LAN property for some fractional type Brownian motion

Serge Cohen, Fabrice Gamboa, Céline Lacaux and Jean-Michel Loubes

Institut de Mathématiques de Toulouse
UMR 5219
31062 Toulouse, Cedex 9, France
E-mail address: Serge.Cohen@math.univ-toulouse.fr
URL: http://perso.math.univ-toulouse.fr/cohen/

Institut de Mathématiques de Toulouse
UMR 5219
31062 Toulouse, Cedex 9, France
E-mail address: Fabrice.Gamboa@math.univ-toulouse.fr
URL: http://www.math.univ-toulouse.fr/~gamboa/

Institut Élie Cartan,
Université de Lorraine, IECL, UMR 7502,
BP 70239, F-54506 Vandoeuvre-lès-Nancy, France.
Inria, Villers-lès-Nancy, F-54600, France.
E-mail address: Celine.Lacaux@univ-lorraine.fr
URL: http://www.iecn.u-nancy.fr/~lacaux/

Institut de Mathématiques de Toulouse
UMR 5219
31062 Toulouse, Cedex 9, France
E-mail address: Jean-Michel.Loubes@math.univ-toulouse.fr
URL: http://www.math.univ-toulouse.fr/~loubes/

Abstract. We study asymptotic expansion of the likelihood of a certain class of Gaussian processes characterized by their spectral density f_θ. We consider the case where $f_\theta(x) \sim_{x \to 0} |x|^{-\alpha(\theta)} L_\theta(x)$ with L_θ a slowly varying function and $\alpha(\theta) \in (-\infty, 1)$. We prove LAN property for these models which include in particular fractional Brownian motion or ARFIMA processes.

1. Introduction

Local asymptotic normality (LAN) property is a fundamental concept in asymptotic statistics. Originated by Wald (1943) and developed by Le Cam (1960), it relies on the idea of approximating a sequence of statistical models by a family of Gaussian distributions. Its consequence is that the initial model is approximately normal and thus inherits, in an asymptotic sense, the simple structure of normal models. Among the many applications in mathematical statistics, local asymptotic normality is essential in asymptotic optimality theory and also explains the asymptotic normality of certain estimators such as the maximum likelihood estimator for instance. We refer for instance to Pollard (1984) or in van der Vaart (1998) for applications of LAN property. When dealing with inference on the parameter, LAN property will enable to assess optimality of any estimation procedure for this parameter which governs the behavior of the random process. Hence LAN property is a powerful framework to understand probabilistic properties of a stochastic model.

Many works have been done to prove LAN property for a large number of observation models such as i.i.d sequences of random variables parametrized by a parameter or Gaussian processes in van der Vaart (1998) or more complicated random processes such as multifractal processes in Loubes and Paindaveine (2011); Gamboa and Loubes (2007), AR or ARMA based models in Garel and Hallin (1995); Lai et al. (2002) or extreme models in Falk (2011) for instance. We focus in this paper on statistical inference for empirical estimation of the parameters of spectral density of a certain class of Gaussian processes. We consider a stationary centered Gaussian process X_n whose spectral density is indexed by a parameter θ and satisfies the condition

$$f_\theta(x) \sim_{x \to 0} |x|^{-\alpha(\theta)} L_\theta(x)$$

with L_θ a slowly varying function and $\alpha(\theta) \in (-\infty, 1)$. More precisely, we aim at proving Local Asymptotic Normality (LAN) for the model where we observe a sample of n observations $\mathbf{X}_n = (X_1, \ldots, X_n)$ by studying an asymptotic expansion of the log likelihood. For this, a precise control over the asymptotic behavior of some Toeplitz matrices linked with f_θ will be required. It relies on the results in Lieberman et al. (2011).

In particular, our assumptions (see section 2) are fulfilled by fractional Gaussian noises, which are defined as increments of fractional Brownian motions (see Kolmogoroff (1940); Mandelbrot and Van Ness (1968)). From the LAN property fulfilled by fractional Gaussian noises, we deduce the LAN property when the observation model is a time-discretized fractional Brownian motion, which is not any more a stationary model. Moreover observation models of autoregressive fractionally integrated moving average processes (ARFIMA(p,d,q)), defined as a fractionally differenced ARMA processes in Granger and Joyeux (1980); Hosking (1981), satisfy also our assumptions and are covered by our results. Note that, when $d \le -1$, ARFIMA(p,d,q) are non-invertible processes (see Bondon and Palma (2007)).

The paper falls into the following parts. Section 2 states the general LAN property for the considered processes. Then Section 3 is devoted to recall some basic properties of Toepliz matrices which are useful to prove our main result. Section 4 is devoted to two examples that undergo the required assumptions (fractional Brownian noises and ARFIMA processes) and the LAN property for the non stationary

model provided by the fractional Brownian motion. Most of the proofs are postponed to Section A.

2. LAN property for a certain class of random processes

Let $X_n, n \in \mathbb{N}$ be a centered Gaussian stationary process with law P_θ parametrized by $\theta = (\theta_1, \ldots, \theta_m)' \in \Theta$ an open subset of $\mathbb{R}^m$ and associated with the 2π-periodic even spectral density f_θ. Then, under P_θ,

$$\mathbf{E}(X_n X_{n+k}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(ikx) f_\theta(x) dx = c_k(f_\theta).$$

As usual, for $\theta \neq \eta$, the set $\{x \in [-\pi, \pi], f_\theta(x) \neq f_\eta(x)\}$ is assumed to have positive Lebesgue measure. This assumption is not needed to obtain the LAN property but is a standard background assumption in statistics. Actually, if this condition is not fulfilled, the model is not identifiable, preventing any estimation issues.

In practice, we observe the vector $\mathbf{X}_n = (X_1, \ldots, X_n)$, with $n \in \mathbb{N}\backslash\{0\}$, whose law is denoted by P_θ^n. Under P_θ^n, the covariance matrix of $\mathbf{X}_n$ is then the symmetric Toeplitz matrix

$$\Gamma_{\mathbf{X_n}} = \frac{1}{2\pi} T(f_\theta)$$

where $T(f)$ denotes the real Toeplitz matrix

$$T_n(f) = \left(\int_{-\pi}^{\pi} \exp\left(i(k-j)x\right) f(x) dx \right)_{1 \leq k, j \leq n}. \tag{2.1}$$

for any integrable symmetric function $f : [-\pi, \pi] \to \overline{\mathbb{R}}$. Note also that the Fisher information of the model is the matrix

$$I(\theta) = \frac{1}{4\pi} \left(\int_{-\pi}^{\pi} \frac{\partial \log f_\theta(x)}{\partial \theta_k} \frac{\partial \log f_\theta(x)}{\partial \theta_j} dx \right)_{1 \leq k, j \leq m}.$$

The LAN property of the model is proved under the following assumption.

Assumption 2.1.

(A.1) *For any $x \in [-\pi, \pi]\backslash\{0\}$, the function $\theta \mapsto f_\theta(x)$ is three times continuously differentiable on Θ. In addition, for any $0 \leq \ell \leq 3$ and $1 \leq k_1, \ldots, k_\ell \leq m$, the partial derivative*

$$(\theta, x) \to \frac{\partial^\ell}{\partial \theta_{k_1} \ldots \partial \theta_{k_\ell}} f_\theta(x)$$

is continuous on $\Theta \times [-\pi, \pi]\backslash\{0\}$, continuously differentiable with respect to x on $[-\pi, \pi]\backslash\{0\}$ and its partial derivative

$$(\theta, x) \to \frac{\partial^{\ell+1}}{\partial x \partial \theta_{k_1} \ldots \partial \theta_{k_\ell}} f_\theta(x)$$

is continuous on $\Theta \times [-\pi, \pi]\backslash\{0\}$.

