



HAL
open science

Local Asymptotic Normality For Linear Functional AR Processes with multi dimensional Parameter

Nesrine Kara-Terki, Tahar Mourid

► **To cite this version:**

Nesrine Kara-Terki, Tahar Mourid. Local Asymptotic Normality For Linear Functional AR Processes with multi dimensional Parameter. *Annales de l'ISUP*, 2019, 63 (1), pp.41-72. hal-03604383

HAL Id: hal-03604383

<https://hal.science/hal-03604383>

Submitted on 10 Mar 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Local Asymptotic Normality For Linear Functional AR Processes with multi dimensional Parameter

Nesrine Kara-Terki¹ and Tahar Mourid²

*Laboratoire de Statistiques et Modélisations Aléatoires
Centre Universitaire Belhadj Bouchaib -Ain Témouchent - Algerie ¹
Université Abou Bekr Belkaid -Tlemcen- 13000 - Algerie ²*

Abstract

We consider Local Asymptotic Normality (LAN) and Uniform Local Asymptotic Normality (ULAN) conditions for Linear Functional Autoregressive Processes with multi dimensional parameter. The correlation operator depends on an unknown multi-dimensional parameter and we establish the LAN and ULAN conditions. Then we derive with asymptotic properties of the conditional maximum likelihood and Bayes estimators giving Hajek minimax bound, consistency, asymptotic normality and efficiency of these estimators yielding their optimality.

Keywords. *Linear Functional AR processes- Local Asymptotic Normality- Maximum likelihood estimators-Hajek bound- Efficiency.*

1. Introduction

Nowadays in many fields such that economy, biology, medicine etc., a huge quantities of data is collected and used to draw modeling evolution problems and prediction information. Doing so, generated models can be described by continuous smooth dynamics allowing often to accurate estimates of parameters in the phase analysis, or with effective data noise reduction through curve smoothing leading to data with irregular time sampling schedules. In order to describe conveniently such continuous flow of data, practitioners and statisticians often consider them as "high dimensional" vectors resulting from realizations of random curves or a sample of a valued function space random variable. Then they may apply statistical concepts coming from multivariate data analysis. Now their statistical analysis is carried out through the well-known Functional Data Analysis (FDA) that we may find out some basic concepts in Ramsay [26] with strong arguments for its wide

scoop. Since this area becomes more attractive giving a central topic of studies of modern statistics. There exist many textbooks and we may cite without be exhaustive the monographs by Bosq [4], Ferraty and Vieu [5], Ramsay and Silverman [26] (and references therein), Horvath and Kokozka [9], Hsing and Eubank [30] where generalizations of existing models and classical inference procedures are well developed. We also mention recent surveys in FDA realised by different authors, we cite (Cuevas [7], Goia [6],...)

We are interested by the class of Linear Functional Autoregressive Processes. This class of processes has received more attention and investigated by many authors for predicting continuous time random processes. We may cite a few Bosq [4], Mas [21], Antoniadis and Sapatinas [2], [23] etc... We also refer to the papers by [19] and references therein. We quote the article by Arkoun and Pergamenchchikov [3] and references therein, concerning recent progress in nonparametric functional time series and Vincent Lefieux's thesis [31] for semi-parametric functional approaches. Various applications have been successfully performed in a many areas. The examples range from prediction of electricity consumption, road traffic, El-Nino temperature to concentration levels in air pollution (see Bosq [4]). An application to environmental data processing was given by Omatu et al. [24] predicting the concentration levels of pollutants in air pollution analyzed by Bucy-Kalman filter system. In the same framework, Loges [20] treated the case of the unobservable rv's are ruled by a Hilbert space valued autoregressive equation AR(1). we also cite the recent works of [1], [14] and references therein very interesting papers with regard to the functional time series process, especially autoregressive models.

We deal with the Local Asymptotic Normality (LAN) condition introduced by LeCam which became a fundamental concept in parametric estimation. In a previous paper (Kara and Mourid [13]), we show the LAN condition for the class of functional autoregressive processes when the correlation operator depends on one-dimensional parameter. We then obtain Hajek minimax bound, consistency, asymptotic normality and efficiency of the conditional maximum likelihood and Bayes estimators. In this paper, we extend these results when the correlation operator depends on a multi-dimensional parameter namely $\theta \in \Theta \subset \mathbb{R}^p$, for $p \geq 1$. From this point of view, the proofs of the LAN condition in multi-dimensional case requires more technics and are not direct especially in the functional setting. An example of operator depending on a multi-dimensional parameter is given by Loges [20] where $\rho_\theta = \sum_{i=1}^p \theta_i \rho_i$, $\theta_i \in]-1, 1[$, and ρ_i a bounded linear operator with $\|\rho_i\|_{\mathcal{L}} < 1$. For a one-dimensional parameter, recall Ornstein-Uhlenbeck process autoregressive representation (1) (see Bosq [4] Ex. 3.4 p. 76) where the operator ρ_θ is explicit and defined by $\rho_\theta(f)(t) = e^{-\theta t} f(1)$, $f \in L^2_{[0,1]}$, $\theta > 0$, $t \in [0, 1]$. Recall that under LAN condition, local experiments converge weakly, in the LeCam distance, to Gaussian shift experiment and lower bounds for the law of a class of

regular estimators (Hajek convolution theorem) and their efficiency are derived. (see Ibragimov and Hasminski [11], LeCam and Yang [18], Kutoyants [16] [17]). Since LAN condition is satisfied by many statistical models under mild conditions. It mainly relies on checking mean square differentiability property with respect to the parameter of probability measures $(P_\theta, \theta \in \Theta)$. For ergodic Markov process, this condition is investigated by Philippou and Roussas [25] and Roussas [27]. In autoregressive time series framework, Swensen [29] and Kreiss [15] showed that the LAN condition also holds (see also [22]) and Hwang and Basawa [10] discussed this condition for a \mathbb{R}^d -valued stationary Markov non-linear time series.

Our aim is to establish LAN and Uniform LAN conditions for functional autoregressive processes when $\theta \in \Theta \subset \mathbb{R}^p$, for $p \geq 1$. We show that the log-likelihood ratio has the typical LAN decomposition with the same scale $1/\sqrt{n}$, therefore the local experiments converge weakly, in the LeCam distance, to Gaussian shift experiment. Then using general results by Ibragimov and Hasminski [11], we derive asymptotic properties of the conditional maximum likelihood (MLE) and Bayes estimators and hence set their optimality when $\theta \in \Theta \subset \mathbb{R}^p$ only for $p = 1, 2, 3$. Nevertheless, the properties of conditional maximum likelihood estimators MLE derived in Theorem 3 still remain difficult to extend in the full generality when $p \geq 4$. Our objective in the future will be to find suitable techniques to overcome varied levels of difficulty.

This paper is organized as follows. Section 2 introduces notations and definitions. Section 3 deals with main results. Section 4 is devoted to examples. Section 5 contains simulations results. Auxiliary lemmas are in Section 6 and technical proofs are postponed to Appendix.

2. Notations and Definitions

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a complete probability space and (H, \mathcal{H}) a real separable Hilbert space equipped with the inner product $\langle \cdot, \cdot \rangle$, the associate norm $\|\cdot\|$ and its Borel σ -field \mathcal{H} . We denote by $\mathcal{L}(H)$ the space of bounded linear operators defined on H into H , and by $\|\cdot\|_{\mathcal{L}}$ the usual norm of bounded linear operators. A strong H -Gaussian white noise $(\varepsilon_n, n \in \mathbb{Z})$ on H , defined on $(\Omega, \mathcal{A}, \mathbb{P})$, with zero mean and covariance operator C_ε , is a Gaussian sequence of H -valued i.i.d. rv's. A process $X = (X_n, n \in \mathbb{N})$ defined on $(\Omega, \mathcal{A}, \mathbb{P})$ with values in (H, \mathcal{H}) is a Hilbert space valued autoregressive process (ARH process) if

$$X_n = \rho_\theta(X_{n-1}) + \varepsilon_n \quad (1)$$

where $\rho_\theta \in \mathcal{L}(H)$ depending on a parameter θ and $(\varepsilon_n, n \in \mathbb{Z})$ is a strong H -Gaussian white noise. We suppose that the parameter $\theta \in \Theta$ where the parameter space is an open set of $\mathbb{R}^p, p \geq 1$. In Introduction we cited the example

of the Ornstein-Uhlenbeck process that admits the representation (1) with an explicit operator ρ_θ as well as the example given by Loges [20] both depending on a parameter θ . Our goal is to estimate the multi-dimensional parameter θ in the LAN setting based on the observations of (1). We consider the family of probabilities $(P_{n,\theta}, \theta \in \Theta)$ of the conditional probability law of (X_1, \dots, X_n) given the initial value $X_0 : P_{n,\theta} = P_{X_1, \dots, X_n}^{X_0}$. We denote by $Z_{n,\theta}(h) := dP_{n,\theta_n}/dP_{n,\theta}$ the derivative of the absolute continuous part of P_{n,θ_n} with respect to $P_{n,\theta}$ where $\theta_n = \theta + \varphi_n(\theta)h$, $\varphi_n(\theta)$ is a sequence of definite positive matrices and $h \in \mathbb{R}^p$. The basic tool in this study is the local asymptotic normality structure of the family of the conditional likelihood ratio $Z_{n,\theta}(h)$ associated to (1), in the vicinity of θ . We first recall the Local Asymptotic Normality condition.

Definition 1. *The family of parametric probability measures $(P_{n,\theta}, \theta \in \Theta)$ is said to satisfy Local Asymptotic Normality condition (LAN condition) at a point θ if there exists a sequence of definite positive matrices $(\varphi_n(\theta))$, $\varphi_n(\theta) \rightarrow 0$, a sequence of rv's $(S_n(\theta))$ such that for all $h \in \mathbb{R}^p$ the log-likelihood ratio $\ln Z_{n,\theta}(h)$ satisfies the following representation under $P_{n,\theta}$:*

$$\ln Z_{n,\theta}(h) = hS_n(\theta) - 1/2\|h\|^2 + o_{P_{n,\theta}}(1), \quad \text{as } n \rightarrow \infty, \quad (2)$$

where

$$S_n(\theta) \Rightarrow_{n \rightarrow \infty} \mathcal{N}(0, I)$$

where I is the identity matrix in \mathbb{R}^p .

The family of parametric probability measures $(P_{n,\theta}, \theta \in \Theta)$ is said to satisfy Uniform Local Asymptotic Normality condition (ULAN condition) if we have LAN condition (2) and moreover for all $M > 0$ under $P_{n,\theta}$:

$$\sup_{|h| \leq M} |\ln Z_{n,\theta}(h) - hS_n(\theta) + 1/2\|h\|^2| = o_{P_{n,\theta}}(1). \quad (3)$$

Throughout the paper we need to use the following functional matters. Let $\{e_k, k \geq 1\}$ be a complete orthonormal system in H consisting of eigenvectors of the covariance operator C_ε of the rv's ε_0 and $\{\sigma_k^2, k \geq 1\}$ the corresponding eigenvalues where $\sigma_k^2 > 0$. For $x \in H$ we denote by $a_k(x) := \langle x, e_k \rangle$ the Fourier coefficient of x with respect to the vector e_k . The continuity and differentiability are with respect to the parameter θ . If a function $g(\theta)$ is differentiable we denote by $\dot{g}(\theta) = \nabla g(\theta)$ the gradient vector of g and $Jg(\theta)$ the matrix defined by $Jg(\theta) = \nabla g(\theta) \nabla^T g(\theta)$ where T designs the transpose of a vector.

3. Main Results

In general the study of LAN condition requires technical conditions and we make the following assumptions which essentially carry on the operator ρ_θ and the

Fourier coefficient $a_k(x)$. The parameter space Θ is a convex open set of \mathbb{R}^p , $p \geq 1$.

Assumptions

A0. $\exists C > 0$, $\sup_{\theta \in \Theta} \|\rho_\theta\|_{\mathcal{L}} = C < 1$.

A1. For all $\theta \in \Theta$ and for all $x \in H$,

$$\sum_{k=1}^{\infty} \sigma_k^{-2} a_k^2(\rho_\theta x) < \infty.$$

A2. For all $k \geq 1$ and $i \geq 1$, Fourier coefficients $a_k(\rho_\theta X_i)$ are differentiable and for all $1 \leq j \leq p$, the function

$$\sum_{k=1}^{\infty} \sigma_k^{-2} \frac{\partial}{\partial \theta_j} (a_k^2(\rho_\theta X_{i-1}) - 2a_k(X_i)a_k(\rho_\theta X_{i-1}))$$

is finite and continuous for all $\theta \in \Theta$, P_θ -a.s..

A3.

a. For all $1 \leq j \leq p$

$$E_\theta \left(\sup_{u \in \Theta} \sum_{k=1}^{\infty} \sigma_k^{-2} \left[\frac{\partial}{\partial u_j} a_k(\rho_u X_1) \right]^2 \right) < \infty.$$

b. The matrix

$$\Gamma(\theta) = \sum_{k=1}^{\infty} \sigma_k^{-2} E_\theta [\nabla(a_k \rho_\theta X_1) \nabla^T(a_k \rho_\theta X_1)],$$

is positive definite.

