Available at
http://pvamu.edu/aam
Appl. Appl. Math.
ISSN: 1932-9466
Vol. 11, Issue 2 (December 2016), pp. 527 – 558

Applications and Applied Mathematics:
An International Journal
(AAM)

Exact asymptotic errors of the hazard conditional rate kernel for functional random fields

[1]Abbes Rabhi & [2]Sara Soltani

[1]Laboratory of Mathematics
University of Sidi Bel Abbes
P.O.Box. 89
Sidi Bel Abbes 22 000, Algeria
rabhi_abbes@yahoo.fr
&
[2]Stochastic Models, Statistics and Applications Laboratory
Moulay Tahar University of Saida
Algeria
s.assala@hotmail.com

Received: February 22, 2015; Accepted: October 19, 2016

Abstract

We consider the problem of nonparametric estimation of the kernel type estimators for the conditional cumulative distribution function and the successive derivatives of the conditional density for spatial data. More precisely, given a strictly stationary random field $Z = (X, Y)$, we investigate a kernel estimate of the conditional hazard function of univariate response variable Y given the functional variable X. The principal aim of this paper is to give the mean squared convergence rate of the proposed estimator. Finally, we apply these theoretical results to the estimation of the conditional hazard function where we give the mean squared convergence rate of the proposed estimator.

Keywords: Conditional distribution; Kernel conditional hazard function; Kernel estimation; Spatial process; Functional data

MSC 2010 No.: 62G05, 62G20, 62N02

1. Introduction

The statistical problems involved in the modelization of spatial data have received an increasing interest in the literature. The infatuation for this topic is linked with many fields of applications in which the data are collected in the spatial order. Key references on spatial statistic are Ripley (1981) or Cressie (1991). The nonparametric treatment of such data is relatively recent. The first results have been obtained by Tran (1990). For relevant works on the nonparametric modelization of spatial data, see Lu and Chen (2004), Biau and Cadre (2004), Carbon et al. (2007), or Li et al. (2009) (for a list of references). In this paper, we are interested in the nonparametric estimation of the conditional hazard function when the covariates are of functional nature.

The nonparametric estimation of the hazard and/or the conditional hazard function is quite important in a variety of fields including medicine, reliability, survival analysis, or in seismology. The literature on this model in multivariate statistics is abundant. Historically, the hazard estimate was introduced by Watson and Leadbetter (1964). Since then, several results have been added; for example, see Roussas (1989) for previous works, and see Li and Tran (2007) for recent advances and references.

The literature is strictly limited in the case where the data is of functional nature (a curve). The first result in this context was given by Ferraty et al. (2008). They established the almost complete convergence of the kernel estimate of the conditional hazard function in the i.i.d. case. Their results have been extended to the dependent case by Quintetla (2008). The latter has stated, under α-mixing condition, the almost couplet convergence, the mean quadratic convergence, and the asymptotic normality of this estimate. Recently, Laksaci and Mechab (2010) consider the spatial case. They studied the almost complete convergence of an adapted estimate of this model. More recently Bouchentouf et al. (2014) give the uniform version of the almost complete convergence rate in the i.i.d. case from a model to a single-index functional. Djebouri et al. (2015) studied the mean quadratic convergence and asymptotic normality under α-mixing condition of this estimate. In practice, the interest of our study comes mainly from the fact that the main fields of application of functional statistical methods related to the analysis of continuously indexed spatial processes. It should be noted that the modelization of the functional spatial data has been selected by Ramsay (2008). Among the recent papers on the functional statistic of spatial data, we refer to Dabo-Niang et al. (2011).

The main aim of this paper is to study, under general conditions, the asymptotic proprieties of the functional spatial kernel estimate of the conditional hazard function introduced by Laksaci and Mechab (2010). More precisely, we treat the L^2-convergence rate by giving the exact expression involved in the leading terms of the quadratic error of the construct estimator. We point out that our asymptotic results are useful in some statistical problems such as in risk analysis. The present work extends to the spatial case the result of Quintela (2008), given in the functional time series case. We note that one of the main difficulties that arise in the analysis of spatial data comes from the fact that points in the N-dimensional space do not have a linear order. Thus, extending classical nonparametric statistic results for functional random fields is far from being trivial.