(A.2) *There exists a continuous function $\alpha : \Theta \to (-\infty, 1)$ such that for any $\delta > 0$ and any compact set $\Theta^* \subset \Theta$, the following conditions hold for every $(\theta, x) \in \Theta^* \times [-\pi, \pi]\backslash\{0\}$.*

  (a) $c_{1,\delta,\Theta^*} |x|^{-\alpha(\theta)+\delta} \leq f_\theta(x) \leq c_{2,\delta,\Theta^*} |x|^{-\alpha(\theta)-\delta}$

  (b) $\left| \dfrac{\partial}{\partial x} f_\theta(x) \right| \leq c_{2,\delta,\Theta^*} |x|^{-\alpha(\theta)-1-\delta}$

(c) for any $\ell \in \{1, 2, 3\}$, and any $k \in \{1, \ldots, m\}^\ell$,

$$\left| \frac{\partial^\ell}{\partial \theta_{k_1} \ldots \partial \theta_{k_\ell}} f_\theta(x) \right| \leq c_{2,\delta,\Theta^*} |x|^{-\alpha(\theta)-\delta}.$$

with c_{i,δ,Θ^*} some finite positive constants which only depend on δ and Θ^*.

This assumption implies that $1/f_\theta$ is well-defined on $[-\pi, \pi]\backslash\{0\}$ and corresponds to Assumptions (A1), (A2) and (A4) in Lieberman et al. (2012), except that we impose some smoothness property on the derivative of order three. This assumption, as noted in Lieberman et al. (2012), is an extension and a reformulation of Dahlhaus's ones in Dahlhaus (1989, 2006).

If for the true value θ_0 of the parameter, $\alpha(\theta_0) \in (-1, 1)$, the LAN property (see Theorem 2.4) holds. Nevertheless, if $\alpha(\theta_0) \leq -1$, the LAN property is established under the following additional assumption (which allows to apply Theorem 3.5).

Assumption 2.2. *Let $\Theta^* = B(\theta_0, r) \subset \Theta$, where $B(u, r)$ is the Euclidean closed ball of $\mathbb{R}^m$ centered at u with radius $r > 0$. For any $\delta > 0$, for any $\ell \in \{1, 2, 3\}$, for any $\theta \in \Theta^*$, and $k \in \{1, \ldots, m\}^\ell$,*

$$\left| \frac{\partial^{\ell+1}}{\partial x \partial \theta_{k_1} \ldots \partial \theta_{k_\ell}} f_\theta(x) \right| \leq c_{2,\delta,\Theta^*} |x|^{-\alpha(\theta)-1-\delta}$$

with α and c_{2,δ,Θ^*} given in Assumption 2.1.

Under Assumption 2.1, for each θ, the Toeplitz matrix $T_n(f_\theta)$ is positive and then invertible since the non-negative function f_θ is positive on a non neglectible set. In particular, the covariance matrix $\Gamma_{\mathbf{X_n}} = \frac{1}{2\pi} T_n(f_\theta)$ of the model is invertible for each θ. Then we compare the distribution of the model under P_θ and P_η using the following proposition.

Proposition 2.3. *For any symmetric positive definite matrix Γ on $\mathbb{R}^n$ ($n \in \mathbb{N}\backslash\{0\}$), P_Γ denotes the distribution of a centered Gaussian vector with covariance Γ. Then, for any covariance matrices Γ_1 and Γ_2,*

$$2 \log \frac{dP_{\Gamma_1}}{dP_{\Gamma_2}}(x) = \langle x, (\Gamma_2^{-1} - \Gamma_1^{-1})x \rangle + \log \det(\Gamma_1^{-1}\Gamma_2),$$

where $\langle \cdot \rangle$ denotes the usual Hermitian product in $\mathbb{C}^n$.

We will denote by $\|y\|$ the Hermitian norm of $y \in \mathbb{C}^n$ and A^* denotes the conjugate transpose of the matrix A with complex coefficients.

The following theorem states LAN property for the observation model.

Theorem 2.4 (LAN property)**.** *Let $\theta_0 \in \Theta$. Assume that Assumption 2.1 is fulfilled. If $\alpha(\theta_0) \leq -1$, assume also that Assumption 2.2 is fulfilled. Then under $P_{\theta_0}^n$, for $t \in \mathbb{R}^m$, we get*

$$\log \frac{dP_{\theta_0+t/\sqrt{n}}^n}{dP_{\theta_0}^n} = \langle t, Z_n \rangle - \frac{1}{2}t^* I(\theta_0)t + \psi_{\theta_0}(t, n)$$

where Z_n does not depend on t and converges in distribution, under $P_{\theta_0}^n$, to a centered Gaussian vector with covariance matrix $I(\theta_0)$, while $\psi_{\theta_0}(\cdot, n)$ converges uniformly on each compact to 0 $P_{\theta_0}^n$-almost surely when $n \to +\infty$.

Proof: Let $K \subset \mathbb{R}^m$ be a compact set and consider $r > 0$ such that

$$\Theta^* = B(\theta_0, r) \subset \Theta.$$

Then, we can choose n_0, such that for any integer $n \geq n_0$ and any $t \in K$, $\theta_0 + t/\sqrt{n} \in \Theta^*$.

Let us now consider $n \geq n_0$ and observe that for any $t \in K$, $P_{\theta_0}^n$ and $P_{\theta_0 + t/\sqrt{n}}^n$ are well-defined. Moreover, using Proposition 2.3, for any $t \in K$, we get

$$\log \frac{dP_{\theta_0 + t/\sqrt{n}}^n}{dP_{\theta_0}^n}(\mathbf{x}_n) = F_n\left(\theta_0 + \frac{t}{\sqrt{n}}\right)$$

where $\mathbf{x}_n = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and for $\theta \in B(\theta_0, r) = \Theta^*$,

$$F_n(\theta) = \pi < \mathbf{x}_n, [T_n(f_{\theta_0})^{-1} - T_n(f_\theta)^{-1}]\mathbf{x}_n > + \frac{1}{2}\log\det[T_n(f_\theta)^{-1}T_n(f_{\theta_0})].$$

By Assumption 2.1, F_n is three times continuously differentiable on $B(\theta_0, r)$. Hence, for any $t \in K$, since $F_n(\theta_0) = 0$,

$$\left| F_n\left(\theta_0 + \frac{t}{\sqrt{n}}\right) - \frac{\langle t, \nabla F_n(\theta_0)\rangle}{\sqrt{n}} - \frac{t^* \nabla^2 F_n(\theta_0)t}{2n}\right| \leq$$
$$\frac{M_K^3}{6n^{3/2}} \max_{1 \leq j,k,l \leq m} \sup_{\theta \in B(\theta_0, r)} \left| \frac{\partial^3 F_n}{\partial\theta_j \partial\theta_k \partial\theta_l}(\theta)\right|$$

Let us now introduce the spectral norm

$$\|A\|_{2,n} = \sup_{x \in \mathbb{C}^n} \left(\frac{x^* A^* A x}{x^* x}\right)^{1/2} \tag{2.2}$$

of a $n \times n$ matrix A and recall that for any $x \in \mathbb{C}^n$,

$$x^* A x \leq x^* x \|A\|_{2,n} = \|A\|_{2,n}\|x\|^2, \tag{2.3}$$

see Graybill (1983) for example.

Hence, setting $Z_n = \nabla F_n(\theta_0)/\sqrt{n}$ (which does not depend on $t \in K$), and applying Equation (2.3) with $A = \nabla^2 F_n(\theta_0)/n + I(\theta_0)$, we get

$$\forall t \in K, \ F_n\left(\theta_0 + \frac{t}{\sqrt{n}}\right) = \langle t, Z_n\rangle - \frac{1}{2}t^* I(\theta_0)t + \psi_{\theta_0}(t, n)$$

with

$$\sup_{s \in K} |\psi_{\theta_0}(t, n)| \leq \frac{M_K^2}{2}\left\|\frac{\nabla^2 F_n(\theta_0)}{n} + I(\theta_0)\right\|_{2,m}$$
$$+ \frac{M_K^3}{6n^{3/2}} \max_{1 \leq j,k,l \leq m} \sup_{\theta \in B(\theta_0, r)} \left|\frac{\partial^3 F_n}{\partial\theta_j \partial\theta_k \partial\theta_l}(\theta)\right|.$$

The conclusion follows from the three following lemmas, whose proofs are postponed to the Appendix for sake of clearness. The first lemma deals with the behavior of Z_n.

