Afrika Statistika

ISSN 2316-090X

Afrika Statistika
Vol. 14 (1), 2019, pages 1937 - 1950.
DOI: http://dx.doi.org/10.16929/as/2019.1937.143

Hellinger Distance Estimation of Strongly Dependent Gaussian random fields

Aubin Yao N'dri [1,2,*], Ouagnina Hili[2] and Gueï Cyrille Okou[1]

[1]Laboratoire Sciences et Technologies de l'environnement, Université Jean Lorougnon Guédé, Daloa
[2]Laboratoire de Mathématiques et des Nouvelles Technologies de l'Information, Institut National Polytechnique Félix Houphouët-Boigny de Yamoussoukro

Received on July 30, 2019. Accepted on October 5, 2019

Abstract. The Minimum Hellinger Distance Estimator (MHDE) of density function is investigated for stationary long memory (long-range dependent) random fields observed over a finite set of spatial points. A general result on the consistency of the MHD Estimator is first obtained and then, under some mild assumptions, the asymptotic distribution of this estimator is established.

Résumé. L'estimateur du minimum de distance de Hellinger est étudié pour des champs aléatoires stationnaires à longue mémoire observés sur un ensemble fini de points de l'espace. Un résultat général sur la convergence presque sûre de cet estimateur est d'abord obtenue, puis, sous certaines hypothèses, la distribution asymptotique est établie.

Key words: Asymptotic properties ; Strongly dependence ; Hellinger distance estimation, random field.
AMS 2010 Mathematics Subject Classification Objects : 60G10, 60G15, 60G60, 62F12.

1. Introduction

A random field is a collection of random variables indexed by points of $\mathbb{Z}^d$ where $d \in \mathbb{N}^*$. In the case $d = 1$, a random field reduces to a time series. The case $d \geq 2$ is useful for stochastic phenomena evolving in more than one direction

*Corresponding author: Aubin Yao N'DRI : aubin_ndri@yahoo.com
Ouagnina HILI : o_hili@yahoo.fr
Gueï Cyrille OKOU : okou.guei.cyrille@gmail.com.

and showing spatial interaction. They are important in many applications and are used, for example, to model properties of natural images (biological tissue, stone, sand, grass, moss, foliage); natural phenomena (water surface, clouds), but also in imaging (structure of materials, meteorology, etc). They are also used to model the velocity fields in turbulent flows and permeability coefficients of rocks.

In this paper, we focus on strongly dependent random field which extends the concept of classical strong dependence time series to the spatial domain. The study of strong dependence random fields presents interesting and challenging probabilistic and statistical problems and it is important to many scientific investigations. In fact, the strong dependence phenomenon is more commonly seen in higher dimensional random fields than in one-dimensional time series. For $d \geq 2$, the reader should keep in mind the following two examples given respectively in Lavancier (2005).

Linear random fields : Let $(\xi_{\mathbf{j}})_{\mathbf{j}\in\mathbb{Z}^d}$ be zero-mean random variables with finite second moments. The linear random field $(X_{\mathbf{j}})$ is defined for any $\mathbf{j} \in \mathbb{Z}^d$ by

$$X_{\mathbf{j}} = \sum_{\mathbf{i}\in\mathbb{Z}^d} a_{\mathbf{i}}\xi_{\mathbf{j}-\mathbf{i}}, \text{ where } a_{\mathbf{i}} = \|\mathbf{i}\|^{-\beta}\mathcal{L}(\|\mathbf{i}\|)\widetilde{b}\left(\frac{\mathbf{i}}{\|\mathbf{i}\|}\right), \ \frac{d}{2} < \beta < d, \tag{1}$$

with $\mathcal{L}$ a slowly varying function at infinity and $\widetilde{b}$ a continuous function on the unit sphere in $\mathbb{R}^d$.

The process defined in Equation (1) is a strongly dependent stationary, isotropic random field. The concept of isotropy is that the process covariance function $\gamma(\mathbf{i},\mathbf{j})$ is a function of the Euclidean distance $\|\mathbf{i}-\mathbf{j}\|$.

Indeed, by Lavancier (2005), when $\|\mathbf{i}\| \to +\infty$

$$\gamma(\mathbf{i}) = \|\mathbf{i}\|^{-\alpha}L(\|\mathbf{i}\|)^2\left(b\left(\frac{\mathbf{i}}{\|\mathbf{i}\|}\right)+o(1)\right)$$

where $\alpha = 2\beta - d$ and

$$b(t) = \int_{\mathbb{R}^d}\widetilde{b}\left(\frac{s}{\|s\|}\right)\widetilde{b}\left(\frac{s-t}{\|s-t\|}\right)\|s\|^{-\beta}\|t-s\|^{-\beta}ds.$$

Fractional filtering field : The next random field comes from a fractional filtering tensor product :

$$X_{n_1,n_2} = (1-L_1)^{\alpha_1}(1-L_2)^{\alpha_2}\varepsilon_{n_1,n_2}$$

where L_i is lag operator such that $L_1^j\varepsilon_{n_1,n_2} = \varepsilon_{n_1-j,n_2}$ and $L_2^j\varepsilon_{n_1,n_2} = \varepsilon_{n_1,n_2-j}$, $|\alpha_i| < 1/2$ for $i \in \{1,2\}$ and ε_{n_1,n_2} a white noise.