2. The Model

Consider $Z_{\mathbf{i}} = (X_{\mathbf{i}}, Y_{\mathbf{i}})$, $\mathbf{i} \in \mathbb{N}^N, N \geq 1$, a $\mathcal{F} \times \mathbb{R}$-valued measurable strictly stationary process, defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, where, $(\mathcal{F}, d)$ is a semi-metric space. A point $\mathbf{i}$ denotes an element of $\mathbb{N}^N$, and we shall use the notation $\mathbf{i} = (i_1, \ldots, i_N)$. We define the rectangular region $I_{\mathbf{n}}$ by $I_{\mathbf{n}} = \{\mathbf{i} \in \mathbb{N}^N\ ,\ 1 \leq i_k \leq n_k, k = 1, \ldots, N\}$, for $\mathbf{n} = (n_1, \ldots, n_N) \in \mathbb{N}^N$. We assume that the process under study $Z_{\mathbf{i}}$ is observed on $I_{\mathbf{n}}$. We set $\mathbf{1} = (1, \ldots, 1)$, $\widehat{\mathbf{n}} = n_1 \ldots n_N$, and we write $\mathbf{n} \to \infty$ if $\min\{n_k\} \to \infty$ and $|\frac{n_j}{n_k}| < C$ for a constant $0 < C < \infty$ for all $1 \leq j, k \leq N$.

In the following, x will be a fixed point in $\mathcal{F}$ and N_x will denote a fixed neighborhood of x. We assume that the regular version of the conditional probability of Y given X exists. Moreover, we suppose that for all $z \in N_x$ the conditional distribution function of Y given $X = z$, $F^z(\cdot)$, is j-times continuously differentiable with respect to y on $\mathcal{S}_{\mathbb{R}}$ and we denote by $f(\cdot|z) = F(1)(\cdot|z)$ its conditional density with respect to (w.r.t.) Lebesgue's measure over $\mathbb{R}$. In this paper, we consider the problem of the nonparametric estimation of the successive derivatives of the conditional distribution and the conditional hazard function. Our aim is to build nonparametric estimates of several functions related with the conditional probability distribution (*cond-cdf*) of Y given X. For $x \in \mathcal{F}$, we will denote the *cond-cdf* of Y given $X = x$ by

$$\forall y \in \mathbb{R},\ F(y|x) = \mathbb{P}(Y \leq y|X = x).$$

If this distribution is absolutely continuous with respect to the Lebesgues measure on $\mathbb{R}$, then we will denote by $f(\cdot|x) = F(\cdot|x)^{(1)}$ (*resp.* $f(\cdot|x)^{(j)} = F(\cdot|x)^{(j+1)}$) the conditional density (*resp.* its j^{th} order derivative) of Y given $X = x$. In the following, any real function with an integer in brackets as exponent denotes its derivative with the corresponding order. In Section 3, we will give almost complete convergence results (with rates) for nonparametric estimates of both functions F^x and $f^{x(j)}$. Since $f^x = f^{x(0)}$, we will deduce immediately the convergence of the conditional density estimate from the general results concerning $f^{x(j)}$.

In this work, we will assume that the function random field $(Z_{\mathbf{i}}, \mathbf{i} \in \mathbb{N}^N)$ satisfies the following mixing condition:

$$\left\{ \begin{array}{c} \text{There exists a function } \varphi\left(t\right) \downarrow 0 \text{ as } t \to \infty, \text{ such that} \\ \forall\, E,\ E^{'} \text{ subsets of } \mathbb{N}^N \text{ with finite cardinals} \\ \alpha\left(\mathcal{B}\left(E\right), \mathcal{B}\left(E^{'}\right)\right) = \sup\limits_{B \in \mathcal{B}(E),\, C \in \mathcal{B}\left(E^{'}\right)} \left|\mathbb{P}\left(B \cap C\right) - \mathbb{P}\left(B\right)\mathbb{P}\left(C\right)\right| \\ \leq \psi\left(\text{Card}\left(E\right), \text{Card}\left(E^{'}\right)\right) \varphi\left(\text{dist}\left(E, E^{'}\right)\right), \end{array} \right. \tag{1}$$

