Asymptotic normality of kernel estimates in a regression model for random fields

Mohamed EL MACHKOURI, Radu STOICA

September 25, 2008

Abstract

We establish the asymptotic normality of the regression estimator in a fixed-design setting when the errors are given by a field of dependent random variables. The result applies to martingale-difference or strongly mixing random fields. On this basis, a statistical test that can be applied to image analysis is also presented.

AMS Subject Classifications (2000): 60G60, 60F05, 62G08
Key words and phrases: Nonparametric regression estimation, asymptotic normality, kernel estimator, strongly mixing random field.
Short title: Asymptotic normality of kernel estimates.

1 Introduction and notations

Our aim in this paper is to establish the asymptotic normality of a regression estimator in a fixed-design setting when the errors are given by a stationary field of random variables which show spatial interaction. Let $\mathbb{Z}^d$, $d \geq 1$ denote the integer lattice points in the d-dimensional Euclidean space. By a stationary random field we mean any family $(\varepsilon_k)_{k\in\mathbb{Z}^d}$ of real-valued random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that for any $(k, n) \in \mathbb{Z}^d \times \mathbb{N}^*$ and any $(i_1, ..., i_n) \in (\mathbb{Z}^d)^n$, the random vectors $(\varepsilon_{i_1}, ..., \varepsilon_{i_n})$ and $(\varepsilon_{i_1+k}, ..., \varepsilon_{i_n+k})$ have the same law. The regression model which we are interested in is

$$Y_i = g(i/n) + \varepsilon_i, \quad i \in \Lambda_n = \{1, ..., n\}^d \tag{1}$$

where g is an unknown smooth function and $(\varepsilon_i)_{i\in\mathbb{Z}^d}$ is a zero mean and square-integrable stationary random field. Let K be a probability kernel defined on $\mathbb{R}^d$ and $(h_n)_{n\geq 1}$ a sequence of positive numbers which converges to zero and which satisfies $(nh_n)_{n\geq 1}$ goes to infinity. We estimate the function g by the kernel-type estimator g_n defined for any x in $[0,1]^d$ by

$$g_n(x) = \frac{\displaystyle\sum_{i\in\Lambda_n} Y_i\, K\left(\frac{x-i/n}{h_n}\right)}{\displaystyle\sum_{i\in\Lambda_n} K\left(\frac{x-i/n}{h_n}\right)}. \tag{2}$$

In a previous paper, El Machkouri [10] obtained strong convergence of the estimator $g_n(x)$ with optimal rate. However, most of existing theoretical nonparametric results for dependent random variables pertain to time series (see Bosq [4]) and relatively few generalisations to the spatial domain are available. Key references on this topic are Biau [2], Carbon et al. [5], Carbon et al. [6], Hallin et al. [12], [13], Tran [26], Tran and Yakowitz [27] and Yao [29] who have investigated nonparametric density estimation for random fields and Altman [1], Biau and Cadre [3], Hallin et al. [14] and Lu and Chen [17], [18] who have studied spatial prediction and spatial regression estimation.

Let μ be the law of the stationary real random field $(\varepsilon_k)_{k\in\mathbb{Z}^d}$ and consider the projection f from $\mathbb{R}^{\mathbb{Z}^d}$ to $\mathbb{R}$ defined by $f(\omega)=\omega_0$ and the family of translation operators $(T^k)_{k\in\mathbb{Z}^d}$ from $\mathbb{R}^{\mathbb{Z}^d}$ to $\mathbb{R}^{\mathbb{Z}^d}$ defined by $(T^k(\omega))_i=\omega_{i+k}$ for any $k\in\mathbb{Z}^d$ and any ω in $\mathbb{R}^{\mathbb{Z}^d}$. Denote by $\mathcal{B}$ the Borel σ-algebra of $\mathbb{R}$. The random field $(f\circ T^k)_{k\in\mathbb{Z}^d}$ defined on the probability space $(\mathbb{R}^{\mathbb{Z}^d},\mathcal{B}^{\mathbb{Z}^d},\mu)$ is stationary with the same law as $(\varepsilon_k)_{k\in\mathbb{Z}^d}$, hence, without loss of generality, one can suppose that $(\Omega,\mathcal{F},\mathbb{P})=(\mathbb{R}^{\mathbb{Z}^d},\mathcal{B}^{\mathbb{Z}^d},\mu)$ and $\varepsilon_k=f\circ T^k$. An element A of $\mathcal{F}$ is said to be invariant if $T^k(A)=A$ for any $k\in\mathbb{Z}^d$. We denote by $\mathcal{I}$ the σ-algebra of all measurable invariant sets. On the lattice $\mathbb{Z}^d$ we define the lexicographic order as follows: if $i=(i_1,...,i_d)$ and $j=(j_1,...,j_d)$ are distinct elements of $\mathbb{Z}^d$, the notation $i<_{lex} j$ means that either $i_1<j_1$ or for some p in $\{2,3,...,d\}$, $i_p<j_p$ and $i_q=j_q$ for $1\leq q<p$. Let the sets $\{V_i^k\,;\,i\in\mathbb{Z}^d\,,\,k\in\mathbb{N}^*\}$ be defined as follows:

$$V_i^1=\{j\in\mathbb{Z}^d\,;\,j<_{lex} i\},$$

and for $k\geq 2$

$$V_i^k=V_i^1\cap\{j\in\mathbb{Z}^d\,;\,|i-j|\geq k\}\quad\text{where}\quad |i-j|=\max_{1\leq l\leq d}|i_l-j_l|.$$

For any subset Γ of $\mathbb{Z}^d$ define $\mathcal{F}_\Gamma=\sigma(\varepsilon_i\,;\,i\in\Gamma)$ and set

$$E_{|k|}(\varepsilon_i)=E(\varepsilon_i|\mathcal{F}_{V_i^{|k|}}),\quad k\in V_i^1.$$

Note that Dedecker [8] established the central limit theorem for any stationary square-integrable random field $(\varepsilon_k)_{k\in\mathbb{Z}^d}$ which satisfies the condition

$$\sum_{k\in V_0^1} \|\varepsilon_k E_{|k|}(\varepsilon_0)\|_1 < \infty. \tag{3}$$

A real random field $(X_k)_{k\in\mathbb{Z}^d}$ is said to be a martingale-difference random field if for any m in $\mathbb{Z}^d$, $E(\, X_m \,|\, \sigma(\, X_k \,;\, k <_{lex} m\,)\,) = 0$ a.s. The condition (3) is satisfied by martingale-difference random fields. Nahapetian and Petrosian [21] defined a large class of Gibbs random fields $(\xi_k)_{k\in\mathbb{Z}^d}$ satisfying the stronger martingale-difference property: $E(\, \xi_m \,|\, \sigma(\, \xi_k \,;\, k \neq m\,)\,) = 0$ a.s. for any m in $\mathbb{Z}^d$. Moreover, for these models, phase transition may occur (see [19],[20]).

Given two sub-σ-algebras $\mathcal{U}$ and $\mathcal{V}$, different measures of their dependence have been considered in the literature. We are interested by one of them. The strong mixing (or α-mixing) coefficient has been introduced by Rosenblatt [25] and is defined by

$$\alpha(\mathcal{U},\mathcal{V}) = \sup\{|\mathbb{P}(U\cap V) - \mathbb{P}(U)\mathbb{P}(V)|,\, U\in\mathcal{U},\, V\in\mathcal{V}\}.$$

Denote by $\sharp\Gamma$ the cardinality of any subset Γ of $\mathbb{Z}^d$. In the sequel, we shall use the following non-uniform mixing coefficients defined for any (k,l,n) in $(\mathbb{N}^*\cup\{\infty\})^2\times\mathbb{N}$ by

$$\alpha_{k,l}(n) = \sup\{\alpha(\mathcal{F}_{\Gamma_1},\mathcal{F}_{\Gamma_2}),\, \sharp\Gamma_1\leq k,\, \sharp\Gamma_2\leq l,\, \rho(\Gamma_1,\Gamma_2)\geq n\},$$

where the distance ρ is defined by $\rho(\Gamma_1,\Gamma_2) = \min\{|i-j|,\, i\in\Gamma_1,\, j\in\Gamma_2\}$. We say that the random field $(\varepsilon_k)_{k\in\mathbb{Z}^d}$ is strongly mixing (or α-mixing) if there exists a pair (k,l) in $(\mathbb{N}^*\cup\{\infty\})^2$ such that $\lim_{n\to\infty}\alpha_{k,l}(n) = 0$.
The condition (3) is satisfied by strongly mixing random fields. For example, one can construct stationary Gaussian random fields with a sufficiently large polynomial decay of correlation such that (5) holds ([9], p. 59, Corollary 2).