The expression of X is given by :

$$X_{n_1,n_2} = \sum_{i\geq 0}\sum_{j\geq 0} \psi_{i,1}\psi_{j,2}\varepsilon_{n_1-i,n_2-j}$$

where

$$\psi_{i,1} = \frac{\Gamma(i-\alpha_1)}{\Gamma(-\alpha_1)\Gamma(i+1)} \text{ and } \psi_{j,2} = \frac{\Gamma(j-\alpha_2)}{\Gamma(-\alpha_2)\Gamma(j+1)}.$$

We get

$$\gamma(\mathbf{j}) = \gamma(j_1, j_2) \sim c j_1^{-2\alpha_1-1} j_2^{-2\alpha_2-1}$$

when $j_1 \to +\infty$ and $j_2 \to +\infty$.

To describe the problem more specifically in this paper, we consider $(X_{\mathbf{j}})_{\mathbf{j}\in\mathbb{Z}^d}$ a stationary, long-range dependent gaussian random field with density $f(x,\theta_0)$, where $x \in \mathbb{R}$ and θ_0 is assumed to belong to a compact subset Θ of $\mathbb{R}^p$.
A set of random variables $X_{\mathbf{j}}$, $\mathbf{j} \in \mathbb{Z}^d$ is called a $d-$dimensional stationary Gaussian field, if the random variables $X_{\mathbf{j_1}}, X_{\mathbf{j_2}}, \ldots, X_{\mathbf{j_n}}$ have a joint normal distribution for any $\mathbf{j_1}, \mathbf{j_2}, \ldots, \mathbf{j_n} \in \mathbb{Z}^d$; $E(X_{\mathbf{i}}) = E(X_{\mathbf{j}})$ for any $\mathbf{i}, \mathbf{j} \in \mathbb{Z}^d$ and

$$\begin{aligned}\gamma(\mathbf{j}) &= E(X_{\mathbf{1}}X_{\mathbf{1+j}}) = E(X_{\mathbf{i}}X_{\mathbf{i+j}}) \\ &\sim \|\mathbf{j}\|^{-\alpha}\mathcal{L}(\|\mathbf{j}\|)b\left(\frac{\mathbf{j}}{\|\mathbf{j}\|}\right), \ \|\mathbf{j}\| \to +\infty\end{aligned}$$

where the function $\mathcal{L}$, is a slowly varying function at infinity :

$$\lim_{s\to+\infty} \frac{\mathcal{L}(st)}{\mathcal{L}(s)} = 1, \ t > 0 \ ; \ 0 < \alpha < d$$

and b a continuous function on $S^d = \{x \in \mathbb{R}^d, \ \|x\| = 1\}$.

We assume throughout this paper that $E(X_{\mathbf{1}}) = 0$ and $E(X_{\mathbf{1}}^2) = \sigma^2(\theta_0)$. We also suppose that the long memory parameter satisfies $\alpha = \alpha(\theta_0)$.

The main goal of this paper is to estimate the parameter θ_0 and study its asymptotic properties. This study extends N'dri and Hili (2018) to the random field. We construct an estimator $\hat{\theta}_n$ of the parameter θ_0. The values of this estimator are in the parameter space Θ and minimize the Hellinger distance between $f(.,\theta_0)$ and f_n, the kernel density estimator.

For $d = 1$, the kernel density estimator f_n of f introduced by Rosenblatt (1956) and Parzen (1962) is defined for any positive integer n and any $x \in \mathbb{R}$ by

$$f_n(x) = \frac{1}{nb_n}\sum_{i=1}^{n} K\left(\frac{x-X_i}{b_n}\right), \quad x \in \mathbb{R}.$$

We denote

$$I_n = \left\{ \mathbf{i} = (i_1,\, ...,\, i_d) \in \mathbb{Z}^d, 1 \leq i_l \leq n, l = 1, \ldots, d \right\} \quad \text{for} \quad n \in \mathbb{N}^*.$$

In this case, the density estimator f_n of f is defined for any positive integer n and any x in $\mathbb{R}$ by

$$\begin{aligned} f_n(x) &= \frac{1}{n^d b_n} \sum_{\mathbf{i} \in I_n} K\left(\frac{x - X_{\mathbf{i}}}{b_n} \right) \\ &= \frac{1}{n^d b_n} \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} ... \sum_{i_d=1}^{n} K\left(\frac{x - X_{i_1,\ldots,i_d}}{b_n} \right). \end{aligned}$$

The rest of the paper proceeds as follows. Section 2 presents the assumptions and states our main results (Theorem 1 and Theorem 2). We prove an important result in Theorem 1, the convergence of the kernel density estimator to the true density $f(x, \theta_0)$ in the Hellinger topology. Then, applying this asymptotic property, we show consistency property for the Minimum Hellinger Distance Estimator(MHDE). The asymptotic distribution of this estimator is studied in Theorem 2. The proofs of this Theorems are given in section 3.

