\infty$. The second moment of the first term on the right of (5.17) is

$$\sigma_y^2\sum_{|j|=1}^{N}\left(\widetilde{\psi}_{j,n}-\psi_j\right)^2+2\sum_{1=|j_1|<|j_2|}^{N}\left(\widetilde{\psi}_{j_1,n}-\psi_{j_1}\right)\left(\widetilde{\psi}_{j_2,n}-\psi_{j_2}\right)\gamma_{j_2-j_1}.$$

Because $\gamma_j=O\left(j^{\alpha-1}\right)$, proceeding as in the proof of Theorem 3.3 part (b), and by Lemma 6 $\widetilde{\psi}_{j,n}-\psi_j=O\left(M^{-1}\right)$ after one identifies $\widetilde{\psi}_{j,n}$, $\psi_{j,n}$ and ψ_j as $\widetilde{a}_{v,n}$, $a_{v,n}$ and a_v respectively, the above expression is $O\left(M^{-2-\alpha}N^{1+\alpha}\right)=o\left(1\right)$ by $C.5$. Thus, the left side of $\widetilde{x}_{t|\infty}-x_{t|\infty}=o_p\left(1\right)$. ■

6. TECHNICAL LEMMAS

Lemma 1. *Assuming C.1,* $\widetilde{c}_{u,n}-c_{u,n}=O\left(M^{-2}\right)$.

Proof. Since from the definition of $\widetilde{f}_{x,\ell}$ and Taylor expansion of $\log\left(z\right)$,

$$\widetilde{c}_{u,n}-c_{u,n}=\frac{1}{M}\sum_{\ell=1}^{M}\left(\frac{\widetilde{f}_{x,\ell}-f_{x,\ell}}{f_{x,\ell}}\right)\cos\left(u\widetilde{\lambda}_\ell\right)+\frac{1}{M}\sum_{\ell=1}^{M}\left(\frac{\widetilde{f}_{x,\ell}-f_{x,\ell}}{f_{x,\ell}}\right)^2\cos\left(u\widetilde{\lambda}_\ell\right)\left(1+o\left(1\right)\right),$$

the proof is completed if we show that $f_{x,\ell}^{-1}\left(\widetilde{f}_{x,\ell}-f_{x,\ell}\right)=O\left(M^{-2}\right)$ uniformly in ℓ. But,

$$\begin{aligned}
f_{x,\ell}^{-1}\left(\widetilde{f}_{x,\ell}-f_{x,\ell}\right) &= \frac{g_{x,\ell}^{-1}}{(2m+1)}\sum_j\left(g_x\left(\lambda_{j+2m\ell}\right)-g_x\left(\lambda_{2m\ell}\right)\right)\\
&= g_{x,\ell}^{-1}\frac{\partial}{\partial\lambda}\left(g_x\left(\lambda_{2m\ell}\right)\right)\frac{1}{(2m+1)}\sum_j\frac{2\pi j}{n}\\
&\quad+\frac{g_{x,\ell}^{-1}}{(2m+1)}\sum_j\left(\frac{2\pi j}{n}\right)^2\frac{\partial^2}{\partial\lambda^2}g_x\left(\lambda_{2m\ell}+\xi\lambda_j\right)\\
&= g_{x,\ell}^{-1}\frac{\partial^2}{\partial\lambda^2}\left(g_x\left(\lambda_{2m\ell}\right)\right)\frac{\pi^2}{4M^2}\frac{1}{(2m+1)}\sum_j\left(\frac{j}{m}\right)^2\left(1+o\left(1\right)\right),
\end{aligned}$$

where $\xi=\xi(j)\in(0,1)$ and because $\frac{\partial^2}{\partial\lambda^2}g_x\left(\lambda_{2m\ell}+\lambda_j\right)=\frac{\partial^2}{\partial\lambda^2}g_{x,\ell}\left(1+o\left(1\right)\right)$ for all j by $C.1$. The proof now follows since by Brillinger (1981, $p.15$), $(2m+1)^{-1}\sum_j\left(\frac{j}{m}\right)^2-2^{-1}\int_{-1}^{1}z^2dz=O\left(m^{-1}\right)$, and $\left|g_{x,\ell}^{-1}\frac{\partial^2}{\partial\lambda^2}g_{x,\ell}\right|\leq K$. ■