where, $\mathcal{B}\left(E\right)$ (*resp.* $\mathcal{B}\left(E^{'}\right)$) denotes the Borel σ-field generated by $(Z_{\mathbf{i}},\ \mathbf{i} \in E)$ (*resp.* $\left(Z_{\mathbf{i}},\ \mathbf{i} \in E^{'}\right)$), $\text{Card}(E)$ (*resp.* $\text{Card}\left(E^{'}\right)$) the cardinality of E (*resp.* $E^{'}$), $\text{dist}\left(E, E^{'}\right)$ the Euclidean distance between E and $E^{'}$, and $\psi : \mathbb{Z}^2 \to \mathbb{R}^+$ is a symmetric positive function nondecreasing in each variable such that

$$\psi(n,m) \leq C \min(n,m), \quad \forall n,m \in \mathbb{N}, \tag{2}$$

for some $C > 0$,. We assume also that the process satisfies the following mixing condition

$$\sum_{i=1}^{\infty} i^{\delta} \varphi(i) < \infty, \;\; \delta > 0. \tag{3}$$

Note that condition (2) and (3) are the same as the mixing conditions used by Tran (1990) and Carbon et al. (1996) and are satisfied by many spatial models (see Guyon (1987) for some examples). It should be noted that if $\psi = 1$, then $Z_{\mathbf{i}}$ is called strongly mixing (see Doukhan et al. (1994) for discussion on mixing and examples).

In the following x will be a fixed point in $\mathcal{F}$, $\mathcal{N}_x$ will denote a fixed neighborhood of x. Assume that the Z_i's have the same distribution as (X,Y) and there exists a regular version of the conditional probability of Y given X. Let $F(\cdot|x)$ be the conditional distribution of the variable Y given $X = x$ and we assume that there is some compact subset $\mathcal{S}_{\mathbb{R}} := [\alpha_x, \beta_x]$.

The purpose is to estimate the *cond-cdf* $F(\cdot|x)$. We introduce a kernel type estimator $\widehat{F}(\cdot|x)$ of $F(\cdot|x)$ as follows:

$$\widehat{F}(y|x) = \frac{\sum_{\mathbf{i}\in I_{\mathbf{n}}} K(h_K^{-1} d(x, X_{\mathbf{i}})) H(h_H^{-1}(y - Y_{\mathbf{i}}))}{\sum_{\mathbf{i}\in I_{\mathbf{n}}} K(h_K^{-1} d(x, X_{\mathbf{i}}))}, \qquad \forall y \in \mathbb{R}, \tag{4}$$

where K is a kernel, H is a *cdf*, and $h_K = h_{K,n}$ (resp. $h_H = h_{H,n}$) is a sequence of positive real numbers. Note that if $N = 1$, the same estimators are obtained by Ferraty et al. (2005).

A natural and usual estimator of the j^{th} (resp. $(j-1)^{th}$ order derivative of the conditional distribution $F^{(j)}(\cdot|x)$ (resp. conditional density $f^{(j-1)}(\cdot|x)$) of Y given $X = x$. We propose to define the estimator $\widehat{F}(\cdot|x)^{(j)}$ of $F^{(j)}(\cdot|x)$ (resp. $\widehat{f}^{(j-1)}(\cdot|x)$ of $f^{(j-1)}(\cdot|x)$) as follows for $j \geq 1$:

$$\widehat{f}^{(j-1)}(y|x) = \widehat{F}^{(j)}(y|x) = \frac{h_H^{-j} \sum_{\mathbf{i}\in I_{\mathbf{n}}} K(h_K^{-1} d(x, X_{\mathbf{i}})) H^{(j)}(h_H^{-1}(y - Y_{\mathbf{i}}))}{\sum_{\mathbf{i}\in I_{\mathbf{n}}} K(h_K^{-1} d(x, X_{\mathbf{i}}))}, \qquad \forall y \in \mathbb{R}. \tag{5}$$

3. Main Results

All along the paper, when no confusion is possible, we will denote by C and C' some strictly positive generic constants, and we will denote for all $\mathbf{i} \in I_{\mathbf{n}}$, $K_i = K\left(h_K^{-1} d(x, X_i)\right)$, $H_i(y) = H\left(h_H^{-1}(y - Y_i)\right)$, $H_i^{(j)}(y) = H^{(j)}\left(h_H^{-1}(y - Y_i)\right)$ and $B(x,h) = \{x' \in \mathcal{F} : /d(x,x') < h\}$.