2 Main results

First, we recall the concept of stability introduced by Rényi [22].

Definition. *Let $(X_n)_{n\geq 0}$ be a sequence of real random variables and let X be defined on some extension of the underlying probability space $(\Omega,\mathcal{A},\mathbb{P})$. Let $\mathcal{U}$ be a sub-σ-algebra of $\mathcal{A}$. Then $(X_n)_{n\geq 0}$ is said to converge $\mathcal{U}$-stably to X if for any continuous bounded function φ and any bounded and $\mathcal{U}$-measurable* *variable Z we have* $\lim_{n\to\infty} E\left(\varphi(X_n)Z\right) = E\left(\varphi(X)Z\right)$.

For any $B > 0$, we denote by $\mathcal{C}^1(B)$ the set of real functions f continuously differentiable on $[0,1]^d$ such that

$$\sup_{x\in[0,1]^d} \max_{\alpha\in\mathcal{M}} |D_\alpha(f)(x)| \leq B,$$

where

$$D_\alpha(f) = \frac{\partial^{\hat{\alpha}} f}{\partial x_1^{\alpha_1} \dots \partial x_d^{\alpha_d}} \quad \text{and} \quad \mathcal{M} = \{\alpha = (\alpha_i)_i \in \mathbb{N}^d\,;\, \hat{\alpha} = \sum_{i=1}^d \alpha_i \leq 1\}.$$

In the sequel we denote $\|x\| = \max_{1\leq k\leq d} |x_k|$ for any $x = (x_1, ..., x_d) \in [0,1]^d$. We make the following assumptions on the regression function g and the probability kernel K:

A1) The probability kernel K fulfils $\int K(u)\,du = 1$ and $\int K^2(u)\,du < \infty$. K is also symmetric, non-negative, supported by $[-1,1]^d$ and satisfies a Lipschitz condition $|K(x)-K(y)| \leq r\|x-y\|$ for any $x, y \in [-1,1]^d$ and some $r > 0$. In addition there exists $c, C > 0$ such that $c \leq K(x) \leq C$ for any $x \in [-1,1]^d$.

A2) There exists $B > 0$ such that g belongs to $\mathcal{C}^1(B)$.

We consider also the notations:

$$\sigma^2 = \int_{\mathbb{R}^d} K^2(u)\,du \quad \text{and} \quad \eta = \sum_{k\in\mathbb{Z}^d} E(\varepsilon_0\varepsilon_k|\mathcal{I}).$$

The following proposition (see [10]) gives the convergence of $Eg_n(x)$ to $g(x)$.

Proposition 1 *Assume that the assumption* **A2)** *holds then*

$$\sup_{x\in[0,1]^d} \sup_{g\in\mathcal{C}^1(B)} |Eg_n(x) - g(x)| = O\left[h_n\right].$$

By proposition 3 in [8], we know that under condition (3), the random variable η belongs to L^1. Our main result is the following.

Main theorem. *If $nh_n^{d+1} \to \infty$ and the condition* (3) *holds then for any $k \in \mathbb{N}^*$ and any distinct points $x_1, ..., x_k$ in $[0,1]^d$, the sequence*

$$(nh_n)^{d/2} \begin{pmatrix} g_n(x_1) - Eg_n(x_1) \\ \vdots \\ g_n(x_k) - Eg_n(x_k) \end{pmatrix} \xrightarrow[n\to+\infty]{\mathcal{L}} \sigma\sqrt{\eta} \begin{pmatrix} \tau^{(1)} \\ \vdots \\ \tau^{(k)} \end{pmatrix} \quad \textit{($\mathcal{I}$-stably)}$$

where $\sigma^2 = \int_{\mathbb{R}^d} K^2(u)\,du$ *and* $(\tau^{(i)})_{1\leq i\leq k} \sim \mathcal{N}(0, \mathbb{I}_k)$ *where* $\mathbb{I}_k$ *is the identity matrix. Moreover,* $(\tau^{(i)})_{1\leq i\leq k}$ *is independent of* $\eta = \sum_{k\in\mathbb{Z}^d} E(\varepsilon_0\varepsilon_k|\mathcal{I})$.

As a consequence of this theorem, we obtain the following result for strongly mixing random fields.

Corollary. *Let us consider the following assumption*

$$\sum_{k\in\mathbb{Z}^d} \int_0^{\alpha_{1,\infty}(|k|)} \mathcal{Q}^2_{\varepsilon_0}(u)\,du < \infty \tag{4}$$

where Q_{ε_0} *denotes the cadlag inverse of the function* $H_{\varepsilon_0} : t \to \mathbb{P}(|\varepsilon_0| > t)$. *Then* (4) *implies* (3) *and also the main theorem.*

Remark. If ε_0 is $(2+\delta)$-integrable for some $\delta > 0$ then the condition

$$\sum_{m=1}^{\infty} m^{d-1} \alpha_{1,\infty}^{\delta/(2+\delta)}(m) < \infty \tag{5}$$

is more restrictive than condition (4).

In order to use the main theorem for establishing confidence intervals, one needs to estimate η. It is done by the following result established in [8].

Proposition 2 *Assume that the condition* (3) *holds. For any* $N \in \mathbb{N}^*$, *set* $G_N = \{(i,j) \in \Lambda_n \times \Lambda_n\,;\, |i-j| \leq N\}$. *Let* ρ_n *be a sequence of positive integers satisfying:*

$$\lim_{n\to+\infty} \rho_n = +\infty \quad \textit{and} \quad \lim_{n\to+\infty} \rho_n^{3d} E(\varepsilon_0^2(1 \wedge n^{-d}\varepsilon_0^2) = 0$$

Then

$$\frac{1}{n^d} \max\left(1, \sum_{(i,j)\in G_{\rho_n}} \varepsilon_i\varepsilon_j\right) \xrightarrow[n\to+\infty]{\mathbb{P}} \eta.$$

3 Proofs

3.1 Proof of the main theorem

Let x in $[0,1]^d$ and $n \geq 1$ be fixed. For any i in Λ_n, denote

$$a_i(x) = K\left(\frac{x - i/n}{h_n}\right) \quad \text{and} \quad b_i(x) = \frac{a_i(x)}{\sqrt{\sum_{j\in\Lambda_n} a_j^2(x)}}.$$

Denote also

$$v_n(x) = \sqrt{\frac{(nh_n)^d}{\sum_{i\in\Lambda_n} a_i(x)}} \times \sqrt{\frac{\sum_{i\in\Lambda_n} a_i^2(x)}{\sum_{i\in\Lambda_n} a_i(x)}}.$$

Without loss of generality, we consider the case $k = 2$ and we refer to x_1 and x_2 as x and y. Let λ_1 and λ_2 be two real numbers such that $\lambda_1^2 + \lambda_2^2 = 1$ and let $x, y \in [0,1]^d$ such that $x \neq y$. One can notice that

$$\frac{(nh_n)^{d/2}}{\sigma} \left[\lambda_1(g_n(x) - Eg_n(x)) + \lambda_2(g_n(y) - Eg_n(y))\right] = \sum_{i\in\Lambda_n} \tilde{s}_i(x,y)\,\varepsilon_i$$

where $\tilde{s}_i(x,y) = (\lambda_1 v_n(x) b_i(x) + \lambda_2 v_n(y) b_i(y))/\sigma$.

Lemma 1 *Let $x, y \in [0,1]^d$ be fixed. If $nh_n^{d+1} \to \infty$ then*

$$\lim_{n\to+\infty} \frac{1}{(nh_n)^d} \sum_{i\in\Lambda_n} a_i(x) a_i(y) = \delta_{xy}\,\sigma^2 \tag{6}$$

and

$$\lim_{n\to+\infty} \frac{1}{(nh_n)^d} \sum_{i\in\Lambda_n} a_i(x) = 1 \tag{7}$$

where δ_{xy} equals 1 if $x = y$ and 0 if $x \neq y$.

