

MINIMUM HELLINGER DISTANCE ESTIMATION OF A STATIONARY MULTIVARIATE LONG MEMORY ARFIMA PROCESS

K. STANISLAS MBEKE and OUAGNINA HILI

Laboratory of Mathematics and New Technologies of Information

National Polytechnic Institute

Félix HOUPHOUET-BOIGNY Yamoussoukro

P. O. Box 1911

Ivory Coast

e-mail: o_hili@yahoo.fr

Abstract

In this note, we determine the minimum Hellinger distance estimator (MHDE) of a stationary multivariate long memory ARFIMA (Auto regressive fractionally integrated moving average) process. We establish, under some assumptions, the almost sure convergence of the estimator and its asymptotic normality.

1. Introduction

Granger and Joyeux [4] and Hosking [6] have proposed the ARFIMA (p, d, q) model to define a time series, which presents a character of short or long memory following d . For $-\frac{1}{2} < d < 0$, the process is short memory. For $0 < d < \frac{1}{2}$, the process is long memory. This long memory is

2010 Mathematics Subject Classification: 62F12, 62H12.

Keywords and phrases: minimum Hellinger distance, stationary multivariate ARFIMA process, estimation, long memory.

Received February 16, 2018; Revised March 15, 2018

characterized by a slow decay of the autocorrelation function or the sum of unfinished autocorrelations. The process is non-stationary for $d > \frac{1}{2}$ and stationary for $d < \frac{1}{2}$. In spite of ARFIMA processes, the notion of long memory has been widely discussed by the authors such as Bitty and Hili [2] for linear processes with long memory, N'dri and Hili [10] for strongly dependent multi-dimensional Gaussian processes. Laura Mayoral [9] proposed by minimum distance a new method for estimating the parameters of stationary and non-stationary ARFIMA (p, d_0, q) process for $d_0 > -0.75$. Kamagaté and Hili [7] and [8] estimated by the minimum Hellinger distance method a stationary univariate ARFIMA process and by the quasi maximum likelihood approach a non-stationary multivariate ARFIMA process.

In this paper, we generalize the results of Kamagaté and Hili [7] to the multivariate case. We consider an m -dimensional ARFIMA stationary process $(y_1(t), \dots, y_m(t))$, after inversion of the process, we establish the consistence and asymptotic normality by using the minimum Hellinger distance. The reasons for choosing this estimation technic lie in the fact that these estimators obtained are efficient and robust (cf. Beran [1]).

The paper is organised as following. After some notes about the estimator, in Section 2, we present a multivariate ARFIMA model. Section 3 is devoted to the estimation of parameters including the consistency of the estimator and its asymptotic normality. In Section 4, we establish the main results of this work.

We denote by θ the vector of parameters of interest composed of (d_1, \dots, d_m) and matrix coefficients.

The minimum Hellinger distance estimator of θ is defined by

$$\hat{\theta}_n = \arg \min_{\theta \in \Theta} H_2(\hat{f}_{\theta_n}, f_{\theta_0}),$$

where $H_2(\hat{f}_{\theta_n}, f_{\theta_0})$ is the Hellinger distance defined by

$$H_2(\hat{f}_{\theta_n}, f_{\theta_0}) = \left\{ \int_{\mathbb{R}^m} |\hat{f}_{\theta_n}^{\frac{1}{2}}(x) - f_{\theta_0}^{\frac{1}{2}}(x)|^2 dx \right\}^{\frac{1}{2}}. \quad (1)$$

The minimum Hellinger distance minimizes the Hellinger distance between \hat{f}_{θ_n} and f_{θ_0} . $f_{\theta_0}(\cdot)$ is a theoretical probability density, $\hat{f}_{\theta_n}(\cdot)$ is a random function of $\hat{\underline{\varepsilon}}(t)$ and $\tilde{f}_n(\cdot)$ a non-parametric kernel density estimator of $\underline{\varepsilon}(t)$ defined by:

$$\hat{f}_{\theta_n}(x) = \frac{1}{nb_n^m} \sum_{t=1}^n K\left(\frac{x - \hat{\underline{\varepsilon}}(t)}{b_n}\right), \quad x \in \mathbb{R}^m, \quad (2)$$

$$\tilde{f}_n(x) = \frac{1}{nb_n^m} \sum_{t=1}^n K\left(\frac{x - \underline{\varepsilon}(t)}{b_n}\right), \quad x \in \mathbb{R}^m, \quad (3)$$

where $K : \mathbb{R}^m \rightarrow \mathbb{R}_+$ is a kernel function and (b_n) is a sequence of bandwidths and $f : \mathbb{R}^m \rightarrow \mathbb{R}_+$.

2. Multivariate ARFIMA Models

The multivariate ARFIMA model was introduced by Sowell [13]. We consider an m -dimensional ARFIMA stationary process $(y_1(t), \dots, y_m(t))$ following $d < \frac{1}{2}$ which is generated by

$$A(L)D(L)(y_1(t), \dots, y_m(t))' = B(L)(\varepsilon_1(t), \dots, \varepsilon_m(t))', \quad (4)$$

where M' denote the transpose of the matrix M .

L is the backward shift operator, who, to any element of a time series, associates the previous observation as

$$L^j X_t = X_{t-j}, \quad j \in \mathbb{N},$$

$\{\varepsilon_1(t), \dots, \varepsilon_m(t)\}$ are white noise processes that follow the normal law of mean zero and covariance

$$r\{\varepsilon_i(t), \varepsilon_j(s)\} = \delta(t, s)k_{ij}, \quad i, j = 1, \dots, m.$$

Denote by $\mathbf{K} = (k_{ij})$ the positive definite covariance matrix.

The expression $D(L)$ defined in (4) represents a diagonal $(m \times m)$ -matrix of the fractional difference operators of backward shift defined by

$$D(L) = \begin{pmatrix} (1-L)^{d_1} & 0 & \dots \\ 0 & \ddots & 0 \\ \vdots & 0 & (1-L)^{d_m} \end{pmatrix},$$

with $(1-L)^d = 1 + \sum_{k=1}^{+\infty} \frac{\Gamma(k-d)}{\Gamma(-d)k!} L^k$ and $d_1, \dots, d_m \in \left(-\frac{1}{2}, \frac{1}{2}\right)$. $\Gamma(\cdot)$ is the gamma function such that $\Gamma(j+1) = j!$.