In order to establish our asymptotic results we need the following hypotheses:

(H1) $\forall r > 0, \mathbb{P}(X \in B(x,r)) =: \phi_x(r) > 0$.

(H2) $\forall \mathbf{i} \neq \mathbf{j}$,

$$0 < \sup_{\mathbf{i}\neq\mathbf{j}} \mathbb{P}\left[(X_{\mathbf{i}}, X_{\mathbf{j}}) \in B(x, h_K) \times B(x, h_K)\right] \leq C(\phi_x(h_K))^{(a+1)/a}, \text{ for some } 1 < a < \delta N^{-1}.$$

(H3) The successive derivatives of the conditional cumulative distribution function $F^{(j)}(y|x)$ satisfies the Hölder condition, that is: $\forall (x_1, x_2) \in \mathcal{N}_x \times \mathcal{N}_x$, $\forall (y_1, y_2) \in \mathcal{S}_{\mathbb{R}}^2$,

$$|F^{(j)}(y_1|x_1) - F^{(j)}(y_2|x_2)| \leq C\left(d^{b_1}(x_1, x_2) + |y_1 - y_2|^{b_2}\right), \quad b_1 > 0,\, b_2 > 0,$$

where, for any positive integer l, $F^{(l)}(z|x)$ denotes its $l-th$ derivative (*i.e.*, $\left.\dfrac{\partial^l F(y|x)}{\partial y^l}\right|_{y=z}$).

(H4) K is a function with compact support $(0, 1)$ such that $0 < C < K(t) < C' < \infty$.

(H5) H is a bounded continuous Lipschitz function, such that the first derivative $H^{(1)}$ verifies

$$\int |t|^{b_2} H^{(1)}(t)dt < \infty.$$

(H6) The support of $H^{(1)}$ is compact and $\forall l \geq j$, $H^{(l)}$ exists and is bounded with Lipschitz's condition.

(H7) There exists $0 < \alpha < (\delta - 5N)/3N$ and $\eta_0 > 0$, such that

$$\lim_{\mathbf{n}\to\infty} \widehat{\mathbf{n}}^{\alpha}\, h_H = \infty \quad \text{ and } \quad C\widehat{\mathbf{n}}^{\frac{(5+3\alpha)N-\delta}{\delta}+\eta_0} \leq h_H \phi_x(h), \quad \text{where,} \quad \widehat{\mathbf{n}} = n_1 \ldots n_N.$$

Our assumptions are fairly standard, since the conditions H1-H7 are very similar to those used by Ferraty et al. (2005).

Theorem 3.1.

Under the hypotheses (H1)-(H7), we have

$$\sup_{y\in\mathcal{S}_{\mathbb{R}}} |\widehat{F}^{(j)}(y|x) - F^{(j)}(y|x)| = \mathcal{O}\left(h_K^{b_1} + h_H^{b_2}\right) + \mathcal{O}_{a.co.}\left(\left(\frac{\log \widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\, h_H^{2j-1} \phi_x(h_K)}\right)^{\frac{1}{2}}\right). \tag{6}$$

Proof:

Let, $\widehat{F}_N^x(y)$ (resp. $\widehat{F}_D^x$) be defined as

$$\widehat{F}_D(x) := \frac{1}{\widehat{\mathbf{n}}\mathbb{E}K_{\mathbf{1}}}(x) \sum_{\mathbf{i}\in I_{\mathbf{n}}} K(h_K^{-1} d(x, X_{\mathbf{i}})), \;\; K_{\mathbf{1}}(x) = K(h_K^{-1} d(x, X_{\mathbf{1}}))$$

$$\widehat{F}_N^{(j)}(y|x) := \frac{h_H^{-j}}{\widehat{\mathbf{n}}\mathbb{E}K_{\mathbf{1}}}(x) \sum_{\mathbf{i}\in I_{\mathbf{n}}} K(h_K^{-1} d(x, X_{\mathbf{i}})) H^{(j)}(h_H^{-1}(y - Y_{\mathbf{i}})).$$