Proof of Lemma 1. In the sequel, we denote $\psi(u) = \frac{1}{h_n^d} K\left(\frac{x-u}{h_n}\right) K\left(\frac{y-u}{h_n}\right)$ and $I_n(x,y) = \int_{[0,1]^d} \psi(u)\,du$, we have

$$\begin{aligned} I_n(x,y) &= \sum_{i\in\Lambda_n} \int_{R_{i/n}} \psi(u)\,du \\ &= \sum_{i\in\Lambda_n} \lambda(R_{i/n})\psi(c_i) \text{ with } c_i \in R_{i/n} \\ &= \sum_{i\in\Lambda_n} n^{-d}\psi(c_i) \end{aligned}$$

where $R_{i/n} =](i_1-1)/n, i_1/n] \times \ldots \times](i_d-1)/n, i_d/n]$ and λ is the Lebesgue measure on $\mathbb{R}^d$. Let $\varphi_x(u) = (x-u)/h_n$, for any v in $[0,1]^d$, we have

$$d(K\circ\varphi_x)(u)(v) = \frac{-1}{h_n} \sum_{i=1}^{d} v_i \sum_{j=1}^{d} \frac{\partial K}{\partial u_j}(\varphi_x(u)).$$

Using the assumptions on the kernel K and noting that

$$d\psi(u) = \frac{1}{h_n^d}\left[d(K\circ\varphi_x)(u) \times K(\varphi_y(u)) + d(K\circ\varphi_y)(u) \times K(\varphi_x(u))\right]$$

we derive that there exists $c > 0$ such that $\sup_{u\in[0,1]^d} \|d\psi(u)\| \le c h_n^{-(d+1)}$. So, it follows that

$$
\begin{aligned}
\left| \frac{1}{(nh_n)^d} \sum_{i\in\Lambda_n} a_i(x)a_i(y) - I_n(x,y) \right| &= \left| \sum_{i\in\Lambda_n} n^{-d}(\psi(i/n) - \psi(c_i)) \right| \\
&\le \sup_{u\in[0,1]^d} \|d\psi(u)\| \sum_{i\in\Lambda_n} n^{-d}\|i/n - c_i\|_\infty \\
&\le \frac{c}{h_n^{d+1}} \sum_{i\in\Lambda_n} n^{-(d+1)} \\
&= \frac{c}{nh_n^{d+1}} \xrightarrow[n\to+\infty]{} 0.
\end{aligned}
$$

Moreover,

$$
\begin{aligned}
I_n(x,y) &= \int_{[0,1]^d} \frac{1}{h_n^d} K\left(\frac{x-u}{h_n}\right) K\left(\frac{y-u}{h_n}\right) du \\
&= \int_{\varphi_x([0,1]^d)} K(u)K\left(u + \frac{y-x}{h_n}\right) du.
\end{aligned}
$$

So, by the dominated convergence theorem, we obtain

$$
\lim_{n\to+\infty} I_n(x,y) = \delta_{xy}\,\sigma^2
$$

and consequently (6) holds. The proof of (7) follows the same lines. The proof of Lemma 1 is complete. $\square$

Using Lemma 1 and denoting $\kappa_{xy}^2 = (\lambda_1+\lambda_2)^2\delta_{xy} + 1 - \delta_{xy}$, we derive

$$
\lim_{n\to+\infty} \sum_{i\in\Lambda_n} \tilde{s}_i^2(x,y) = \kappa_{xy}^2 = 1 \quad (\text{since } x \neq y).
$$

So, denoting

$$
s_i(x,y) = \frac{\tilde{s}_i(x,y)}{\sqrt{\sum_{j\in\Lambda_n} \tilde{s}_j^2(x,y)}},
$$

it suffices to prove the convergence $\mathcal{I}$-stably of $\sum_{i\in\Lambda_n} s_i(x,y)\,\varepsilon_i$ to $\sqrt{\eta}\tau_0$ where $\tau_0 \sim \mathcal{N}(0,2)$. In fact, we are going to adapt the proof of the central limit theorem by Dedecker [8].

For any i in $\mathbb{Z}^d$, let us define the tail σ-algebra $\mathcal{F}_{i,-\infty} = \cap_{k\in\mathbb{N}^*}\mathcal{F}_{V_i^k}$ (we are going to note $\mathcal{F}_{-\infty}$ in place of $\mathcal{F}_{0,-\infty}$) and consider the following proposition established in [8].

Proposition *The σ-algebra $\mathcal{I}$ is included in the $\mathbb{P}$-completion of $\mathcal{F}_{-\infty}$.*

Let f be a one to one map from $[1,N]\cap\mathbb{N}^*$ to a finite subset of $\mathbb{Z}^d$ and $(\xi_i)_{i\in\mathbb{Z}^d}$ a real random field. For all integers k in $[1,N]$, we denote

$$S_{f(k)}(\xi)=\sum_{i=1}^{k}\xi_{f(i)} \quad \text{and} \quad S^c_{f(k)}(\xi)=\sum_{i=k}^{N}\xi_{f(i)}$$

with the convention $S_{f(0)}(\xi)=S^c_{f(N+1)}(\xi)=0$. To describe the set $\Lambda_n=\{1,...,n\}^d$, we define the one to one map f_n from $[1,n^d]\cap\mathbb{N}^*$ to Λ_n by: f_n is the unique function such that for $1\leq k<l\leq n^d$, $f(k)<_{lex} f(l)$. From now on, we consider two independent fields $(\tau_i^{(1)})_{i\in\mathbb{Z}^d}$ and $(\tau_i^{(2)})_{i\in\mathbb{Z}^d}$ of i.i.d. random variables independent of $(\varepsilon_i)_{i\in\mathbb{Z}^d}$ and $\mathcal{I}$ such that $\tau_0^{(1)}$ and $\tau_0^{(2)}$ have the standard normal law $\mathcal{N}(0,1)$. We introduce the two sequences of fields $X_i=s_i(x,y)\varepsilon_i$ and $\gamma_i=s_i(x,y)\tau_i\sqrt{\eta}$ where $\tau_i=\tau_i^{(1)}+\tau_i^{(2)}\sim\mathcal{N}(0,2)$. Let h be any function from $\mathbb{R}$ to $\mathbb{R}$. For $0\leq k\leq l\leq n^d+1$, we introduce $h_{k,l}(X)=h(S_{f(k)}(X)+S^c_{f(l)}(\gamma))$. With the above convention we have that $h_{k,n^d+1}(X)=h(S_{f(k)}(X))$ and also $h_{0,l}(X)=h(S^c_{f(l)}(\gamma))$. In the sequel, we will often write $h_{k,l}$ instead of $h_{k,l}(X)$ and s_i instead of $s_i(x,y)$. We denote by $B_1^4(\mathbb{R})$ the unit ball of $C_b^4(\mathbb{R})$: h belongs to $B_1^4(\mathbb{R})$ if and only if it belongs to $C^4(\mathbb{R})$ and satisfies $\max_{0\leq i\leq 4}\|h^{(i)}\|_\infty\leq 1$.

3.1.1 Lindeberg's decomposition

Let Z be a $\mathcal{I}$-measurable random variable bounded by 1. It suffices to prove that for all h in $B_1^4(\mathbb{R})$,

$$\lim_{n\to+\infty} E\left(Zh(S_{f(n^d)}(X))\right)=E\left(Zh\left((\lambda_1\tau_0^{(1)}+\lambda_2\tau_0^{(2)})\sqrt{\eta}\right)\right).$$

We use Lindeberg's decomposition:

$$\begin{aligned} E\left(Z\left[h(S_{f(n^d)}(X))-h\left((\lambda_1\tau_0^{(1)}+\lambda_2\tau_0^{(2)})\sqrt{\eta}\right)\right]\right) &= E\left(Z\left[h(S_{f(n^d)}(X))-h\left(S_{f(n^d)}(\gamma)\right)\right]\right)\\ &= E\left(Z[h_{n^d,n^d+1}-h_{0,1}]\right)\\ &= \sum_{k=1}^{n^d}E\left(Z[h_{k,k+1}-h_{k-1,k}]\right). \end{aligned}$$

Now,

$$h_{k,k+1}-h_{k-1,k}=h_{k,k+1}-h_{k-1,k+1}+h_{k-1,k+1}-h_{k-1,k}.$$

Applying Taylor's formula we get that:

$$h_{k,k+1} - h_{k-1,k+1} = X_{f(k)} h'_{k-1,k+1} + \frac{1}{2} X^2_{f(k)} h''_{k-1,k+1} + R_k$$

and

$$h_{k-1,k+1} - h_{k-1,k} = -\gamma_{f(k)} h'_{k-1,k+1} - \frac{1}{2} \gamma^2_{f(k)} h''_{k-1,k+1} + r_k$$

where $|R_k| \leq X^2_{f(k)}(1 \wedge |X_{f(k)}|)$ and $|r_k| \leq \gamma^2_{f(k)}(1 \wedge |\gamma_{f(k)}|)$. Since $(X, \tau_i)_{i \neq f(k)}$ is independent of $\tau_{f(k)}$, it follows that

$$E\left(Z\gamma_{f(k)} h'_{k-1,k+1}\right) = 0 \quad \text{and} \quad E\left(Z\gamma^2_{f(k)} h''_{k-1,k+1}\right) = E\left(Z s^2_{f(k)} \eta h''_{k-1,k+1}\right)$$