Let $A(\cdot)$ and $B(\cdot)$ be matrix polynomials in L of degrees p and q , respectively, defined as hereinafter by:

$$A(L) = I - A_1L - \dots - A_pL^p,$$

$$B(L) = I + B_1L + \dots + B_qL^q,$$

where I represents the $(m \times m)$ identity matrix. The $\det A(L) \neq 0$ and $\det B(L) \neq 0$ are, respectively, the characteristic polynomial of the matrix polynomials $A(\cdot)$ and $B(\cdot)$. We assume that the roots of characteristic polynomial are all outside the unit disk.

Odaki [11] and Hosking [6] showed that the process is invertible for $d > -1$ and stationary for $d < \frac{1}{2}$. Taking into account the conditions on the polynomials, the process (4) is invertible and causal and admits a representation of an autoregressive process of infinite order as following:

$$\begin{aligned}
B(L) \begin{pmatrix} \varepsilon_1(t) \\ \vdots \\ \varepsilon_m(t) \end{pmatrix} &= A(L)D(L) \begin{pmatrix} y_1(t) \\ \vdots \\ y_m(t) \end{pmatrix}, \\
\begin{pmatrix} \varepsilon_1(t) \\ \vdots \\ \varepsilon_m(t) \end{pmatrix} &= (B(L))^{-1} A(L)D(L) \begin{pmatrix} y_1(t) \\ \vdots \\ y_m(t) \end{pmatrix}. \tag{5}
\end{aligned}$$

Let $(B(L))^{-1} A(L) = C(L)$, the equality (5) can be written as follows:

$$\begin{pmatrix} \varepsilon_1(t) \\ \vdots \\ \varepsilon_m(t) \end{pmatrix} = \begin{pmatrix} c_{11}(L, \theta) & \cdots & c_{1m}(L, \theta) \\ \vdots & \cdots & \vdots \\ c_{m1}(L, \theta) & \cdots & c_{mm}(L, \theta) \end{pmatrix} D(L) \begin{pmatrix} y_1(t) \\ \vdots \\ y_m(t) \end{pmatrix},$$

and

$$\begin{pmatrix} \varepsilon_1(t) \\ \vdots \\ \varepsilon_m(t) \end{pmatrix} = \sum_{j=0}^{\infty} \Psi_j(\theta) \begin{pmatrix} y_1(t-j) \\ \vdots \\ y_m(t-j) \end{pmatrix},$$

where $\theta = (A_i, d_l, B_j) \in \Theta \subset \mathbb{R}^n$ be the vector parameters of interest. For $0 \leq i \leq p, 1 \leq l \leq m$ and $0 \leq j \leq q$. Θ is a compact set. $\mathbf{n} = m[m(p+q)+1]$. $\{\Psi_j(\theta)\}_{j=0}^{\infty}$ are $(m \times m)$ -matrix associated with the entire series development of the matrix polynomial $C(L)D(L)$ in power of L such as $\sum_{j \in \mathbb{Z}} |\Psi_{r,s}(j)| < \infty$ for $1 \leq r, s \leq m$.

Let y_i for $1 \leq i \leq n$ be the observations. The innovations $(\varepsilon_1(t), \dots, \varepsilon_m(t))'$ are not observable, they are estimated by

$$\begin{pmatrix} \hat{\varepsilon}_1(t) \\ \vdots \\ \hat{\varepsilon}_m(t) \end{pmatrix} = \sum_{j=0}^n \Psi_j(\theta) \begin{pmatrix} y_1(t-j) \\ \vdots \\ y_m(t-j) \end{pmatrix}.$$

3. Parameter Estimation

To establish the consistency and limit law of the parameter, we need the following assumptions:

Assumption (A1)

(1) $E(|\underline{\varepsilon}_t|^s) < +\infty$ for $s \geq 1$.

For all $(u, \nu) \in \mathbb{R}^{2m}$, we have:

(2) $\int_{\mathbb{R}^m} K^2(u) du < \infty$, $\int_{\mathbb{R}^m} u_i K(u) du = 0$ for $1 \leq i \leq m$;

(3) $\int_{\mathbb{R}^m} u_i u_j K(u) du = 0$, $\int_{\mathbb{R}^m} u_i^2 K(u) du < \infty$ for $1 \leq j \leq m$;

(4) There exists $C > 0$ such that $\sup_{u \in \mathbb{R}^m} |K(u + \nu) - K(u)| \leq C|\nu|$.

Assumption (A2)

For each $\theta \in \Theta$ and each $x \in \mathbb{R}^m$, the functions $x \mapsto f_\theta(x)$ and $\theta \mapsto f_\theta^{\frac{1}{2}}(x)$ are continuously differentiable and $\underline{\varepsilon}_t$'s admits a density absolutely continuous with respect to the Lebesgue measure \mathbb{R}^m , positive in a neighbourhood of the origin.

Assumption (A3)

For each $x \in \mathbb{R}^m$, the functions $\theta \mapsto \frac{\partial}{\partial \theta_j} f_{\theta}^{\frac{1}{2}}(x)$, for $1 \leq j \leq q$ and $\theta \mapsto \frac{\partial^2}{\partial \theta_j \partial \theta_k} f_{\theta}^{\frac{1}{2}}(x)$, for $1 \leq j, k \leq q$

are finished, continuous and defined in $L^2(\mathbb{R}^q)$.

Assumption (A4)

$b_n = n^\alpha \mathcal{L}(n)$, where $-1 < \alpha < 0$ with \mathcal{L} a slowly varying function.

$$\lim_{n \rightarrow +\infty} b_n = 0, \quad \lim_{n \rightarrow +\infty} n b_n = +\infty, \quad \lim_{n \rightarrow +\infty} \frac{\mathcal{L}(an)}{\mathcal{L}(n)} = 1, \quad a > 0.$$

For each $\theta \in \Theta$, $\sup_{x \in \mathbb{R}^m} \left| \frac{\partial^i f_{\theta_0}}{\partial x_k^i}(x) \right| < \infty$, $i = 0, 1, 2, \dots$ and $k = 1, \dots, m$.

Assumption (A5)

For $\theta, \theta' \in \Theta$, $\theta \neq \theta'$ implies that $\{x \in \mathbb{R}^m / f_{\theta}(x) \neq f_{\theta'}(x)\}$ is a set of positive Lebesgue measure.

Assumption (A6)

We suppose that there is a constant M such as $\sup_{x \in \mathbb{R}^m} f(x) \leq M < \infty$.