This proof is based on the decomposition

$$\begin{aligned}\widehat{F}^{(j)}(y|x) - F^{(j)}(y|x) &= \frac{1}{\widehat{F}_D(x)}\left\{\left(\widehat{F}_N^{(j)}(y|x) - \mathbb{E}\widehat{F}_N^{(j)}(y|x)\right)\right.\\ &\quad \left.-\left(F^{(j)}(y|x) - \mathbb{E}\widehat{F}_N^{(j)}(y|x)\right)\right\}\\ &\quad +\frac{F^{(j)}(y|x)}{\widehat{F}_D(x)}\left\{\mathbb{E}\widehat{F}_D(x) - \widehat{F}_D(x)\right\}, \qquad (7)\end{aligned}$$

and on the following intermediate results.

Lemma 3.2.

Under the hypotheses (H1)-(H2), (H4) and (H7), we have

$$\widehat{F}_D(x) - \mathbb{E}\widehat{F}_D(x) = \mathcal{O}_{a.co.}\left(\left(\frac{\log \widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,\phi(h_K)}\right)^{\frac{1}{2}}\right) \quad \text{and} \quad \sum_{\mathbf{n}\in\mathbb{N}^N} \mathbb{P}\left(\widehat{F}_D(x) < 1/2\right) < \infty. \qquad (8)$$

Lemma 3.3.

Under the hypotheses (H1), (H3)-(H7), we have

$$\sup_{y\in\mathcal{S}_{\mathbb{R}}} |F^{(j)}(y|x) - \mathbb{E}\widehat{F}_N^{(j)}(y|x)| = \mathcal{O}\left(h_K^{b_1}\right) + \mathcal{O}\left(h_H^{b_2}\right). \qquad (9)$$

Lemma 3.4.

Under the hypotheses of Theorem 3.1, we have

$$\sup_{y\in\mathcal{S}_{\mathbb{R}}} |\widehat{F}_N^{(j)}(y|x) - \mathbb{E}\widehat{F}_N^{(j)}(y|x)| = \mathcal{O}_{a.co}\left(\sqrt{\frac{\log \widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\, h_H^{2j-1}\,\phi_x(h_K)}}\right). \qquad (10)$$

□

The proofs of Lemma 3.3 and Lemma 3.4 are postponed to Section 5.

Now, let $\widehat{F}(\cdot|x) = \widehat{F}^{(0)}(\cdot|x)$ and $\widehat{f}(\cdot|x) = \widehat{f}^{(1)}(\cdot|x)$; it is obvious that the previous theorem allows us to get the two following corollaries.

Corollary 3.5.

Suppose that hypothesis (H3) are verified for $j = 0$ and under the assumptions (H1)-(H2), (H4)-(H5) and (H7), we have

$$\sup_{y\in\mathcal{S}_{\mathbb{R}}} |\widehat{F}(y|x) - F(y|x)| = \mathcal{O}\left(h_K^{b_1} + h_H^{b_2}\right) + \mathcal{O}_{a.co.}\left(\left(\frac{\log \widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,\phi_x(h_K)}\right)^{\frac{1}{2}}\right). \qquad (11)$$

Corollary 3.6

Suppose that hypothesis (H3) is verified for $j = 1$. Then, if the hypotheses of Corollary 3.1. are satisfied, we have

$$\sup_{y\in\mathcal{S}_{\mathbb{R}}} |\widehat{f}(y|x) - f(y|x)| = \mathcal{O}\left(h_K^{b_1} + h_H^{b_2}\right) + \mathcal{O}_{a.co.}\left(\left(\frac{\log \widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\, h_H \phi_x(h_K)}\right)^{\frac{1}{2}}\right). \tag{12}$$

Remark 3.7.