Lemma 2. *Assuming C.1-C.4,*

$$\widetilde{f}_{x,\ell}^{-1}\left(\check{f}_{x,\ell}-\widetilde{f}_{x,\ell}\right)=O_p\left(m^{-1/2}\right).$$

Proof. By definition of $\widetilde{f}_{x,j}$ and $\check{f}_{x,j}$, the left side of the last displayed equality is

$$\left((2m+1)^{-1}\sum_j g_x^{-1}\left(\lambda_{j+2m\ell}\right)\right)^{-1}(2m+1)^{-1}\sum_j g_x\left(\lambda_{j+2m\ell}\right)\left(\frac{I\left(\lambda_{j+2m\ell}\right)}{f_x\left(\lambda_{j+2m\ell}\right)}-1\right).$$

But, the first factor by $C.1$ is bounded and bounded away from zero whereas the second factor is $O_p\left(m^{-1/2}\right)$ by routine extension of Robinson (1995) since $g_x\left(\lambda_{j+2m\ell}\right)$ is a differentiable function by $C.1$. ∎

Lemma 3. *Assuming C.1-C.4,*

$$\widetilde{f}_{x,\ell}^{-1}\left(\widehat{f}_{x,\ell}-\check{f}_{x,\ell}\right)=\frac{(\widehat{\alpha}-\alpha)}{2}g_{x,\ell}\int_{-1}^{1}\log\left(1+\frac{v}{2\ell}\right)dv+O_p\left(m^{-1}+m^{-1/2}M^{-1}\ell^{-1}\right). \tag{6.1}$$

Proof. By definition of $\widetilde{f}_{x,\ell}$, $\widehat{f}_{x,\ell}$ and $\check{f}_{x,\ell}$ and elementary algebra, the left side of (6.1) is

$$\left(\frac{1}{(2m+1)}\sum_j g_x\left(\lambda_{j+2m\ell}\right)\right)^{-1}\times$$
$$\Bigg\{\left(\lambda_{2m\ell}^{-(\widehat{\alpha}-\alpha)}-1\right)\frac{1}{(2m+1)}\sum_j\lambda_{2m\ell+j}^{\alpha}I\left(\lambda_{j+2m\ell}\right)\left(\lambda_{2m\ell+j}^{(\widehat{\alpha}-\alpha)}-1\right) \tag{6.2}$$
$$+\frac{1}{(2m+1)}\sum_j\lambda_{2m\ell+j}^{\alpha}I\left(\lambda_{j+2m\ell}\right)\left\{\left(\lambda_{2m\ell}^{-(\widehat{\alpha}-\alpha)}-1\right)+\left(\lambda_{2m\ell+j}^{(\widehat{\alpha}-\alpha)}-1\right)\right\}\Bigg\}.$$

By $C.1$, the first factor in (6.2) is bounded uniformly in ℓ. Next, by Lemma 2,

$$\begin{aligned}\frac{1}{(2m+1)}\sum_j\lambda_{2m\ell+j}^{\alpha}I\left(\lambda_{j+2m\ell}\right) &= \frac{1}{(2m+1)}\sum_j g_x\left(\lambda_{2m\ell+j}\right)\left(\frac{I\left(\lambda_{j+2m\ell}\right)}{f\left(\lambda_{j+2m\ell}\right)}-1\right)\\ &\quad+\frac{1}{(2m+1)}\sum_j g_x\left(\lambda_{2m\ell+j}\right) \qquad (6.3)\\ &= g_{x,\ell}+O\left(M^{-2}\right)+O_p\left(m^{-1/2}\right),\end{aligned}$$

since by $C.1$ since $g_x(\lambda)$ is twice continuously differentiable, $\sum_j j=0$ and Robinson (1995) for the first term on the right of (6.3).

The first term inside the curly brackets of (6.2) is $O_p\left(m^{-1}\right)$ since $(\widehat{\alpha}-\alpha)=O_p\left(m^{-1/2}\right)$ by Robinson (1995) and $\left|\lambda_{2m\ell}^{-(\widehat{\alpha}-\alpha)}-1\right|\leq K\left|\widehat{\alpha}-\alpha\right|$. So, it remains to examine the contribution from the second term inside the curly brackets of (6.2), which by a Taylor expansion is