Now we can state the LAN condition (2) of the family of probabilities $(P_{n,\theta}, \theta \in \Theta)$ satisfying (1).

Theorem 1. Under A0, A1, A2, A3, the family of probability measures $(P_{n,\theta}, \theta \in \Theta)$ of (1) satisfies the LAN condition (2) at each point $\theta \in \Theta$ where the matrix $\varphi_n(\theta) = \Gamma^{-1/2}(\theta)/\sqrt{n}$, the random vector

$$S_n(\theta) = n^{-1} \Gamma^{-1/2}(\theta) \sum_{i=1}^n \sum_{k=1}^{\infty} \frac{1}{\sigma_k^2} \nabla(\rho_\theta X_{i-1})_k [a_k(X_i) - a_k(\rho_\theta X_{i-1})]$$

and the matrix

$$\Gamma(\theta) = \sum_{k=1}^{\infty} \sigma_k^{-2} E_\theta [\nabla(a_k \rho_\theta X_1) \nabla^T(a_k \rho_\theta X_1)],$$

where $\Gamma^{-1/2}(\theta)$ is the inverse matrix of $\Gamma^{1/2}(\theta)$ (the square root matrix of $\Gamma(\theta)$) and ∇ is the operator gradient.

To establish the ULAN condition (3), we need to add the following conditions.

A4. For all $k \geq 1$, Fourier coefficients $a_k(\rho_\theta(X_1))$ are three times differentiable functions on Θ a.s. and $\forall u \in \Theta$.

- a. For all $1 \leq j \leq p$ $\sup_{\theta, u \in \Theta} E_u \sum_{k=1}^{+\infty} \sigma_k^{-2} [\frac{\partial}{\partial \theta_j} a_k(\rho_\theta(X_1))]^2 < \infty$.
- b. For all $1 \leq j, r \leq p$ $\sup_{\theta, u \in \Theta} E_u \sum_{k=1}^{+\infty} \sigma_k^{-2} [\frac{\partial^2}{\partial \theta_j \partial \theta_r} a_k(\rho_\theta(X_1))]^2 < \infty$.
- c. For all $1 \leq j, r, s \leq p$, $\sup_{\theta, u \in \Theta} E_u \sum_{k=1}^{+\infty} \sigma_k^{-2} [\frac{\partial^3}{\partial \theta_j \partial \theta_r \partial \theta_s} a_k(\rho_\theta(X_1))]^2 < \infty$.

A5. For all $x \in H$ and for all $k \geq 1$, the Fourier coefficients $a_k(\rho_\theta(x))$ are differentiable in $\theta \in \Theta$ and $\forall 1 \leq j \leq p$,

$$\frac{\partial}{\partial \theta_j} \langle \rho_\theta x, e_k \rangle = \langle \rho_\theta^{(j)} x, e_k \rangle = \langle x, (\rho_\theta^{(j)})^* e_k \rangle,$$

where $\rho_\theta^{(j)}$ defined by $\rho_\theta^{(j)} x = \frac{\partial}{\partial \theta_j} \rho_\theta x$ and its adjoint $(\rho_\theta^{(j)})^*$ are bounded and linear operators.

A6.

a.

$$\sup_{\theta \in \Theta} \sum_{k=1}^{\infty} \sigma_k^{-2} \|(\rho_\theta^{(j)})^* e_k\|^2 < \infty.$$

b.

$$\sup_{\theta \in \Theta} \sum_{k=1}^{\infty} \sigma_k^{-1} \|(\rho_\theta^{(j)})^* e_k\| < \infty.$$

The following result establish the ULAN condition (3) of the family of probabilities $(P_{n,\theta})$ satisfying (1).

Theorem 2. Under **A0**, ..., **A6(a)**, the family of probability measures $(P_{n,\theta}, \theta \in \Theta)$ of (1) satisfies the ULAN condition (3) where $S_n(\theta)$ and $\Gamma(\theta)$ are defined in Theorem 1.

In particular for any bounded continuous loss function l such that the sets $\{l(x) \leq a\}$, $a > 0$, are convex and symmetric and for all estimator $\hat{\theta}_n$, we have Hajek bound

$$\lim_{\delta \rightarrow 0} \liminf_{n \rightarrow \infty} \sup_{|\theta - \theta_0| < \delta} E_\theta [l(\sqrt{n} \Gamma^{1/2}(\theta_0)(\hat{\theta}_n - \theta))] \geq E l(\zeta),$$

where $\zeta \hookrightarrow \mathcal{N}(0, I)$ in R^p .

Remark 1. *A0* ensures a strictly stationary ergodic Markov process (X_n) $f(1)$ with transition $Q_\theta(x, dy) = P_{\varepsilon_1 + \rho_\theta(x)}(dy)$ (see Bosq [4] Th. 3.1 p. 74 and Jourid [23] Chap. 1). *A1* ensures the equivalence of Gaussian measures via a Kakutani Theorem. The rest of assumptions mainly ensure finiteness of series and moments, differentiability, interchanging between sum and integral signs, etc..

2. Recall that an estimator is said to be efficient in Hajek sense if we have equality in the Hajek bound of Theorem 2.

In this part we consider the parameter space Θ an open set of \mathbb{R}^p , $p = 1, 2, 3$. We define the conditional maximum likelihood estimator MLE $\hat{\theta}_n$ by

$$\hat{\theta}_n = \arg \max_{\theta \in \Theta} \ln \frac{dP_{n,\theta}}{dP_{n,\theta_0}},$$

where θ_0 is a fixed value in $\Theta \in \mathbb{R}^p$ where $p \in \{1, 2, 3\}$.

In order to establish properties of the conditional maximum likelihood estimators ($\hat{\theta}_n, n \geq 1$), we make the following assumptions.

A7. For all $\theta \in \Theta$, ρ_θ is a Hilbert-Schmidt symmetric operator commuting with the covariance operator C_ε of ε_0 .

A8. The eigenvalues $(\lambda_{k,\theta}, k \geq 1)$ of ρ_θ are such that $\theta \rightarrow \lambda_{k,\theta}$ are C^1 -class functions

and $\inf_{\theta \in \Theta} \mu_\theta > 0$ where μ_θ is the smallest eigenvalue of the matrix $\sum_{k=1}^{\infty} \nabla \lambda_{k,\theta} \nabla^T \lambda_{k,\theta}$.

The properties of the conditional ML estimators are summarized in the following theorem.

Theorem 3. Under *A0*, ..., *A5*, *A6(a)*, *A7*, *A8*, the conditional maximum likelihood estimator (MLE) $\hat{\theta}_n$ satisfies :

1. For all compact interval $K \subset \Theta$ and $\delta > 0$, we have

$$\sup_{\theta \in K} P_\theta(\|\hat{\theta}_n - \theta\| > \delta) \rightarrow_{n \rightarrow \infty} 0.$$

2. Under *P0*

$$\sqrt{n} \Gamma^{1/2}(\theta)(\hat{\theta}_n - \theta) \Rightarrow_{n \rightarrow \infty} \zeta,$$

where $\zeta \hookrightarrow \mathcal{N}(0, I)$ and the convergence is uniform on any compact set $K \subset \Theta$. The matrix $\Gamma(\theta)$ is defined in Theorem 1.

3. The conditional maximum likelihood estimator MLE $\hat{\theta}_n$ is efficient in Hajek sense.

Remark 2. From the limit law of the maximum likelihood estimator, it is quite possible to test hypotheses and determine confidence bands. For the prediction of the process, from relation (1) we know that $\rho_\theta(X_n)$ is the best probabilistic predictor of X_{n+1} . Hence we define a predictor of X_{n+1} by $\rho_{\hat{\theta}_n}(X_n)$, where $\hat{\theta}_n$ is the conditional maximum likelihood estimators MLE. We can show the behavior of predictors and give the values of both errors RMAE and MSE knowing that the quality of the predictor $\rho_{\hat{\theta}_n}(X_n)$ calculated at points $t_i, i = 1; \dots; m$, is measured by relative mean absolute error RMAE and mean-squared error MSE defined by: $RMAE = 1/m \sum_{i=1}^m |\hat{X}_{n+1}(t_i) - X_{n+1}(t_i)|/|X_{n+1}(t_i)|$, $MSE = 1/m \sum_{i=1}^m (\hat{X}_{n+1}(t_i) - X_{n+1}(t_i))^2$.

For a particular form of the Hilbert-Schmidt operator ρ_θ especially when only the eigenvalues depend on the parameter, we have the following result.

Corollary 1. Let $\theta \in \Theta \subset \mathbb{R}^p, p=1, 2, 3$. The operator ρ_θ of (1) is given by : $\rho_\theta x = \sum_{k=1}^\infty \lambda_{\theta,k} < x, e_k > e_k$. Suppose

1. $\exists c > 0$ such that $\sup_{\theta \in \Theta} \sup_k |\lambda_{\theta,k}| \leq c < 1$ and $\forall k$, the function $\theta \rightarrow \lambda_{\theta,k}$ has all partial derivatives of order 1, 2 and 3.

2. $\exists c_i > 0, i = 0, \dots, 3$, such that for all $1 \leq j, r, s \leq p$,

$$\sup_{k \geq 1} \sup_{\theta \in \Theta} \frac{|\lambda_{\theta,k}|^2}{\sigma_k^2} \leq c_0, \sup_{k \geq 1} \sup_{\theta \in \Theta} \sigma_k^{-2} \left| \frac{\partial}{\partial \theta_j} \lambda_{\theta,k} \right|^2 \leq c_1,$$

$$\sup_{k \geq 1} \sup_{\theta \in \Theta} \sigma_k^{-2} \left| \frac{\partial^2}{\partial \theta_j \partial \theta_r} \lambda_{\theta,k} \right|^2 \leq c_2, \sup_{k \geq 1} \sup_{\theta \in \Theta} \sigma_k^{-2} \left| \frac{\partial^3}{\partial \theta_j \partial \theta_r \partial \theta_s} \lambda_{\theta,k} \right|^2 \leq c_3.$$

3. $\sup_{\theta \in \Theta} \sum_{k=1}^\infty \sigma_k^{-2} \left| \frac{\partial}{\partial \theta_j} \lambda_{\theta,k} \right|^2 < \infty$

4. $\inf_{\theta \in \Theta} \mu_\theta > 0$ where μ_θ is the smallest eigenvalue of the matrix $\sum_{k=1}^\infty \nabla \lambda_{k,\theta} \nabla^T \lambda_{k,\theta}$.

Then the conditional MLE $\hat{\theta}_n$ satisfy all the properties of Theorem 3.

4. Examples

We give two examples where the conditions **A0**, ..., **A8** are satisfied and then we deduce the LAN and ULAN conditions.

Example 1. Let H be a real separable Hilbert space and $X = (X_n, n \in \mathbb{N})$ a H -valued AR process defined by (1) where ρ_θ is given by [20]:

$$\rho_\theta = \theta_1 \rho^{(1)} + \theta_2 \rho^{(2)}$$

where $\theta = (\theta_1, \theta_2) \in \Theta = B(0, 1)$ (the unit ball in \mathbb{R}^2) and $\rho^{(1)}, \rho^{(2)}$ are two symmetric Hilbert-Schmidt operators commuting with the covariance operator C_ε . Let $(\sigma_k^2, e_k, k \geq 1)$ be the eigen-elements of the covariance operator C_ε where $e_k(t) = \sqrt{2} \sin((k - 1/2)\pi t)$ and $\sigma_k^2 = 1/(k - 1/2)^2 \pi^2$. The three operators $\rho^{(1)}, \rho^{(2)}$ and C_ε share the same set of eigenvectors $(e_k, k \geq 1)$. The eigenvalues of $\rho^{(1)}$ and $\rho^{(2)}$ are denoted by $(\lambda_k(\rho^{(1)}), (\lambda_k(\rho^{(2)}))$ respectively. We take $\lambda_k(\rho^{(1)}) = (k + 1/2)^{-2}$, and $\lambda_k(\rho^{(2)}) = (k + 1)^{-3}$.

Example 2. Let $X = (X_n, n \in \mathbb{N})$ be a H -valued AR process defined by (1) where ρ_θ is given by $\rho_\theta x = \sum_{k=1}^{\infty} \lambda_{\theta, k} \langle x, e_k \rangle e_k$ and $\theta = (\theta_1, \theta_2) \in \Theta = B(0, 1)$ (the unit ball of \mathbb{R}^2) and $\lambda_{\theta, k} = e^{-k(\theta_1+1)}(k)^{-5} + e^{-k(\theta_2+1)}(k)^{-4}$. The covariance operator C_ε has the same eigen-elements as in Example 1.

In the two examples, the process $(X_n, n \geq 0)$ satisfies **A0**, ..., **A8**. So by Theorem 2 and 3, the corresponding family of probabilities $(P_{n, \theta}, \theta \in \Theta)$ satisfies ULAN condition (3) and ML and Bayes estimators are consistent, asymptotically Gaussian and efficient.