The following notations are adopted throughout the paper : $\sim$, $\longrightarrow^P$, $\longrightarrow^{\mathcal{D}}$ denote respectively asymptotically equivalent, convergence in probability and convergence in distribution. For every site $\mathbf{i}$, $\mathbf{j} \in I_n \subset \mathbb{N}^d$, $\|\mathbf{i} - \mathbf{j}\|$ is the Euclidean distance between sites $\mathbf{i}$ and $\mathbf{j}$. The integer m is the Hermite rank of a family defined in the subsection 2.3.

2. Statement of assumptions and results

2.1. Assumptions

Assumptions A

(A1) Suppose that $0 < m\alpha(\theta_0) < d$, $n^d b_n \to +\infty$ and there is a positive real δ such as

$$\lim_{n \to +\infty} n^{\delta - \alpha(\theta_0) \wedge (d - m\alpha(\theta_0))} = 0.$$

(A2) The bandwidth is such that : $b_n := n^{-3\tau}$ for some $\tau > \frac{m\alpha(\theta_0)}{12}$.

(A3) The kernel K is bounded with compact support, such that $\int_{\mathbb{R}} uK(u)du = 0$ and $0 < \int_{\mathbb{R}} u^2 K(u)du < +\infty$.

Assumptions B

(B1) For each $\theta \in \Theta$, the function $x \mapsto f(x, \theta)$ is twice continuously differentiable.

(B2) For each $x \in \mathbb{R}$, the function $\theta \mapsto f(x, \theta)$ is continuous.

(B3) For each $x \in \mathbb{R}$, the function $\theta \mapsto \frac{\partial}{\partial \theta_j} f^{1/2}(x, \theta)$, for $1 \leq j \leq q$ is continuous and for every j, the function $\theta \mapsto \frac{\partial}{\partial \theta_j} f^{1/2}(x, \theta)$ is in $L^2(\mathbb{R}^q)$.

(B4) For each $x \in \mathbb{R}$, the function $\theta \mapsto \frac{\partial^2}{\partial \theta_j \partial \theta_k} f^{1/2}(x, \theta),\ 1 \leq j, k \leq q$ is continuous and for every $j,\ k$, the function $\theta \mapsto \frac{\partial^2}{\partial \theta_j \partial \theta_k} f^{1/2}(x, \theta)$ is in $L^2(\mathbb{R}^q)$.

(B5) For $\theta,\ \theta' \in \Theta,\ \theta \neq \theta'$ implies that $\{x \in \mathbb{R}\ f(x, \theta) \neq f(x, \theta')\}$ is a set of positive Lebesgue measure.

2.2. Minimum Hellinger Distance Estimators (MHDE)

In this subsection, we will briefly discuss minimum Hellinger distance estimation. Let $f(x)$ and $g(x)$ be any two densities ; the Hellinger distance between $f(x)$ and $g(x)$ is defined as the L_2-norm of the difference between square root of density functions, i.e.

$$HD^2(f, g) = \int_{-\infty}^{+\infty} \left[(f(x))^{1/2} - (g(x))^{1/2} \right]^2 dx.$$

Let $X_1,\ X_2,\ \cdots,\ X_n$ be strongly dependent random field with density belonging to a specified parametric family $\{f(., \theta) : \theta \in \Theta\}$. To motivate the MHDE, replace f by $f(., \theta)$ and g by f_n, a non-parametric estimator of the density. Therefore, the Hellinger distance in our question becomes the distance between the true density $f(., \theta_0)$ and the non-parametric density estimator of the X_i's, which can be expressed as follows :

$$HD^2(f(., \theta), f_n) = \int_{-\infty}^{+\infty} \left[(f(x, \theta))^{1/2} - (f_n(x))^{1/2} \right]^2 dx\,. \tag{2}$$

The Minimum Hellinger distance estimator of θ is defined to be the value $\widehat{\theta_n}$ (in the parameter space Θ), if it exists, that minimizes the relation (2), namely

$$\hat{\theta_n} = arg \min_{\theta \in \Theta} HD^2(f(., \theta), f_n).$$

Beran (1977) has shown that the MHDE is more robust than maximum likelihood estimator when data contaminations are present. Furthermore MHDE is known to be asymptotically efficient under a specified parametric family of densities and is minimax robust in a small Hellinger metric neighborhood of the given family (see Beran (1977)).

Let $\mathcal{G}$ denote the class of densities metrized by the L_1 distance. We define the Minimum Hellinger Distance Functional (MHDF) to be the functional $T : \mathcal{G} \to \Theta$ such that

$$T(g) = arg \min_{\theta \in \Theta} HD^2(f(.,\theta), g).$$

By applying Assumption (B2) and (B5), Beran (1977) shows the existence of MHDE for Θ compact and discusses the extension of the result for noncompact Θ. Moreover $T(f(.,\theta))$ may have multiple values, so we shall assume that it stands for any one of those values.

Remark 1. The assumption (B5) is the identifiability assumption on the parametrization.

2.3. Main Results

In this subsection, we study in Theorem 1, the efficiency property of the MHD estimator. For the proof of this Theorem, we use the continuity of the functional T (see Beran (1977)). The study of asymptotic distribution property in Theorem 2 is very important because it is useful in the selection criteria of the estimators. Knowing the asymptotic distribution of the estimator can allow to solve the estimation problem to the interval confidence construction or the hypotheses tests.