$$
\begin{aligned}
& -\frac{(\widehat{\alpha}-\alpha)}{(2m+1)}\sum_{j}\lambda_{2m\ell+j}^{\alpha}I\left(\lambda_{j+2m\ell}\right)\left(\log\left(\lambda_{2m\ell+j}\right)-\log\left(\lambda_{2m\ell}\right)\right)+O_{p}\left(m^{-1}\right) \\
= & \ (\widehat{\alpha}-\alpha)\,g_{x,\ell}\frac{1}{(2m+1)}\sum_{j}\log\left(1+\frac{j}{2m\ell}\right)+O_{p}\left(m^{-1}+m^{-1/2}M^{-1}\ell^{-1}\right) \\
= & \ \frac{1}{2}(\widehat{\alpha}-\alpha)\,g_{x,\ell}\int_{-1}^{1}\log\left(1+\frac{v}{2\ell}\right)dv+O_{p}\left(m^{-1}+m^{-1/2}M^{-1}\ell^{-1}\right),
\end{aligned}
$$

proceeding as with (6.3) and using that $\log\left(1+j/\left(2m\ell\right)\right)=O\left(\ell^{-1}\right)$ and $g_{x}\left(\lambda_{2m\ell+j}\right)-g_{x}\left(\lambda_{2m\ell}\right)=O\left(M^{-1}\right)$, for the first equality and for the second one that by Brillinger (1981, p.15)

$$
\left|(2m+1)^{-1}\sum_{j}\log\left(1+\frac{j}{2m\ell}\right)-\frac{1}{2}\int_{-1}^{1}\log\left(1+\frac{v}{2\ell}\right)dv\right|\leq Km^{-1}. \qquad \square
$$

Lemma 4.

$$
\frac{1}{M^{2}}\sum_{p=1}^{M}\left(\sum_{u_{1}=1}^{M}\cos\left(u_{1}\widetilde{\lambda}_{p}\right)e^{-iu_{1}\widetilde{\lambda}_{j_{1}}}\right)\left(\sum_{u_{2}=1}^{M}\cos\left(-u_{2}\widetilde{\lambda}_{p}\right)e^{iu_{2}\widetilde{\lambda}_{j_{2}}}\right)=2^{-1}\phi_{j_{1}-j_{2}}+O\left(\frac{1}{M}\right). \tag{6.4}
$$

Proof. Because $\cos x\cos y=\left(\cos(x-y)+\cos(x+y)\right)/2$, the left side of (6.4) is

$$
\frac{1}{2M}\sum_{u_{1}=1}^{M}\sum_{u_{2}=1}^{M}e^{-iu_{1}\widetilde{\lambda}_{j_{1}}}e^{iu_{2}\widetilde{\lambda}_{j_{2}}}\frac{1}{M}\sum_{p=1}^{M}\left(\cos\left(\left(u_{1}+u_{2}\right)\widetilde{\lambda}_{p}\right)+\cos\left(\left(u_{1}-u_{2}\right)\widetilde{\lambda}_{p}\right)\right).
$$

But, since $\widetilde{\lambda}_{p}=\left(\pi p\right)/M$, the only term different than zero is when $u_{1}=u_{2}$, in which case it is

$$
\begin{aligned}
\frac{1}{2M}\sum_{u=1}^{M}e^{-iu\widetilde{\lambda}_{j_{1}-j_{2}}} &= \frac{1}{2}\left(\delta_{j_{1}-j_{2}}+\frac{1}{\pi\left(j_{1}-j_{2}\right)}\left(1-\cos\left(\left(j_{2}-j_{1}\right)\pi\right)\right)\right)+O\left(\frac{1}{M}\right) \\
&= 2^{-1}\phi_{j_{1}-j_{2}}+O\left(M^{-1}\right),
\end{aligned}
$$

using standard Euler-McLauring approximation of sums by integrals.

Lemma 5.

$$
\sum_{s=M+1}^{\infty}c_{s}e^{-is\lambda_{2mj}}=O\left(\lambda_{2mj}^{-1}M^{-1}\mathcal{I}\left(\alpha>0\right)+M^{-1}\mathcal{I}\left(\alpha=0\right)\right). \tag{6.5}
$$

Proof. >From the definition of c_{s} in (2.10) and C.1, $c_{s}=c_{s1}+c_{s2}$ where

$$
c_{s1}=-\frac{\alpha}{\pi}\int_{0}^{\pi}\left(\log\lambda\right)\cos\left(s\lambda\right)d\lambda
$$

and

$$
c_{s2}=\frac{1}{\pi}\int_{0}^{\pi}\log\left(g\left(\lambda\right)\right)\cos\left(s\lambda\right)d\lambda.
$$