In this section, we have obtained a rate of convergence of the form

$$\mathcal{O}\left(h_K^{b_1} + h_H^{b_2}\right) + \mathcal{O}_{a.co.}\left(\left(\frac{\log \widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\, h_H^{2j-1} \phi_x(h_K)}\right)^{\frac{1}{2}}\right),$$

where $\mathcal{O}\left(h_K^{b_1} + h_H^{b_2}\right)$ is the rate of the bias of the estimator which only depends on the regularity of $F^{(j)}$. Note that, by using the same approach as in this paper, one can easily show that if $F^{(j)}$ satisfies a Lipschitz condition:

$\forall (x_1, x_2) \in \mathcal{N}_x \times \mathcal{N}_x$, $\forall (y_1, y_2) \in \mathbb{R}^2$,

$$|F^{(j)}(y_1|x_1) - F^{(j)}(y_2|x_2)| \leq C\left(d(x_1, x_2) + |y_1 - y_2|\right),$$

for some $C > 0$, then the rate of convergence is the following:

$$\mathcal{O}\left(h_K + h_H\right) + O_{a.co.}\left(\left(\frac{\log \widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\, h_H^{2j-1} \phi_x(h_K)}\right)^{\frac{1}{2}}\right).$$

A. Mean squared convergence

The first result concerns the L^2-consistency of $\widehat{F}^{(j)}(y|x)$ and $\widehat{h'}^{x}(y)$. In order to establish our asymptotic results we need the following hypotheses.

(H8) For $l \in \{0, 2\}$, the functions

$$\begin{cases} \Psi_l(s) = \mathbb{E}\left(\frac{\partial^l F(y|X)}{\partial y^l} - \frac{\partial^l F(y|x)}{\partial y^l}\middle| d(x,X) = s\right), \quad \text{and} \\ \Phi_l(s) = \mathbb{E}\left(\frac{\partial^l f(y|X)}{\partial y^l} - \frac{\partial^l f(y|x)}{\partial y^l}\middle| d(x,X) = s\right), \end{cases}$$

are derivable at $s = 0$.

(H9) The bandwidth h_K satisfies:

$$h_K \downarrow 0, \quad \forall t \in [0,1] \quad \lim_{h_K \to 0} \frac{\phi_x(t h_K)}{\phi_x(h_K)} = \beta_x(t) \;\; \text{and} \;\; \widehat{\mathbf{n}} h_H \phi_x(h_K) \to \infty \;\; \text{as } \mathbf{n} \to \infty.$$

(H10) The kernel K from $\mathbb{R}$ into $\mathbb{R}^+$ is a differentiable function supported on $[0,1]$. Its derivative K' exists and is such that there exist two constants C and C' with $-\infty < C < K'(t) < C' < 0$ for $0 \leq t \leq 1$.

These conditions are very standard in this context. Indeed, assumptions (H8) are a regularity condition which characterize the functional space of our model and are needed to evaluate the bias. The hypotheses (H9) and (H10) are technical conditions and are also similar to those considered in Ferraty and Vieu (2006) for the regression case.

Theorem 3.8.

Under the hypotheses (H1)-(H2), (H5) and (H8)-(H10), we have

$$\begin{aligned}\mathbb{E}\left(\widehat{F}^{(j)}(y|x) - F^{(j)}(y|x)\right)^2 &= B^2(H,F^{(j)})h_H^4 + B^2(K,F)h_K^2 + \frac{V_{HK}(x,y)}{\widehat{\mathbf{n}}h_H^{2j-1}\phi(h_K)} \\ &\quad + o(h_H^4) + o(h_K^2) + o\left(\frac{1}{\widehat{\mathbf{n}}h_H^{2j-1}\phi(h_K)}\right), \qquad (13)\end{aligned}$$

with

$$B(H,F^{(j)}) = \frac{1}{2}\frac{\partial^2 F^{(j)}(y|x)}{\partial y^2}\int t^2 H'(t)dt,$$

$$B(K,F) = h_K \Psi_0'(0)\frac{\left(K(1) - \int_0^1 (sK(s))'\beta_x(s)ds\right)}{\left(K(1) - \int_0^1 K'(s)\beta_x(s)ds\right)},$$

and

$$V_{HK}(x,y) = \frac{\beta_2 F^{(j)}(y|x)(1-F^{(j)}(y|x))}{(\beta_1^2(1-F^x(y)))} \ \text{(with } \beta_j = K^j(1) - \int_0^1 (K^j)'(s)\beta_x(s)ds, \text{ for } j = 1,\ 2).$$

Proof:

By using the same decomposition used in Theorem 3.1 and Theorem 3.2 in Rabhi et al. (2013), pp. 409, we show that the proof of Theorem 3.8 can be deduced from the following intermediate results:

Lemma 3.9.