Lemma 2.5. *Under $P_{\theta_0}^n$, Z_n converges in distribution, as $n \to +\infty$, to a centered Gaussian random vector whose covariance matrix is the Fisher information $I(\theta_0)$.*

Let us now state the asymptotic of $\nabla^2 F_n(\theta_0)$.

Lemma 2.6. *Under $P_{\theta_0}^n$, $\nabla^2 F_n(\theta_0)/n$ converges almost surely to $-I(\theta_0)$, as $n \to +\infty$. Hence,*

$$\left\| \frac{\nabla^2 F_n(\theta_0)}{n} + I(\theta_0) \right\|_{2,m}$$

converges almost surely to 0 as $n \to +\infty$.

The next lemma deals with the behavior of the partial derivative of F_n of order three.

Lemma 2.7. *For r small enough, for any $1 \leq j, k, l \leq m$, under $P_{\theta_0}^n$*

$$\frac{1}{n^{3/2}} \sup_{B(\theta_0, r)} \left| \frac{\partial^3 F_n}{\partial \theta_j \partial \theta_k \partial \theta_l} \right|$$

converges almost surely to 0.

Conbining Lemmas 2.6 and 2.7, we get

$$\lim_{n \to +\infty} \sup_{s \in K} |\psi_{\theta_0}(s, n)| = 0 \quad P_{\theta_0}^n\text{-almost surely,}$$

which concludes the proof.   □

The proofs of Lemmas 2.5, 2.6 and 2.7 rely on the behavior of the spectral norm and of the trace of some products of Toeplitz matrices. These behaviors are recalled in the next section.

3. Technical results on Toeplitz matrices

Before stating some results on Toeplitz matrices, let us recall that if the $n \times n$ matrix A is nonnegative and Hermitian, hence the matrix $A^{1/2}$ defined as a solution of $A = (A^{1/2})^2$, exists and is a nonnegative Hermitian matrix.

Under some assumptions on functions f and g, we can then consider products of the form

$$T_n(f)^{-1/2} T_n(g)^{1/2},$$

where $T_n(h)$ is defined by (2.1). To prove Lemmas 2.5, 2.6 and 2.7, one of the main tools we use is the following lemma, which gives a bound for the spectral norm of theses products. This lemma, given in Lieberman et al. (2011) (full version of Lieberman et al. (2012)), generalizes Lemma 5.3 in Dahlhaus (1989).

Lemma 3.1. *Let f and g be nonnegative symmetric functions defined on $[-\pi, \pi]$. Assume that there exist some constants $c_1, c_2 \in (0, +\infty)$ and $\beta_1, \beta_2 \in (-\infty, 1)$ such that for any $x \in [-\pi, \pi] \backslash \{0\}$,*

$$f(x) \geq c_1 |x|^{-\beta_1} \quad \text{and} \quad g(x) \leq c_2 |x|^{-\beta_2}. \tag{3.1}$$

Then, there exists a constant K which only depends on $(c_1, c_2, \beta_1, \beta_2)$ such that for any integer $n \geq 1$,

$$\left\| T_n(f)^{-1/2} T_n(g)^{1/2} \right\|_{2,n} = \left\| T_n(g)^{1/2} T_n(f)^{-1/2} \right\|_{2,n} \leq K n^{\max((\beta_2 - \beta_1)/2, 0)},$$

with $\|\cdot\|_{2,n}$ the spectral norm given by (2.2).

Remark 3.2. In the previous lemma, observe that the assumption on f ensures that $T_n(f)^{-1/2}$ exists. Moreover, one can choose the constant K independently of (β_1, β_2) and such that the conclusion holds for any $\beta_1, \beta_2 \in [a, b]$.

In our framework, f depends on an unknown parameter θ and is the spectral density of a centered Gaussian stationary sequence $(X_n)_n$. This spectral density will be denoted f_θ and is assumed to be such that

$$f_\theta(x) \sim_{x \to 0} |x|^{-\alpha(\theta)} L_\theta(x)$$

with L_θ a slowly varying function and $\alpha(\theta) \in (-\infty, 1)$. Then next theorem deals with the uniform behavior in θ as $n \to +\infty$ of

$$\text{tr}\left[\prod_{\ell=1}^{p} (T_n(f_\theta))^{-1} T_n(g_{\theta,\ell}) \right],$$

where $g_{\theta,\ell}$ denotes a spectral density or one of its derivatives which undergoes some technical assumptions. If the true value θ_0 of the parameter is such that $\alpha(\theta_0) \in (-1,1)$, this theorem is one of the main tools we use to obtain the LAN property. It allows us to consider a process $(X_n)_n$ which admits antipersistence ($\alpha(\theta_0) < 0$), short memory ($\alpha(\theta_0) = 0$) or long memory ($\alpha(\theta_0) \in (0,1)$). This theorem is stated as Theorem 5 in Lieberman et al. (2011) (full version of Lieberman et al. (2012)). It generalizes Theorem 2 in Lieberman et al. (2003), which is already a uniform version of Theorem 1.a Fox and Taqqu (1987) and Theorem 5.1 in Dahlhaus (1989).

Theorem 3.3. *Let $\Theta^* \subset \mathbb{R}^m$ be a compact set and $p \in \mathbb{N}\backslash\{0\}$. For any $1 \leq \ell \leq p$, consider $f_\ell : \Theta^* \times [-\pi, \pi] \to [0, \infty]$ and $g_\ell : \Theta^* \times [-\pi, \pi] \to \overline{\mathbb{R}}$ two symmetric functions with respect to their second variable. In the following,*

$$f_{\theta,\ell} = f_\ell(\theta, \cdot) \quad \text{and} \quad g_{\theta,\ell} = g_\ell(\theta, \cdot).$$

Assume that the following conditions hold.

  (1) For any $1 \leq \ell \leq p$, for any $\theta \in \Theta^*$, $f_{\theta,\ell}$ and $g_{\theta,\ell}$ are differentiable on $[-\pi, \pi]\backslash\{0\}$. Moreover, for any $1 \leq \ell \leq p$, f_ℓ, $\frac{\partial}{\partial x} f_\ell$, g_ℓ and $\frac{\partial}{\partial x} g_\ell$ are continuous on $\Theta^* \times [-\pi, \pi]\backslash\{0\}$.
  (2) There exist two continuous functions $\alpha : \Theta^* \to (-1,1)$ and $\beta : \Theta^* \to (-\infty, 1)$ such that for any $\delta > 0$, for every $(\theta, x) \in \Theta^* \times [-\pi, \pi]\backslash\{0\}$ and any $1 \leq \ell \leq p$
    (a) $c_{1,\delta,\Theta^*} |x|^{-\alpha(\theta)+\delta} \leq f_\ell(\theta, x) \leq c_{2,\delta,\Theta^*} |x|^{-\alpha(\theta)-\delta}$
    (b) $\left| \frac{\partial}{\partial x} f_\ell(\theta, x) \right| \leq c_{2,\delta,\Theta^*} |x|^{-\alpha(\theta)-1-\delta}$
    (c) and $|g_\ell(\theta, x)| \leq c_{2,\delta,\Theta^*} |x|^{-\beta(\theta)-\delta}$,
    with c_{i,δ,Θ^*}, $i \in \{1,2\}$ some finite positive constants which only depend on δ and Θ^*.
  (3) For any $\theta \in \Theta^*$, $p(\beta(\theta) - \alpha(\theta)) < 1$.