Now, since $g(\lambda)$ is continuously differentiable by $C.1$, $c_{s2} = O\left(s^{-2}\right)$, whereas c_{s1} is by integration by parts

$$-\frac{\alpha}{\pi}\int_0^{\pi}(\log\lambda)\cos(s\lambda)\,d\lambda = -\frac{\alpha}{\pi s}\sin(s\lambda)\,\log\lambda|_0^{\pi} + \frac{\alpha}{\pi s}\int_0^{\pi}\frac{\sin(s\lambda)}{\lambda}d\lambda.$$

The first term on the right is zero, whereas the second term is

$$\frac{\alpha}{2s} + O\left(s^{-2}\right)$$

since by Courant and John (1974, sect. 8.4.c),

$$\left|\int_0^{\pi}\frac{\sin(s\lambda)}{\lambda}d\lambda - \int_0^{\infty}\frac{\sin(\lambda)}{\lambda}d\lambda\right| \leq Ks^{-2}$$

as $s\to\infty$ and $\int_0^{\infty}\lambda^{-1}\sin(\lambda)\,d\lambda = \pi/2$. Thus, the left side of (6.5) is bounded in absolute value by

$$\begin{aligned}\left|\sum_{s=M+1}^{\infty}c_s e^{-is\lambda_{2mj}}\right| &\leq \left|\sum_{s=M+1}^{\infty}c_{s2}e^{-is\lambda_{2mj}}\right| + \left|\sum_{s=M+1}^{\infty}c_{s1}e^{-is\lambda_{2mj}}\right| \\ &\leq O\left(M^{-1}\right) + \frac{\alpha}{2}\left|\sum_{s=M+1}^{\infty}s^{-1}e^{-is\lambda_{2mj}}\right|.\end{aligned}$$

But the second term on the right is by Abel summation by parts

$$\alpha\left|\sum_{s=M+1}^{\infty}\left(s^{-1}-(s+1)^{-1}\right)\sum_{\ell=M+1}^{s}e^{-i\ell\lambda_{2mj}}\right| \leq \alpha\lambda_{2mj}^{-1}\sum_{s=M+1}^{\infty}s^{-2}.$$

>From here the conclusion of the lemma is obvious. ■

Lemma 6. $a_{u,n}-a_u = O\left(M^{-1}\right)$.

Proof. $a_{u,n}-a_u$ is

$$\frac{1}{2M}\sum_{j=-M+1}^{M-1}\left(A_{j,n}-A_j^*\right)e^{iu\lambda_{2jm}} + \frac{1}{2M}\sum_{j=-M+1}^{M-1}\left(A_j^*-A_j\right)e^{iu\lambda_{2jm}}$$
$$+\left(\frac{1}{2M}\sum_{j=-M+1}^{M-1}A_j e^{iu\lambda_{2jm}} - \frac{1}{2\pi}\int_{-\pi}^{\pi}A(\lambda)\,e^{iu\lambda}d\lambda\right),$$

where $A_j^* = \exp\left(-\sum_{s=1}^{M}c_s e^{-is\lambda_{2mj}}\right)$. The third term of the last displayed expression is $O\left(M^{-1}\right)$ by Brillinger (1981, p.15) since $|A_j|\cos(u\lambda)$ and $|A_j|\sin(u\lambda)$ have an integrable derivative.

Assume that $\alpha > 0$, the proof for $\alpha = 0$ is identical. The second term is

$$\frac{1}{2M}\sum_{j=-M+1}^{M-1}A_j\left(1-\exp\left\{-\sum_{s=M}^{\infty}c_s e^{-is\lambda_{2mj}}\right\}\right)e^{iu\lambda_{2jm}}$$

which is, since $A(0)=0$, bounded in absolute value by

$$\frac{K}{2M}\sum_{j=-M+1;j\neq 0}^{M-1}|A_j|\left|\sum_{s=M+1}^{\infty}c_s e^{-is\lambda_{2mj}}\right|(1+O(1))$$
$$\leq \frac{1}{2M^2}\sum_{j=-M+1;j\neq 0}^{M-1}\lambda_{2mj}^{-1}|A_j| = O\left(M^{-1}\right)$$

since $|A_j| = O\left((j/M)^{\alpha/2}\right)$ by C.1 and $\sum_{s=M+1}^{\infty}c_s e^{-is\lambda_{2mj}} = O\left(\lambda_{2mj}^{-1}M^{-1}\right)$ by Lemma 5.