Hence, we obtain

$$\begin{aligned} E\left(Z\left[h(S_n(X)) - h\left((\lambda_1 \tau_0^{(1)} + \lambda_2 \tau_0^{(2)})\sqrt{\eta}\right)\right]\right) &= \sum_{k=1}^{n^d} E(ZX_{f(k)} h'_{k-1,k+1}) \\ &+ \sum_{k=1}^{n^d} E\left(Z\left(X^2_{f(k)} - s^2_{f(k)}\eta\right)\frac{h''_{k-1,k+1}}{2}\right) \\ &+ \sum_{k=1}^{n^d} E\left(R_k + r_k\right). \end{aligned}$$

Arguing as in Rio [24], it is easily proved that

$$\lim_{n \to +\infty} \sum_{k=1}^{n^d} E\left(|R_k| + |r_k|\right) = 0.$$

Let us denote $C_N = [-N, N]^d \cap \mathbb{Z}^d$ for any positive integer N. If we define $\eta_N = \sum_{k \in C_{N-1}} E\left(\varepsilon_0 \varepsilon_k | \mathcal{I}\right)$, the upper bound $E|\eta - \eta_N| \leq 2 \sum_{k \in V_0^N} E|E\left(\varepsilon_0 \varepsilon_k | \mathcal{I}\right)|$ holds. Hence according to condition (3) and the above proposition, we derive $\lim_{N \to +\infty} E|\eta - \eta_N| = 0$ and consequently we have only to show

$$\lim_{N \to +\infty} \limsup_{n \to +\infty} \sum_{k=1}^{n^d} \left(E(ZX_{f(k)} h'_{k-1,k+1}) + E\left(Z\left(X^2_{f(k)} - s^2_{f(k)}\eta_N\right)\frac{h''_{k-1,k+1}}{2}\right)\right) = 0. \tag{8}$$

3.1.2 First reduction

First, we focus on $\sum_{k=1}^{n^d} E\left(ZX_{f(k)} h'_{k-1,k+1}\right)$. For all N in $\mathbb{N}^*$ and all integer k in $[1, n^d]$, we define

$$E_k^N = f([1,k] \cap \mathbb{N}^*) \cap V_{f(k)}^N \quad \text{and} \quad S_{f(k)}^N(X) = \sum_{i \in E_k^N} X_i.$$

For any function Ψ from $\mathbb{R}$ to $\mathbb{R}$, we define $\Psi_{k-1,l}^{N} = \Psi(S_{f(k)}^{N}(X) + S_{f(l)}^{c}(\gamma))$ (we shall apply this notation to the successive derivatives of the function h). Our aim is to show that

$$\lim_{N\to+\infty} \limsup_{n\to+\infty} \sum_{k=1}^{n^d} E\left(Z\left(X_{f(k)}h'_{k-1,k+1} - X_{f(k)}\left(S_{f(k-1)}(X) - S_{f(k)}^{N}(X)\right)h''_{k-1,k+1}\right)\right) = 0. \tag{9}$$

First, we use the decomposition

$$X_{f(k)}h'_{k-1,k+1} = X_{f(k)}h'^{N}_{k-1,k+1} + X_{f(k)}\left(h'_{k-1,k+1} - h'^{N}_{k-1,k+1}\right).$$

We consider a one to one map m from $[1, |E_k^N|] \cap \mathbb{N}^*$ to E_k^N and such that $|m(i) - f(k)| \le |m(i-1) - f(k)|$. This choice of m ensures that $S_{m(i)}(X)$ and $S_{m(i-1)}(X)$ are $\mathcal{F}_{V_{f(k)}^{|m(i)-f(k)|}}$-measurable. The fact that γ is independent of X together with proposition 3 in [8] imply that

$$E\left(ZX_{f(k)}h'\left(S_{f(k+1)}^{c}(\gamma)\right)\right) = E\left(h'\left(S_{f(k+1)}^{c}(\gamma)\right)\right)E\left(ZE\left(X_{f(k)}|\mathcal{F}_{-\infty}\right)\right) = 0.$$

Therefore $|E\left(ZX_{f(k)}h'^{N}_{k-1,k+1}\right)|$ equals

$$\left|\sum_{i=1}^{|E_k^N|} E\left(ZX_{f(k)}\left[h'\left(S_{m(i)}(X) + S_{f(k+1)}^{c}(\gamma)\right) - h'\left(S_{m(i-1)}(X) + S_{f(k+1)}^{c}(\gamma)\right)\right]\right)\right|.$$

Since $S_{m(i)}(X)$ and $S_{m(i-1)}(X)$ are $\mathcal{F}_{V_{f(k)}^{|m(i)-f(k)|}}$-measurable, we can take the conditional expectation of $X_{f(k)}$ with respect to $\mathcal{F}_{V_{f(k)}^{|m(i)-f(k)|}}$ in the right hand side of the above equation. On the other hand the function h' is 1-Lipschitz, hence

$$|h'\left(S_{m(i)}(X) + S_{f(k+1)}^{c}(\gamma)\right) - h'\left(S_{m(i-1)}(X) + S_{f(k+1)}^{c}(\gamma)\right)| \le |X_{m(i)}|.$$

Consequently, the term

$$\left|E\left(ZX_{f(k)}\left[h'\left(S_{m(i)}(X) + S_{f(k+1)}^{c}(\gamma)\right) - h'\left(S_{m(i-1)}(X) + S_{f(k+1)}^{c}(\gamma)\right)\right]\right)\right|$$

is bounded by

$$E|X_{m(i)}E_{|m(i)-f(k)|}\left(X_{f(k)}\right)|$$

and

$$|E\left(ZX_{f(k)}h'^{N}_{k-1,k+1}\right)| \le \sum_{i=1}^{|E_k^N|} E|X_{m(i)}E_{|m(i)-f(k)|}(X_{f(k)})|.$$

Hence,

$$\left|\sum_{k=1}^{n^d} E\left(ZX_{f(k)}h'^{N}_{k-1,k+1}\right)\right| \le \sum_{k=1}^{n^d}|s_{f(k)}|\sum_{i=1}^{|E_k^N|}|s_{m(i)}|E|\varepsilon_{m(i)}E_{|m(i)-f(k)|}(\varepsilon_{f(k)})|$$
$$\le \sum_{k=1}^{n^d}|s_{f(k)}|\sum_{j\in V_0^N}|s_{j+f(k)}|E|\varepsilon_j E_{|j|}(\varepsilon_0)|$$
$$\le A\sum_{j\in V_0^N}\|\varepsilon_j E_{|j|}(\varepsilon_0)\|_1 < +\infty \quad (A\in\mathbb{R}_+^*)$$

where (by Lemma 1) we used the fact that

$$\sup_{i\in\Lambda_n}|s_i| = O\left(\frac{1}{(nh_n)^{d/2}}\right) \tag{10}$$

and

$$\sum_{i\in\Lambda_n}|s_i| = O\left((nh_n)^{d/2}\right). \tag{11}$$

Since (3) is satisfied, this last term is as small as we wish by choosing N large enough. Applying again Taylor's formula, it remains to consider

$$X_{f(k)}(h'_{k-1,k+1}-h'^{N}_{k-1,k+1}) = X_{f(k)}(S_{f(k-1)}(X)-S^N_{f(k)}(X))h''_{k-1,k+1}+R'_k,$$

where $|R'_k|\le 2|X_{f(k)}(S_{f(k-1)}(X)-S^N_{f(k)}(X))(1\wedge|S_{f(k-1)}(X)-S^N_{f(k)}(X)|)|$. It follows that

$$\sum_{k=1}^{n^d}E|R'_k| \le 2\left(\sum_{k=1}^{n^d}|s_{f(k)}|\right)E\left(|\varepsilon_0|\left(\sum_{i\in\Lambda_N}|s_i||\varepsilon_i|\right)\left(1\wedge\sum_{i\in\Lambda_N}|s_i||\varepsilon_i|\right)\right)$$
$$\le 2A\,E\left(|\varepsilon_0|\left(\sum_{i\in\Lambda_N}|\varepsilon_i|\right)\left(1\wedge\sum_{i\in\Lambda_N}|s_i||\varepsilon_i|\right)\right)\quad (A\in\mathbb{R}_+^*).$$

Keeping in mind that $s_i\to 0$ as $n\to\infty$ and applying the dominated convergence theorem, this last term converges to zero as n tends to infinity and (9) follows.