Theorem 1. *Supposing that Assumptions (A1)-(A6) are satisfied. Then $\hat{\theta}_n$ converges almost surely to θ_0 for all $x \in \mathbb{R}^m$.*

We denote by:

$$g_{\theta_0} = f_{\theta_0}^{\frac{1}{2}}, \quad \dot{g}_{\theta_0} = \frac{\partial g_{\theta_0}}{\partial \theta_0}, \quad \ddot{g}_{\theta_0} = \frac{\partial^2 g_{\theta_0}}{\partial \theta_0 \partial \theta_0^t}, \quad V_{\theta_0} = \left[\int_{\mathbb{R}^m} \dot{g}_{\theta_0}(x) \dot{g}_{\theta_0}^t(x) dx \right]^{-1} \dot{g}_{\theta_0}(x).$$

when these quantities exist.

Theorem 2. *Supposing that Assumptions (A1)-(A6) are satisfied. If*

Condition C1: *The components of \dot{g}_{θ_0} and \ddot{g}_{θ_0} are in L_2 and if the norms of these components are continuous functions of θ_0 .*

Condition C2: $\int_{\mathbb{R}^m} \ddot{g}_{\theta_0}(x)g_{\theta_0}(x)dx$ *is a non-singular $(n \times n)$ -matrix, then the limit distribution of $\sqrt{n}(\hat{\theta}_n - \theta_0)$ is $N(0, \Sigma^2)$, where*

$$\Sigma^2 = \frac{1}{4} \left[\int_{\mathbb{R}^m} \dot{g}_{\theta_0}(x)\dot{g}_{\theta_0}^t(x)dx \right]^{-1} \int_{\mathbb{R}^m} K^2(u)du.$$

4. Proof of the Theorems

We need the following lemma to prove the Theorem 1.

Lemma 1. *Supposing that Assumptions (A1) and (A2) are satisfied.*

Then

$$\hat{f}_{\theta_n}(x) - f_{\theta_0}(x) \rightarrow 0 \quad \text{a.s. when } n \rightarrow +\infty.$$

Proof. By triangular inequality, we have

$$\sup_{x \in \mathbb{R}^m} |\hat{f}_{\theta_n}(x) - f_{\theta_0}(x)| \leq (a) + (b) + (c),$$

where

$$(a) = \sup_{x \in \mathbb{R}^m} |\hat{f}_{\theta_n}(x) - \tilde{f}_n(x)|;$$

$$(b) = \sup_{x \in \mathbb{R}^m} |\tilde{f}_n(x) - E\tilde{f}_n(x)|;$$

$$(c) = \sup_{x \in \mathbb{R}^m} |E\tilde{f}_n(x) - f_{\theta_0}(x)|.$$

We demonstrate Lemma 1 in three steps.

Step 1. The convergence of (a) to zero after inversion of the process (4).

Considering the conditions on the polynomial functions $A(L)$ and $B(L)$, the process (4) is invertible and can be rewritten as a representation of a autoregressive process of infinite order.

We consider two density functions $\hat{f}_{\theta_n}(\cdot)$ and $\tilde{f}_n(\cdot)$, respectively of $\hat{\underline{\varepsilon}}_t$ and $\underline{\varepsilon}_t$. By Assumption (A1), we have

$$\sup_{x \in \mathbb{R}^m} |\hat{f}_{\theta_n}(x) - \tilde{f}_n(x)| \leq \frac{C}{nb_n^{m+1}} \sum_{t=1}^n \|\underline{\eta}_t\|,$$

where

$$\underline{\eta}_t = \begin{pmatrix} \eta_1(t) \\ \vdots \\ \eta_m(t) \end{pmatrix} = \begin{pmatrix} \varepsilon_1(t) - \hat{\varepsilon}_1(t) \\ \vdots \\ \varepsilon_m(t) - \hat{\varepsilon}_m(t) \end{pmatrix},$$

$\underline{\eta}_t$ represents the rest when truncating the series from n ,

$$\underline{\eta}_t = \sum_{j=n+1}^{\infty} \Psi_j(\theta) y_{t-j}.$$

We adapt the notion of invertibility according to Granger and Andersen [3] as below

$$\lim_{t \rightarrow +\infty} E(\underline{\varepsilon}_t - \hat{\underline{\varepsilon}}_t)^2 = \Sigma^2 \sum_{j=n+1}^{+\infty} \Psi_{j,n-1}^2 < \infty.$$

Let's examine the following expression:

$$E \left(\frac{1}{nb_n^{m+1}} \sum_{t=1}^n \|\underline{\eta}_t\| \right)^2.$$

Denote by

$$\chi = \frac{1}{nb_n^{m+1}} \sum_{t=1}^n \|\eta_t\|,$$

$$E(\chi)^2 = E\left(\frac{1}{n^2b_n^{2m+2}} \sum_{t=1}^n \|\eta_t\|^2\right) + 2E\left(\frac{1}{n^2b_n^{2m+2}} \sum_{t=1}^n \sum_{\substack{s=1 \\ t < s}}^n \|\eta_t\| \|\eta_s\|\right).$$

For all reals $u > 0$ and $v > 0$,

$$uv \leq \frac{1}{2}(u^2 + v^2), \quad (6)$$

then

$$E(\chi)^2 \leq E\left(\frac{1}{n^2b_n^{2m+2}} \sum_{t=1}^n \|\eta_t\|^2\right) + \frac{1}{n^2b_n^{2m+2}} E\left(\sum_{t=1}^n \|\eta_t\|^2 + \sum_{\substack{s=1 \\ t < s}}^n \|\eta_s\|^2\right),$$

thus

$$\begin{aligned} E(\chi)^2 &\leq \frac{2}{n^2b_n^{2m+2}} E\left(\sum_{t=1}^n \|\eta_t\|^2\right) + \frac{1}{n^2b_n^{2m+2}} E\left(\sum_{\substack{s=1 \\ t < s}}^n \|\eta_s\|^2\right) \\ &\leq \frac{2}{n^2b_n^{2m+2}} E\left(\sum_{t=1}^n \|\eta_t\|^2\right) + \frac{2}{n^2b_n^{2m+2}} E\left(\sum_{\substack{s=1 \\ t < s}}^n \|\eta_s\|^2\right) \\ &\leq \frac{2}{n^2b_n^{2m+2}} \left(E\left(\sum_{t=1}^n \|\eta_t\|^2\right) + E\left(\sum_{\substack{s=1 \\ t < s}}^n \|\eta_s\|^2\right) \right) \\ &\leq \frac{2}{n^2b_n^{2m+2}} (d + e). \end{aligned}$$