To complete the proof we need to examine the first term which by definition of $A_{j,n}$ and A_j^* is

$$-\frac{1}{2M}\sum_{j=-M+1}^{M-1}\left(1-\exp\left\{-\sum_{s=1}^{M-1}(c_{s,n}-c_s)e^{-is\lambda_{2mj}}\right\}\right)A_j^* e^{iu\lambda_{2jm}}. \tag{6.6}$$

Now, by definition of $c_{s,n}$ and c_s and symmetry, $\sum_{1\leq\ell<M}\log(f_\ell)\cos\left(s\widetilde{\lambda}_\ell\right) = 2^{-1}\sum_{1\leq|\ell|<M}\log(f_\ell)e^{is\widetilde{\lambda}_\ell}$, so

$$\sum_{s=1}^{M-1}(c_{s,n}-c_s)e^{-is\widetilde{\lambda}_j} = \frac{1}{2}\sum_{s=1}^{M-1}\left(\frac{1}{M}\sum_{1\leq|\ell|<M}\log(f_\ell)e^{is\widetilde{\lambda}_\ell} - \frac{1}{\pi}\int_{-\pi}^{\pi}(\log f(\lambda))e^{is\lambda}d\lambda\right)e^{-is\widetilde{\lambda}_j}. \tag{6.7}$$

Let $D(\lambda) = \sum_{s=1}^{M-1}e^{is\lambda}$. Because $\int_{-\pi}^{\pi}D\left(\lambda-\widetilde{\lambda}_j\right)d\lambda = 0$ and $\sum_{1\leq|\ell|<M}e^{is\widetilde{\lambda}_\ell} = -1$, the right side of (6.7) is

$$-\log(f_j)\frac{1}{2M}\sum_{s=1}^{M-1}e^{is\widetilde{\lambda}_j} + \frac{1}{2M}\sum_{1\leq|\ell|<M}(\log(f_\ell)-\log(f_j))D\left(\widetilde{\lambda}_\ell-\widetilde{\lambda}_j\right)$$
$$-\frac{1}{2\pi}\int_{-\pi}^{\pi}(\log f(\lambda)-\log(f_j))D\left(\lambda-\widetilde{\lambda}_j\right)d\lambda. \tag{6.8}$$

The contribution of the first term of (6.8) into (6.6) is

$$-\frac{1}{2M}\sum_{j=-M+1}^{M-1}\left(1-\exp\left\{\frac{1}{2M}\log(f_j)\sum_{s=1}^{M}e^{-is\lambda_{2mj}}\right\}\right)A_j^* e^{iu\lambda_{2jm}}$$

which is bounded in absolute value by

$$\frac{K}{2M}\sum_{j=-M+1}^{M-1}j^{-1}\left|A_j^*\right|\log(f_j) = O\left(M^{-1}\right),$$

because $\left|\sum_{s=1}^{M-1}e^{-is\lambda_{2mj}}\right| \leq \lambda_{2mj}^{-1}$. The second and third terms of (6.8) are $O\left(j^{-1}\right)$ by Brillinger (1981, p.15), since by Taylor expansion $(\log(f(\lambda))-\log(f_j)) = O\left(\left(\lambda-\widetilde{\lambda}_j\right)\widetilde{\lambda}_j^{-1}\right)$ and because $\left(\lambda-\widetilde{\lambda}_j\right)D\left(\lambda-\widetilde{\lambda}_j\right)$ has an integrable derivative. So, their contribution into (6.6) is bounded in absolute value by

$$\frac{1}{2M}\sum_{j=-M+1}^{M-1} j^{-1}\left|A_j^*\right| + \frac{1}{2M}\sum_{j=-M+1}^{M-1} j^{-2}\left|A_j^*\right| = O\left(M^{-1}\right)$$

because $\left|A_j^*\right| = O\left(\widetilde{\lambda}_j^{\alpha/2}\right)$ and $\left|1-\exp(z)-z-2^{-1}z^2\right| \leq K$ if z is bounded. So, the lemma is shown. ■

7. CONCLUSIONS AND EXTENSIONS

In this paper we have extended the nonparametric prediction algorithm examined by Bhansali (1974, 77) to any covariance stationary linear process which may exhibit strong dependence. Since we do not impose any particular structure on the underline process of the data, we are thus able to avoid the problem that the misspecification of the model may induce to obtain accurate predictions. In addition, we have illustrated and discussed how the $FLES$ can be adopted to extract the signal from a covariance stationary strong dependent process. One feature, in contrast to previous work on the topic, is that we do not need assume any particular model for the noise. So, we can name the approach semiparametric.