It is readily to check the assumptions **A0**...**A6 a** and **A7**. For the assumption **A8** we have to find the smallest eigenvalue of the matrix $\sum_{k=1}^{\infty} \nabla \lambda_{k, \theta} \nabla^T \lambda_{k, \theta}$. Some calculations give $\mu_\theta = 1/2[(\sum_{k \geq 1} (\lambda_{k, \theta}^{(1)})^2 + \sum_{k \geq 1} (\lambda_{k, \theta}^{(2)})^2) -$

$$\sqrt{(\sum_{k \geq 1} (\lambda_{k, \theta}^{(1)})^2 + \sum_{k \geq 1} (\lambda_{k, \theta}^{(2)})^2 - 4[\sum_{k \geq 1} (\lambda_{k, \theta}^{(1)})^2 \sum_{k \geq 1} (\lambda_{k, \theta}^{(2)})^2 - (\sum_{k \geq 1} \lambda_{k, \theta}^{(1)} \lambda_{k, \theta}^{(2)})^2]}$$

For Example 1, $\lambda_{k, \theta}^{(i)} = \lambda_k(\rho^{(i)})$, for $i = 1, 2$. Hence μ_θ doesn't depend on θ and is strictly positive. So **A8** is satisfied. For Example 2, we have $\lambda_{k, \theta}^{(1)} = -e^{-k(\theta_1+1)}(k)^{-4}$ and $\lambda_{k, \theta}^{(2)} = -e^{-k(\theta_2+1)}(k)^{-3}$. The function $\theta \rightarrow \mu_\theta$ is continuous and so $\inf_{\theta \in \Theta} \mu_\theta = \inf_{\theta \in \bar{\Theta}} \mu_\theta = \mu_{\tilde{\theta}}$ where $\tilde{\theta} \in \bar{B}(0, 1)$. We may easily get $\mu_{\tilde{\theta}} > 0$. Hence the condition **A8** is verified.

5. Numerical Simulations

In this section, we carry out numerical simulations to illustrate the behavior of conditional maximum likelihood estimators MLE in two cases of Eq (1) where it is given explicitly.

The MLE estimator $\hat{\theta}_n = (\hat{\theta}_{n,1}, \hat{\theta}_{n,2})$ is a random variable so it is more correct to write $\hat{\theta}_n = \hat{\theta}_n(\omega)$. In order to simulate the behavior of the estimator around the true value (intersection of the lines $(x = \theta_2)$ and $(y = \theta_1)$) we fixe ω and we

increase n , for each n we have a realization so a couple so we get several pairs since n varies. We consider five different realizations, so 5 different ω , from which the different couples represented obtained. Each realization is represented by a particular type of point, for which reason we find on the figures **fig1** and **fig4** the signs + and \times , the triangles, the circles and the lozenges.

Example a. For this first case, we take again the preceding example **example 1.**

The true value is $\theta = (\theta_1, \theta_2) = (0.3, 0.9)$.

In this case the conditional ML estimators $\hat{\theta}_n = (\hat{\theta}_{n,1}, \hat{\theta}_{n,2})$ is explicit and given by:

$$\hat{\theta}_{n,1} = \left(\frac{b_1}{a_{1,2}} - \frac{b_2}{a_2} \right) / \left(\frac{a_1}{a_{1,2}} - \frac{a_{1,2}}{a_2} \right)$$

$$\hat{\theta}_{n,2} = \left(\frac{b_2}{a_{1,2}} - \frac{b_1}{a_1} \right) / \left(\frac{a_2}{a_{1,2}} - \frac{a_{1,2}}{a_1} \right)$$

where for $j = 1, 2$ we have:

$$a_j = \sum_{i=1}^n \sum_{k \geq 1} \sigma_k^{-2} a_k^2 (\rho^{(j)} X_{i-1})$$

$$a_{1,2} = \sum_{i=1}^n \sum_{k \geq 1} \sigma_k^{-2} a_k (\rho^{(1)} X_{i-1}) a_k (\rho^{(2)} X_{i-1})$$

$$b_j = \sum_{i=1}^n \sum_{k \geq 1} \sigma_k^{-2} a_k (\rho^{(j)} X_{i-1}) a_k (X_i)$$

The following graphic shows the behavior of the conditional ML estimators around the true value when n increase.

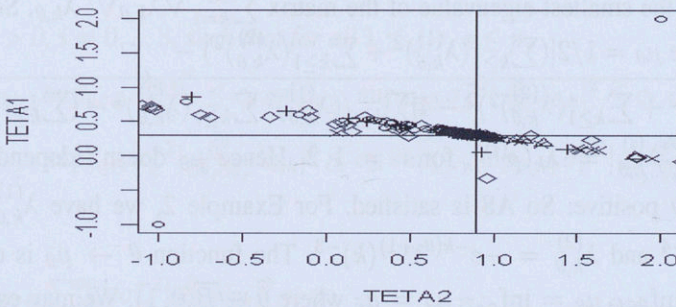


fig 1

The following graphic shows the behavior of the conditional ML estimators $\hat{\theta}_{n,1}$ around the true value when n increase.

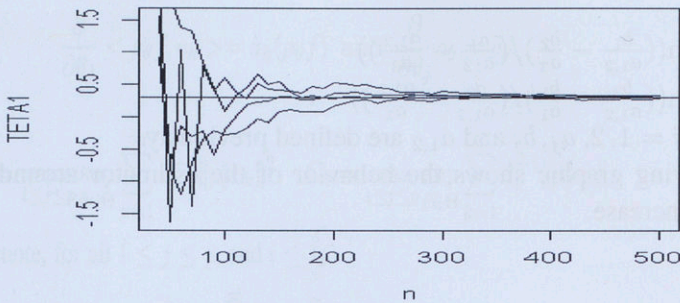


fig2

The following graphic shows the behavior of the conditional ML estimators $\hat{\theta}_{n,2}$ around the true value when n increase.

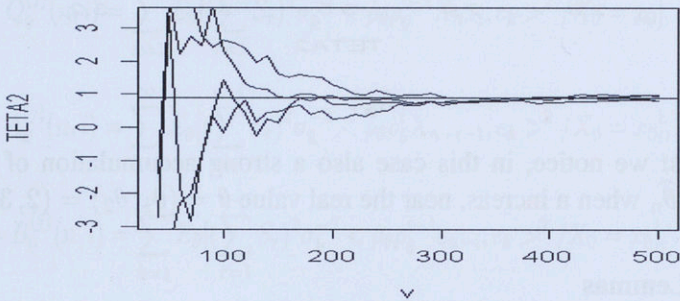


fig3

Comments The two figures **fig2** and **fig3** show good behavior of the estimators $\hat{\theta}_{n,1}$ and $\hat{\theta}_{n,2}$ around true respective values θ_1 and θ_2 , when n increases. The figure **fig1** surely, does not allow conclusions to be drawn about the behavior of the estimator $\hat{\theta}_n$ when n increases, but we notice a strong accumulation of its different realizations, for different n , around the real value $\theta = (\theta_1, \theta_2) = (0.3, 0.9)$

Example b. In this case ρ_θ is given by :

$$\rho_\theta = e^{-\theta_1} \rho^{(1)} + e^{-\theta_2} \rho^{(2)}$$

where the true value is $\theta = (\theta_1, \theta_2) = (2, 3)$ and $\rho^{(1)}x = \sum_{k \geq 1} (k+1)^{-2} \langle x, e_k \rangle e_k$, $\rho^{(2)}(x) = \sum_{k \geq 1} (k+1)^{-3} \langle x, e_k \rangle e_k$. Let $(\sigma_k^2, e_k, k \geq 1)$ be the eigen-elements of the covariance operator C_ε where $e_k(t) = \sqrt{2} \sin((k-1/2)\pi t)$ and $\sigma_k^2 = 1/(k-1/2)^2 \pi^2$.

In this case the conditional ML estimators $\hat{\theta}_n = (\hat{\theta}_{n,1}, \hat{\theta}_{n,2})$ is explicitly given by :

$$\hat{\theta}_{n,1} = -\ln\left(\frac{\frac{b_1}{a_{1,2}} - \frac{b_2}{a_2}}{\frac{a_1}{a_{1,2}} - \frac{a_{1,2}}{a_2}}\right)$$

$$\hat{\theta}_{n,2} = -\ln\left(\frac{\frac{b_2}{a_{1,2}} - \frac{b_1}{a_1}}{\frac{a_2}{a_{1,2}} - \frac{a_{1,2}}{a_1}}\right)$$

where for $j = 1, 2$, a_j, b_j and $a_{1,2}$ are defined previously.

The following graphic shows the behavior of the estimator around the true value when n increase.

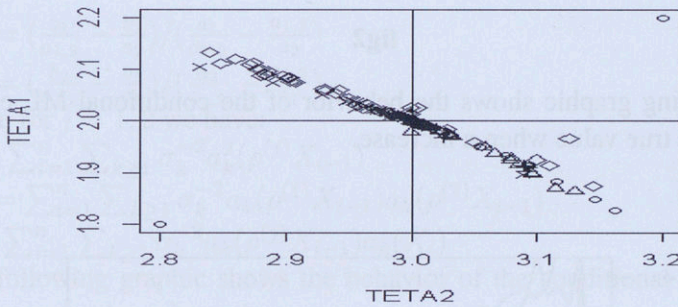


fig4

Comments: we notice, in this case also a strong accumulation of different realizations of $\hat{\theta}_n$ when n increases, near the real value $\theta = (\theta_1, \theta_2) = (2, 3)$

6. Auxiliary Lemmas

In this section we present auxiliary lemmas needed in the proofs of the theorems of Section 3. For commodity we recall some lemmas established in a previous paper (Kara and Mourid [13]). With the help of notations of Section 2, we introduce the following quantities. For $i = 1, \dots, n$, let

$$g_\theta(y/x) = dP_{X_i, \theta}^{X_{i-1}=x} / dP_{\varepsilon_i}(y),$$

$$\phi_i(\theta^*, \theta) = \sqrt{g_{\theta^*}(X_i/X_{i-1}) / g_\theta(X_i/X_{i-1})},$$

$$\psi_{\theta, k}(y, x) = 1/2\sigma_k^{-2} [a_k^2(\rho_\theta y) - 2a_k(x)a_k(\rho_\theta y)]$$

where $\theta, \theta^* \in \Theta$. Define the set

$$U_{n, \theta} = \sqrt{n}\Gamma^{1/2}(\theta)(\Theta - \theta)$$

For all $1 \leq j \leq p$, we consider the following notations: for all $f \in H, k \geq 1$,

$$\frac{\partial}{\partial \theta_j} \langle \rho_\theta f, e_k \rangle = \dot{a}_k(\rho_\theta f) \quad \text{and} \quad \frac{\partial}{\partial \theta_j} \rho_\theta = \dot{\rho}_\theta,$$

$$C_3 = \max_{1 \leq j \leq p} \sup_{\theta \in \Theta} \sum_{k=1}^{\infty} \sigma_k^{-2} \|\dot{\rho}_\theta^* e_k\|^2, \quad C_4 = \max_{1 \leq j \leq p} \sup_{\theta \in \Theta} \sum_{k=1}^{\infty} \sigma_k^{-1} \|\dot{\rho}_\theta^* e_k\|.$$

We also denote, for all $1 \leq j \leq p$, and $i \leq 1$,

$$S_i = S_i^{(j)}(\theta) = \sum_{k=1}^{\infty} \sigma_k^{-2} \frac{\partial}{\partial \theta_j} a_k(\rho_\theta X_{i-1}) a_k(\varepsilon_i),$$

$$F_n(\theta) = F_n^{(j)}(\theta) = E_\theta \left[\left(\sum_{i=1}^n \sum_{k=1}^{\infty} \sigma_k^{-2} \frac{\partial a_k(\rho_\theta X_{i-1})}{\partial \theta_j} a_k(\varepsilon_i) \right)^4 / X_0 = x_0 \right],$$

$$Q(n, i) = Q_\theta^{(j)}(n, i) = \sum_{k=1}^{\infty} E_\theta \left[\left(\sum_{r=1}^{n-i} S_r \right)^2 \sigma_k^{-2} \langle \dot{\rho}_\theta \rho_\theta^{i-1} X_{n-i}, e_k \rangle^2 / X_0 = x_0 \right],$$

$$A(n, i) = A_\theta^{(j)}(n, i) = \sum_{k=1}^{\infty} E_\theta \left[\left(\sum_{r=1}^{n-i} S_r \right)^2 \sigma_k^{-2} \langle \dot{\rho}_\theta \rho_\theta^i X_{n-i-1}, e_k \rangle^2 / X_0 = x_0 \right],$$

$$B(n, i) = B_\theta^{(j)}(n, i) = \sum_{k=1}^{\infty} E_\theta \left[\left(\sum_{r=1}^{n-i} S_r \right)^2 \sigma_k^{-2} \langle \dot{\rho}_\theta \rho_\theta^{i-1} \varepsilon_{n-i}, e_k \rangle^2 / X_0 = x_0 \right],$$

$$E(n, i) = E_\theta^{(j)}(n, i) = \sum_{k=1}^{\infty} E_\theta \left[\left(\sum_{r=1}^{n-i} S_r \right)^2 \sigma_k^{-1} \langle \dot{\rho}_\theta \rho_\theta^{i-1} \varepsilon_{n-i}, e_k \rangle \sigma_k^{-1} \langle \dot{\rho}_\theta \rho_\theta^i X_{n-i-1}, e_k \rangle / X_0 = x_0 \right],$$

$$I_i(\theta) = \begin{cases} E_\theta \left[\left(\sum_{r=1}^i S_r \right)^2 / X_0 = x_0 \right], & i \geq 1, \\ 0, & i = 0. \end{cases}$$

The expression of the likelihood ratio is given by the following lemma.