Theorem 1 (almost sure convergence). *Let assumptions* (A1)-(A3), (B1), (B2) *and* (B5) *be fulfilled. If* θ_0 *is in the interior of* Θ, *then* $\hat{\theta}_n$ *almost surely converges to* θ_0 *when* $n \to +\infty$.

For the following theorem,
$S(.,\theta) = f^{1/2}(.,\theta)$, $\dot{S}(.,\theta) = \left(\frac{\partial}{\partial\theta_1} f^{1/2}(.,\theta), \cdots, \frac{\partial}{\partial\theta_q} f^{1/2}(.,\theta)\right)^T$, $\dot{S}(.,\theta)^T$ is the transpose of $\dot{S}(.,\theta)$ and

$$\rho(x,\theta) = \left[\int_{-\infty}^{+\infty} \dot{S}(x,\theta)\dot{S}(x,\theta)^T dx\right]^{-1} \dot{S}(x,\theta).$$

Consider I_d the identity function, $I\{.\}$ the indicator function, $F(.)$ the continuous marginal distribution function of the sequence $(X_{\mathbf{i}})_{\mathbf{i}}$ and $F_n(x) = n^{-d} \sum_{\mathbf{i} \in I_n} I\{X_{\mathbf{i}} \leq x\}$, $x \in \mathbb{R}$, the sample distribution function. The integer m is the Hermite rank of the family

$$\{I\{I_d(.) \leq x\} - F(x) : \ x \in \mathbb{R}\} \text{ where } F(x) = \mathbb{P}(X_{\mathbf{i}} \leq x).$$

Let $F_n(x) - F(x) = n^{-d} \sum_{\mathbf{i} \in I_n} B_x(X_{\mathbf{i}})$, where $B_x(.) = I\{I_d(.) \leq x\} - F(x)$.

In $L^2(\phi)$, let's consider the Fourier-Hermite expansion

$$B_x(.) = \sum_{k=m}^{+\infty} J_k(x) \frac{H_k(.)}{k!}$$

where

$$H_k(z) = (-1)^k \exp(z^2/2) \frac{d^k}{dz^k} \left(\exp(-z^2/2)\right), \; z \in \mathbb{R}.$$

is the $k-th$ Hermite polynomial and $J_k(x) = E(H_k(.)B_x(.)), \; k = 0, 1, \ldots.$

Denote by

$$d^2_{m,n} = Var\left(\sum_{\mathbf{i}\in I_n} H_m(X_{\mathbf{i}})\right)$$

,

$$J'_m(x) = \frac{dJ_m(x)}{dx}$$

and

$$Y_{m,n} = \frac{1}{d_{m,n}} \sum_{\mathbf{i}\in I_n} \frac{H_m(X_{\mathbf{i}})}{m!}.$$

The long-range dependence condition for the sequence $(X_{\mathbf{i}})_{\mathbf{i}}$ becomes

$$0 < m\alpha(\theta_0) < d.$$

In addition, we have

$$n^d d^{-1}_{m,n} \sim \sqrt{\frac{(1 - m\alpha(\theta_0))(2 - m\alpha(\theta_0))}{2m!}} \frac{n^{m\alpha(\theta_0)/2}}{\mathcal{L}^{m/2}(n)}.$$

Theorem 2 (asymptotic distribution). *Let assumptions* (A1)-(A3) *and* (B1)-(B5) *be fulfilled.If* θ_0 *lies in the interior of* Θ *and* $\int_{-\infty}^{+\infty} \dot{S}(x,\theta_0)\dot{S}(x,\theta_0)^T dx$ *is a non singular* $(q \times q)$*-matrix, then the limiting distribution of* $n^d d^{-1}_{m,n}(\hat{\theta}_n - \theta_0)$ *is* $(Z_m(1)/m!) \int_{-\infty}^{+\infty} \left(\rho(x,\theta_0)/2f^{1/2}(x,\theta_0)\right) J'_m(x)dx$ *where* $(Z_m(t))_{0\le t\le 1}$ *is the Hermite process.*

3. Proofs

Proof of Theorem 1. First we state the following lemma required in proof.

Lemma *If the assumptions (A1)-(A3), (B1), (B2) and (B5) are satisfied, then* f_n *almost surely converges to* $f(.,\theta_0)$ *in the Hellinger topology.*

Proof. We have $f_n(x) - f(x,\theta_0) = f_n(x) - Ef_n(x) + (Ef_n(x) - f(x,\theta_0))$.
Now, we show

$$f_n(x) - Ef_n(x) = \frac{1}{n^d b_n} \sum_{\mathbf{i} \in I_n} \left(K\left(\frac{x - X_{\mathbf{i}}}{b_n}\right) - EK\left(\frac{x - X_{\mathbf{i}}}{b_n}\right)\right) \to 0 \text{ a.s., } n \to +\infty.$$

For $k \in \mathbb{N}$ let's defined

$$B_{k,n} := \{|f_n(x) - Ef_n(x)| \geq \varepsilon\} \text{ with } n \in [2^k, 2^{k+1}[\,.$$