We will now focus on the expressions of (d) and (e).

$$d = E \left(\sum_{t=1}^n \left\| \sum_{j=n+1}^{\infty} \Psi_j(\theta) \underline{y}_{t-j} \right\|^2 \right) \quad e = E \left(\sum_{s=1}^n \left\| \sum_{j=n+1}^{\infty} \Psi_j(\theta) \underline{y}_{t-j} \right\|^2 \right).$$

Let's consider $\sum_{j=n+1}^{\infty} \Psi_j(\theta) \underline{y}_{t-j}$ a vectorial series in \mathbb{R}^m vectorial space.

$$\sum_{j=n+1}^{\infty} \Psi_j(\theta) \underline{y}_{t-j} = \lim_{N \rightarrow \infty} \sum_{j=n+1}^N \Psi_j(\theta) \underline{y}_{t-j}.$$

By triangular inequality, we have

$$d \leq E \left(\sum_{t=1}^n \lim_{N \rightarrow \infty} \left(\sum_{j=n+1}^N \left\| \Psi_j(\theta) \underline{y}_{t-j} \right\| \right)^2 \right),$$

therefore

$$d \leq E \left(\sum_{t=1}^n \lim_{N \rightarrow \infty} \left(\sum_{j=n+1}^N \left\| \Psi_j(\theta) \underline{y}_{t-j} \right\|^2 + 2 \sum_{J=n+1}^N \sum_{\substack{l=n+1 \\ j < l}}^N \left\| \Psi_j(\theta) \underline{y}_{t-j} \right\| \left\| \Psi_l(\theta) \underline{y}_{t-l} \right\| \right) \right).$$

Using inequality (6), we have

$$2 \sum_{J=n+1}^N \sum_{\substack{l=n+1 \\ j < l}}^N \left\| \Psi_j(\theta) \underline{y}_{t-j} \right\| \left\| \Psi_l(\theta) \underline{y}_{t-l} \right\| = \sum_{J=n+1}^N \left\| \Psi_j(\theta) \underline{y}_{t-j} \right\|^2 + \sum_{\substack{l=n+1 \\ j < l}}^N \left\| \Psi_l(\theta) \underline{y}_{t-l} \right\|^2,$$

consequently,

$$\begin{aligned} d &\leq 2E \left(\sum_{t=1}^n \lim_{N \rightarrow \infty} \left(\sum_{j=n+1}^N \left\| \Psi_j(\theta) \underline{y}_{t-j} \right\|^2 + \sum_{\substack{l=n+1 \\ j < l}}^N \left\| \Psi_l(\theta) \underline{y}_{t-l} \right\|^2 \right) \right) \\ &\leq 2E \left(\sum_{t=1}^n \lim_{N \rightarrow \infty} \left(\sum_{j=n+1}^N \left(\sum_{i=1}^m \Psi_{i,tj}^2 \right) + \sum_{\substack{l=n+1 \\ j < l}}^N \left(\sum_{i=1}^m \Psi_{i,tl}^2 \right) \right) \right), \end{aligned} \quad (7)$$

where $\Psi_{i,tj}$ and $\Psi_{i,tl}$ are the coefficients of the vector $\Psi_j(\theta) \underline{y}_{t-j}$.

Using the same argument as in (d), we obtain

$$\begin{aligned}
e &= E \left(\sum_{s=1}^n \left\| \sum_{j=n+1}^{\infty} \Psi_j(\theta) y_{s-j} \right\|^2 \right) \\
&\leq 2E \left(\sum_{s=1}^n \lim_{N \rightarrow \infty} \left(\sum_{j=n+1}^N \|\Psi_j(\theta) y_{s-j}\|^2 + \sum_{\substack{l=n+1 \\ j < l}}^N \|\Psi_l(\theta) y_{s-l}\|^2 \right) \right) \\
&\leq 2E \left(\sum_{s=1}^n \lim_{N \rightarrow \infty} \left(\sum_{j=n+1}^N \left(\sum_{i=1}^m \Psi_{i,sj}^2 \right) + \sum_{\substack{l=n+1 \\ j < l}}^N \left(\sum_{i=1}^m \Psi_{i,sl}^2 \right) \right) \right), \tag{8}
\end{aligned}$$

where $\Psi_{i,sj}$ and $\Psi_{i,sl}$ are the coefficients of the vector $\Psi_j(\theta) y_{s-j}$.

Odaki [11] characterizes the invertibility of the process by a function $f_n(d)$ defined as following:

$$f_n(d) = \begin{cases} 1/n & \text{for } 0 < |d| < \frac{1}{2}, \\ \frac{\log(n)}{n} & \text{for } d = -\frac{1}{2}, \\ 1/n^{2(1+d)} & \text{for } d < -\frac{1}{2}. \end{cases}$$

Odaki [11] shows that the order of magnitude of the sum of squares of these coefficients is

$$\sum_{i=0}^{+\infty} \Psi_{i,tj}^2 = o\left(\frac{1}{n}\right) \text{ if } d \in (-1/2; 1/2). \tag{9}$$

By (7), (8) and (9), we have

$$\begin{aligned}
\frac{C}{n^2 b_n^{2m+2}} o\left(\frac{1}{n}\right) &= o\left(\frac{1}{n^3 b_n^{2m+2}}\right) \\
&= o\left(\frac{1}{n^{2m\alpha+2\alpha+3} \mathcal{L}^{2m+2}(n)}\right).
\end{aligned}$$

Then

$$\sup_{x \in \mathbb{R}^m} |\hat{f}_{\theta_n}(x) - \tilde{f}_n(x)| = o\left(\frac{1}{n^{2m\alpha+2\alpha+3} \mathcal{L}^{2m+2}(n)}\right) \text{ when } n \rightarrow \infty,$$

$$\sup_{x \in \mathbb{R}^m} n^{\frac{1}{4}} |\hat{f}_{\theta_n}(x) - \tilde{f}_n(x)| = o\left(\frac{1}{n^{2m\alpha+2\alpha+\frac{11}{4}} \mathcal{L}^{2m+2}(n)}\right) \text{ when } n \rightarrow \infty.$$

Hence the convergence of (a) to $o\left(\frac{1}{n^{2m\alpha+2\alpha+\frac{11}{4}} \mathcal{L}^{2m+2}(n)}\right)$, where

$$2m\alpha + 2\alpha + \frac{11}{4} > 0.$$

Step 2. We will now prove the almost sure convergence of (b) using the Prakasa-Rao's inequality [12].