An alternative method to predict x_t or recover the signal is via the estimation of the spectral density function by fitting an autoregressive model where the order of the AR polynomial increases with the sample size, see Berk (1974) or An *et al.* (1982) among others. However, as Bhansali (1978) showed for weakly dependent data, this method is asymptotically equivalent to that describe in Section 2, at least for prediction purposes. It thus seems to be of interest to examine whether the results hold the same under strong dependence.

References

[1] An, H-Z., Chen, Z-G. and Hannan, E.J. (1982), Autocorrelation, autoregression and autoregressive approximation. Ann. of Statist., 10, 926-936.

[2] Beran, J. (1994), Statistics for long memory processes. New York: Chapman and Hall.

[3] Berk, K.N. (1974), Consistent autoregressive spectral estimates. Ann. of Statist., 2, 489-502.

[4] Bhansali, R.J. (1974), Asymptotic properties of the Wiener-Kolgomorov predictor. I. Journal of the Royal Statist. Soc., B, 36, 61-73.

[5] Bhansali, R.J. (1977), Asymptotic properties of the Wiener-Kolgomorov predictor. II. Journal of the Royal Statist. Soc., B, 39, 66-72.

[6] Bhansali, R.J. (1978), Linear prediction by autoregressive model fitting in the time domain. Ann. of Statist., 6, 224-231.

[7] Bloomfield, P. (1973), An exponential model for the spectrum of a scalar time series. Biometrika, 60, 217-226.

[8] Brillinger, D.R. (1981), Time series, data analysis and theory. San Francisco: Holden Day.

[9] Courant, R. and John, F. (1974), Introducción al Análisis Matemático. Vol. 1. Limusa, México.

[10] Crato, N. and Ray, B.K. (1996), Model selection and forecasting for long-range dependent processes. Journal of Forecasting, 15, pp. 107-125.

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

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

[13] Giraitis, L., Hidalgo, J. and Robinson, P.M. (1999), Gaussian estimation of parametric spectral density with unknown pole. Preprint.

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

[15] Hannan, E.J. (1967), The estimation of a lagged regression relation. Biometrika, 54, 409-418.

[16] Hannan, E.J. (1970), Multiple time series. Wiley, New York.

[17] Hidalgo, J. (1998), Nonparametric test for causality with long range dependence. Preprint.

[18] Hidalgo, J. and Yajima, Y. (1999), Semiparametric estimation of the long-range parameter. Preprint.

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

[20] Mandelbrot, B.B. and Van Ness, J.W. (1968), Fractional Brownian motions, fractional noises and applications. SIAM Review, 10, 422-437.

[21] Peiris, M.S. and Perera, B.J.C. (1988), On prediction with fractionally differenced ARIMA models. Journal of Time Ser. Analysis, 9, 215-220.

[22] Porter-Hudak, S. (1990), An application of the seasonal fractionally differenced model to the monetary aggregates. Journal of the Amer. Statist. Assoc., 85, 338-344.

[23] Press, H. and Tukey, J.W. (1956), Power Spectral methods of analysis and their application to problems in airplane dynamics. Flight Test Manual, NATO, Advisory Group for Aeronautical Research and Development, Vol. IV-C, 1-41. Reprinted as Bell System Monograph No. 2606.

[24] Ray, B.K. (1993), Long range forecasting of IBM product revenues using a seasonal fractional differenced ARMA model. International Journal of Forecasting, 9, 255-269.

[25] Robinson, P.M. (1994), Time series with strong dependence. In C.A. Sims, ed., Advances in Econometrics: Sixth World Congress, Vol. 1, 47-95. Cambridge University Press, Cambridge.

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

[27] Serfling, R.J. (1980), Approximation Theorems of mathematical Statistics. John Wiley and Sons. New York.

[28] Sinai, Y.G. (1976), Self-similar probability distribution. Theo. Probability and its Applications, 21, 64-80.

[29] Whittle, P. (1963), Prediction and regulation by linear least squares methods. Princeton: Van Nostram.

[30] Zygmund, A. (1977), Trigonometric Series. Cambridge Univ. Press.