We have

$$\begin{aligned}
B_k &= \bigcup_{2^k \leq n < 2^{k+1}} B_{k,n} = \bigcup_{2^k \leq n < 2^{k+1}} \{|f_n(x) - Ef_n(x)| \geq \varepsilon\} \\
&= \bigcup_{2^k \leq n < 2^{k+1}} \left\{ \left| \frac{1}{n^{d-1} b_n} \sum_{\mathbf{i} \in I_n} \left(K\left(\frac{x - X_{\mathbf{i}}}{b_n}\right) - EK\left(\frac{x - X_{\mathbf{i}}}{b_n}\right)\right) \right| \geq n\varepsilon \right\} \\
&\subset \bigcup_{2^k \leq n < 2^{k+1}} \left\{ \left| \frac{1}{n^{d-1} b_n} \sum_{\mathbf{i} \in I_n} \left(K\left(\frac{x - X_{\mathbf{i}}}{b_n}\right) - EK\left(\frac{x - X_{\mathbf{i}}}{b_n}\right)\right) \right| \geq 2^k\varepsilon \right\} \\
&\subset \left\{ \max_{2^k \leq n \leq 2^{k+1}} \left| \frac{1}{n^{d-1} b_n} \sum_{\mathbf{i} \in I_n} \left(K\left(\frac{x - X_{\mathbf{i}}}{b_n}\right) - EK\left(\frac{x - X_{\mathbf{i}}}{b_n}\right)\right) \right| \geq 2^k\varepsilon \right\}
\end{aligned}$$

Thus,

$$\mathbb{P}(B_k) \leq \mathbb{P}\left(\max_{2^k \leq n \leq 2^{k+1}} \left| \frac{1}{n^{d-1} b_n} \sum_{\mathbf{i} \in I_n} \left(K\left(\frac{x - X_{\mathbf{i}}}{b_n}\right) - EK\left(\frac{x - X_{\mathbf{i}}}{b_n}\right)\right) \right| \geq 2^k\varepsilon \right).$$

Fix an x and let,

$$M_k = \max_{2^k \leq n \leq 2^{k+1}} \left| \frac{1}{n^{d-1} b_n} \sum_{\mathbf{i} \in I_n} \left(K\left(\frac{x - X_{\mathbf{i}}}{b_n}\right) - EK\left(\frac{x - X_{\mathbf{i}}}{b_n}\right)\right) \right|.$$

Then, there exists $j' \in [2^k,\, 2^{k+1}]$ such that

$$\begin{aligned}
M_k &= \left| \frac{1}{j'^{d-1} b_{j'}} \sum_{\mathbf{i} \in I_{j'}} \left(K\left(\frac{x - X_{\mathbf{i}}}{b_{j'}}\right) - EK\left(\frac{x - X_{\mathbf{i}}}{b_{j'}}\right)\right) \right| \\
&\leq \frac{1}{2^{k(d-1)} b_{2^k}} \left| \sum_{\mathbf{i} \in I_{2^k}} \left(K\left(\frac{x - X_{\mathbf{i}}}{b_{2^k}}\right) - EK\left(\frac{x - X_{\mathbf{i}}}{b_{2^k}}\right)\right) \right| \\
&+ \frac{1}{2^{k(d-1)} b_{2^k}} \left| \sum_{\mathbf{i} \in I_{2^{k+1}} - I_{2^k}} \left(K\left(\frac{x - X_{\mathbf{i}}}{b_{2^k}}\right) - EK\left(\frac{x - X_{\mathbf{i}}}{b_{2^k}}\right)\right) \right|
\end{aligned}$$

and

$$E\left(M_k^2\right) \le \frac{2}{2^{2k(d-1)}b_{2^k}^2}E\left(\sum_{\mathbf{i}\in I_{2^k}}\left(K\left(\frac{x-X_{\mathbf{i}}}{b_{2^k}}\right)-EK\left(\frac{x-X_{\mathbf{i}}}{b_{2^k}}\right)\right)\right)^2 + \frac{2}{2^{2k(d-1)}b_{2^k}^2}E\left(\sum_{\mathbf{i}\in I_{2^{k+1}}-I_{2^k}}\left(K\left(\frac{x-X_{\mathbf{i}}}{b_{2^k}}\right)-EK\left(\frac{x-X_{\mathbf{i}}}{b_{2^k}}\right)\right)\right)^2.$$

Let $\varepsilon > 0$, by Markov inequality

$$\mathbb{P}\left(M_k \ge 2^k\varepsilon\right) \le \frac{2}{\varepsilon^2 2^{2kd}b_{2^k}^2}Var\sum_{\mathbf{i}\in I_{2^k}}\left(K\left(\frac{x-X_{\mathbf{i}}}{b_{2^k}}\right)-EK\left(\frac{x-X_{\mathbf{i}}}{b_{2^k}}\right)\right) + \frac{2}{\varepsilon^2 2^{2kd}b_{2^k}^2}(Card(I_{2^{k+1}}-I_{2^k}))^2E\left(K\left(\frac{x-X_{\mathbf{i}}}{b_{2^k}}\right)-EK\left(\frac{x-X_{\mathbf{i}}}{b_{2^k}}\right)\right)^2.$$