Then,

$$\lim_{n \to +\infty} \sup_{\theta \in \Theta^*} \left| \frac{1}{n}\text{tr}\left[\prod_{\ell=1}^{p} (T_n(f_{\theta,\ell}))^{-1} T_n(g_{\theta,\ell}) \right] - \frac{1}{2\pi} \int_{-\pi}^{\pi} \prod_{j=1}^{p} (f_{\theta,\ell}(x))^{-1} g_{\theta,\ell}(x) dx \right| = 0.$$

Remark 3.4. Observe that under Conditions 1. and 2., $f_\ell^{-1} = 1/f_\ell$ is continuous on $\Theta^* \times [-\pi, \pi]\backslash\{0\}$, as assumed in Theorem 5 of Lieberman et al. (2011).

If the true value θ_0 of the parameter is such that $\alpha(\theta_0) \leq -1$, the previous theorem can not be applied. However, the following theorem, which is a simple consequence of Lemma 8 in Lieberman et al. (2011), provides a sufficient property to establish the LAN property. In particular, it allows us to study the LAN property

for ARFIMA models whose order of differentiability are lower than $1/2$, which includes some non invertible models.

Theorem 3.5. *Let $\Theta^* = B(\theta_0, r) \subset \mathbb{R}^m$ be the closed Euclidean ball centered at θ_0 with radius r and let $p \in \mathbb{N}\backslash\{0\}$. Consider $f : \Theta^* \times [-\pi, \pi] \to [0, \infty]$ and for $1 \leq \ell \leq p$, $g_\ell : \Theta^* \times [-\pi, \pi] \to \mathbb{R}$ some symmetric functions in their second variable. In the following,*

$$f_\theta = f(\theta, \cdot) \quad and \quad g_{\theta,\ell} = g_\ell(\theta, \cdot).$$

Assume that the following conditions hold.

(1) The functions f and g_ℓ satisfy assumption 1 of Theorem 3.3.

(2) There exists a continuous function $\alpha : \Theta^* \to (-\infty, 1/2)$ such that for any $\delta > 0$, for every $(\theta, x) \in \Theta^* \times [-\pi, \pi]\backslash\{0\}$ and any $1 \leq \ell \leq p$, assertion 3(a), 3(b) of Theorem 3.3 are fulfilled (with $f_\ell = f$), assertion 3(c) of Theorem 3.3 holds with $\beta = \alpha$ and

$$\left| \frac{\partial}{\partial x} g_{\theta,\ell}(x) \right| \leq c_{2,\delta,\Theta^*} |x|^{-\alpha(\theta)-1-\delta}.$$

Then, for r small enough,

$$\lim_{n \to +\infty} \sup_{\theta \in \Theta^*} \left| \frac{1}{n} \mathrm{tr} \left[\prod_{\ell=1}^p (T_n(f_\theta))^{-1} T_n(g_{\theta,\ell}) \right] - \frac{1}{2\pi} \int_{-\pi}^\pi f_\theta^{-p}(x) \prod_{j=1}^p g_{\theta,\ell}(x) dx \right| = 0.$$

4. Application to Fractional Gaussian noises and ARFIMA processes

Here we consider two particular cases where the LAN property can be proved.

Fractional Gaussian noises:

Let $(B_H(t))_{t \geq 0}$ be a fractional Brownian motion (see Kolmogoroff (1940); Mandelbrot and Van Ness (1968)) with Hurst index $H \in (0, 1)$. In other words, B_H is a centered Gaussian random process whose covariance function is given by

$$\mathbf{E}(B_H(t)B_H(s)) = \frac{\sigma^2}{2} \left[|t|^{2H} - 2|t-s|^{2H} + |s|^{2H} \right]. \tag{4.1}$$

The parameter σ^2 corresponds to the variance of $B_H(1)$. Let us now consider the centered stationary Gaussian sequence $(X_n)_{n \geq 1}$, called the fractional Gaussian noise of index H, defined by

$$\text{for } n \geq 1, \quad X_n = B_H(n) - B_H(n-1).$$

The law of $(X_n)_{n \geq 1}$ is parametrized by $\theta = (\sigma^2, H) \in (0, +\infty) \times (0, 1)$. According to Samorodnitsky and Taqqu (1994), its spectral density $f_{\sigma^2, H}$ is given by

$$f_{\sigma^2, H}(x) = \frac{\sigma^2 |e^{ix} - 1|^2}{C_2^2(H)} \sum_{k \in \mathbb{Z}} \frac{1}{|x + 2k\pi|^{2H+1}}, \quad x \in [-\pi, \pi]\backslash\{0\}, \tag{4.2}$$

where

$$C_2^2(\alpha) = \frac{\pi}{\alpha \Gamma(2\alpha) \sin(\alpha\pi)}.$$

Then, the model satisfies Assumption 2.1 with $\alpha(\theta) = 2H - 1$. Since the range of α is $(-1, 1)$, the Assumption 2.2 is not needed in this example.

Next proposition establishes the LAN property when the observation are modeled by

$$\mathbf{B}_n = \{B_H(1), \ldots, B_H(n)\}$$

with B_H the fractional Brownian motion whose covariance function is given by (4.1). This model is not a stationary one but its log-likelihood can be linked to those of the fractional Gaussian noise

$$\mathbf{X}_n = \{B_H(1), B_H(2) - B_H(1), \ldots, B_H(n) - B_H(n-1)\},$$

which fulfills Assumption 2.1. The law of $\mathbf{B}_n$ is parametrized by $(\sigma^2, H) \in (0, +\infty) \times (0,1) = \Theta$ and denoted by $Q^n_{\sigma^2, H}$.

Proposition 4.1. *Let I be the Fisher information of the fractional Gaussian noise $\mathbf{X}_n$, that is the Fisher information associated with the spectral density $f_{\sigma^2, H}$ defined by (4.2). Then, under $Q^n_{\sigma_0^2, H_0}$, for $t \in \mathbb{R}^2$ and n large enough*

$$\log \frac{dQ^n_{(\sigma_0^2, H_0) + t/\sqrt{n}}}{dQ^n_{(\sigma_0^2, H_0)}} = \langle t, Z_n \rangle - \frac{1}{2}t^* I(\sigma_0^2, H_0)t + \psi_{\sigma_0^2, H_0}(t, n)$$

where Z_n does not depend on t and converges in distribution, under $Q^n_{(\sigma_0^2, H_0)}$, to a centered Gaussian vector with covariance matrix $I(\sigma_0^2, H_0)$, while $\psi_{\sigma_0^2, H_0}(\cdot, n)$ converges uniformly on each compact to 0 $Q^n_{(\sigma_0^2, H_0)}$-almost surely when $n \to +\infty$.

Proof: Let $\theta_0 = (\sigma_0^2, H_0)$. As previously, P^n_θ denotes the law of $\mathbf{X}_n$. Observe that

$$\log \frac{dQ^n_{\theta_0 + t/\sqrt{n}}}{dQ^n_{\theta_0}}(\mathbf{b}_n) = \log \frac{dP^n_{\theta_0 + t/\sqrt{n}}}{dP^n_{\theta_0}}(\mathbf{x}_n)$$

where $\mathbf{b}_n = (b_1, \ldots, b_n)'$ and $\mathbf{x}_n = (b_1, b_2 - b_1, \ldots, b_n - b_{n-1})'$. Since under $Q^n_{\sigma_0^2, H_0}$, the law of $\mathbf{x}_n$ is $P^n_{\theta_0}$, the conclusion follows from Theorem 2.4. $\qquad\square$

ARFIMA processes:

ARFIMA processes have been introduced in Granger and Joyeux (1980); Hosking (1981). We also refer to Beran (1994) for general properties of ARFIMA(p, d, q).