3.1.3 The second order terms

It remains to control

$$W_1 = E\left(Z\sum_{k=1}^{n^d}h''_{k-1,k+1}\left(\frac{X^2_{f(k)}}{2}+X_{f(k)}\left(S_{f(k-1)}(X)-S^N_{f(k)}(X)\right)-\frac{s^2_{f(k)}\eta_N}{2}\right)\right). \tag{12}$$

We consider the following sets:

$$\Lambda_n^N = \{i \in \Lambda_n\,;\, d(i, \partial\Lambda_n) \geq N\} \quad \text{and} \quad I_n^N = \{1 \leq i \leq n^d\,;\, f(i) \in \Lambda_n^N\},$$

and the function Ψ from $\mathbb{R}^{\mathbb{Z}^d}$ to $\mathbb{R}$ such that

$$\Psi(\varepsilon) = \varepsilon_0^2 + \sum_{i \in V_0^1 \cap C_{N-1}} 2\varepsilon_0\varepsilon_i.$$

For k in $[1, n^d]$, we set $D_k^N = \eta_N - \Psi \circ T^{f(k)}(\varepsilon)$. By definition of Ψ and of the set I_n^N, we have for any k in I_n^N

$$\Psi \circ T^{f(k)}(\varepsilon) = \varepsilon_{f(k)}^2 + 2\varepsilon_{f(k)}(S_{f(k-1)}(\varepsilon) - S_{f(k)}^N(\varepsilon)).$$

Therefore for k in I_n^N

$$s_{f(k)}^2 D_k^N = s_{f(k)}^2 \eta_N - X_{f(k)}^2 - 2X_{f(k)}(S_{f(k-1)}(X) - S_{f(k)}^N(X)).$$

Since $\lim_{n\to+\infty} n^{-d}|I_n^N| = 1$, it remains to prove that

$$\lim_{N\to+\infty} \limsup_{n\to+\infty} E\left(Z \sum_{k=1}^{n^d} s_{f(k)}^2 h''_{k-1,k+1} D_k^N\right) = 0. \tag{13}$$

3.1.4 Conditional expectation with respect to the tail σ-algebra

Now, we are going to replace D_k^N by $E\left(D_k^N | \mathcal{F}_{f(k),-\infty}\right)$. We introduce the expression

$$H_n^N = \sum_{k=1}^{n^d} E\left(s_{f(k)}^2 Z h''_{k-1,k+1}[\Psi \circ T^{f(k)}(\varepsilon) - E(\Psi \circ T^{f(k)}(\varepsilon)|\mathcal{F}_{f(k),-\infty})]\right).$$

For sake of brevity, we have written $h''_{k-1,k+1}$ instead of $h''_{k-1,k+1}(X)$. Using the stationarity of the field we get that

$$H_n^N = \sum_{k=1}^{n^d} E\left(s_{f(k)}^2 Z (h''_{k-1,k+1} \circ T^{-f(k)})(X)[\Psi(\varepsilon) - E(\Psi(\varepsilon)|\mathcal{F}_{-\infty})]\right).$$

For any positive integer p, we decompose H_n^N in two parts

$$H_n^N = \sum_{k=1}^{n^d} J_k^1(p) + \sum_{k=1}^{n^d} J_k^2(p),$$

where

$$J_k^1(p) = E\left(s_{f(k)}^2 Z(h_{k-1,k+1}^{''p} \circ T^{-f(k)})[\Psi(\varepsilon) - E(\Psi(\varepsilon)|\mathcal{F}_{-\infty})]\right)$$

and $J_k^2(p)$ equals to

$$E\left(s_{f(k)}^2 Z[h_{k-1,k+1}^{''} \circ T^{-f(k)} - h_{k-1,k+1}^{''p} \circ T^{-f(k)}](X)[\Psi(\varepsilon) - E(\Psi(\varepsilon)|\mathcal{F}_{-\infty})]\right).$$

From the definition of $h_{k-1,k+1}^{''p}$, we infer that the variable $h_{k-1,k+1}^{''p} \circ T^{-f(k)}(X)$ is $\mathcal{F}_{V_0^p}$-measurable. Therefore, we can take the conditional expectation of $\Psi(\varepsilon) - E(\Psi(\varepsilon)|\mathcal{F}_{-\infty})$ with respect to $\mathcal{F}_{V_0^p}$ in the expression of $J_k^1(p)$. Now, the backward martingale limit theorem implies that

$$\lim_{p\to+\infty} E|E(\Psi(\varepsilon)|\mathcal{F}_{V_0^p}) - E(\Psi(\varepsilon)|\mathcal{F}_{-\infty})| = 0$$

and consequently

$$\lim_{p\to+\infty} \limsup_{n\to+\infty} \left|\sum_{k=1}^{n^d} J_k^1(p)\right| = 0.$$

On the other hand

$$\left|\sum_{k=1}^{n^d} J_k^2(p)\right| \leq E\left[\left(2 \wedge \sum_{|i|<p} s_{f(i)}^2 |\varepsilon_i|\right) |\Psi(\varepsilon) - E(\Psi(\varepsilon)|\mathcal{F}_{-\infty})|\right].$$

Hence, applying the dominated convergence theorem, we conclude that H_n^N tends to zero as n tends to infinity. It remains to consider

$$W_2 = E\left(Z \sum_{k=1}^{n^d} h_{k-1,k+1}^{''} s_{f(k)}^2 E(D_k^N|\mathcal{F}_{f(k),-\infty})\right).$$

3.1.5 Truncation

For any integer k in $[1, n^d]$ and any M in $\mathbb{R}^+$ we introduce the two sets

$$B_k^N(M) = E(D_k^N|\mathcal{F}_{f(k),-\infty})\, \mathbb{1}_{|\eta_N - E(\Psi\circ T^{f(k)}(\varepsilon)|\mathcal{F}_{f(k),-\infty})|\leq M}$$

and

$$\overline{B}_k^N(M) = E(D_k^N|\mathcal{F}_{f(k),-\infty}) - B_k^N(M).$$

The stationarity of the field ensures that $E|\overline{B}_k^N(M)| = E|\overline{B}_1^N(M)|$ for any k in $[1, n^d]$. Now, applying the dominated convergence theorem, we have $\lim_{M\to+\infty} E|\overline{B}_1^N(M)| = 0$. It follows that

$$\lim_{M\to+\infty} \sum_{k=1}^{n^d} E\left(h''_{k-1,k+1} s_{f(k)}^2 \overline{B}_k^N(M)\right) = 0.$$

Therefore instead of W_2 it remains to consider

$$W_3 = E\left(Z \sum_{k=1}^{n^d} h''_{k-1,k+1} s_{f(k)}^2 B_k^N(M)\right).$$

3.1.6 An ergodic lemma

The next result is the central point of the proof.

Lemma 2 *For all M in $\mathbb{R}^+$, we introduce*

$$\beta_N(M) = E\left([\eta_N - E(\Psi(\varepsilon)|\mathcal{F}_{-\infty})] \mathbb{1}_{|\eta_N - E(\Psi(\varepsilon)|\mathcal{F}_{-\infty})| \le M} \middle| \mathcal{I}\right).$$

Then

$$\lim_{M\to+\infty} \beta_N(M) = 0 \quad a.s. \quad and \quad \lim_{n\to+\infty} E\left|\beta_N(M) - \sum_{k=1}^{n^d} s_{f(k)}^2 B_k^N(M)\right| = 0.$$

Proof of Lemma 2. Let

$$u(\varepsilon) = [\eta_N - E(\Psi(\varepsilon)|\mathcal{F}_{-\infty})] \mathbb{1}_{|\eta_N - E(\Psi(\varepsilon)|\mathcal{F}_{-\infty})| \le M}.$$

Using the function u, we write $\beta_N(M) = E(u(\varepsilon)|\mathcal{I})$. The fact that $\beta_N(M)$ tends to zero as M tends to infinity follows from the dominated convergence theorem. In fact

$$\lim_{M\to\infty} u(\varepsilon) = \eta_N - E(\Psi(\varepsilon)|\mathcal{F}_{-\infty})$$

and $u(\varepsilon)$ is bounded by $|\eta_N - E(\Psi(\varepsilon)|\mathcal{F}_{-\infty})|$ which belongs to L^1. This implies that

$$\lim_{M\to\infty} \beta_N(M) = E\left(\eta_N - E(\Psi(\varepsilon)|\mathcal{F}_{-\infty}) \middle| \mathcal{I}\right) \quad \text{a.s.}$$