Let

$$U_i = \frac{1}{b_{2^k}}\left(K\left(\frac{x-X_{\mathbf{i}}}{b_{2^k}}\right)-EK\left(\frac{x-X_{\mathbf{i}}}{b_{2^k}}\right)\right).$$

First,

$$\begin{aligned} Var\sum_{\mathbf{i}\in I_{2^k}}\frac{1}{b_{2^k}}\left(K\left(\frac{x-X_{\mathbf{i}}}{b_{2^k}}\right)-EK\left(\frac{x-X_{\mathbf{i}}}{b_{2^k}}\right)\right) &= \sum_{\mathbf{i},\mathbf{j}\in I_{2^k}}\left(EU_{\mathbf{i}}U_{\mathbf{j}}-E(U_{\mathbf{i}})E(U_{\mathbf{j}})\right)\\ &\le \sum_{\mathbf{i},\mathbf{j}\in I_{2^k}}\int_{\mathbb{R}^2}|u_{\mathbf{i}}u_{\mathbf{j}}|\left(f_{\mathbf{i},\mathbf{j}}(x_{\mathbf{i}},x_{\mathbf{j}})-f(x_{\mathbf{i}})f(x_{\mathbf{j}})\right)dx_{\mathbf{i}}dx_{\mathbf{j}}\\ &\le \sum_{\mathbf{i},\mathbf{j}\in I_{2^k}}\int_{\mathbb{R}^2}|u_{\mathbf{i}}u_{\mathbf{j}}|f(x_{\mathbf{i}})f(x_{\mathbf{j}})\|\gamma(\mathbf{j}-\mathbf{i})\|dx_{\mathbf{i}}dx_{\mathbf{j}}\\ &\le C_1\sum_{i,j\in I_{2^k}}\|\gamma(\mathbf{j}-\mathbf{i})\| \quad \text{(where } C_1 \text{ is a constant)}. \end{aligned}$$

By the relation (6.1.5) in (Lavancier (2005), pp. 127)

$$\sum_{i,j\in I_{2^k}}\|\gamma(\mathbf{j}-\mathbf{i})\| \sim 2^{k(2d-\alpha)}\mathcal{L}(2^k).$$

So

$$\sum_{k=1}^{+\infty}\frac{2\varepsilon^{-2}}{2^{2kd}}Var\sum_{\mathbf{i}\in I_{2^k}}\frac{1}{b_{2^k}}\left(K\left(\frac{x-X_{\mathbf{i}}}{b_{2^k}}\right)-EK\left(\frac{x-X_{\mathbf{i}}}{b_{2^k}}\right)\right) \sim \sum_{k=1}^{+\infty}\frac{1}{2^{k\alpha}} < +\infty.$$

On the other hand,

$$\sum_{k=1}^{+\infty} \frac{(Card(I_{2^{k+1}} - I_{2^k}))^2}{2^{2kd}} E\left(\frac{1}{b_{2^k}} K\left(\frac{x - X_{\mathbf{i}}}{b_{2^k}}\right) - EK\left(\frac{x - X_{\mathbf{i}}}{b_{2^k}}\right)\right)^2 = \sum_{k=1}^{+\infty} \frac{C_2(Card(I_{2^{k+1}} - I_{2^k}))^2}{2^{2kd}}.$$

We get

$$\sum_{k=1}^{+\infty} \frac{C_2(Card(I_{2^{k+1}} - I_{2^k}))^2}{2^{2kd}} = \sum_{k=1}^{+\infty} \frac{d^2 C_2}{2^{2k(d-1)}} < +\infty.$$

Borel-Cantelli lemma implies that $M_k \to 0$ a.s., which in turn implies that

$$f_n(x) - E(f_n(x)) \to \ 0 \ \text{ a.s.}$$

On the other hand, $Ef_n(x) - f(x, \theta_0) = \int_{\mathbb{R}} K(u)\left(f(x - b_n u, \theta_0) - f(x, \theta_0)\right) du$ and for each $x \in \mathbb{R}$, $K(u)\left|f(x - b_n u, \theta_0) - f(x, \theta_0)\right| \leq cK(u)$.

By the continuity of the density and by the dominated convergence theorem, we conclude that $Ef_n(x) - f(x, \theta_0) \to 0$ as $n \to +\infty$ for each $x \in \mathbb{R}$.
Then, for all $x \in \mathbb{R}$, $f_n(x)$ almost surely (a.s.) converges to $f(x, \theta_0)$ and

$$\mathbb{P}\left(\lim_{n\to+\infty} f_n^{1/2}(x) = f^{1/2}(x, \theta_0), \ \forall x \in \mathbb{R}\right) = 1.$$

Furthermore, since $\int_{\mathbb{R}} f_n(x)dx = \int_{\mathbb{R}} f(x, \theta_0)dx = 1$, hence

$$\lim_{n\to+\infty} \left(\int_{\mathbb{R}} |f_n^{1/2}(x) - f^{1/2}(x, \theta_0)|^2 dx\right)^{1/2} = 0 \ \ a.s.$$

Therefore $f_n \to f$ a.s. when $n \to +\infty$ in the Hellinger topology. From the continuity of the functional T (see Theorem 1 in Beran (1977)), we deduce that $\hat{\theta}_n = T(f_n) \to T(f(., \theta)) = \theta_0$ a.s. as $n \to +\infty$.
This completes the proof. □