Lemma 1. *Suppose A0, A1. Then*

$$g_\theta(y/x) = \prod_{k=1}^{\infty} \exp(1/2 \sigma_k^{-2} [2a_k(y) a_k(\rho_\theta x) - a_k^2(\rho_\theta x)])$$

and the likelihood ratio is given by

$$\frac{dP_{X_0=x_0}^{(X_1, \dots, X_n)}}{dP_{(\varepsilon_1, \dots, \varepsilon_n)}}(x_1 \dots x_n) = \prod_{i=1}^n \prod_{k=1}^{\infty} \exp[1/2 \sigma_k^{-2} (2a_k(x_i) a_k(\rho_\theta x_{i-1}) - a_k^2(\rho_\theta x_{i-1}))].$$

Lemma 2. Suppose **A0**, ..., **A3**. For all $i \geq 1$ and $M > 0$, we have

$$\lim_{\lambda \rightarrow 0} \sup_{|h| \leq M} \lambda^{-2} E_{\theta} [\phi_i(\theta + \lambda h, \theta) - 1]^2 \leq E_{\theta} [h^T \dot{\phi}_i(\theta)]^2$$

and

$$\lim_{\lambda \rightarrow 0} \sup_{|h| \leq M} E_{\theta} [(\lambda^{-1} \phi_i(\theta + h\lambda, \theta) - 1) - h^T \dot{\phi}_i(\theta)]^2 = 0$$

where

$$\dot{\phi}_i(\theta) = \sum_{k=1}^{\infty} 1/2 \sigma_k^{-2} a_k(\rho_{\theta} X_{i-1}) [a_k(X_i) - a_k(\rho_{\theta} X_{i-1})].$$

The following lemmas give the tightness of the likelihood ratio process $(Z_{n,\theta}(u))$ in the space $C_0(\mathbb{R}^p)$.

Lemma 3. Suppose **A0**, **A1**, **A7**, **A8**. For all $0 < a < 1/2$, $u \in U_{n,\theta}$ and $\theta \in K$, a compact set of Θ , we have for $n \geq n_0$

$$E_{\theta} Z_{n,\theta}^a(u) < \exp(-g_n(|u|))$$

where $g_n(\|u\|) = 2(a - 2a^2) \frac{\inf_{\theta \in \Theta} \mu_{\theta}}{\sup_{\theta \in K} \|\Gamma^{1/2}(\theta)\|^2} \|u\|^2$.

Lemma 4. Suppose **A0**, **A1**, **A2**, **A5**, **A6(b)**. For all $3 < \beta \leq 4$ and compact set $K \subset \Theta$, there exists $B = B(K)$ and $a = a(K) = 4 - \beta$, such that

$$\sup_{\theta \in K} \sup_{u, v \in U_{n,\theta}, |u| < R, |v| < R} |u - v|^{-\beta} E_{\theta} |Z_{n,\theta}^{1/4}(u) - Z_{n,\theta}^{1/4}(v)|^4 < B(1 + R^a),$$

where $B = \frac{8C_0 p^2 \sup_{\theta \in \Theta} \sum_{r=0}^4 E_{\theta} \|X_0\|^r}{1-C} \frac{1}{\inf_{\theta \in K} \|\Gamma(\theta)\|^2}$

$C_0 = \max(2C_3^2 \sqrt{2(\sum_{r \geq 1} \sigma_r^2) C_2}; 2C_3^2 (\sum_{r \geq 1} \sigma_r^2) + 2C_3^2 \sqrt{8C_1 (\sum_{r \geq 1} \sigma_r^2)}; 16C_3 C_4^2 \sqrt{(\sum_{r \geq 1} \sigma_r^2)} + 8C_3^2 \sqrt{(\sum_{r \geq 1} \sigma_r^2)}; 8C_3^2; C_2 C_3^2 + C_1 C_3^2 (\sum_{r \geq 1} \sigma_r^2) + 2C_3 C_4^2 C_2^{3/4} \sqrt{(\sum_{r \geq 1} \sigma_r^2)} + 2C_3^2 \sqrt{C_1 C_2 (\sum_{r \geq 1} \sigma_r^2)})$ and $C_1 = 2 \frac{\sum_{k \geq 1} \sigma_k^2}{(1-C)^2}$, $C_2 = 8 \frac{E(\|\varepsilon_1\|^4)}{(1-C)^4}$, C is given in **A0** and C_3, C_4 are given in the beginning of this section.

Remark 3. Under **A0**, **A1**, **A2**, **A5**, **A6(a)** and **A7**, we may have the same results of Lemma 4 with $C_0 = \max(2C_3^2 \sqrt{2(\sum_{r \geq 1} \sigma_r^2) C_2}; 2C_3^2 (\sum_{r \geq 1} \sigma_r^2) + 2C_3^2 \sqrt{8C_1 (\sum_{r \geq 1} \sigma_r^2)}; 8C_3^2 \sqrt{(\sum_{r \geq 1} \sigma_r^2)}; 8C_3^2; C_2 C_3^2 + C_1 C_3^2 (\sum_{r \geq 1} \sigma_r^2) + 2C_3^2 \sqrt{C_1 C_2 (\sum_{r \geq 1} \sigma_r^2)})$.

Lemma 5. Suppose A0, A1, A2. Then

$$\dot{\phi}_i(\theta) := \frac{\partial \phi_i}{\partial \theta}(\theta, \theta^*)|_{\theta^*=\theta} = \sum_{k=1}^{\infty} 1/2\sigma_k^{-2} \dot{a}_k(\rho_\theta X_{i-1}) [a_k(X_i) - a_k(\rho_\theta X_{i-1})] \quad \text{a.s.}$$

Lemma 6. For $y \in H$, let $\{f_n^y, n \geq 1\}$ be a sequence of functions defined by

$$f_n^y(x) = \sum_{k=1}^n 1/2\sigma_k^{-2} \dot{a}_k(\rho_\theta y) (a_k(x) - a_k(\rho_\theta y)), \quad x \in H.$$

Under A0, ..., A3, we have

i) $\{f_n^y, n \geq 1\}$ is a Cauchy sequence in $L^2(H, P_{X_i, \theta}^{X_{i-1}=y})$.

ii) $\lim_{n \rightarrow \infty} f_n^y = f^y$ in $L^2(H, P_{X_i, \theta}^{X_{i-1}=y})$ where

$$f^y(x) = \sum_{k=1}^{\infty} 1/2\sigma_k^{-2} \dot{a}_k(\rho_\theta y) (a_k(x) - a_k(\rho_\theta y)).$$

Moreover

$$\begin{aligned} \lim_{n \rightarrow +\infty} \int (f_n^y(x))^2 P_{X_i, \theta}^{X_{i-1}=y}(dx) &= \int (f^y(x))^2 P_{X_i, \theta}^{X_{i-1}=y}(dx) \\ &= \sum_{k=1}^{\infty} 1/4\sigma_k^{-2} [\dot{a}_k(\rho_\theta y)]^2. \end{aligned}$$

Lemma 7. Let $\{X_n(h, \omega), n \geq 1\}$ be a sequence of real rv's in L^2 continuous in $h \in \mathbb{R}^p$ and X a \mathbb{R}^p -valued rv in L^2 . Suppose that for all $h \in \mathbb{R}^p$

1. $X_n(h, \omega) \rightarrow_{n \rightarrow \infty} h^T X(\omega)$ P -a.s.,

2. for all $M > 0$, $\sup_{|h| \leq M} |E(X_n^2(h, \omega)) - E(h^T X(\omega))^2| \rightarrow_{n \rightarrow \infty} 0$.

Then

$$\sup_{|h| \leq M} E(X_n(h, \omega) - h^T X(\omega))^2 \rightarrow_{n \rightarrow \infty} 0.$$

Lemma 8. Suppose A0 and A4(b). Then for all $1 \leq r, j \leq p$ and $\beta > 1/2$,

$$n^{1/4} (\ln(n))^{-\beta} (n^{-1} \sum_{i=1}^n \sum_{k=1}^{\infty} \sigma_k^{-2} \frac{\partial^2}{\partial \theta_j \partial \theta_r} a_k(\rho_\theta X_{i-1}) a_k(\varepsilon_i)) \rightarrow_{n \rightarrow \infty} 0 \quad P_\theta - \text{a.s.}$$

Under A4(a), A5, A6, we have

$$n^{-1} \sum_{i=1}^n \sum_{k=1}^{\infty} \sigma_k^{-2} \left[\frac{\partial}{\partial \theta_r} a_k(\rho_\theta X_{i-1}) \right]^2 \rightarrow_{n \rightarrow \infty} E_\theta \left(\sum_{k=1}^{\infty} \sigma_k^{-2} \left[\frac{\partial}{\partial \theta_r} a_k(\rho_\theta X_1) \right]^2 \right) \text{ in } P_\theta\text{-probability.}$$

Lemma 9. Let (X_n) be a ARH process defined by (1) and suppose A0. Then

$$E_\theta(\|X_n\|^2/X_0) \leq 2C^m \|X_0\|^2 + C_1,$$

and

$$E_\theta(\|X_n\|^4/X_0) \leq 8C^m \|X_0\|^4 + C_2,$$

where $C_1 = 2 \frac{\sum_{k \geq 1} \sigma_k^2}{(1-C)^2}$ and $C_2 = 8 \frac{E(\|\varepsilon_1\|^4)}{(1-C)^4}$. C is given in A0.

For the quantities S_i , $Q(n, i)$, $A(n, i)$, $B(n, i)$, $C(n, i)$, $E(n, i)$, I_i defined in the beginning of Section we have the following results.

Lemma 10. We have for all $n \geq 1$

$$E_\theta(S_n^3(\sum_{i=1}^{n-1} S_i)/X_0 = x_0) = 0 = E_\theta(S_n(\sum_{i=1}^{n-1} S_i)^3/X_0 = x_0).$$

Lemma 11. We have for all $n \geq 1$

$$F_n(\theta) = \sum_{i=1}^n E_\theta(S_i^4/X_0 = x_0) + 6 \sum_{i=1}^{n-1} E_\theta(S_{i+1}^2(\sum_{r=1}^i S_r)^2/X_0 = x_0).$$

Lemma 12. Under A0 A5 A6(a), we have

$$\sum_{i=1}^n E_\theta(S_i^4/X_0 = x_0) \leq 3n(\sum_{k=1}^{\infty} \sigma_k^{-2} \|(\rho_\theta^{(j)})^* e_k\|^2)^2 (8\|x_0\|^4 + C_2)$$

where C_2 is defined in Lemma 9.

Lemma 13. Under A0 A5 A6(a), we have for all $i \geq 1$

$$\sup_{\theta \in \Theta} I_i(\theta) \leq iC_3(2\|x_0\|^2 + C_1),$$

where C_3 is in Lemma 4 and C_1 in Lemma 9.

Lemma 14. Under A0 A5 A6(b), we have for all $i \geq 1$

1. $A(n, i) \leq Q(n, i+1) + C^i C_3^2 (8\|x_0\|^4 + C_2)$.
2. $B(n, i) \leq C^{i-1} C_3^2 (2\|x_0\|^2 + C_1) [(\sum_{r \geq 1} \sigma_r^2)(n-i+3)]$.
3. $E(n, i) \leq 2(C^{i-1})^2 (\sum_{k \geq 1} \sigma_k^2)^{1/2} [C_3 C_4^2 (8\|x_0\|^4 + C_2)^{3/4} + \sqrt{n-i-1} C_3^2 (2\|x_0\|^2 + C_1)^{1/2} (8\|x_0\|^4 + C_2)^{1/2}]$.