By Prakasa-Rao's inequality [12], we have

$$\mathbb{P}\left(|f_n(x) - Ef_n(x)| > \varepsilon \sqrt{\frac{s_m \log m}{m}}\right) \leq 2 \exp\left(-\frac{n \frac{s_m \log m}{m} \varepsilon^2}{8c_0 M}\right).$$

Let's consider a sequence $(s_n)_{n \in \mathbb{N}^*}$ defined by

$$S_n = nb_n,$$

where b_n is a sequence of bandwidths satisfying Assumption (A4). Let choose (b_n) such that

$$b_n = n^\alpha \ln(n); s_n = n^{1+\alpha} \ln(n) \text{ with } -1 < \alpha < 0.$$

$$-1 < \alpha < 0 \Leftrightarrow 0 < 1 + \alpha < 1$$

$$\Leftrightarrow 0 < n^{1+\alpha} < n.$$

Then for $n > 1$, the general term sequence s_n is positive.

Let's examine the limit of

$$S_n = n^{1+\alpha} \ln(n) \text{ and } \frac{S_n \ln(n)}{n},$$

$$0 < 1 + \alpha < 1 \Rightarrow \lim_{n \rightarrow +\infty} S_n = +\infty,$$

$$0 < 1 + \alpha < 1 \Rightarrow \lim_{n \rightarrow +\infty} \frac{S_n \ln(n)}{n} = 0.$$

Taking $n = m$ and $\delta_m(x, \underline{\varepsilon}_t) = \frac{1}{b_n^m} K\left(\frac{x - \underline{\varepsilon}_t}{b_n}\right)$, there is a positive constant c_0 such as

$$\sup_{x \in \mathbb{R}^m} \frac{1}{b_n^m} K\left(\frac{x - \underline{\varepsilon}_t}{b_n}\right) \leq c_0 s_n \rightarrow +\infty.$$

We can rewrite the Prakasa-Rao's inequality as following:

$$\mathbb{P}\left(|\tilde{f}_n(x) - E\tilde{f}_n(x)| > \varepsilon \sqrt{\frac{s_n \ln(n)}{n}}\right) \leq 2 \exp\left(-\frac{s_n \ln(n) \varepsilon^2}{8c_0 M}\right),$$

$$\tilde{f}_n(x) = \frac{1}{n b_n^m} \sum_{t=1}^n K\left(\frac{x - \underline{\varepsilon}_t}{b_n}\right), \quad x \in \mathbb{R}^m.$$

By Assumption (A6), we have

$$Q = 8c_0 M < \infty,$$

and it follows:

$$\mathbb{P}\left(|\tilde{f}_n(x) - E\tilde{f}_n(x)| > n^{\frac{\alpha}{2}} \varepsilon \ln(n)\right) \leq 2 \exp\left(-\frac{\varepsilon^2 n^{\alpha+1} \ln^2(n)}{Q}\right),$$

$$\mathbb{P}\left(n^{\frac{1}{4}} \sup_{x \in \mathbb{R}^m} |\tilde{f}_n(x) - E\tilde{f}_n(x)| > n^{\frac{2\alpha+1}{4}} \varepsilon \ln(n)\right) \leq 2 \exp\left(-\frac{\varepsilon^2 n^{\frac{2\alpha+3}{2}} \ln^2(n)}{Q}\right).$$

(10)

We will now dominate the next expression

$$\mathbb{P}\left(n^{\frac{1}{4}} \sup_{x \in \mathbb{R}^m} |\tilde{f}_n(x) - E\tilde{f}_n(x)| > n^{\frac{2\alpha+1}{4}} \varepsilon \ln(n)\right).$$

By Assumption (A4), we obtain

$$\begin{aligned} -1 < \alpha < 0 &\Leftrightarrow 1 < 2\alpha + 3 < 3 \\ &\Leftrightarrow 1 < n^{\frac{2\alpha+3}{2}} < \infty \\ &\Leftrightarrow 0 < \frac{\varepsilon^2 n^\mu \ln^2(n)}{Q} < \infty, \end{aligned}$$

where $\mu = \frac{2\alpha + 3}{2} > 0$.

There exists a sequence (V_n) such that $V_n = \beta \ln(n)$ with $\beta > 1$ for a certain rank

$$\frac{\varepsilon^2 n^\mu \ln^2(n)}{Q} > \beta \ln(n),$$

$$\text{under the constraints } \begin{cases} \varepsilon > 1 \\ n > 1 \\ Q > 0 \\ \mu > 0 \end{cases}.$$

Then

$$\exp\left(\frac{\varepsilon^2 n^\mu \ln^2(n)}{Q}\right) > n^\beta,$$

therefore

$$\exp\left(-\frac{\varepsilon^2 n^\mu \ln^2(n)}{Q}\right) < \frac{1}{n^\beta}.$$

By inequality (10), we have

$$\begin{aligned} & \mathbb{P} \left(n^{\frac{1}{4}} \sup_{x \in \mathbb{R}^m} |\tilde{f}_n(x) - E\tilde{f}_n(x)| > n^{\frac{2\alpha+1}{4}} \varepsilon \ln(n) \right) \leq \frac{2}{n^\beta}, \\ & \sum_{n \geq 1} \mathbb{P} \left(n^{\frac{1}{4}} \sup_{x \in \mathbb{R}^m} |\tilde{f}_n(x) - E\tilde{f}_n(x)| > n^{\frac{2\alpha+1}{4}} \varepsilon \ln(n) \right) \leq \sum_{n \geq 1} \frac{2}{n^\beta}, \\ & \sum_{n \geq 1} \mathbb{P} \left(n^{\frac{1}{4}} \sup_{x \in \mathbb{R}^m} |\tilde{f}_n(x) - E\tilde{f}_n(x)| > n^{\frac{2\alpha+1}{4}} \varepsilon \ln(n) \right) < +\infty. \end{aligned}$$

Using the Borel-Cantelli's lemma

$$n^{\frac{1}{4}} \sup_{x \in \mathbb{R}^m} |\tilde{f}_n(x) - E\tilde{f}_n(x)| = 0 \text{ a.s. when } n \rightarrow \infty.$$

Hence the almost sure convergence of $(\tilde{f}_n(x) - E\tilde{f}_n(x))$ to zero.