Let $p, q \in \mathbb{N}$. Then, a stationary ARFIMA process $(X_n)_n$ is parametrized by $\theta = (\sigma^2, d, \Phi_1, \ldots, \Phi_p, \Psi_1, \ldots, \Psi_q)$ where $d \in (-\infty, 1)$ is the order of differentiability and the polynoms

$$\Phi(X) = 1 + \sum_{j=1}^{p} \Phi_j X^j \quad \text{and} \quad \Psi(X) = 1 + \sum_{j=1}^{q} \Psi_j X^j$$

have no zeros in the unit circle and no zeros in common. Then, its spectral density is given by

$$f_\theta(x) = \sigma^2 \left| e^{ix} - 1 \right|^{-2d} \left| \frac{\Psi(e^{ix})}{\Phi(e^{ix})} \right|^2.$$

Then Assumptions 2.1 and 2.2 are fulfilled with $\alpha(\theta) = 2d$. Theorem 2.4 also implies LAN property for this model.

Appendix A.

A.1. *Proof of Lemma 2.5.* In this appendix, for $\ell \in \{1, 2, 3\}$ and $k \in \{1, \ldots, m\}^\ell$, $\partial_k^\ell f_\theta$ denotes the partial derivative of $(\theta, x) \mapsto f_\theta(x)$ with respect to $(\theta_{k_1}, \ldots, \theta_{k_\ell})$, that is

$$\partial_k^\ell f_\theta(x) = \frac{\partial^\ell f_\theta}{\partial \theta_{k_1} \cdots \partial \theta_{k_\ell}}(x). \tag{A.1}$$

By definition of F_n, for any $\theta \in B(\theta_0, r)$, and any integer $1 \leq k \leq m$,

$$\frac{\partial F_n}{\partial \theta_k}(\theta) = \pi \langle \mathbf{x}_n, T_n(f_\theta)^{-1} T_n(\partial_k f_\theta) T_n(f_\theta)^{-1} \mathbf{x}_n \rangle - \frac{1}{2} \text{tr}\big(T_n(\partial_k f_\theta) T_n(f_\theta)^{-1}\big) \quad \text{(A.2)}$$

Let us fix $u \in \mathbb{R}^m$ and study the asymptotic behavior, under $P_{\theta_0}^n$, of $\langle u, \nabla F_n(\theta_0) \rangle$, that is by (A.2) of

$$\langle u, \nabla F_n(\theta_0) \rangle = \pi \langle \mathbf{x}_n, T_n(f_{\theta_0})^{-1} T_n(g_{\theta_0}^u) T_n(f_{\theta_0})^{-1} \mathbf{x}_n \rangle - \frac{1}{2} \text{tr}\big(T_n(g_{\theta_0}^u) T_n(f_{\theta_0})^{-1}\big)$$

with

$$g_\theta^u = \sum_{k=1}^m u_k \, \partial_k f_\theta.$$

To achieve this goal we use the following result on Gaussian random vectors.

Lemma A.1. *Assume that $Y = (Y_1, \ldots, Y_n)'$ is a centered Gaussian random vector with covariance matrix Γ and consider A a real symmetric matrix of order n. Then,*

$$\langle Y, AY \rangle \overset{(d)}{=} \sum_{j=1}^n \lambda_{j,n} \chi_{j,n}$$

where $\overset{(d)}{=}$ stands for equality in distribution, $(\lambda_{j,n})_{1 \leq j \leq n}$ are the eigenvalues of the real symmetric matrix $\Gamma^{1/2} A \Gamma^{1/2}$ and $(\chi_{j,n})_{j,n}$ are i.i.d. random variables with distribution $\chi^2(1)$. Moreover,

$$\mathbf{E}(\langle Y, AY \rangle) = \text{tr}(A\Gamma) = \text{tr}(\Gamma A) \quad \text{and} \quad \text{Var}(\langle Y, AY \rangle) = 2 \sum_{j=1}^n \lambda_{j,n}^2 = 2\text{tr}\big((A\Gamma)^2\big).$$

Observe that under $P_{\theta_0}^n$, $\mathbf{x}_n$ is a centered Gaussian random variable with covariance $\Gamma_n = \frac{1}{2\pi} T_n(f_{\theta_0})$. Then, since $T_n(f_{\theta_0})^{-1} T_n(g_{\theta_0}^u) T_n(f_{\theta_0})^{-1}$ is a real symmetric matrix, under $P_{\theta_0}^n$,

$$\langle u, \nabla F_n(\theta_0) \rangle \overset{(d)}{=} \sum_{j=1}^n \lambda_{j,n}^u (\chi_{j,n} - 1),$$

where $(\lambda_{j,n}^u)_{j=1,\ldots,n}$ are the eigenvalues of

$$B_{\theta_0}^u = \frac{1}{2} T_n(f_{\theta_0})^{-1/2} T_n\big(g_{\theta_0}^u\big) T_n(f_{\theta_0})^{-1/2}.$$

Therefore, under $P_{\theta_0}^n$

$$\langle u, Z_n \rangle \overset{(d)}{=} \sum_{j=1}^n \frac{\sqrt{2}\lambda_{j,n}^u}{\sqrt{n}} \xi_{j,n}$$

where $\xi_{j,n} = (\chi_{j,n} - 1)/\sqrt{2}$ $(1 \leq j \leq n, n \geq 1)$ are i.i.d. centered random variables having unitary variance. To obtain the convergence of Z_n, we use the following Lemma, which is an obvious corollary of Lindenberg theorem (see Billingsley (1995) for instance).

Lemma A.2. *Let $(\xi_{j,n})_{n \geq 1, 1 \leq j \leq n}$ be a sequence of i.i.d centered random variables having unitary variance and let $(v_{j,n})_{n \geq 1, 1 \leq j \leq n}$ be a triangular array of real numbers. Assume further that*

(1) $\lim_{n \to +\infty} \sup_{1 \leq j \leq n} |v_{j,n}| = 0$,
(2) $\lim_{n \to +\infty} \sum_{j=1}^n v_{j,n}^2 = \tau^2 > 0$.

Then, as $n \to +\infty$, $\sum_{j=1}^{n} v_{j,n} \xi_{j,n}$ converges in distribution to a centered Gaussian distribution with variance τ^2.

We first check Condition 1 for the sequence $v_{j,n} = \sqrt{2} n^{-1/2} \lambda_{j,n}^{u}$. Since $B_{\theta_0}^{u}$ is an Hermitian matrix whose eigenvalues are $\left(\lambda_{j,n}^{u} \right)_{1 \leq j \leq n}$, its spectral radius

$$\rho_n(u) := \sup_{1 \leq j \leq n} |\lambda_{j,n}^{u}|$$

is given by

$$\rho_n(u) = \sup_{x \in \mathbb{C}^n \setminus \{0\}} \frac{\left| x^* B_{\theta_0}^{u} x \right|}{x^* x} = \frac{1}{2} \sup_{x \in \mathbb{C}^n \setminus \{0\}} \frac{\left| x^* T_n(f_{\theta_0})^{-1/2} T_n(g_{\theta_0}^{u}) T_n(f_{\theta_0})^{-1/2} x \right|}{x^* x}.$$

Observe that for any $y \in \mathbb{C}^n$, and any integrable symmetric function h,

$$y^* T_n(h) y = \int_{-\pi}^{\pi} \left| \sum_{k=1}^{n} e^{ikx} y_k \right|^2 h(x) dx$$

and therefore that

$$|y^* T_n(h) y| \leq y^* T_n(|h|) y. \tag{A.3}$$

Then, we get

$$\rho_n(u) \leq \frac{1}{2} \sup_{x \in \mathbb{C}^n \setminus \{0\}} \frac{x^* T_n(f_{\theta_0})^{-1/2} T_n(|g_{\theta_0}^{u}|) T_n(f_{\theta_0})^{-1/2} x}{x^* x},$$

which can be written as

$$\rho_n(u) \leq \frac{1}{2} \left\| T_n \left(|g_{\theta_0}^{u}| \right)^{1/2} T_n(f_{\theta_0})^{-1/2} \right\|_{2,n}^2,$$

since $\left| g_{\theta_0}^{u} \right|$ is a nonnegative symmetric function on $[-\pi, \pi]$.