Proof of Theorem 2. We follow the proof of Theorem 2 in N'dri and Hili (2018). Using Theorem 3 in Beran (1977), we obtain :

$$\begin{aligned} \hat{\theta}_n - \theta_0 &= \int_{-\infty}^{+\infty} \rho(x, \theta_0)(f_n^{1/2}(x) - f^{1/2}(x, \theta_0))\, dx \\ &+ \ \mathcal{V}_n \int_{-\infty}^{+\infty} \dot{S}(x, \theta_0)(f_n^{1/2}(x) - f^{1/2}(x, \theta_0))\, dx, \end{aligned}$$

where $\mathcal{V}_n$ is a real $p \times p$ matrix which tends to zero as $n \to +\infty$.

Hence it suffices to prove that the limiting distribution of

$$\frac{n^d}{d_{m,n}}\int_{-\infty}^{+\infty}\rho(x,\theta_0)(f_n^{1/2}(x)-f^{1/2}(x,\theta_0))\,dx$$

is

$$\frac{Z_m(1)}{m!}\int_{-\infty}^{+\infty}\frac{\rho(x,\theta_0)}{2f^{1/2}(x,\theta_0)}J'_m(x)dx.$$

We have,

$$\begin{aligned}&\frac{n^d}{d_{m,n}}\int_{-\infty}^{+\infty}\rho(x,\theta_0)(f_n^{1/2}(x)-f^{1/2}(x,\theta_0))\,dx \\ &=\frac{n^d}{d_{m,n}}\int_{-\infty}^{+\infty}\frac{\rho(x,\theta_0)}{2f^{1/2}(x,\theta_0)}(f_n(x)-f(x,\theta_0))\,dx+\ \mathcal{R}_n(\theta_0)\end{aligned} \tag{3}$$

where

$$\begin{aligned}|\ \mathcal{R}_n(\theta_0)\ |\ \le\ &\frac{n^d}{d_{m,n}}\int_{-\infty}^{+\infty}\left|\frac{\sigma(x,\theta_0)}{f^{3/2}(x,\theta_0)}\right|(f_n(x)-f(x,\theta_0))^2\,dx \\ \le\ &C_3\left\{\frac{n^d}{d_{m,n}}\int_{-\infty}^{+\infty}|\sigma(x,\theta_0)|\,(f_n(x)-Ef_n(x))^2\,dx\right\} \\ &+C_3\left\{\frac{n^d}{d_{m,n}}\int_{-\infty}^{+\infty}|\sigma(x,\theta_0)|\,(Ef_n(x)-f(x,\theta_0))^2\,dx\right\}\end{aligned}$$

and C_3 a positive constant.

By using the same arguments as in the proof of Theorem 1, we get

$$\begin{aligned}&\frac{n^d}{d_{m,n}}\,|\rho(x,\theta_0)|\,(f_n(x)-Ef_n(x))^2 \\ &\le\ \mathcal{L}^{-m/2}(n)\left(2^{\alpha(\theta_0)\left(\frac{m(k+1)}{2}-2k\right)}+d^2 2^{\left(\frac{(k+1)m\alpha(\theta_0)}{2}-4k(d-1)\right)}\right)|\rho(x,\theta_0)|\to\ 0\end{aligned}$$

when $k\to+\infty$ and for, $m\in\{1,2,3\}$.

By using Assumption (B3), (B4) and the dominated convergence Theorem, we obtain

$$\frac{n^d}{d_{m,n}}\int_{-\infty}^{+\infty}|\rho(x,\theta_0)|\,(f_n(x)-Ef_n(x))^2dx\to 0\ \ a.s.,\ \ n\to+\infty.$$

By assumptions (A3) and (B1), Taylor's formula in one variable gives for x such as $|u^*-x|<|b_n u|$,

$$
\begin{aligned}
Ef_n(x) - f(x,\theta_0) &= \int_{-\infty}^{+\infty} K(u)(f(x-b_nu,\theta_0) - f(x,\theta_0))du \\
&= b_n f(x,\theta_0)\int_{-\infty}^{+\infty} uK(u)du \\
&+ \frac{b_n^2}{2} f^{(2)}(u^*,\theta_0)\int_{-\infty}^{+\infty} u^2K(u)du.
\end{aligned}
$$

We get,

$$
|Ef_n(x) - f(x,\theta_0)| \leq \frac{b_n^2}{2}\sup_x |f^{(2)}(u^*,\theta_0)| \int_{-\infty}^{+\infty} |u|^2K(u)du < +\infty.
$$

By using assumption (A2) we deduce,

$$
\lim_{n\to+\infty} n^{m\alpha(\theta_0)/2}(Ef_n(x) - f(x,\theta_0))^2 = \lim_{n\to+\infty} n^{m\alpha(\theta_0)/2} b_n^4 \zeta^2(\theta_0) = 0
$$

where

$$
\zeta^2(\theta_0) = \sup_x \left(\frac{|f^{(2)}(u^*,\theta_0)|}{2}\right)\int_{-\infty}^{+\infty} |u|^2K(u)du < +\infty.
$$

Moreover, by combining the relation (3), Assumptions (B3) and (B4), we obtain

$$
\frac{n^d}{d_{m,n}}\int_{-\infty}^{+\infty} |\rho(x,\theta_0)|\,(Ef_n(x) - f(x,\theta_0))^2dx \to 0 \;\; as \;\; n\to+\infty.
$$

Consequently, $\mathcal{R}_n(\theta_0) \to 0$ when $n \to +\infty$.