Step 3. The convergence of the bias (c).

By (3), we have

$$\begin{aligned} E(\tilde{f}_n(x)) &= \frac{1}{nb_n^m} E \left(\sum_{i=1}^n K \left(\frac{x - \varepsilon_1}{b_n} \right) \right) \\ &= \frac{1}{b_n^m} E \left(K \left(\frac{x - \varepsilon_1}{b_n} \right) \right) \\ &= \frac{1}{b_n^m} \int_{\mathbb{R}^m} K \left(\frac{x - z}{b_n} \right) f_{\theta_0}(z) dx \\ &= \int_{\mathbb{R}^m} K(v) f_{\theta_0}(x - vb_n) dv. \end{aligned}$$

The Taylor's expansion in a neighbourhood of x and under 3 of (A1), we obtain

$$f_{\theta_0}(x - \nu b_n) = f_{\theta_0}(x) + \sum_{k=1}^m \frac{\partial f_{\theta_0}}{\partial x_k}(x)(-b_n)\nu_k + \frac{b_n^2}{2} \sum_{j=1}^m \sum_{k=1}^m \frac{\partial^2 f_{\theta_0}}{\partial x_j \partial x_k}(x)\nu_j \nu_k + o(b_n^2)$$

$$\begin{aligned} E(\tilde{f}_n(x)) - f_{\theta_0}(x) &= \int_{\mathbb{R}^m} K(\nu) f_{\theta_0}(x - \nu b_n) d\nu - f_{\theta_0}(x) \int_{\mathbb{R}^m} K(\nu) d\nu \\ &= \int_{\mathbb{R}^m} K(\nu) [f_{\theta_0}(x - \nu b_n) - f_{\theta_0}(x)] d\nu \end{aligned}$$

$$\begin{aligned} \sup_{x \in \mathbb{R}^m} |E\tilde{f}_n(x) - f_{\theta_0}(x)| &\leq \int_{\mathbb{R}^m} K(\nu) \left[\frac{b_n^2}{2} \sum_{k=1}^m \left| \frac{\partial^2 f_{\theta_0}}{\partial x_k^2}(x) \right| \nu_k^2 + o(b_n^2) \right] d\nu \\ &\leq \Delta b_n^2 \int_{\mathbb{R}^m} K(\nu) [\nu_k^2 + o(1)] d\nu, \end{aligned}$$

where

$$\Delta = \frac{1}{2} \sum_{k=1}^m \left| \frac{\partial^2 f_{\theta_0}}{\partial x_k^2}(x) \right|.$$

Under (A1) and (A4), $\sup_{x \in \mathbb{R}^m} |E\tilde{f}_n(x) - f_{\theta_0}(x)| \rightarrow 0$ a.s. when $n \rightarrow \infty$.

The convergence of (a), (b), and (c) implies Lemma 1. □

We need the following lemma to prove the Theorem 1.

Lemma 2. *Beran [1] and Hili [5] consider \mathbb{F} the set of all densities with respect Lebesgue measure on \mathbb{R} .*

We define the functional $T : \mathbb{F} \rightarrow \Theta$ as following Let be $g \in \mathbb{F}$ we pose

$$A(g) = \left\{ \theta \in \Theta : H_2(g, f_\theta) = \min_{\theta \in \Theta} H_2(g, f_\theta) \right\},$$

where H_2 is the Hellinger distance.

If $A(g)$ is reduced to an unique element, then $T(g)$ is defined as the value of this element. Elsewhere, they choose an arbitrary but unique element of $A(g)$ and call it $T(g)$.

Proof. See Beran [1] and Hili [5] for proof. \square

Proof (Almost surely convergence).

Theorem 1 is a consequence of Lemmas 1 and 2. We have

$$\begin{aligned} \sup_{x \in \mathbb{R}^m} |\widehat{f}_{\theta_n}(x) - f_{\theta_0}(x)| &\leq \sup_{x \in \mathbb{R}^m} |\widehat{f}_{\theta_n}(x) - \widetilde{f}_n(x)| + \sup_{x \in \mathbb{R}^m} |\widetilde{f}_n(x) - E\widetilde{f}_n(x)| \\ &\quad + \sup_{x \in \mathbb{R}^m} |E\widetilde{f}_n(x) - f_{\theta_n}(x)|. \end{aligned}$$

By Lemma 1

$$|\widehat{f}_{\theta_n}(x) - f_{\theta_0}(x)| \rightarrow 0 \text{ a.s. when } n \rightarrow +\infty.$$

Then

$$\mathbb{P} \left\{ \lim_{n \rightarrow +\infty} \widehat{f}_{\theta_n}^{\frac{1}{2}}(x) = f_{\theta_0}^{\frac{1}{2}}(x) \text{ for all } x \right\} = 1.$$

Since

$$\int_{\mathbb{R}^m} \widehat{f}_{\theta_n}(x) dx = \int_{\mathbb{R}^m} f_{\theta_0}(x) dx = 1,$$

consequently

$$H_2(\widehat{f}_{\theta_n}, f_{\theta_0}) = \left\{ \int_{\mathbb{R}^m} |\widehat{f}_{\theta_n}^{\frac{1}{2}}(x) - f_{\theta_0}^{\frac{1}{2}}(x)|^2 dx \right\}^{\frac{1}{2}} \rightarrow 0 \text{ a.s. when } n \rightarrow \infty.$$

By Lemma 2, $T(f_{\theta_0}) = \theta_0$ inequality on Θ , then the functional T is continuous at f_{θ_0} in the Hellinger topology. Therefore

$$\widehat{\theta}_n = T(\widehat{f}_{\theta_n}(x)) \rightarrow T(f_{\theta_0}(x)) = \theta_0,$$

almost surely when $n \rightarrow \infty$.

This achieves the proof of the Theorem 1. \square

Proof of Theorem 2 (Asymptotic normality).

The following Lemma 3 and Lemma 4 were, respectively, proved by Beran [1] and by Wu and Mielniczuk [14].