By Assumption 2.1, the functions $f = f_{\theta_0}$ and $g = \left| g_{\theta_0}^{u} \right|$ satisfy Equation (3.1) with $\beta_1 = \alpha(\theta_0) - \delta$ and $\beta_2 = \alpha(\theta_0) + \delta$ (for any $\delta > 0$). Then, applying Lemma 3.1, for any $\delta > 0$, we get

$$\rho_n(u) < K_\delta n^{2\delta}$$

where the finite positive constant K_δ does not depend on n. This implies that

$$\lim_{n \to +\infty} \sup_{1 \leq j \leq n} |v_{j,n}| = \lim_{n \to +\infty} \frac{\sqrt{2} \rho_n(u)}{\sqrt{n}} = 0.$$

This proves that Condition 1 of Lemma A.2 is fulfilled. Let us now study the asymptotic of

$$\sum_{j=1}^{n} v_{j,n}^2 = \frac{2}{n} \sum_{j=1}^{n} \left(\lambda_{j,n}^{u} \right)^2.$$

By definition of the $\lambda_{j,n}^{u}$, we get

$$\begin{aligned} \sum_{j=1}^{n} v_{j,n}^2 &= \tfrac{1}{2n} \mathrm{tr} \left(\left[T_n(f_{\theta_0})^{-1/2} T_n(g_{\theta_0}^{u}) T_n(f_{\theta_0})^{-1/2} \right]^2 \right) \\ &= \tfrac{1}{2n} \mathrm{tr} \left(\left[T_n(f_{\theta_0})^{-1} T_n(g_{\theta_0}^{u}) \right]^2 \right). \end{aligned}$$

Observe that if $\alpha(\theta_0) > -1$, $f_{\theta,1} = f_{\theta,2} = f_\theta$ and $g_{\theta,1} = g_{\theta,2} = g_\theta^{u}$ satisfy the assumptions of Theorem 3.3 with $\beta = \alpha$ on $\Theta^* = B(\theta_0, r)$ for r chosen small enough.

Otherwise, for r small enough, f_θ and $g_{\theta,1} = g_{\theta,2} = g_\theta^u$ satisfy the assumptions of Theorem 3.5 on $\Theta^* = B(\theta_0, r)$. Hence, applying one of these theorems, we get

$$\lim_{n \to +\infty} \sum_{j=1}^n v_{j,n}^2 = \frac{1}{4\pi} \int_{-\pi}^{\pi} \frac{g_{\theta_0}^u(x)^2}{f_{\theta_0}(x)^2} dx.$$

that is by definition of $g_{\theta_0}^u$ and $I(\theta_0)$,

$$\lim_{n \to +\infty} \sum_{j=1}^n v_{j,n}^2 = u^* I(\theta_0) u.$$

Hence, by Lemma A.2, under $P_{\theta_0}^n$, $\langle u, Z_n \rangle$ converges in distribution, to a centered Gaussian random variable whose variance is $u^* I(\theta_0) u$. In other words, under $P_{\theta_0}^n$, $\langle u, Z_n \rangle$ converges in distribution to $\langle u, G \rangle$ with G a centered Gaussian random vector whose covariance matrix is $I(\theta_0)$. Since this holds for any $u \in \mathbb{R}^m$, under $P_{\theta_0}^n$, Z_n converges in distribution to G. The proof of Lemma 2.5 is then complete.

A.2. *Proof of Lemma 2.6.* Let us consider two integers $1 \le j, k \le m$. We recall that $\partial_k^\ell f_\theta$ is defined by (A.1) and set

$$A_{n,\theta}(g) = T_n(f_\theta)^{-1} T_n(g). \tag{A.4}$$

Then, since $\frac{\partial F_n}{\partial \theta_k}$ is given by (A.2), for any $\theta \in B(\theta_0, r)$,

$$\frac{\partial^2 F_n}{\partial \theta_j \partial \theta_k}(\theta) = G_{n,1}(\theta) + G_{n,2}(\theta) + G_{n,3}(\theta)$$

where

$$G_{n,1}(\theta) = \pi \langle \mathbf{x}_n, A_{n,\theta}(\partial_{j,k}^2 f_\theta) T_n(f_\theta)^{-1} \mathbf{x}_n \rangle - \frac{1}{2} \mathrm{tr}\big(A_{n,\theta}(\partial_{j,k}^2 f_\theta)\big),$$

$$G_{n,2}(\theta) = -\pi \langle \mathbf{x}_n, [A_{n,\theta}(\partial_j f_\theta) A_{n,\theta}(\partial_k f_\theta) + A_{n,\theta}(\partial_k f_\theta) A_{n,\theta}(\partial_j f_\theta)] T_n(f_\theta)^{-1} x_n \rangle,$$

and

$$G_{n,3}(\theta) = \frac{1}{2} \mathrm{tr}\big(T_n(\partial_k f_\theta) T_n(f_\theta)^{-1} T_n(\partial_j f_\theta) T_n(f_\theta)^{-1}\big).$$

By Lemma A.1, under $P_{\theta_0}^n$, $G_{n,1}(\theta_0)$ is a centered square integrable random variable and

$$\mathbf{E}_{P_{\theta_0}^n}\big(G_{n,1}^2(\theta_0)\big) = \mathrm{Var}_{P_{\theta_0}^n} G_{n,1}(\theta_0) = \frac{1}{2} \mathrm{tr}\Big(A_{n,\theta_0}\big(\partial_{j,k}^2 f_{\theta_0}\big)^2\Big).$$

By Assumption 2.1, if $\alpha(\theta_0) > -1$, $f_{\theta,1} = f_{\theta,2} = f_\theta$ and $g_{\theta,1} = g_{\theta,2} = \partial_{j,k}^2 f_\theta$ satisfy the assumptions of Theorem 3.3 (up to a proper choice of a smaller r) with $\beta = \alpha$. Moreover if $\alpha(\theta_0) \le -1$, by Assumptions 2.1 and 2.2, f_θ and $g_{\theta,1} = g_{\theta,2} = \partial_{j,k}^2 f_\theta$ satisfy the assumptions of Theorem 3.5. Hence, by definition of A_{n,θ_0},

$$\lim_{n \to +\infty} \frac{1}{n^2} \mathbf{E}_{P_{\theta_0}^n}\big(G_{n,1}^2(\theta_0)\big) = 0,$$

which implies that $G_{n,1}(\theta_0)/n$ converges $P_{\theta_0}^n$-almost surely to 0.