4. $Q_\theta(n, i)^{(j)} \leq Q_\theta^{(j)}(n, i + 1) + C^{i-1}(n - i + 3)C_0(1 + \|x_0\| + \|x_0\|^2 + \|x_0\|^3 + \|x_0\|^4)$.
5. $\max_{1 \leq j \leq p} \sup_{\theta \in \Theta} Q_\theta(n, i)^{(j)} \leq K(\sum_{r=0}^{n-i-1} C^{i+r-1}(n-i-r+3))(\sum_{r=0}^4 \|x_0\|^r)$.

where

$$C_0 = \max(2C_3^2 \sqrt{2(\sum_{r \geq 1} \sigma_r^2)} C_2; 2C_3^2(\sum_{r \geq 1} \sigma_r^2) + 2C_3^2 \sqrt{8C_1(\sum_{r \geq 1} \sigma_r^2)}; 16C_3 C_4^2 \sqrt{(\sum_{r \geq 1} \sigma_r^2)} + 8C_3^2 \sqrt{(\sum_{r \geq 1} \sigma_r^2)}; 8C_3^2; C_2 C_3^2 + C_1 C_3^2(\sum_{r \geq 1} \sigma_r^2) + 2C_3 C_4^2 C_2^{3/4} \sqrt{(\sum_{r \geq 1} \sigma_r^2)} + 2C_3^2 \sqrt{C_1 C_2(\sum_{r \geq 1} \sigma_r^2)}), C \text{ is in } \mathbf{A0} \text{ and } C_i, i=1..4 \text{ are given in the previous lemmas.}$$

7. Appendix. Proofs

Proof of Theorem 1. Under **A0**, (X_n) defined by (1) is a stationary ergodic Markov chain. From relation $\lim_{\lambda \rightarrow 0} \sup_{|h| \leq M} E_\theta([\lambda^{-1} \phi_i(\theta + h\lambda, \theta) - 1] - h^T \dot{\phi}_i(\theta))^2 = 0$ where $\dot{\phi}_i(\theta) = \sum_{k=1}^{\infty} 1/2\sigma_k^{-2} a_k(\rho_\theta X_{i-1})[a_k(X_i) - a_k(\rho_\theta X_{i-1})]$. (see Lemma 2), we deduce for all $i \geq 1$ that the root of the ratio of transition densities ϕ_i is differentiable with respect to θ in quadratic mean. From a Roussas Theorem (see [10]), we have that the family of probability measures $(P_{n,\theta}, \theta \in \Theta)$ of (1) satisfies the LAN condition (2). ■

Proof of Theorem 2. Let $\Lambda_n = \ln Z_{n,\theta}(h)$ where we can take a change of variable $\theta_n = \theta + n^{-1/2}h$ (see LAN condition [11]). Under **A1**, **A2**, **A3** we have the LAN condition from Theorem 1. For ULAN condition it suffices to verify that under P_θ and for all $M > 0$

$$\sup_{|h| \leq M} |\Lambda_n - h^T \Gamma(\theta)^{1/2} S_n(\theta) + 1/2 h^T \Gamma(\theta) h| = o_p(1).$$

To do we use a Taylor series expansion gives P_θ -a.s.

For more details see the appendix (proof of Theorem 2bis) ■

Proof of Theorem 3. Let $p \in \{1, 2, 3\}$.

We follow the general theory in Ibragimov and Hasminski [11] and apply Theorem 1.1 p. 174. By Theorem 2, we have the convergence of the finite dimensional laws of the likelihood process. For all $0 < a < 1/2$, $u \in U_{n,\theta}$ and $\theta \in K$, a compact set of Θ , we have for $n \geq n_0$ $E_\theta Z_{n,\theta}^a(u) < \exp(-g_n(|u|))$ where $g_n(\|u\|) = 2(a - 2a^2) \frac{\inf_{\theta \in \Theta} \mu_\theta}{\sup_{\theta \in K} \|\Gamma^{1/2}(\theta)\|^2} \|u\|^2$.

and For all $3 < \beta \leq 4$ and compact set $K \subset \Theta$, there exists $B = B(K)$ and $a = a(K) = 4 - \beta$, such that $\sup_{\theta \in K} \sup_{u, v \in U_{n,\theta}, |u| < R, |v| < R} |u - v|^{-\beta} E_\theta |Z_{n,\theta}^{1/4}(u) - Z_{n,\theta}^{1/4}(v)|^4 < B(1 + R^a)$,

Previous relations, all together (see Lemma 4 and Lemma 3) give that the sequence of the laws of the likelihood process is tight in the space $\mathcal{C}_0(\mathbb{R})$. We then deduce all the properties of the conditional ML estimators listed in the theorem. ■

The proofs of Lemmas 1, 5, 6 and 8 are given in [13]. Note that Lemma 7 is well known and is a generalized version of a Levy Lemma.

Proof of Lemma 2. We have $E_\theta[\phi_i(\theta + \lambda h, \theta) - 1]^2 = E_\theta[\sqrt{g_{\theta + \lambda h}(X_i/X_{i-1})/g_\theta(X_i/X_{i-1})} - 1]^2$. So

$$E_\theta[(\sqrt{g_{\theta + \lambda h}(X_i/X_{i-1})/g_\theta(X_i/X_{i-1})} - 1)^2 / X_{i-1} = y] \leq \lambda^2 \int_0^1 l(\theta + \lambda th) dt, \quad (4)$$

where $l(\theta + \lambda th) = \int [\frac{1}{2} h^T \nabla g_{\theta + \lambda th}(x/y) g_{\theta + \lambda th}^{-1/2}(x/y)]^2 P_{\varepsilon_i}(dx)$.

On the other hand

$$\begin{aligned} [h^T \nabla g_{\theta + \lambda th}(x/y)]^2 &= g_{\theta + \lambda th}^2(x/y) [h^T \sum_{k=1}^{\infty} \sigma_k^{-2} \nabla a_k(\rho_{\theta + \lambda th} y) (a_k(\rho_{\theta + \lambda th} y) - a_k(x))]^2 \\ &= g_{\theta + \lambda th}^2(x/y) [h_j \sum_{j=1}^p f_{j, \theta + \lambda th}(x/y)]^2, \end{aligned}$$

where

$$f_{j, \theta + \lambda th}(x/y) = \sum_{k=1}^{\infty} (-\sigma_k^{-2}) \frac{\partial}{\partial \theta_j} a_k(\rho_{\theta + \lambda th} y) (a_k(\rho_{\theta + \lambda th} y) - a_k(x)).$$

The previous notations, the inequality $(\sum_{j=1}^p a_j)^2 \leq 2^{p-1} \sum_{j=1}^p a_j^2$ and Lemma 6 yield

$$l(\theta + \lambda th) \leq 2^{p-1} M^2 \sum_{j=1}^p \sup_{\theta \in \Theta} \sum_{k=1}^{\infty} 1/4 \sigma_k^{-2} [\frac{\partial}{\partial \theta_j} a_k(\rho_\theta y)]^2. \quad (5)$$

Let us show that

$$\lim_{\lambda \rightarrow 0} l(\theta + \lambda th) = l(\theta).$$

Set $F_\lambda(x/y) = \frac{1}{4} g_{\theta + \lambda th}(x/y) [h_j \sum_{j=1}^p f_{j, \theta + \lambda th}(x/y)]^2$. We have

$$l(\theta + \lambda th) = E(F_\lambda(\varepsilon_1/y)). \quad (6)$$

Notice that

$$F_\lambda(\varepsilon_1/y) \rightarrow_{\lambda \rightarrow 0} F_0(\varepsilon_1/y) \quad P - a.s., \quad (7)$$

$$F_\lambda(\varepsilon_1/y) \leq G_\lambda(\varepsilon_1/y), \quad (8)$$

where

$$G_\lambda(\varepsilon_1/y) := 2^{p-1} M^2 \frac{1}{4} \sum_{j=1}^p \frac{1}{4} [f_{j,\theta+\lambda th}(\varepsilon_1/y)]^2 g_{\theta+\lambda th}(\varepsilon_1/y).$$

It is clear that

$$G_\lambda(\varepsilon_1/y) \rightarrow_{\lambda \rightarrow 0} G_0(\varepsilon_1/y) \quad P - a.s. \quad (9)$$

By Lemma 6, we may write

$$E(G_\lambda(\varepsilon_1/y)) = 2^{p-1} \sum_{j=1}^p h_j^2 \sum_{k=1}^{\infty} 1/4 \sigma_k^{-2} \left[\frac{\partial}{\partial \theta_j} a_k(\rho_{\theta+\lambda th} y) \right]^2.$$

Under condition **A2** we conclude that

$$\lim_{\lambda \rightarrow 0} E(G_\lambda(\varepsilon_1/y)) = E(G_0(\varepsilon_1/y)). \quad (10)$$

We have $E(G_\lambda(\varepsilon_1/y))$ and $E(G_0(\varepsilon_1/y))$ are bounded by quantities independent of λ .

Then (9), (10) and Bressis Lieb's Theorem imply

$$\lim_{\lambda \rightarrow 0} E(|G_\lambda(\varepsilon_1/y) - G_0(\varepsilon_1/y)|) = 0. \quad (11)$$

The results (11), (8), (7) and the dominated convergence theorem yield

$$\lim_{\lambda \rightarrow 0} E(|F_\lambda(\varepsilon_1/y) - F_0(\varepsilon_1/y)|) = 0.$$

Hence

$$\begin{aligned} \lim_{\lambda \rightarrow 0} E(F_\lambda(\varepsilon_1/y)) &= E(F_0(\varepsilon_1/y)) \\ &= E\left(\frac{1}{4} g_\theta(\varepsilon_1/y) \left[h_j \sum_{j=1}^p f_{j,\theta}(\varepsilon_1/y) \right]^2\right) \\ &= E_\theta([h^T \dot{\phi}_i(\theta)]^2 / X_{i-1} = y). \end{aligned}$$

From equation (6) we have

$$\lim_{\lambda \rightarrow 0} l(\theta + \lambda th) = E_\theta([h^T \dot{\phi}_i(\theta)]^2 / X_{i-1} = y). \quad (12)$$

The results (5), (12) and the dominated convergence theorem give

$$\lim_{\lambda \rightarrow 0} \int_0^1 l(\theta + \lambda th) dt = E_\theta([h^T \dot{\phi}_i(\theta)]^2 / X_{i-1} = y). \quad (13)$$

From (13) and (4), we have

$$\lim_{\lambda \rightarrow 0} \lambda^{-2} E_\theta[(\phi_i(\theta + \lambda h, \theta) - 1)^2 / X_{i-1} = y] \leq E_\theta([h^T \dot{\phi}_i(\theta)]^2 / X_{i-1} = y).$$

Hence

$$E_\theta[\lim_{\lambda \rightarrow 0} \lambda^{-2} E_\theta[(\phi_i(\theta + \lambda h, \theta) - 1)^2 / X_{i-1} = y]] \leq E_\theta(E_\theta([h^T \dot{\phi}_i(\theta)]^2 / X_{i-1} = y)).$$

Condition **A3**, inequalities (4), (5) and the dominated convergence theorem again yield

$$\lim_{\lambda \rightarrow 0} \lambda^{-2} E_\theta[(\phi_i(\theta + \lambda h, \theta) - 1)^2] \leq E_\theta([h^T \dot{\phi}_i(\theta)]^2). \quad (14)$$

On the other hand, we have for any real vector h :

$$\lambda^{-1}(\phi_i(\theta + \lambda h, \theta) - 1) \rightarrow_{\lambda \rightarrow 0} h^T \dot{\phi}_i(\theta) \quad P_\theta - \text{a.s.} \quad (15)$$

This limit is also uniform over compact sets because l is a continuous function. Hence by Fatou Lemma

$$E_\theta[(h^T \dot{\phi}_i(\theta))^2] \leq \liminf_{\lambda \rightarrow 0} E_\theta \lambda^{-2} (\phi_i(\theta + \lambda h, \theta) - 1)^2. \quad (16)$$

From this and (14), we deduce

$$\lim_{\lambda \rightarrow 0} E_\theta \lambda^{-2} (\phi_i(\theta + \lambda h, \theta) - 1)^2 = E_\theta[(h^T \dot{\phi}_i(\theta))^2].$$

Similarly as above, we show that the function $|E_\theta \lambda^{-2} (\phi_i(\theta + \lambda h, \theta) - 1)^2 - E_\theta[(h^T \dot{\phi}_i(\theta))^2]|$ is continuous.

Hence

$$\lim_{\lambda \rightarrow 0} \sup_{|h| \leq M} |E_\theta \lambda^{-2} (\phi_i(\theta + \lambda h, \theta) - 1)^2 - E_\theta[(h^T \dot{\phi}_i(\theta))^2]| = 0.$$

From this, result (15) and Lemma 7, we conclude

$$\sup_{|h| \leq M} E_\theta(\lambda^{-1}(\phi_i(\theta + \lambda h, \theta) - 1) - h^T \dot{\phi}_i(\theta))^2 \rightarrow_{\lambda \rightarrow 0} 0.$$

■

Proof of Lemma 3. From the proof of Lemma 9 of [13], we have for all $0 < a < 1/2$

$$E_{\theta}(Z_{n,\theta}^a(u)) \leq \prod_{k=1}^{\infty} [1 + 2(a - 2a^2)[\lambda_{k,\theta_n^*} - \lambda_{k,\theta}]^2]^{-(n-1)/4},$$

where $\theta_n^* = \theta + (n\Gamma^{-1/2}(\theta))u$.