We are now studying the first term on the right hand side of the relation (3). Using assumptions (A1)-(A3) and (B3)-(B4), we show that

$$
\frac{n^d}{d_{m,n}}\int_{-\infty}^{+\infty} |\rho(x,\theta_0)|\,(Ef_n(x) - f(x,\theta_0))dx \to 0 \;\; as \;\; n\to+\infty.
$$

As in Csörgő and Mielniczuk (1995), we get

$$
\begin{aligned}
D_{m,n}(x) &= \frac{n^d}{d_{m,n}}\left(f_n(x) - Ef_n(x)\right) \\
&= \frac{1}{b_n}\int_{-\infty}^{+\infty} K\left(\frac{x-y}{b_n}\right) d\left(\frac{1}{d_{m,n}}\sum_{\mathbf{i}\in I_n}\left[I_{\{G(X_{\mathbf{i}})\leq y\}} - F(y) - \frac{J_m(y)H_m(X_{\mathbf{i}})}{m!}\right]\right) \qquad (4)\\
&+ \frac{1}{d_{m,n}}\sum_{\mathbf{i}\in I_n}\frac{H_m(X_{\mathbf{i}})}{m!} \times \frac{1}{b_n}\int_{-\infty}^{+\infty} K\left(\frac{x-y}{b_n}\right) J'_m(y)dy.
\end{aligned}
$$

By using integration by parts in the relation (4) and the relation (6.1.7) in Lavancier (2005) :

$$\sup_{y\in\mathbb{R}}\left|\frac{1}{d_{m,n}}\sum_{\mathbf{i}\in I_n}\left[I_{\{G(X_{\mathbf{i}})\le y\}}-F(y)-\frac{J_m(y)}{m!}H_m(X_{\mathbf{i}})\right]\right|\to^P 0,$$

we show that

$$\frac{n^d}{d_{m,n}}\int_{-\infty}^{+\infty}\frac{\rho(x,\theta_0)}{2f^{1/2}(x,\theta_0)}(f_n(x)-Ef_n(x))\,dx$$

and

$$\frac{1}{d_{m,n}}\sum_{\mathbf{i}\in I_n}\frac{H_m(X_{\mathbf{i}})}{m!}\times\int_{-\infty}^{+\infty}\frac{\rho(x,\theta_0)}{2f^{1/2}(x,\theta_0)}J'_m(x)dx$$

have the same distribution.

Then, using Theorem 1' in Dobrushin and Major (1979), we conclude that

$$Y_{mn}\int_{-\infty}^{+\infty}\frac{\rho(x,\theta_0)}{2f^{1/2}(x,\theta_0)}J'_m(x)\,dx\ \to^{\mathcal{D}}\ \frac{Z_m(1)}{m!}\int_{-\infty}^{+\infty}\frac{\rho(x,\theta_0)}{2f^{1/2}(x,\theta_0)}J'_m(x)\,dx$$

where $Y_{mn}=\frac{1}{d_{m,n}}\sum_{\mathbf{i}\in I_n}\frac{H_m(X_{\mathbf{i}})}{m!}$.

Therefore

$$\frac{n^d}{d_{m,n}}(\hat{\theta}_n-\theta_0)\ \to^{\mathcal{D}}\ (Z_m(1)/m!)\int_{-\infty}^{+\infty}\frac{\rho(x,\theta_0)}{2f^{1/2}(x,\theta_0)}J'_m(x)dx.$$

The proof of Theorem 2 is complete.

References

Beran R. (1977) Minimum Hellinger distance estimates for parametric models. *Ann. Statist.*, **5**, 3, 445-463.

Casti J. L. (1985) Nonlinear system theory. *Math. Sc. En. Academic Press Inc.*, Orlando, FL (175).

Csörgő S. and Mielniczuk J. (1995) Density estimation under long-range dependence. *Ann. Statist.* **23**, 3, 990-999.

Dobrushin R. L. and Major P. (1979) Non-central limit theorems for non-linear functionals of Gaussian fields. *Z. Wahr. Verw. Gebiete.* **50**, 27-52.

Lavancier F. (2005). *Les champs aléatoires à longue mémoire*. USTL.

N'dri Y. A. and Hili 0. (2018) Minimum Hellinger distance estimates of long-range dependence gaussian processes. *Journal of Mathematical Sciences : Advances and Applications*. **49**, 59-79.

Parzen E. (1962) On the estimation of a probability density and the mode. *Ann. Math. Statist.* **33**, 1965-1976.

Rosenblatt M. (1956) A central limit theorem and a strong mixing condition. *Proc. Nat. Acad. Sci. USA,* **42**, 43-47.

Rugh W. J. (1981) *Nonlinear system theory.* Johns Hopkins Series in Information Sciences and Systems. Johns Hopkins University Press, Baltimore, Md.