Moreover, again by Lemma A.1, under $P_{\theta_0}^n$, $G_{n,2}(\theta_0)/n$ is a square integrable random variable with mean

$$
\begin{aligned}
m_n &= \frac{1}{n} \mathbf{E}_{P_{\theta_0}^n}\big(G_{n,2}(\theta_0)\big) \\
&= -\frac{1}{2n} \mathrm{tr}(A_{n,\theta_0}(\partial_j f_{\theta_0}) A_{n,\theta_0}(\partial_k f_{\theta_0}) + A_{n,\theta_0}(\partial_k f_{\theta_0}) A_{n,\theta_0}(\partial_j f_{\theta_0})) \\
&= -\frac{1}{n} \mathrm{tr}(A_{n,\theta_0}(\partial_j f_{\theta_0}) A_{n,\theta_0}(\partial_k f_{\theta_0}))
\end{aligned}
$$

and variance

$$\sigma_n^2 = \frac{1}{n^2}\text{tr}\Big([A_{n,\theta_0}(\partial_j f_{\theta_0})A_{n,\theta_0}(\partial_k f_{\theta_0})]^2\Big) + \frac{1}{n^2}\text{tr}\big(A_{n,\theta_0}(\partial_j f_{\theta_0})^2 A_{n,\theta_0}(\partial_k f_{\theta_0})^2\big).$$

As previously, if $\alpha(\theta_0) > -1$ (respectively $\alpha(\theta_0) \leq -1$), we can apply Theorem 3.3 (respectively Theorem 3.5). These theorems prove that

$$\lim_{n\to+\infty} m_n = -\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\partial_k f_{\theta_0}(x)\partial_j f_{\theta_0}(x)}{f_{\theta_0}(x)^2}dx = -2I(\theta_0)_{kj} \quad \text{and} \quad \lim_{n\to+\infty}\sigma_n^2 = 0.$$

This implies that $G_{n,2}(\theta_0)/n$ converges $P_{\theta_0}^n$-almost surely to $-2I(\theta_0)_{kj}$. Since $G_{n,3}(\theta_0)/n = -m_n/2$,

$$\lim_{n\to 0}\frac{1}{n}\frac{\partial^2 F_n}{\partial\theta_k\partial\theta_j}(\theta_0) = -I(\theta_0)_{kj}.$$

Since this holds for any $1 \leq j, k \leq m$, $\nabla^2 F_n(\theta_0)/n$ converges $P_{\theta_0}^n$-almost surely to $-I(\theta_0)$, which concludes the proof of Lemma 2.6.

A.3. *Proof of Lemma 2.7.* Let us focus on $\frac{\partial^3 F_n}{\partial\theta_k^3}$, where $1 \leq k \leq m$. We recall that $A_{n,\theta}$ is defined in (A.4) and set for the sake of simplicity,

$$\partial_k^2 f_\theta = \partial_{k,k}^2 f_\theta \quad \text{and} \quad \partial_k^3 f_\theta = \partial_{k,k,k}^2 f_\theta$$

where $\partial_{(k_1,k_2,k_3)}^\ell f_\theta$ is given by (A.1). Then, for any $\theta \in B(\theta_0, r)$, we get

$$\frac{\partial^3 F_n}{\partial\theta_k^3}(\theta) = H_{n,1}(\theta) + H_{n,2}(\theta) + H_{n,3}(\theta) + H_{n,4}(\theta)$$

where

$$H_{n,1}(\theta) = \pi\langle \mathbf{x}_n, A_{n,\theta}(\partial_k^3 f_\theta)T_n(f_\theta)^{-1}\mathbf{x}_n\rangle,$$

$$H_{n,2}(\theta) = -3\pi\langle \mathbf{x}_n, \big[A_{n,\theta}(\partial_k^2 f_\theta)A_{n,\theta}(\partial_k f_\theta) + A_{n,\theta}(\partial_k f_\theta)A_{n,\theta}(\partial_k^2 f_\theta)\big]T_n(f_\theta)^{-1}x_n\rangle,$$

$$H_{n,3}(\theta) = 6\pi\langle \mathbf{x}_n, A_{n,\theta}(\partial_k f_\theta)^3 T_n(f_\theta)^{-1}x_n\rangle$$

and

$$H_{n,4}(\theta) = -\text{tr}\Big(A_{n,\theta}(\partial_k f_\theta)^3\Big) + \frac{3}{2}\text{tr}\big(A_{n,\theta}(\partial_k f_\theta)A_{n,\theta}(\partial_k^2 f_\theta)\big) - \frac{1}{2}\text{tr}\big(A_{n,\theta}(\partial_k^3 f_\theta)\big)$$

Control of $H_{n,1}$

Observe that since $T_n(f_{\theta_0})$ is an Hermitian matrix,

$$H_{n,1}(\theta) = \pi\langle T_n(f_{\theta_0})^{-1/2}x_n, T_n(f_{\theta_0})^{1/2}A_{n,\theta}(\partial_k^3 f_\theta)T_n(f_\theta)^{-1}x_n\rangle.$$

Therefore, by (2.3)

$$|H_{n,1}(\theta)| \leq \pi\big\|T_n(f_{\theta_0})^{-1/2}x_n\big\|^2\big\|T_n(f_{\theta_0})^{1/2}A_{n,\theta}(\partial_k^3 f_\theta)T_n(f_\theta)^{-1}T_n(f_{\theta_0})^{1/2}\big\|_{2,n}$$

Since $\|\cdot\|_{2,n}$ is a multiplicative norm, that is since for any $n \times n$ matrices A, B,

$$\|AB\|_{2,n} \leq \|A\|_{2,n}\|B\|_{2,n}, \tag{A.5}$$

we get

$$|H_{n,1}(\theta)| \leq$$

$$\pi\big\|T_n(f_{\theta_0})^{-1/2}x_n\big\|^2\big\|T_n(f_\theta)^{-1/2}T_n(f_{\theta_0})^{1/2}\big\|_{2,n}^2\big\|T_n(|\partial_k^3 f_\theta|)^{1/2}T_n(f_\theta)^{-1/2}\big\|_{2,n}^2.$$
$$\tag{A.6}$$

Let us now consider $\varepsilon > 0$. Then, by continuity of α, we can choose r sufficiently small so that

$$\alpha(\theta_0) - \varepsilon \leq \alpha(\theta) \leq \alpha(\theta_0) + \varepsilon$$

for any $\theta \in B(\theta_0, r) = \Theta^* \subset \Theta$. Then, Assumption 2.1 implies that $f = f_\theta$ satisfies (3.1) with $\beta_1 = \alpha(\theta_0) - 2\varepsilon$ and a constant c_1 which does not depend on $\theta \in B(\theta_0, r)$. Note also that $g = f_{\theta_0}$ satisfies (3.1) with $\beta_2 = \alpha(\theta_0) + \varepsilon > \beta_1$. Then, by Lemma 3.1, we get

$$\forall \theta \in B(\theta_0, r), \ \left\| T_n(f_\theta)^{-1/2} T_n(f_{\theta_0})^{1/2} \right\|_{2,n}^2 \leq K n^{3\varepsilon}$$

where the finite constant $K = K_{\theta_0, r, \varepsilon}$ does not depend on n and θ.

Moreover, $g = \left| \partial_k^3 f_\theta \right|$ is a nonnegative symmetric function which satisfies (3.1) with $\beta_2 = \alpha(\theta_0) + 2\varepsilon > \beta_1$ and a constant c_2 which does not depend on $\theta \in B(\theta_0, r)$. Therefore by Lemma 3.1,

$$\forall \theta \in B(\theta_0, r), \ |H_{n,1}(\theta)| \leq K' n^{7\varepsilon} \left\| T_n(f_{\theta_0})^{-1/2} x_n \right\|^2$$

where the finite constant $K' = K'_{\theta_0, r, \varepsilon}$ does not depend on n and θ. Therefore,

$$\frac{1}{n^{3/2}} \sup_{\theta \in B(\theta_0, r)} |H_{n,1}(\theta)| \leq K' n^{7\varepsilon - 1/2} \frac{\left\| T_n(f_{\theta_0})^{-1/2} x_n \right\|^2}{n}.$$

Under $P_{\theta_0}^n$, $T_n(f_{\theta_0})^{-1/2} x_n$ is a centered Gaussian random vector with $\frac{1}{2\pi} \mathrm{Id}_n$ as covariance matrix. Then, by the strong law of large numbers,

$$\frac{\left\| T_n(f_{\theta_0})^{-1/2} x_n \right\|^2}{n} \xrightarrow[n \to +\infty]{P_{\theta_0}^n\text{-a.s.}} \frac{1}{\pi}.$$