Since $\mathcal{I}$ is included in the $\mathbb{P}$-completion of $\mathcal{F}_{-\infty}$ (see the above proposition) and keeping in mind that η_N is $\mathcal{I}$-measurable, it follows that

$$\lim_{M\to\infty} \beta_N(M) = \eta_N - E(\Psi(\varepsilon)|\mathcal{I}) \quad \text{a.s.}$$

By stationarity of the random field, we know that $E(\varepsilon_0\varepsilon_k|\mathcal{I}) = E(\varepsilon_0\varepsilon_{-k}|\mathcal{I})$ which implies that

$$E(\Psi(\varepsilon)|\mathcal{I}) = \sum_{k\in C_{N-1}} E(\varepsilon_0\varepsilon_k|\mathcal{I}) = \eta_N$$

and the result follows.
We are going to prove the second point of Lemma 2. First note that

$$\begin{aligned} B_k(M) &= [\eta_N - E(\Psi\circ T^{f(k)}(\varepsilon)|\mathcal{F}_{f(k),-\infty})]\,\mathbb{1}_{|\eta_N - E(\Psi\circ T^{f(k)}(\varepsilon)|\mathcal{F}_{f(k),-\infty})|\le M} \\ &= u\circ T^{f(k)}(\varepsilon). \end{aligned}$$

Consequently

$$\sum_{k=1}^{n^d} s_{f(k)}^2 B_k^N(M) = \sum_{i\in\Lambda_n} s_i^2\, u\circ T^i(\varepsilon).$$

Finally, the proof of lemma 2 is completed by the following lemma which is proved in Section 5.

Lemma 3

$$\lim_{n\to\infty}\left\|\sum_{i\in\Lambda_n} s_i^2\, u\circ T^i(\varepsilon) - E(u(\varepsilon)|\mathcal{I})\right\|_2 = 0.$$

As a direct application of lemma 2, we see that

$$\left|E\left(Z\sum_{k=1}^{n^d} h''_{k-1,k+1}s_{f(k)}^2\beta_N(M)\right)\right| \le E|\beta_N(M)|$$

is as small as we wish by choosing M large enough. So instead of W_3 we consider

$$W_4 = E\left(Z\sum_{k=1}^{n^d} h''_{k-1,k+1}s_{f(k)}^2[B_k^N(M)-\beta_N(M)]\right).$$

3.1.7 Abel transformation

In order to control W_4, we use the Abel transformation:

$$\begin{aligned} W_4 = E\Bigg[\sum_{k=1}^{n^d}\left(\sum_{i=1}^{k} s_{f(i)}^2[B_i^N(M)-\beta_N(M)]\right) Z(h''_{k-1,k+1}-h''_{k,k+2})\Bigg] \\ + E\left(Zh''_{n^d,n^d+2}\sum_{k=1}^{n^d} s_{f(k)}^2[B_k^N(M)-\beta_N(M)]\right). \end{aligned}$$

Now

$$\left|E\left(Zh''_{n^d,n^d+2}\sum_{k=1}^{n^d}s_{f(k)}^2[B_k^N(M)-\beta_N(M)]\right)\right| \le E\left|\beta_N(M)-\sum_{k=1}^{n^d}s_{f(k)}^2B_k^N(M)\right|.$$

Then applying lemma 2, we obtain

$$\lim_{n\to+\infty}\left|E\left(Zh''_{n^d,n^d+2}\sum_{k=1}^{n^d}s_{f(k)}^2[B_k^N(M)-\beta_N(M)]\right)\right|=0.$$

Therefore it remains to prove that for any positive integer N and any positive real M,

$$\lim_{n\to+\infty}E\Bigg[\sum_{k=1}^{n^d}\left(\sum_{i=1}^{k}s_{f(i)}^2[B_i^N(M)-\beta_N(M)]\right)Z(h''_{k-1,k+1}-h''_{k,k+2})\Bigg]=0.$$

3.1.8 Last reductions

We are going to finish the proof. We use the same decomposition as before:

$$h''_{k,k+2}-h''_{k-1,k+1}=h''_{k,k+2}-h''_{k,k+1}+h''_{k,k+1}-h''_{k-1,k+1}.$$

Applying Taylor's formula

$$h''_{k,k+2}-h''_{k,k+1}=-\gamma_{f(k+1)}h'''_{k,k+2}+t_k$$

and

$$h''_{k,k+1}-h''_{k-1,k+1}=X_{f(k)}h'''_{k-1,k+1}+T_k$$

where $|t_k|\le\gamma_{f(k+1)}^2$ and $|T_k|\le X_{f(k)}^2$. To examine the remainder terms, we consider:

$$E\left(\sum_{k=1}^{n^d}s_{f(k)}^2\left(\sum_{i=1}^{k}s_{f(i)}^2[B_i^N(M)-\beta_N(M)]\right)Z\varepsilon_{f(k)}^2\right).$$

The definition of $B_i^N(M)$ and of $\beta_N(M)$ enables us to write for all integer k in $[1,n^d]$,

$$\sum_{i=1}^{k}s_{f(i)}^2|B_i^N(M)-\beta_N(M)|\le 2M.$$

Therefore

$$E\left|\sum_{k=1}^{n^d}\left(\sum_{i=1}^{k} s_{f(i)}^2[B_i^N(M)-\beta_n(M)]\right) s_{f(k)}^2 Z\varepsilon_{f(k)}^2\,\mathbb{1}_{|\varepsilon_{f(k)}|>K}\right| \le 2ME\left(\varepsilon_0^2\,\mathbb{1}_{|\varepsilon_0|>K}\right)$$

and applying the dominated convergence theorem this last term is as small as we wish by choosing K large enough. Now, for all K in $\mathbb{R}^+$, Lemma 2 ensures that

$$\lim_{n\to+\infty} E\left(\sum_{k=1}^{n^d} s_{f(k)}^2\left(\sum_{i=1}^{k} s_{f(i)}^2[B_i^N(M)-\beta_N(M)]\right) Z\varepsilon_{f(k)}^2\,\mathbb{1}_{|\varepsilon_{f(k)}|\le K}\right)=0.$$

So, we have proved that

$$\lim_{n\to+\infty} E\left(\sum_{k=1}^{n^d}\left(\sum_{i=1}^{k} s_{f(i)}^2[B_i^N(M)-\beta_N(M)]\right) ZT_k\right)=0.$$

In the same way, we obtain that

$$\lim_{n\to+\infty} E\left(\sum_{k=1}^{n^d}\left(\sum_{i=1}^{k} s_{f(i)}^2[B_i^N(M)-\beta_N(M)]\right) Zt_k\right)=0.$$

Moreover since $(\varepsilon,(\tau_i)_{i\neq f(k+1)})$ is independent of $\tau_{f(k+1)}$ we have

$$E\left(\sum_{i=1}^{k} s_{f(i)}^2[B_i^N(M)-\beta_N(M)]\gamma_{f(k+1)} Z h_{k,k+2}^{'''}\right)=0.$$

Finally, it remains to consider

$$W_5 = E\Bigg[\sum_{k=1}^{n^d}\left(\sum_{i=1}^{k} s_{f(i)}^2[B_i^N(M)-\beta_N(M)]\right) ZX_{f(k)} h_{k-1,k+1}^{'''}\Bigg].$$

Let p be a fixed positive integer. Since $h^{'''}$ is 1-Lipschitz, we have the upper bound $|h_{k-1,k+1}^{'''}-h_{k-1,k+1}^{'''p}| \le |S_{f(k-1)}(X)-S_{f(k)}^p(X)|$. Now, we can apply the same truncation argument as before: first we choose the level of our truncation by applying the dominated convergence theorem and then we use Lemma 2. So, it follows that

$$\lim_{n\to+\infty} E\Bigg[\sum_{k=1}^{n^d}\left(\sum_{i=1}^{k} s_{f(i)}^2[B_i^N(M)-\beta_N(M)]\right) ZX_{f(k)}(h_{k-1,k+1}^{'''}-h_{k-1,k+1}^{'''p})\Bigg]=0.$$

Therefore, to prove our theorem it is enough to show that

$$\lim_{p\to+\infty}\limsup_{n\to+\infty} E\Bigg[\sum_{k=1}^{n^d}\left(\sum_{i=1}^{k} s_{f(i)}^2[B_i^N(M)-\beta_N(M)]\right) Z X_{f(k)} h_{k-1,k+1}^{'''p}\Bigg]=0. \tag{14}$$