By Taylor series expansion, there exists $\tilde{\theta}, \tilde{\theta} = t \frac{\Gamma^{-1/2}(\theta)u}{\sqrt{n}} + \theta$, $0 < t < 1$ such that

$$[\lambda_{k,\theta_n^*} - \lambda_{k,\theta}]^2 = n^{-1}((\Gamma^{-1/2}(\theta)u)^T \nabla \lambda_{k,\tilde{\theta}})^2.$$

Hence for n large enough, we get

$$\begin{aligned} E_{\theta}(Z_{n,\theta}^a(u)) &\leq \prod_{k=1}^{\infty} [1 + 2(a - 2a^2) \frac{1}{n} ((\Gamma^{-1/2}(\theta)u)^T \nabla \lambda_{k,\tilde{\theta}})^2]^{-(n-1)/4} \\ &\leq e^{-2(a-2a^2) \sum_{k=1}^{\infty} ((\Gamma^{-1/2}(\theta)u)^T \nabla \lambda_{k,\tilde{\theta}})^2}. \end{aligned}$$

Set

$$A(\theta, \tilde{\theta}) = \Gamma^{-1/2}(\theta) \left(\sum_{k=1}^{\infty} \nabla \lambda_{k,\tilde{\theta}} \nabla^T \lambda_{k,\tilde{\theta}} \right) \Gamma^{-1/2}(\theta).$$

Let $\{\nu_{A,i}(\theta, \tilde{\theta}), 1 \leq i \leq p\}$ be the eigenvalues of $A(\theta, \tilde{\theta})$. We have

$\sum_{k=1}^{\infty} ((\Gamma^{-1/2}(\theta)u)^T \nabla \lambda_{k,\tilde{\theta}})^2 = n^{-1} \langle A(\theta, \tilde{\theta})u, u \rangle \geq \|u\|^2 \inf_{(\theta, \tilde{\theta}) \in K \times \Theta} \min_{1 \leq i \leq p} \nu_{A,i}(\theta, \tilde{\theta})$
and

$$\begin{aligned} \inf_{(\theta, \tilde{\theta}) \in K \times \Theta} \min_{1 \leq i \leq p} \nu_{A,i}(\theta, \tilde{\theta}) &= \inf_{(\theta, \tilde{\theta}) \in K \times \Theta} \min_{1 \leq i \leq p} [\nu_{A^{-1},i}(\theta, \tilde{\theta})]^{-1} \\ &= \left[\sup_{(\theta, \tilde{\theta}) \in K \times \Theta} \max_{1 \leq i \leq p} \nu_{A^{-1},i}(\theta, \tilde{\theta}) \right]^{-1} \\ &= \left[\sup_{(\theta, \tilde{\theta}) \in K \times \Theta} \|(A(\theta, \tilde{\theta}))^{-1}\| \right]^{-1} \\ &= \left[\sup_{(\theta, \tilde{\theta}) \in K \times \Theta} \|\Gamma^{1/2}(\theta) \left(\sum_{k \geq 1} \nabla \lambda_{k,\tilde{\theta}} \nabla^T \lambda_{k,\tilde{\theta}} \right)^{-1} \Gamma^{1/2}(\theta)\| \right]^{-1} \\ &\geq \left[\sup_{\theta \in K} \|\Gamma^{1/2}(\theta)\|^2 \sup_{\tilde{\theta} \in \Theta} \left\| \left(\sum_{k \geq 1} \nabla \lambda_{k,\tilde{\theta}} \nabla^T \lambda_{k,\tilde{\theta}} \right)^{-1} \right\| \right]^{-1} \\ &\geq \left[\sup_{\theta \in K} \|\Gamma^{1/2}(\theta)\|^2 \frac{1}{\inf_{\tilde{\theta} \in \Theta} \mu_{\tilde{\theta}}} \right]^{-1}, \end{aligned}$$

where μ_θ is the smallest eigenvalue of $\sum_{k \geq 1} \nabla \lambda_{k,\theta} \nabla^T \lambda_{k,\theta}$.
Therefore, for all $0 < a < 1/2$, we have

$$\begin{aligned} E_\theta(Z_{n,\theta}^a(u)) &\leq e^{-2(a-2a^2)[\sup_{\theta \in K} \|\Gamma^{1/2}(\theta)\|^2 \frac{1}{\inf_{\theta \in \Theta} \mu_\theta}]}^{-1} \|u\|^2 \\ &\leq e^{-g_n(\|u\|)}, \end{aligned}$$

where $g_n(\|u\|) = 2(a - 2a^2) \frac{\inf_{\theta \in \Theta} \mu_\theta}{\sup_{\theta \in K} \|\Gamma^{1/2}(\theta)\|^2} \|u\|^2$. ■

Proof of Lemma 4. Let $\theta = (\theta_1, \theta_2, \dots, \theta_p)$ and set $\theta^{(1)} = \theta + \frac{u^T(\Gamma(\theta))^{-1/2}}{\sqrt{n}}$,
 $\theta^{(2)} = \theta + \frac{v^T(\Gamma(\theta))^{-1/2}}{\sqrt{n}}$ and $h = \theta^{(1)} - \theta^{(2)}$.

We have

$$E_\theta |Z_{n,\theta}(u)^{1/4} - Z_{n,\theta}(v)^{1/4}|^4 = E_\theta (E_\theta(|Z_{n,\theta}(u)^{1/4} - Z_{n,\theta}(v)^{1/4}|^4 / X_0)).$$

Set $\varepsilon_0 = X_0 = x_0$. With the previous notations, Taylor series expansion with integral remainder, Cauchy-Schwarz inequality and with the convexity of Θ , all yield :

$$\begin{aligned} &E_\theta(|Z_{n,\theta}(u)^{1/4} - Z_{n,\theta}(v)^{1/4}|^4 / X_0 = x_0) \\ &= E([(\prod_{i=1}^n g_{\theta^{(2)}+h}(\varepsilon_i/\varepsilon_{i-1}))^{1/4} - (\prod_{i=1}^n g_{\theta^{(2)}}(\varepsilon_i/\varepsilon_{i-1}))^{1/4}]^4) \\ &= E[\int_0^1 \nabla^T \exp(-1/4 \sum_{i=1}^n \sum_{k=1}^\infty \psi_{t\theta^{(1)}+(1-t)\theta^{(2)},k}(\varepsilon_{i-1}, \varepsilon_i)) (\theta^{(1)} - \theta^{(2)}) dt]^4 \\ &\leq 2^{-8} \|\theta^{(1)} - \theta^{(2)}\|^4 \int_0^1 E_{t\theta^{(1)}+(1-t)\theta^{(2)}} ([\sum_{j=1}^p (\frac{\partial}{\partial \theta_j} \sum_{i=1}^n \sum_{k=1}^\infty \psi_{t\theta^{(1)}+(1-t)\theta^{(2)},k}(X_{i-1}, X_i))^2]^2 / X_0 = x_0) dt \end{aligned}$$

$$\leq 2^{-8} p^2 \|\theta^{(1)} - \theta^{(2)}\|^4 \max_{1 \leq j \leq p} \sup_{\tau \in \Theta} F_n^{(j)}(\tau).$$

The rv's ε_i are independent of X_s , $\forall i \geq s+1$ and $\forall i$, $\{a_k(\varepsilon_i), k \geq 1\}$ is a sequence of independent Gaussian r.v's $N(0, \sigma_k^2)$, with lemmas 11; 12, 14 relation (17), and monotone convergence theorem, all these imply

$$\begin{aligned} &E_\theta(|Z_{n,\theta}(u)^{1/4} - Z_{n,\theta}(v)^{1/4}|^4 / X_0 = x_0) \\ &\leq 2^{-8} p^2 \|\theta^{(1)} - \theta^{(2)}\|^4 \max_{1 \leq j \leq p} \sup_{\tau \in \Theta} \sum_{i=1}^n E_\tau((S_i^{(j)}(\tau))^4 / X_0 = x_0) + \\ &6 \sum_{i=1}^{n-1} E_\tau((S_{i+1}^{(j)}(\tau))^2 (\sum_{r=1}^i (S_r^{(j)}(\tau)))^2 / X_0 = x_0) \\ &\leq 2^{-8} p^2 \|\theta^{(1)} - \theta^{(2)}\|^4 \max_{1 \leq j \leq p} \sup_{\tau \in \Theta} [3n (\sum_{k=1}^\infty \sigma_k^{-2} \|\frac{\partial}{\partial \theta_j} (\rho_\tau^* e_k)\|^2) (8\|x_0\|^4 + \\ &8E\|\varepsilon_1\|^4 (1-C)^{-4}) \end{aligned}$$

$$\begin{aligned}
 &+ 6 \sum_{i=1}^{n-1} \sum_{k \geq 1} E_{\tau}((\sum_{r=1}^i (S_r^{(j)}(\tau)))^2 \sigma_k^{-2} (\frac{\partial}{\partial \theta_j} a_k(\rho_{\tau} X_i)^2) / X_0 = x_0) \\
 &\leq 2^{-8} p^2 \|\theta^{(1)} - \theta^{(2)}\|^4 \max_{1 \leq j \leq p} \sup_{\tau \in \Theta} [3n(\sum_{k=1}^{\infty} \sigma_k^{-2} \|\frac{\partial}{\partial \theta_j} (\rho_{\tau}^* e_k)\|^2) (8\|x_0\|^4 + \\
 &8E\|\varepsilon_1\|^4 (1-C)^{-4}) + 6 \sum_{i=1}^{n-1} Q_{\tau}^{(j)}(i+1, 1)] \\
 &\leq \sum_{r=0}^4 \|x_0\|^r \|\Gamma(\theta)\|^{-2} (2Kp^2)(1-C)^{-1} \|u-v\|^4.
 \end{aligned}$$

Hence

$$E_{\theta}(|Z_{n,\theta}(u)^{1/4} - Z_{n,\theta}(v)^{1/4}|^4) \leq (8Kp^2)(1-C)^{-1} (\sup_{\theta \in \Theta} \sum_{r=0}^4 E_{\theta} \|X_0\|^r) (\inf_{\theta \in \Theta} \|\Gamma(\theta)\|^2)^{-1} R^4.$$

■

Proof of Lemma 9. From $X_n = \rho_{\theta}(X_{n-1}) + \varepsilon_n$ we have $X_n = \rho_{\theta}^n(X_0) + \sum_{i=0}^{n-1} \rho_{\theta}^i \varepsilon_{n-i}$. Since $\{\varepsilon_i, i \geq 1\}$ is a strong Gaussian noise independent of X_0 , using Holder inequality and $(a+b)^2 \leq 2a^2 + 2b^2$, we have

$$\begin{aligned}
 E_{\theta}(\|X_n\|^4 / X_0) &\leq 8\|\rho_{\theta}^n(X_0)\|^4 + 8 \sum_{i_1, i_2, i_3, i_4=0}^{n-1} C^{i_1+i_2+i_3+i_4} E(\|\varepsilon_{n-i_1}\| \|\varepsilon_{n-i_2}\| \|\varepsilon_{n-i_3}\| \|\varepsilon_{n-i_4}\|) \\
 &\leq 8C^n \|X_0\|^4 + 8E(\|\varepsilon_1\|^4) ((1-C)^{-4}).
 \end{aligned}$$

We deduce a similar result for the first inequality from $E(\|\varepsilon_1\|^2) = \sum_{k \geq 1} \sigma_k^2$. ■

Proof of Lemma 10. We have

$$E_{\theta}(S_n^3(\sum_{i=1}^{n-1} S_i) / X_0) = E_{\theta}((\sum_{i=1}^{n-1} S_i) E_{\theta}(S_n^3 / \sigma(X_0, \dots, X_{n-1}, \varepsilon_1, \dots, \varepsilon_{n-1})) / X_0),$$

But

$$E_{\theta}(S_n^3 / (X_0 = x_0, \dots, X_{n-1} = x_{n-1}, \varepsilon_1 = \varepsilon_1, \dots, \varepsilon_{n-1} = \varepsilon_{n-1})) = E_{\theta}(\sum_{k=1}^{\infty} \sigma_k^{-2} \frac{\partial}{\partial \theta_j} a_k(\rho_{\theta} x_{n-1}) a_k(\varepsilon_n))^3 = 0,$$

because the exponent is odd (=3) and

$$\forall n \geq 1, \sum_{k=1}^{\infty} \sigma_k^{-2} \frac{\partial}{\partial \theta_j} a_k(\rho_{\theta} x_{n-1}) a_k(\varepsilon_n) \hookrightarrow N(0, \sum_{k=1}^{\infty} \sigma_k^{-2} (\frac{\partial}{\partial \theta_j} a_k(\rho_{\theta} x_{n-1}))^2), \tag{17}$$

(which is the a.s. limit of a sequence of Gaussian rv's). We proceed similarly for the 2nd equality. ■

Proof of Lemma 11. By the definition of $F_2(\theta)$ and using Lemma 10, we have

$$\begin{aligned} F_2(\theta) &= E_\theta(S_1^4 + S_2^4 + 6S_2^2 S_1^2 + 4S_2^3 S_1 + 4S_2 S_1^3 / X_0 = x_0) \\ &= E_\theta(S_1^4 + S_2^4 + 6S_2^2 S_1^2 / X_0 = x_0). \end{aligned}$$

Again using Lemma 10 and a recurrence on n , we obtain

$$\begin{aligned} F_{n+1}(\theta) &= F_n(\theta) + E_\theta(S_{n+1}^4 + 6S_{n+1}^2 (\sum_{i=1}^n S_i)^2 / X_0 = x_0) \\ &= \sum_{i=1}^{n+1} E_\theta(S_i^4 / X_0 = x_0) + 6 \sum_{i=1}^n E_\theta(S_{i+1}^2 (\sum_{r=1}^i S_r)^2 / X_0 = x_0). \end{aligned}$$

■

Proof of Lemma 12. We have

$$E_\theta(S_i^4 / X_0) = E_\theta(E_\theta[(\sum_{k=1}^{\infty} \sigma_k^{-2} \frac{\partial}{\partial \theta_j} a_k(\rho_\theta X_{i-1}) a_k(\varepsilon_i))^4 / X_0, X_{i-1}] / X_0).$$

From the relation (17) and the fact that if $Y \hookrightarrow N(0, \sigma^2)$ then $E(Y^{2m}) = \frac{(2m)!}{2^m m!} \sigma^{2m}$, we imply

$$E_\theta[(\sum_{k=1}^{\infty} \sigma_k^{-2} \frac{\partial}{\partial \theta_j} a_k(\rho_\theta X_{i-1}) a_k(\varepsilon_i))^4 / (X_0, X_{i-1}) = (x_0, x_{i-1})] = 3[\sum_{k=1}^{\infty} \sigma_k^{-2} (\frac{\partial}{\partial \theta_j} a_k(\rho_\theta x_{i-1}))^2]^2.$$

Hence from Cauchy-Schwarz inequality and Lemma 9

$$E_\theta(S_i^4 / X_0 = x_0) \leq 3(\sum_{k=1}^{\infty} \sigma_k^{-2} \|(\rho_\theta^{(j)})^* e_k\|^2)^2 [8C^{i-1} \|x_0\|^4 + 8E\|\varepsilon_1\|^4 (1-C)^{-4}].$$

As $C < 1$ we have the result. ■

Proof of Lemma 13. As in the proof of Lemma 9, we may obtain

$$I_i(\theta) \leq I_{i-1}(\theta) + C_3(2\|x_0\|^2 + C_1).$$

Hence

$$I_i(\theta) \leq I_{i-1}(\theta) + C_3(2\|x_0\|^2 + C_1).$$

We have the result by recurrence on i . ■

Proof of Lemma 14.