Lemma 3. *Let's suppose that Assumptions (A2) and (A5) and the conditions C1 and C2 of Theorem 2 are satisfied and that θ_0 lies in interior of Θ . So for any density sequence $\{\widehat{f}_{\theta_n}\}$ convergent to f_{θ_0} in the Hellinger metric, we have*

$$\begin{aligned} T(\widehat{f}_{\theta_n}(x)) &= \theta_0 + \int_{\mathbb{R}^m} V_{\theta_0}(x) \left[\widehat{f}_{\theta_n}^{\frac{1}{2}}(x) - f_{\theta_0}^{\frac{1}{2}}(x) \right] dx \\ &\quad + A_n \int_{\mathbb{R}^m} \dot{g}_{\theta_0}(x) \left[\widehat{f}_{\theta_n}^{\frac{1}{2}}(x) - f_{\theta_0}^{\frac{1}{2}}(x) \right] dx, \end{aligned}$$

where A_n is a non-singular $m[m(p+q)+1]$ matrix whose components of $\sqrt{n}A_n$ tends to zero when $n \rightarrow +\infty$.

Lemma 4. *Let's suppose that Assumptions (A1), (A2), and (A4) are satisfied, then $N\left(0, f_{\theta_0}(x) \int_{\mathbb{R}^m} K^2(u) du\right)$ is the limit distribution of $\sqrt{nb_n} [\widehat{f}_{\theta_n}(x) - f_{\theta_0}(x)]$.*

Proof. Let's focus now on the proof of Theorem 2, referring to the above lemmas. By the Lemma 3, we have

$$\begin{aligned} T(\widehat{f}_{\theta_n}(x)) &= \theta_0 + \int_{\mathbb{R}^m} V_{\theta_0}(x) \left[\widehat{f}_{\theta_n}^{\frac{1}{2}}(x) - f_{\theta_0}^{\frac{1}{2}}(x) \right] dx \\ &\quad + A_n \int_{\mathbb{R}^m} \dot{g}_{\theta_0}(x) \left[\widehat{f}_{\theta_n}^{\frac{1}{2}}(x) - f_{\theta_0}^{\frac{1}{2}}(x) \right] dx. \end{aligned}$$

Since $T(\hat{f}_{\theta_n}(x)) = \hat{\theta}_n$ and by multiplying the equation above by \sqrt{n} , we have

$$\begin{aligned} \sqrt{n}(\hat{\theta}_n - \theta_0) &= \sqrt{n} \int_{\mathbb{R}^m} V_{\theta_0}(x) \left[\hat{f}_{\theta_n}^{\frac{1}{2}}(x) - f_{\theta_0}^{\frac{1}{2}}(x) \right] dx \\ &\quad + \sqrt{n} A_n \int_{\mathbb{R}^m} \dot{g}_{\theta_0}(x) \left[\hat{f}_{\theta_n}^{\frac{1}{2}}(x) - f_{\theta_0}^{\frac{1}{2}}(x) \right] dx. \end{aligned}$$

The components of $\sqrt{n}A_n \rightarrow 0$, when $n \rightarrow \infty$. Then

$$\sqrt{n}(\hat{\theta}_n - \theta_0) = B_n + o_p(1),$$

where

$$B_n = \sqrt{n} \int_{\mathbb{R}^m} V_{\theta_0}(x) \left[\hat{f}_{\theta_n}^{\frac{1}{2}}(x) - f_{\theta_0}^{\frac{1}{2}}(x) \right] dx.$$

Let's examine the limit law of B_n to deduce the limit law of $\sqrt{n}(\hat{\theta}_n - \theta_0)$,

where $V_{\theta_0}(x) \in L_2$ and $V_{\theta_0} \perp f_{\theta_0}^{\frac{1}{2}}$, where \perp means orthogonality in L_2 .

By Assumption (A2), $\hat{f}_{\theta_n}^{\frac{1}{2}}(x) > 0$ and the following algebraic equality, we can rewrite B_n .

$$\begin{aligned} \hat{f}_{\theta_n}^{\frac{1}{2}}(x) - f_{\theta_0}^{\frac{1}{2}}(x) &= \frac{\hat{f}_{\theta_n}(x) - f_{\theta_0}(x)}{2f_{\theta_0}^{\frac{1}{2}}(x)} - \frac{(\hat{f}_{\theta_n}(x) - f_{\theta_0}(x))^2}{2f_{\theta_0}^{\frac{1}{2}}(x) \left(\hat{f}_{\theta_n}^{\frac{1}{2}}(x) + f_{\theta_0}^{\frac{1}{2}}(x) \right)^2}. \\ B_n &= \sqrt{n} \int_{\mathbb{R}^m} V_{\theta_0}(x) \left[\frac{\hat{f}_{\theta_n}(x) - f_{\theta_0}(x)}{2f_{\theta_0}^{\frac{1}{2}}(x)} - \frac{(\hat{f}_{\theta_n}(x) - f_{\theta_0}(x))^2}{2f_{\theta_0}^{\frac{1}{2}}(x) \left(\hat{f}_{\theta_n}^{\frac{1}{2}}(x) + f_{\theta_0}^{\frac{1}{2}}(x) \right)^2} \right] dx, \end{aligned} \tag{11}$$

by distributivity in (11), we have

$$B_n = \sqrt{n} \int_{\mathbb{R}^m} V_{\theta_0}(x) \frac{\widehat{f}_{\theta_n}(x) - f_{\theta_0}(x)}{2f_{\theta_0}^{\frac{1}{2}}(x)} dx + C_n, \quad (12)$$

where

$$C_n = -\sqrt{n} \int_{\mathbb{R}^m} \frac{V_{\theta_0}(x)}{2f_{\theta_0}^{\frac{1}{2}}(x) \left(\widehat{f}_{\theta_n}^{\frac{1}{2}}(x) + f_{\theta_0}^{\frac{1}{2}}(x) \right)^2} \left(\widehat{f}_{\theta_n}(x) - f_{\theta_0}(x) \right)^2 dx.$$