We consider a one to one map m from $[1,|E_k^p|]\cap\mathbb{N}^*$ to E_k^p and such that $|m(i)-f(k)|\le|m(i-1)-f(k)|$. Now, we use the same argument as before:

$$\begin{aligned} h_{k-1,k+1}^{'''p}-h^{'''}(S_{f(k)}^c(\gamma)) &= \sum_{i=1}^{|E_k^p|} h^{'''}(S_{m(i)}(X)+S_{f(k)}^c(\gamma))-h^{'''}(S_{m(i-1)}(X)+S_{f(k)}^c(\gamma)) \\ &\le \sum_{i=1}^{|E_k^p|}|X_{m(i)}|. \end{aligned}$$

Here recall that $B_i^N(M)$ is $\mathcal{F}_{f(i),-\infty}$-measurable and $\beta_N(M)$ is $\mathcal{I}$-measurable. We have $E(\varepsilon_{f(k)}|\mathcal{I})=0$, $E(\varepsilon_{f(k)}|\mathcal{F}_{f(k),-\infty})=0$ and $E(\varepsilon_{f(k)}|\mathcal{F}_{f(i),-\infty})=0$ for any positive integer i such that $i<k$. Consequently, for any positive integer i such that $i\le k$, we have

$$E\left(s_{f(i)}^2[B_i^N(M)-\beta_N(M)]Z s_{f(k)}\varepsilon_{f(k)}h^{'''}(S_{f(k)}^c(\gamma))\right)=0.$$

Therefore using the conditional expectation, we find

$$\begin{aligned} & E\Bigg[\sum_{k=1}^{n^d}\left(\sum_{i=1}^{k} s_{f(i)}^2[B_i^N(M)-\beta_N(M)]\right) Z X_{f(k)} h_{k-1,k+1}^{'''p}\Bigg] \\ &\le 2M\sum_{k=1}^{n^d}|s_{f(k)}|\sum_{i=1}^{|E_k^p|}|s_{m(i)}|E|\varepsilon_{m(i)}E_{|m(i)-f(k)|}(\varepsilon_{f(k)})| \\ &\quad = 2M\sum_{k=1}^{n^d}|s_{f(k)}|\sum_{j\in V_0^p}|s_{j+f(k)}|E|\varepsilon_j E_{|j|}(\varepsilon_0)| \\ &\qquad \le 2AM\sum_{j\in V_0^p}E|\varepsilon_j E_{|j|}(\varepsilon_0)| \quad (A\in\mathbb{R}_+^*) \quad \text{by (10) and (11).} \end{aligned}$$

Since (3) is realised the last term is as small as we wish by choosing p large enough, hence W_4 is handled. Finally, the main theorem is proved. □

3.2 Proof of the corollary

As observed in [8], the proof of the corollary is a direct consequence of Theorem 1.1 in Rio [23]. In fact, for any k in V_0^1, we have

$$\begin{aligned} E|\varepsilon_k E_{|k|}(\varepsilon_0)| &= Cov\left(|\varepsilon_k|\left(\mathbb{1}_{E_{|k|}(\varepsilon_0)\geq 0} - \mathbb{1}_{E_{|k|}(\varepsilon_0)\leq 0}\right), \varepsilon_0\right) \\ &\leq 4\int_0^{\alpha_{1,\infty}(|k|)} Q^2_{\varepsilon_0}(u)\, du. \end{aligned}$$

The proof of the corollary is complete. $\square$

4 Application

The direct consequence of our result is that it allows the construction of statistical tests able to quantify the estimation error. For this purpose, we show the construction of such a test that can be used in image denoising [11, 16, 28]. In the context given by the model (1), let us consider the following situation : a true image g is affected by a correlated additive noise ϵ, that gives Y for the observed image.
For the original function the classical Lena image is used. This image is a gray level image with pixels values in the interval $[0, 255]$. The size of the image is 256×256 pixels. The correlated noise we consider is a Gaussian field $(\varepsilon_k)_{k\in\mathbb{Z}^2}$ built using an exponential covariance function

$$C(k) = E(\varepsilon_0\varepsilon_k) = \text{Cst} \times \exp\{-\frac{|k|}{a}\}.$$

The choice of such random field ensures the validity of the projective criterion (3)(see [9], p.59 Corollary 2). There exist several methods for simulating such a random field, here we have opted for the spectral method [15]. In order to obtain an important visual effect of how the noise affects the original image Cst was set to 200 and $a = 1$. The noisy image is obtained by adding pixel by pixel the original image to the simulated noise. The estimator of the original image is computed using the Epanechnikov kernel

$$K(x) = \frac{3}{8}(1-|x|^2)\mathbb{I}_{\{|x|\leq 1\}},\ x = (x_1, x_2) \in \mathbb{R}^2.$$

In order to compute the expectation of the estimated function, several realisation of the noisy image are needed. Here we have considered 50 such images, constructed by adding the original Lena image with a noise realisation. Using (2), for each noisy image, an estimate g_n of the original function

g was computed using the kernel K defined above. The expectation $E(g_n)$ is computed by just taking the pixel by pixel arithmetical means corresponding to the images previously restored.

Clearly, it is now possible to estimate the difference $g_n - E(g_n)$. Following our theoretical result, the normalised square of this difference follows a χ^2 distribution with one degree of freedom. Since this quantity is observable, p-values pixel by pixel can be computed.

The original image and the realisation of a noisy image are shown in Figure 1a and 1b, respectively. It can be noticed that in the "dirty" picture, spots are formed, due to the noise correlation. The expectation of the estimated original images in Figure 1c exhibits almost no such spots. Furthermore, the visual quality of this restored image is close to the original. A more quantitative evaluation of this result is given by the image of p-values of the proposed statistical test given in Figure 1d. The light-coloured pixels represent p-values close to 1, whereas the dark-coloured pixels indicate values close to 0. We have counted 83% of the pixels for which we have obtained a p-value higher than 0.01. This ratio is quite a reliable indicator concerning the restored image. Together with the visual analysis of the results, it provides a detailed description of the obtained result. We conclude that, under these considerations, the theoretical results developed in this paper may be used as a basis for the development of practical tools in image analysis.

5 Annexe

In this section we prove lemma 3.

Proof of Lemma 3. In fact, for any u in L^2, we can write $u = w + E(u|\mathcal{I})$ where $w = u - E(u|\mathcal{I})$, hence it suffices to prove that

$$\lim_{n\to+\infty} \left\| \sum_{k\in\Lambda_n} s_k^2 \, w \circ T^k \right\|_2 = 0.$$

Let us consider the transformations $T_1, T_2, ..., T_d$ defined by $T_1^i = T^{(i,\dots,0)}$, $T_2^i = T^{(0,i,\dots,0)}, ..., T_d^i = T^{(0,\dots,i)}$ for any integer i. It is well known (cf. [7]) that the space

$$H = \left\{ h_1 - h_1 \circ T_1 - (h_2 - h_2 \circ T_2) - ... - (h_d - h_d \circ T_d)\,;\, h_1, .., h_d \in L^2 \right\}$$

is dense in the space $G = \{g \in L^2\,;\, E(g|\mathcal{I}) = 0\}$. Let $\varepsilon > 0$ be fixed, there exist $h_1, ..., h_d \in L^2$ such that $\left\| w - [(h_1 - h_1 \circ T_1) - ... - (h_d - h_d \circ T_d)] \right\|_2 \leq \varepsilon$.