1. Similarly as in the proof of Lemma 9, we may have

$$\begin{aligned}
 A(n, i) &= \sum_{k=1}^{\infty} E_{\theta} \left[\left(\sum_{r=1}^{n-i} S_r \right)^2 \sigma_k^{-2} a_k^2 (\dot{\rho}_{\theta} \rho_{\theta}^i X_{n-i-1}) / X_0 = x_0 \right] \\
 &= \sum_{k=1}^{\infty} E_{\theta} \left[\left(\sum_{r=1}^{n-i-1} S_r \right)^2 \sigma_k^{-2} a_k^2 (\dot{\rho}_{\theta} \rho_{\theta}^i X_{n-i-1}) / X_0 = x_0 \right] + E_{\theta} \left[(S_{n-i})^2 \sigma_k^{-2} a_k^2 (\dot{\rho}_{\theta} \rho_{\theta}^i X_{n-i-1}) / X_0 = x_0 \right] \\
 &\quad + 2E_{\theta} \left[\left(\sum_{r=1}^{n-i-1} S_r \right) S_{n-i} \sigma_k^{-2} a_k^2 (\dot{\rho}_{\theta} \rho_{\theta}^i X_{n-i-1}) / X_0 = x_0 \right] \\
 &= Q(n, i+1) + \sum_{k=1}^{\infty} E_{\theta} \left[E_{\theta} \left[\left(\sum_{u=1}^{\infty} \sigma_u^{-2} \dot{a}_u (\rho_{\theta} X_{n-i-1}) a_u (\varepsilon_{n-i}) \right)^2 \sigma_k^{-2} a_k^2 (\dot{\rho}_{\theta} \rho_{\theta}^i X_{n-i-1}) / X_0 = x_0, \dots \right. \right. \\
 &\quad \left. \left. , X_{n-i-1} = x_{n-i-1}, \varepsilon_1, \dots, \varepsilon_{n-i-1} \right] / X_0 = x_0 \right] + 2E_{\theta} \left[E_{\theta} \left[\left(\sum_{r=1}^{n-i-1} S_r \right) \sigma_k^{-2} a_k^2 (\dot{\rho}_{\theta} \rho_{\theta}^i X_{n-i-1}) \right. \right. \\
 &\quad \left. \left. \left(\sum_{u=1}^{\infty} \sigma_u^{-2} \dot{a}_u (\rho_{\theta} X_{n-i-1}) a_u (\varepsilon_{n-i}) \right) / X_0 = x_0, \dots, X_{n-i-1} = x_{n-i-1}, \varepsilon_1, \dots, \varepsilon_{n-i-1} \right] / X_0 = x_0 \right].
 \end{aligned}$$

As for all $n \geq 1$, the random variable $\sum_{u=1}^{\infty} \sigma_u^{-2} \dot{a}_u (\rho_{\theta} x_{n-1}) a_u (\varepsilon_n) \hookrightarrow N(0, \sum_{u=1}^{\infty} \sigma_u^{-2} \dot{a}_u^2 (\rho_{\theta} x_{n-1}))$ (a.s. limit of Gaussian rv's), Cauchy-Schwartz inequality, Lemme 9 and assumption **A0**, all this imply

$$\begin{aligned}
 A(n, i) &= Q(n, i+1) + \sum_{k=1}^{\infty} E_{\theta} \left[\left(\sum_{u=1}^{\infty} \sigma_u^{-2} \dot{a}_u^2 (\rho_{\theta} X_{n-i-1}) \right) \sigma_k^{-2} < \dot{\rho}_{\theta} \rho_{\theta}^i X_{n-i-1}, e_k >^2 / X_0 = x_0 \right] + 0 \\
 &\leq Q(n, i+1) + C^i \left(\sum_{k=1}^{\infty} \sigma_k^{-2} \|\dot{\rho}_{\theta}^* e_k\|^2 \right) (8\|x_0\|^4 + C_2).
 \end{aligned}$$

Hence the result.

2. With the same arguments as in 1, Fatou inequality, Lemmas 13 and 9, we may deduce

$$\begin{aligned}
 B(n, i) &= \sum_{k=1}^{\infty} E_{\theta} \left[\left(\sum_{r=1}^{n-i} S_r \right)^2 \sigma_k^{-2} < \dot{\rho}_{\theta} \rho_{\theta}^{i-1} \varepsilon_{n-i}, e_k >^2 / X_0 = x_0 \right] \\
 &\leq C^{i-1} C_3^2 \left(\sum_{k=1}^{\infty} \sigma_k^2 \right) (2\|x_0\|^2 + C_1)(n-i+3).
 \end{aligned}$$

3. With the same arguments as in 1, Lebesgue Theorem, **A6(b)**, Cauchy-Schwartz inequality, $C_{\varepsilon} e_u = \sigma_u^2 e_u$, Holder inequality, Lemma 9 and Lemma 13, we may

have

$$\begin{aligned}
 E(n, i) &= \sum_{k=1}^{\infty} E_{\theta} \left[\left(\sum_{r=1}^{n-i} S_r \right)^2 \sigma_k^{-1} \langle \dot{\rho}_{\theta} \rho_{\theta}^{i-1} \varepsilon_{n-i}, e_k \rangle \sigma_k^{-1} \langle \dot{\rho}_{\theta} \rho_{\theta}^i X_{n-i-1}, e_k \rangle / X_0 = x_0 \right] \\
 &\leq \sqrt{3} \|\rho_{\theta}^{i-1}\|^2 \left(\sum_{k \geq 1} \sigma_k^{-2} \|\dot{\rho}_{\theta}^* e_k\|^2 \right) \left(\sum_{k \geq 1} \sigma_k^2 \right)^{1/2} \left(\sum_{u \geq 1} \sigma_u^{-1} \|\dot{\rho}_{\theta}^* e_u\|^2 \right) E_{\theta} [\|X_{n-i-1}\|^3 / X_0 = x_0] \\
 &\quad + 2 \|\rho_{\theta}^{i-1}\|^2 (I_{n-i-1}(\theta))^{1/2} \left(\sum_{u \geq 1} \|\rho_{\theta}^* e_u\| \right) (\sigma_k^{-2} \|\dot{\rho}_{\theta}^* e_k\|^2) (E_{\theta} [\|X_{n-i-1}\|^4 / X_0 = x_0])^{1/2} \\
 &\leq 2(C^{i-1})^2 C_3 C_4^2 \left(\sum_{k \geq 1} \sigma_k^2 \right)^{1/2} (8\|x_0\|^4 + C_2)^{3/4} + \\
 &\quad 2(C^{i-1})^2 \sqrt{n-i-1} C_3^2 \left(\sum_{k \geq 1} \sigma_k^2 \right)^{1/2} (2\|x_0\|^2 + C_1)^{1/2} (8\|x_0\|^4 + C_2)^{1/2}.
 \end{aligned}$$

4. The inequality of 4 follows from the relation $Q_{\theta}^{(j)}(n, i) = A_{\theta}^{(j)}(n, i) + B_{\theta}^{(j)}(n, i) + E_{\theta}^{(j)}(n, i)$ and the previous inequalities.

5. From a recurrence in the inequality of 4 we obtain

$$Q_{\theta}^{(j)}(n, i) \leq Q_{\theta}^{(j)}(n, i + (n-i-1)) + C_0 \left(\sum_{r=0}^{n-i-2} C^{i+r-1} (n-i-r+3) \right) \left(\sum_{r=0}^4 \|x_0\|^r \right), \quad (18)$$

But

$$\begin{aligned}
 Q_{\theta}^{(j)}(n, i + (n-i-1)) &= Q_{\theta}^{(j)}(n, n-1) \\
 &\leq 2A(n, n-1) + 2B(n, n-1).
 \end{aligned}$$

Similarly as previously we may have

$$\begin{aligned}
 2A(n, n-1) &\leq 2C_3^2 (C^{n-1})^2 \|x_0\|^4 \\
 &\leq C_0 C^{n-2} (\|x_0\|^4)
 \end{aligned}$$

and

$$\begin{aligned}
 2B(n, n-1) &\leq 2\sqrt{3} C_3^2 (C^{n-2})^2 \left(\sum_{k \geq 1} \sigma_k^2 \right) \|x_0\|^2 \\
 &\leq C_0 C^{n-2} (\|x_0\|^2).
 \end{aligned}$$

Hence

$$Q_{\theta}^{(j)}(n, n-1) \leq C^{n-2} K \left(\sum_{r=0}^4 \|x_0\|^r \right).$$

Plugging the last result in equality (18) and since C_0 and C are not depending of θ and j , we have the desired result. ■

Poof of Theorem 2bis. Let $\Lambda_n = \ln Z_{n,\theta}(h)$ where we can take a change of variable $\theta_n = \theta + n^{-1/2}h$ (see LAN condition [11]). Under **A1**, **A2** **A3** we have the LAN condition from Theorem 1. For ULAN condition it suffices to verify that under P_θ and for all $M > 0$

$$\sup_{|h| \leq M} |\Lambda_n - h^T \Gamma(\theta)^{1/2} S_n(\theta) + 1/2 h^T \Gamma(\theta) h| = o_p(1).$$

A Taylor series expansion gives P_θ -a.s.: $\forall i, k \geq 1, \exists \theta_{n,k}^*, \exists \theta_{n,k}^{**}$ between θ and $\theta_n = \theta + n^{-1/2}h$

$$\Lambda_n = h^T \Gamma(\theta)^{1/2} S_n(\theta) + (2n)^{-1} h^T R_{\theta_n^*, \theta_n^{**}} h,$$

where

$$h^T R_{\theta_n^*, \theta_n^{**}} h = \sum_{i=1}^n \sum_{k=1}^{\infty} -\sigma_k^{-2} \{h^T J a_k(\rho_{\theta_n^*, k} X_{i-1}) h a_k(\rho_{\theta_n^*, k} X_{i-1}) + [h^T \nabla a_k(\rho_{\theta_n^*, k} X_{i-1})]^2 - h^T J a_k(\rho_{\theta_n^{**}, k} X_{i-1}) h a_k(X_i)\}.$$

$\theta_n^* = \{\theta_{n,k}^*, k \geq 1\}$ and $\theta_n^{**} = \{\theta_{n,k}^{**}, k \geq 1\}$ are given by

$$\theta_{n,k}^* = r_k \theta_n + (1 - r_k) \theta = \theta + n^{-1/2} r_k h \quad \text{and } 0 < \|r_k\| < 1,$$

$$\theta_{n,k}^{**} = s_k \theta_n + (1 - s_k) \theta = \theta + n^{-1/2} s_k h \quad \text{and } 0 < \|s_k\| < 1.$$

Notice that, for all k the matrix r_k and s_k are diagonal.