Therefore, choosing ε and r small enough, we get

$$\frac{1}{n^{3/2}} \sup_{\theta \in B(\theta_0, r)} |H_{n,1}(\theta)| \xrightarrow[n \to +\infty]{P_{\theta_0}^n\text{-a.s.}} 0. \tag{A.7}$$

Control of $H_{n,2}$ and $H_{n,3}$

Proceeding as for $H_{n,1}$, one check that for r small enough, for $\ell \in \{2, 3\}$,

$$\frac{1}{n^{3/2}} \sup_{\theta \in B(\theta_0, r)} |H_{n,\ell}(\theta)| \xrightarrow[n \to +\infty]{P_{\theta_0}^n\text{-a.s.}} 0. \tag{A.8}$$

Control of $H_{n,4}$

Assume first that $\alpha(\theta_0) > -1$. For $1 \leq \ell \leq 3$, consider $f_{\theta,\ell} = f_\theta$ and $g_{\theta,\ell} = \partial_k^3 f_\theta$. Then, by Assumptions 2.1, these functions satisfy assumptions of Theorem 3.3 on the compact set $\Theta^* = B(\theta_0, r)$ (choosing r small enough) with $\beta = \alpha$. Then, applying this theorem, we get that

$$\sup_n \sup_{\theta \in B(\theta_0, r)} \frac{1}{n} \left| \mathrm{tr}\left(\prod_{\ell=1}^3 T_n(f_\theta)^{-1} T_n(\partial_k^3 f_\theta) \right) \right| < +\infty,$$

that is

$$\sup_n \sup_{\theta \in B(\theta_0, r)} \frac{1}{n} \left| \mathrm{tr}\left(A_{n,\theta}(\partial_k^3 f_\theta) \right) \right| < +\infty$$

Assumption 2.1 also allows us to control the two other terms of $H_{n,4}$ by applying Theorem 3.3. This leads to

$$\sup_n \frac{1}{n} \sup_{\theta \in B(\theta_0, r)} \frac{1}{n} |H_{n,4}(\theta)| < +\infty. \tag{A.9}$$

If $\alpha(\theta_0) \leq -1$, applying Theorem 3.5 instead of Theorem 3.3, we see that Equation (A.9) still holds.

Control of $\frac{\partial^3 F_n}{\partial \theta_k^3}$

Equations (A.7), (A.8) and (A.9) leads to

$$\frac{1}{n^{3/2}} \sup_{\theta \in B(\theta_0, r)} \left| \frac{\partial^3 F_n}{\partial \theta_k^3} \right| \xrightarrow[n \to +\infty]{P_{\theta_0}^n \text{-a.s}} 0$$

for r small enough.

Computing $\frac{\partial^3 F_n}{\partial \theta_j \partial \theta_k \partial \theta_l}$ and then using the same arguments as for $j = k = l$, one obtains that

$$\frac{1}{n^{3/2}} \sup_{\theta \in B(\theta_0, r)} \left| \frac{\partial^3 F_n}{\partial \theta_j \partial \theta_k \partial \theta_l} \right| \xrightarrow[n \to +\infty]{P_{\theta_0}^n \text{-a.s}} 0,$$

which concludes the proof.

References

J. Beran. *Statistics for long-memory processes*, volume 61 of *Monographs on Statistics and Applied Probability*. Chapman and Hall, New York (1994). ISBN 0-412-04901-5. MR1304490.

P. Billingsley. *Probability and measure*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, third edition (1995). ISBN 0-471-00710-2. A Wiley-Interscience Publication. MR1324786.

P. Bondon and W. Palma. A class of antipersistent processes. *J. Time Ser. Anal.* **28** (2), 261–273 (2007). MR2345660.

R. Dahlhaus. Efficient parameter estimation for self-similar processes. *Ann. Statist.* **17** (4), 1749–1766 (1989). MR1026311.

R. Dahlhaus. Extended correction note for: "efficient parameter estimation for self-similar processes". *Ann. Statist.* **34** (2), 1045–1047 (2006). MR2283403.

M. Falk. Local asymptotic normality in a stationary model for spatial extremes. *J. Multivariate Anal.* **102** (1), 48–60 (2011). MR2729419.

R. Fox and M. S. Taqqu. Central limit theorems for quadratic forms in random variables having long-range dependence. *Probab. Theory Related Fields* **74** (2), 213–240 (1987). MR871252.

F. Gamboa and J.-M. Loubes. Estimation of parameters of a multifractal process. *TEST* **16** (2), 383–407 (2007). MR2393660.

B. Garel and M. Hallin. Local asymptotic normality of multivariate ARMA processes with a linear trend. *Ann. Inst. Statist. Math.* **47** (3), 551–579 (1995). MR1364260.

C. W. J. Granger and R. Joyeux. An introduction to long-memory time series models and fractional differencing. *J. Time Ser. Anal.* **1** (1), 15–29 (1980). MR605572.

F. A. Graybill. *Matrices with applications in statistics.* Wadsworth Statistics/Probability Series. Wadsworth Advanced Books and Software, Belmont, Calif., second edition (1983). ISBN 0-534-98038-4. MR682581.

J. R. M. Hosking. Fractional differencing. *Biometrika* **68** (1), 165–176 (1981). MR614953.

A. N. Kolmogoroff. Wienersche Spiralen und einige andere interessante Kurven im Hilbertschen Raum. *C. R. (Doklady) Acad. Sci. URSS (N.S.)* **26**, 115–118 (1940). MR0003441.

D. Lai, X. Wang and J. Wiorkowski. Local asymptotic normality for multivariate nonlinear AR processes. *Stoch. Stoch. Rep.* **72** (3-4), 277–287 (2002). MR1897918.

L. Le Cam. Locally asymptotically normal families of distributions. Certain approximations to families of distributions and their use in the theory of estimation and testing hypotheses. *Univ. california Publ. Statist.* **3**, 37–98 (1960). MR0126903.

O. Lieberman, R. Rosemarin and J. Rousseau. Asymptotic theory for maximum likelihood estimation of the memory parameter in stationary Gaussian processes (full version) (2011). Available at hal.archives-ouvertes.fr/hal-00641474.

O. Lieberman, R. Rosemarin and J. Rousseau. Asymptotic theory for maximum likelihood estimation of the memory parameter in stationary Gaussian processes. *Econometric Theory* **28** (2), 457–470 (2012). doi: 10.1017/ S0266466611000399.

O. Lieberman, J. Rousseau and D. M. Zucker. Valid asymptotic expansions for the maximum likelihood estimator of the parameter of a stationary, Gaussian, strongly dependent process. *Ann. Statist.* **31** (2), 586–612 (2003). Dedicated to the memory of Herbert E. Robbins. MR1983543.

J.-M.l Loubes and D. Paindaveine. Local asymptotic normality property for lacunar wavelet series multifractal model. *ESAIM Probab. Stat.* **15**, 69–82 (2011). MR2870506.

Benoit B. Mandelbrot and John W. Van Ness. Fractional Brownian motions, fractional noises and applications. *SIAM Rev.* **10**, 422–437 (1968). MR0242239.

D. Pollard. *Convergence of stochastic processes.* Springer Series in Statistics. Springer-Verlag, New York (1984). ISBN 0-387-90990-7. MR762984.

G. Samorodnitsky and M. S. Taqqu. *Stable non-Gaussian random processes.* Stochastic Modeling. Chapman & Hall, New York (1994). ISBN 0-412-05171-0. Stochastic models with infinite variance. MR1280932.

A. W. van der Vaart. *Asymptotic statistics*, volume 3 of *Cambridge Series in Statistical and Probabilistic Mathematics.* Cambridge University Press, Cambridge (1998). ISBN 0-521-49603-9; 0-521-78450-6. MR1652247.

A. Wald. Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Trans. Amer. Math. Soc.* **54**, 426–482 (1943). MR0012401.