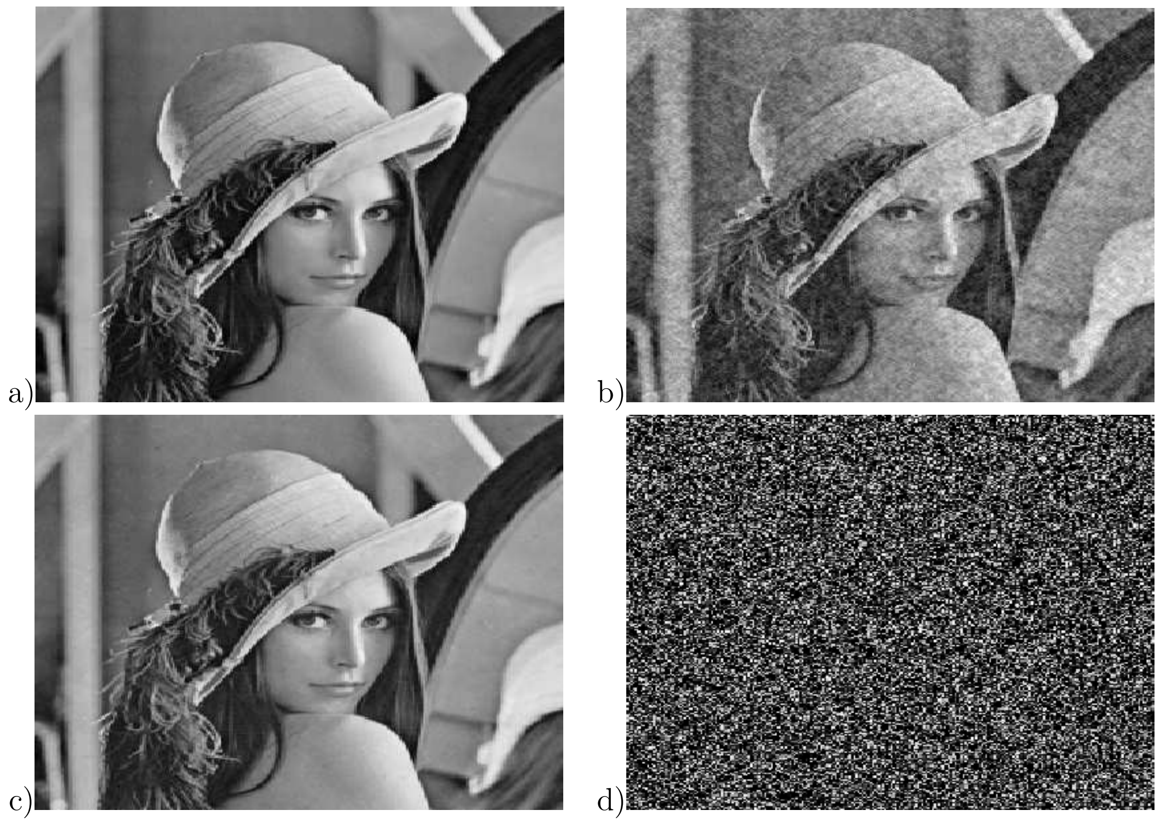

Figure 1: Results of the image restoration procedure : a) original Lena image, b) realisation of a noisy image, c) expectation of the restored images, d) obtained $p-$values as a gray level image (white pixels represent values close to 1, whereas black pixels indicate values close to 0).

So, we derive

$$\Big\| \sum_{k \in \Lambda_n} s_k^2 \, w \circ T^k \Big\|_2 \le \varepsilon + \sum_{j=1}^{d} \Big\| \sum_{k \in \Lambda_n} s_k^2 \left(h_j \circ T^k - h_j \circ T^{k(j)} \right) \Big\|_2$$

where $k(j) = (k_1, ..., k_{j-1}, k_j + 1, k_{j+1}, ..., k_d)$. Using Lemma 1 and keeping in mind that K is a bounded and Lipschitzian kernel, one can check that

$$s^2_{(k_1,..,k_d)} = O\left(\frac{1}{(nh_n)^d}\right) \quad \text{and} \quad s^2_{(k_1,..,k_j,..,k_d)} - s^2_{(k_1,..,k_j-1,..,k_d)} = O\left(\frac{1}{(nh_n)^{d+1}}\right)$$

and consequently, we obtain

$$\sum_{j=1}^{d} \Big\| \sum_{k \in \Lambda_n} s_k^2 \left(h_j \circ T^k - h_j \circ T^{k(j)} \right) \Big\|_2 = O\left(\frac{1}{nh_n^{d+1}}\right).$$

Finally, keeping n sufficiently large, we obtain

$$\Big\| \sum_{k \in \Lambda_n} s_k^2 \, w \circ T^k \Big\|_2 \le 2\varepsilon.$$

The proof of lemma 3 is complete. $\square$

References

[1] N. Altman. Krige, smooth, both or neither ? *Aust. N. Z. J. Stat.*, 42:441–461, 2000.

[2] G. Biau. Spatial kernel density estimation. *Math. Methods Statist.*, 12(4):371–390, 2003.

[3] G. Biau and B. Cadre. Nonparametric spatial prediction. *Stat. Inference Stoch. Process.*, 7(3):327–349, 2004.

[4] D. Bosq. *Nonparametric Statistics for Stochastic Processes-Estimation and Prediction-2nde Edition.* Lecture Notes in Statistics, Springer Verlag, New York, 1998.

[5] M. Carbon, M. Hallin, and B. Wu. Kernel density estimation for random fields: the l_1 theory. *Journal of nonparametric Statistics*, 6:157–170, 1996.

[6] M. Carbon, L.T. Tran, and B. Wu. Kernel density estimation for random fields. *Statist. Probab. Lett.*, 36:115–125, 1997.

[7] J.P. Conze. Personal communication, 2005.

[8] J. Dedecker. A central limit theorem for stationary random fields. *Probab. Theory Relat. Fields*, 110:397–426, 1998.

[9] P. Doukhan. *Mixing: properties and examples*, volume 85. Lecture Notes in Statistics, Berlin, 1994.

[10] M. El Machkouri. Nonparametric regression estimation for random fields in a fixed-design. *Stat. Inference Stoch. Process.*, 10(1):29–47, 2007.

[11] R. C. Gonzalez and R. E. Woods. *Digital Image Processing*. Addison-Wesley, 1992.

[12] M. Hallin, Z. Lu, and L.T. Tran. Density estimation for spatial linear processes. *Bernoulli*, 7:657–668, 2001.

[13] M. Hallin, Z. Lu, and L.T. Tran. Density estimation for spatial processes: the l_1 theory. *J. Multivariate Anal.*, 88(1):61–75, 2004.

[14] M. Hallin, Z. Lu, and L.T. Tran. Local linear spatial regression. *Ann. Statist.*, 32(6):2469–2500, 2004.

[15] C. Lantuejoul. *Geostatistical Simulation*. Springer-Verlag Berlin, 2002.

[16] S. Z. Li. *Markov Random Field Modeling in Image Analysis (3rd Edition)*. Springer, 2008.

[17] Z. Lu and X. Chen. Spatial nonparametric regression estimation: Non-isotropic case. *Acta Mathematicae Applicatae Sinica, English series*, 18:641–656, 2002.

[18] Z. Lu and X. Chen. Spatial kernel regression estimation: weak consistency. *Statist. Probab. Lett.*, 68:125–136, 2004.

[19] D. H. Martirosian. Phase transitions for martingale-difference gibbs lattice models. *J. Contemp. Math. Anal.*, 29:76–80, 1994.

[20] B. Nahapetian. Billingsley-Ibragimov theorem for martingale-difference random fields and its applications to some models of classical statistical physics. *C.R. Acad. Sci. Paris*, 320, série I:1539–1544, 1995.

[21] B. Nahapetian and A. N. Petrosian. Martingale-difference Gibbs random fields and central limit theorem. *Ann. Acad. Sci. Fenn., Series A-I Math.*, 17:105–110, 1992.

[22] A. Rényi. On stable sequences of events. *Sankhya Ser. A*, 25:189–206, 1963.

[23] E. Rio. Covariance inequalities for strongly mixing processes. *Ann. Inst. H. Poincaré Probab. Statist.*, 29:587–597, 1993.

[24] E. Rio. About the Lindeberg method for strongly mixing sequences. *ESAIM*, 1:35–61, 1995.

[25] M. Rosenblatt. A central limit theorem and a strong mixing condition. *Proc. Nat. Acad. Sci. USA*, 42:43–47, 1956.

[26] L.T. Tran. Kernel density estimation on random fields. *J. Multivariate Anal.*, 34:37–53, 1990.

[27] L.T. Tran and Yakowitz S. Nearest neighbor estimators for random fields. *J. Multivariate Anal.*, 44:23–46, 1993.

[28] G. Winkler. *Image Analysis, Random Fields and Markov Chain Monte Carlo Methods*. Springer, 2003.

[29] Q. Yao. Exponential inequalities for spatial processes and uniform convergence rates for density estimation. *In Development of Modern Statistics and Related Topics - In Celebration of Prof. Yaoting Zhang's 70th Birthday, H. Zhang and J. Huang (edit.), World Scientific, Singapore*, pages 118–128, 2003.

Mohamed EL MACHKOURI, Radu STOICA
Laboratoire de Mathématiques Paul Painlevé
UMR CNRS 8524, Université Lille 1
U.F.R. de Mathématiques Pures et Appliquées
59655 Villeneuve d'Ascq Cedex
Mohamed.El-Machkouri@math.univ-lille1.fr,
Radu.Stoica@math.univ-lille1.fr