We have

$$\Lambda_n - h S_n(\theta) + 1/2 h^T \Gamma(\theta) h = (2n)^{-1} h^T R_{\theta_n^*, \theta_n^{**}} h + 1/2 h^T \Gamma(\theta) h. \quad (19)$$

Therefore to obtain the result it remains to show that under P_θ :

$$\sup_{h, |h| \leq M} |(2n)^{-1} h^T R_{\theta_n^*, \theta_n^{**}} h + 1/2 h^T \Gamma(\theta) h| = o_p(1).$$

We get

$$\begin{aligned} |(2n)^{-1} h^T R_{\theta_n^*, \theta_n^{**}} h + 1/2 h^T \Gamma(\theta) h| &\leq |(2n)^{-1} h^T R_{\theta_n^*, \theta_n^{**}} h - (2n)^{-1} h^T R_{\theta, \theta} h| \\ &\quad + |(2n)^{-1} h^T R_{\theta, \theta} h + 1/2 h^T \Gamma(\theta) h| \quad (20) \end{aligned}$$

and

$$\begin{aligned} |(2n)^{-1} h^T R_{\theta, \theta} h + 1/2 h^T \Gamma(\theta) h| &\leq 1/2 M^2 \|n^{-1} \sum_{i=1}^n \sum_{k=1}^{\infty} -\sigma_k^{-2} \{J a_k(\rho_\theta X_{i-1}) a_k(\varepsilon_i)\| \\ &\quad + \|n^{-1} \sum_{i=1}^n \sum_{k=1}^{\infty} -\sigma_k^{-2} [\nabla a_k(\rho_\theta X_{i-1}) \nabla^T a_k(\rho_\theta X_{i-1})]\| \\ &\quad + \sum_{k=1}^{\infty} \sigma_k^{-2} E_\theta [\nabla a_k(\rho_\theta X_{i-1}) \nabla^T a_k(\rho_\theta X_{i-1})]. \end{aligned}$$

From the equality $2ab = (a+b)^2 - a^2 - b^2$, we obtain under **A5**, for all $1 \leq r, s \leq p$

$$\begin{aligned} n^{-1} \sum_{i=1}^n \sum_{k=1}^{\infty} \sigma_k^{-2} \left[\frac{\partial}{\partial \theta_r} a_k(\rho_\theta X_{i-1}) \frac{\partial}{\partial \theta_s} a_k(\rho_\theta X_{i-1}) \right] &= (2n)^{-1} \sum_{i=1}^n \sum_{k=1}^{\infty} \sigma_k^{-2} \\ &< X_{i-1}, \left(\frac{\partial}{\partial \theta_r} \rho_\theta \right)^* + \left(\frac{\partial}{\partial \theta_s} \rho_\theta \right)^* e_k >^2 \\ &- (2n)^{-1} \sum_{i=1}^n \sum_{k=1}^{\infty} \sigma_k^{-2} < X_{i-1}, \left(\frac{\partial}{\partial \theta_r} \rho_\theta \right)^* e_k >^2 \\ &- (2n)^{-1} \sum_{i=1}^n \sum_{k=1}^{\infty} \sigma_k^{-2} < X_{i-1}, \left(\frac{\partial}{\partial \theta_s} \rho_\theta \right)^* e_k >^2. \end{aligned}$$

By Lemma 8 we deduce

$$n^{-1} \sum_{i=1}^n \sum_{k=1}^{\infty} \sigma_k^{-2} \left[\frac{\partial}{\partial \theta_r} a_k(\rho_\theta X_{i-1}) \frac{\partial}{\partial \theta_s} a_k(\rho_\theta X_{i-1}) \right] \rightarrow_{n \rightarrow \infty} \Gamma_{r,s},$$

with

$$\Gamma_{r,s} = E \sum_{k=1}^{\infty} \sigma_k^{-2} \left[\frac{\partial}{\partial \theta_r} a_k(\rho_\theta X_{i-1}) \frac{\partial}{\partial \theta_s} a_k(\rho_\theta X_{i-1}) \right].$$

Hence

$$\| n^{-1} \sum_{i=1}^n \sum_{k=1}^{\infty} -\sigma_k^{-2} [\nabla a_k(\rho_\theta X_{i-1}) \nabla a_k(\rho_\theta X_{i-1})^T] + \sum_{k=1}^{\infty} \sigma_k^{-2} E_\theta [\nabla a_k(\rho_\theta X_{i-1}) \nabla a_k(\rho_\theta X_{i-1})^T] \| \rightarrow_{n \rightarrow \infty} 0.$$

By Lemma 8 and under **A0** and **A4(b)**, we have for all $1 \leq r, s \leq p$, under P_θ

$$n^{-1} \sum_{i=1}^n \sum_{k=1}^{\infty} \sigma_k^{-2} \frac{\partial^2}{\partial \theta_r \partial \theta_r} a_k(\rho_\theta X_{i-1}) a_k(\varepsilon_i) \rightarrow_{n \rightarrow \infty}.$$

Therefore, under P_θ

$$\sup_{\|h\| \leq M} |(2n)^{-1} h^T R_{\theta, \theta} h + 1/2 h^T \Gamma(\theta) h| \rightarrow_{n \rightarrow \infty} 0. \quad (21)$$

On the other part,

$$\begin{aligned} |(2n)^{-1} h^T R_{\theta_n^*, \theta_n^*} h - (2n)^{-1} h^T R_{\theta, \theta} h| &\leq 1/2 M^2 |I_{n,1}(h) + I_{n,2}(h) + I'_{n,2}(h) \\ &+ I_{n,3}(h) + I_{n,4}(h) + I_{n,5}(h)|, \quad (22) \end{aligned}$$

where

$$I_{n,1}(h) = n^{-1} \sum_{i=1}^n \sum_{k=1}^{\infty} -\sigma_k^{-2} \frac{\partial^2}{\partial \theta_r \partial \theta_s} a_k(\rho_{\theta_n^*, k} X_{i-1}) [a_k(\rho_{\theta_n^*, k} X_{i-1}) - a_k(\rho_{\theta} X_{i-1})],$$

$$I_{n,2}(h) = n^{-1} \sum_{i=1}^n \sum_{k=1}^{\infty} -\sigma_k^{-2} \frac{\partial}{\partial \theta_s} a_k(\rho_{\theta_n^*, k} X_{i-1}) \left[\frac{\partial}{\partial \theta_r} a_k(\rho_{\theta_n^*, k} X_{i-1}) - \frac{\partial}{\partial \theta_r} a_k(\rho_{\theta} X_{i-1}) \right],$$

$$I'_{n,2}(h) = n^{-1} \sum_{i=1}^n \sum_{k=1}^{\infty} -\sigma_k^{-2} \frac{\partial}{\partial \theta_r} a_k(\rho_{\theta, k} X_{i-1}) \left[\frac{\partial}{\partial \theta_s} a_k(\rho_{\theta_n^*, k} X_{i-1}) - \frac{\partial}{\partial \theta_s} a_k(\rho_{\theta} X_{i-1}) \right],$$

$$I_{n,3}(h) = n^{-1} \sum_{i=1}^n \sum_{k=1}^{\infty} \sigma_k^{-2} \left[\frac{\partial^2}{\partial \theta_r \partial \theta_s} a_k(\rho_{\theta_n^{**}, k} X_{i-1}) - \frac{\partial^2}{\partial \theta_r \partial \theta_s} a_k(\rho_{\theta} X_{i-1}) \right] a_k(\varepsilon_i),$$

$$I_{n,4}(h) = n^{-1} \sum_{i=1}^n \sum_{k=1}^{\infty} -\sigma_k^{-2} a_k(\rho_{\theta} X_{i-1}) \left[\frac{\partial^2}{\partial \theta_r \partial \theta_s} a_k(\rho_{\theta_n^*, k} X_{i-1}) - \frac{\partial^2}{\partial \theta_r \partial \theta_s} a_k(\rho_{\theta} X_{i-1}) \right],$$

$$I_{n,5}(h) = n^{-1} \sum_{i=1}^n \sum_{k=1}^{\infty} \sigma_k^{-2} a_k(\rho_{\theta} X_{i-1}) \left[\frac{\partial^2}{\partial \theta_r \partial \theta_s} a_k(\rho_{\theta_n^{**}, k} X_{i-1}) - \frac{\partial^2}{\partial \theta_r \partial \theta_s} a_k(\rho_{\theta} X_{i-1}) \right].$$

By the same technics used in the proof of Theorem 2 in ([13]), we may obtain that for all $j = 1, \dots, 5$, $I_{n,j}(h)$ converges to 0 in P_{θ} probability uniformly in $0 < h \leq M$.

From the last result and inequalities (22), (22) and (20), we conclude that under P_{θ} and for all $M > 0$

$$\sup_{|h| \leq M} |\Lambda_n - h^T \Gamma(\theta)^{1/2} S_n(\theta) + 1/2 h^T \Gamma(\theta) h| = o_p(1).$$

This ends the proof. ■

Acknowledgments

The authors would like to sincerely thank the anonymous Referee, The Editor and The Associated Editor for insightful comments which led to improve the presentation of this paper.

References

- [1] Aue A, Norinho DD, Hormann S, On the prediction of stationary functional time series, Journal of the American Statistical Association 110, 378-392 (2015).

- [2] Antoniadis A, Sapatinas T, Wavelet methods for continuous-time prediction using hilbert-valued autoregressive processes. *Journal of Multivariate Analysis*, 87(1): 133-158, 2003.
- [3] Arkoun O, Pergamenchtchikov S, Nonparametric Estimation for an Autoregressive Model. *Vestnik of Tomsk State University, Ser. Mathematics et Mechanics* 2 (3), 20 - 30, 2008.
- [4] Bosq D., *Linear Processes in Function Spaces*, Springer New-York (2000).
- [5] Ferraty F, Vieu P, *Nonparametric Functional Data Analysis: Theory and Practice*. Springer-Verlag, New York (2006).
- [6] Goia A, Vieu P, An introduction to recent advances in high/infinite dimensional statistics, *JMVA*, 146, 1-6 (2016).
- [7] Cuevas A, A partial overview of the theory of statistics with functional data, *JSPI*, 147, 1-23 (2014).
- [8] Hajek J, Local asymptotic Minimax and Admissibility in Estimation, *Proc. Sixth Berkeley Symp. on Math. Statist. and Prob (Univ. of Calif. Press)*, 1, 175-194 (1972).
- [9] Horváth L, Kokoszka P., *Inference for Functional Data with Applications*, Springer-Verlag New York (2012).
- [10] Hwang SY, Basawa IV, Asymptotic optimal inference for a class of nonlinear time series models, *Stoch. Proc. Appl*, 46, 91-113 (1993).
- [11] Ibragimov IA, Hasminski RZ, *Statistical Estimation. Asymptotic Theory*. Springer, New York (1981).
- [12] Ibragimov IA, Hasminski RZ, Asymptotically Normal Families of Distributions and Efficient Estimation, *Ann. Statist*, 19, 1681-1724 (1991).
- [13] Kara Terki N, Mourid T, On Local Asymptotic Normality For Functional Autoregressive Processes, *J. Multivariate Anal*, 148, 120-140 (2016).
- [14] Kowal DR, Matteson DS, Ruppert D, A Bayesian multivariate functional dynamic linear model, *Journal of the American Statistical Association* 112, 733-744 (2017).
- [15] Kreiss JP, Local asymptotic normality for autoregression with infinite order, *J. Stat. Plan. Inference*, 26, 185-219 (1990).

- [16] Kutoyants Yu A, *Parameter Estimation For Stochastic Processes*. Berlin, Helldermann (1984).
- [17] Kutoyants Yu A, *Statistical Inference for Ergodic Diffusion Processes*. Springer Series in Statistics (2004).
- [18] LeCam L, Yang GL, *Asymptotics in Statistics*. Second ed., Springer (2000).
- [19] Liébana J.A, Bosq D, Medina M.D.R, Consistency of the plug-in functional predictor of the Ornstein-Uhlenbeck process in Hilbert and Banach spaces. *Statistics and Probability Letters* 117 (2016) 12-22.
- [20] Loges W, Estimation of Parameters for Hilbert Space-Valued Partially Observable Stochastic Processes, *J. Multivariate Anal*, 20, 161-174 (1986).
- [21] Mas.A, Estimation d'opérateurs de corrélation de processus fonctionnels: lois limites, tests, déviations modérées. PhD thesis, Université Paris 6, 2000.
- [22] Milhaud X, Raugi A, Etude de l'estimateur du maximum de vraisemblance dans le cas d'un processus autorégressif: convergence, normalité asymptotique, vitesse de convergence, *Annales de l'I.H.P. Probabilités et Statistiques*, 25, 383-428 (1989).
- [23] Mourid T, *Contribution à la statistique des processus autorégressifs à temps continu*. Thèse de Doctorat D'Etat Es Sciences. Université Paris 6 (1995).
- [24] Omatu S, Nagamine H, Soeda T, *Optimal filter for a discret-time distributed parameter system and its application to environmental data processing*. In Application of Information and Control Systems, Vol.III.2nd International. Conf. Inform.Sci. Systems, Patras,Greece (1979).
- [25] Philippou AN, Roussas GG, Asymptotic distribution of the likelihood function in the independent not identically distributed case, *Ann. Statist*, 1, 454-471 (1973).
- [26] Ramsay JO, Silverman BW, *Functional Data Analysis*. Springer, New-York (2005).
- [27] Roussas GG, Asymptotic distribution of the log-likelihood function for stochastic processes, *Z.Wahrsch.Verw. Gebiete*, 47, 31-46 (1979).
- [28] Rozanov JU A, *Infinite-Dimensional Gaussian Distributions*. American Mathematical Society, Providence, Rhode Island (1971).

- [29] Swensen AR, The asymptotic distribution of the likelihood ratio for autoregressive time series with a regression trend, *J. Multivariate Anal*, 16, 54-70 (1985).
- [30] Tailen H, Randall E., *Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators* , Wiley (2015).
- [31] Vincent Lefieux. Modèles semi-paramétriques appliqués à la prévision des séries temporelles. Thèse de Doctorat. Université rennes 2(2007).