Using inequality

$$2f_{\theta_0}^{\frac{1}{2}} \left(\widehat{f}_{\theta_n}^{\frac{1}{2}} + f_{\theta_0}^{\frac{1}{2}} \right)^2 = 2f_{\theta_0}^{\frac{3}{2}} + \gamma > 2f_{\theta_0}^{\frac{3}{2}} \quad \text{with } \gamma > 0,$$

and posing $\delta = \inf_{x \in \mathbb{R}^m} f(x)$, we can take C_n in absolute value as following:

$$\begin{aligned} |C_n| &\leq \sqrt{n} \int_{\mathbb{R}^m} \frac{|V_{\theta_0}(x)|}{2f_{\theta_0}^{\frac{1}{2}}(x) \left(\widehat{f}_{\theta_n}^{\frac{1}{2}}(x) + f_{\theta_0}^{\frac{1}{2}}(x) \right)^2} \left(\widehat{f}_{\theta_n}(x) - f_{\theta_0}(x) \right)^2 dx. \\ |C_n| &\leq \sqrt{n} \int_{\mathbb{R}^m} \frac{|V_{\theta_0}(x)|}{2\delta^{\frac{3}{2}}} \left(\widehat{f}_{\theta_n}(x) - f_{\theta_0}(x) \right)^2 dx \\ &\leq \frac{1}{2} \delta^{-\frac{3}{2}} \int_{\mathbb{R}^m} |V_{\theta_0}(x)| \sqrt{n} \left(\widehat{f}_{\theta_n}(x) - f_{\theta_0}(x) \right)^2 dx. \end{aligned}$$

By Lemma 1

$$n^{\frac{1}{4}} \sup_{x \in \mathbb{R}^m} \left| \widehat{f}_{\theta_n}(x) - f_{\theta_0}(x) \right| \rightarrow 0 \text{ a.s.} \quad \text{with } n \rightarrow \infty,$$

then

$$\sqrt{n} \left(\hat{f}_{\theta_n}(x) - f_{\theta_0}(x) \right)^2 \rightarrow 0 \text{ a.s. with } n \rightarrow \infty.$$

V_{θ_0} is continuous and bounded (for θ_0 fixed). By applying Vitali's theorem on the sequence

$$W_n(x) = |V_{\theta_0}(x)| \sqrt{n} \left(\hat{f}_{\theta_n}(x) - f_{\theta_0}(x) \right)^2,$$

$|C_n| \rightarrow 0$ in probability when $n \rightarrow \infty$.

Let's consider the first term on the right of Equation (12)

$$\sqrt{n} \int_{\mathbb{R}^m} V_{\theta_0}(x) \frac{\hat{f}_{\theta_n}(x) - f_{\theta_0}(x)}{2f_{\theta_0}^{\frac{1}{2}}(x)} dx. \quad (13)$$

Therefore, by Lemma 4, the limit distribution of (13) is $N(0, \Sigma^2)$, where

$$\begin{aligned} \Sigma^2 &= \int_{\mathbb{R}^m} \left(\frac{V_{\theta_0}(x)}{2f_{\theta_0}^{\frac{1}{2}}(x)} \right) \left(\frac{V_{\theta_0}(x)}{2f_{\theta_0}^{\frac{1}{2}}(x)} \right)^t \int_{\mathbb{R}^m} K^2(u) du f_{\theta_0}(x) dx \\ &= \frac{1}{4} \int_{\mathbb{R}^m} V_{\theta_0}(x) V_{\theta_0}^t(x) dx \int_{\mathbb{R}^m} K^2(u) du \\ &= \frac{1}{4} \left[\int_{\mathbb{R}^m} \dot{g}_{\theta_0}(x) \dot{g}_{\theta_0}^t(x) dx \right]^{-1} \int_{\mathbb{R}^m} K^2(u) du. \end{aligned}$$

Hence the result.

Acknowledgements

We thank the anonymous referee for a careful reading and comments, which helped to improve the quality of the note.

References

- [1] R. Beran, Minimum Hellinger distance estimates for parametric models, *Ann. Statist.* 5(3) (1977), 445-463.
DOI: <http://dx.doi.org/10.1214/aos/1176343842>
- [2] A. L. Bitty and O. Hili, Estimateurs du minimum de distance de Hellinger des processus linéaires à longue mémoire, *Comptes Rendus Mathématique* 348(7-8) (2010), 445-448.
DOI: <https://doi.org/10.1016/j.crma.2010.02.020>
- [3] C. W. J. Granger and A. Andersen, On the invertibility of time series models, *Stochastic Process. Appl.* 8(1) (1978), 87-92.
DOI: [https://doi.org/10.1016/0304-4149\(78\)90069-8](https://doi.org/10.1016/0304-4149(78)90069-8)
- [4] C. W. J. Granger and R. Joyeux, An introduction to long-memory time series models and fractional differencing, *J. Time Series Anal.* 1(1) (1980), 15-29.
DOI: <https://doi.org/10.1111/j.1467-9892.1980.tb00297.x>
- [5] O. Hili, On the estimation of nonlinear time series models, *Stochastics and Stochastic Reports* 52(3-4) (1995), 207-226.
DOI: <http://dx.doi.org/10.1080/17442509508833972>
- [6] J. R. M. Hosking, Fractional differencing, *Biometrika* 68(1) (1981), 165-176.
DOI: <https://doi.org/10.1093/biomet/68.1.165>
- [7] A. Kamagaté and O. Hili, Estimation par le minimum de distance de Hellinger d'un processus ARFIMA, *Comptes Rendus Mathématique* 350(13-14) (2012), 721-725.
DOI: <http://doi.org/10.1016/j.crma.2012.07.005>
- [8] A. Kamagaté and O. Hili, The quasi maximum likelihood approach to statistical inference on a non-stationary multivariate ARFIMA process, *Random Oper. Stoch. Equ.* 21(3) (2013), 305-320.
DOI: <https://doi.org/10.1515/rose-2013-0014>
- [9] L. Mayoral, Minimum distance estimation of stationary and non-stationary ARFIMA processes, *Econom. J.* 10(1) (2007), 124-148.
DOI: <https://doi.org/10.1111/j.1368-423X.2007.00202.x>
- [10] A. N'dri and O. Hili, Hellinger distance estimation of strongly dependent multi-dimensional Gaussian processes, *International Journal of Statistics and Probability* 2(3) (2013), 70-84.
DOI: <https://doi.org/10.5539/ijsp.v2n3p70>

- [11] M. Odaki, On the invertibility of fractionally differenced ARIMA processes, *Biometrika* 80(3) (1993), 703-709.
DOI: <https://doi.org/10.1093/biomet/80.3.703>
- [12] B. L. S. Prakasa Rao, *Nonparametric Functional Estimation*, New York: Academic Press, 1983.
- [13] F. Sowell, *A Decomposition of Block Toeplitz Matrices with Applications to Vector Time Series*, GSIA Mimograph, Carnegie Mellon University, 1989.
- [14] W. B. Wu and J. Mielniczuk, Kernel density estimation for linear process, *The Annals of Statistics* 30(5) (2002), 1441-1459.