Under the hypotheses of Theorem 3.8, we have

$$\mathbb{E}\left(\widehat{F}_N^{(j)}(y|x)\right) - F^{(j)}(y|x) = B(H,F^{(j)})h_H^2 + B(K,F^{(j)})h_K + o(h_H^2) + o(h_K).$$

Remark 3.10.

Observe that the result of this lemma permits to write

$$\left(\mathbb{E}\widehat{F}_N^{(j)}(y|x) - F^{(j)}(y|x)\right) = \mathcal{O}(h_H^2) + \mathcal{O}(h_K).$$

Lemma 3.11.

Under the hypotheses of Theorem 3.8, we have

$$Var\left(\widehat{F}_N^{(j)}(y|x)\right] = \frac{\sigma^2_{F^{(j)}}(x,y)}{\widehat{\mathbf{n}}h_H^{2j-1}\phi_x(h_K)} + o\left(\frac{1}{\widehat{\mathbf{n}}h_H^{2j-1}\phi_x(h_K)}\right),$$

and

$$Var\left(\widehat{F}_D(x)\right) = o\left(\frac{1}{\widehat{\mathbf{n}}h_H\phi_x(h_K)}\right),$$

where $\sigma^2_{F^{(j)}}(x,y) := F^{(j)}(y|x)(1-F^{(j)}(y|x))\int H'^2(t)dt$.

Lemma 3.12.

Under the hypotheses of Theorem 3.8, we have

$$Cov\left(\widehat{F}_N^{(j)}(y|x), \widehat{F}_D(x)\right) = o\left(\frac{1}{\widehat{\mathbf{n}}h_H^{2j-1}\phi_x(h_K)}\right).$$

Remark 3.13.

It is clear that the results of Lemma 3.9 and Lemma 3.12 allow us to write

$$Var\left[\widehat{F}_D(x) - \widehat{F}_N^{(j)}(y|x)\right] = o\left(\frac{1}{\widehat{\mathbf{n}}h_H^{2j-1}\phi_x(h_K)}\right).$$

□

Now, let $\widehat{F}(\cdot|x) = \widehat{F}^{(0)}(\cdot|x)$ and $\widehat{f}(\cdot|x) = \widehat{F}^{(1)}(\cdot|x)$. It is obvious that the previous theorem allows us to get the two following corollaries.

Corollary 3.14.

Under the hypotheses of Theorem 3.8, we have

$$\mathbb{E}\left(\widehat{f}_N(y|x)\right) - f(y|x) = B_H^f(x,y)h_H^2 + B_K^f(x,y)h_K + o(h_H^2) + o(h_K),$$

and

$$\mathbb{E}\left(\widehat{F}_N(y|x)\right) - F(y|x) = B_H^F(x,y)h_H^2 + B_K^F(x,y)h_K + o(h_H^2) + o(h_K).$$

Remark 3.15.

Observe that the result of this lemma permits us to write

$$\left(\mathbb{E}\widehat{F}_N(y|x) - F(y|x)\right) = \mathcal{O}(h_H^2) + \mathcal{O}(h_K),$$

and

$$\left(\mathbb{E}\widehat{f}_N(y|x) - f(y|x)\right) = \mathcal{O}(h_H^2) + \mathcal{O}(h_K).$$

Corollary 3.16.

Under the hypotheses of Theorem 3.8, we have

$$Var\left(\widehat{f}_N(y|x)\right) = \frac{\sigma_f^2(x,y)}{\widehat{\mathbf{n}}h_H\phi_x(h_K)} + o\left(\frac{1}{\widehat{\mathbf{n}}h_H\phi_x(h_K)}\right),$$

where $\sigma_f^2(x,y) := f(y|x)\int H'^2(t)dt$.

Corollary 3.17.

Under the hypotheses of Theorem 3.8, we have

$$Cov\left(\widehat{f}_N(y|x), \widehat{F}_D(x)\right) = o\left(\frac{1}{\widehat{\mathbf{n}}h_H\phi_x(h_K)}\right),$$

and

$$Cov\left(\widehat{f}_N(y|x), \widehat{F}_N(y|x)\right) = o\left(\frac{1}{\widehat{\mathbf{n}}h_H\phi_x(h_K)}\right).$$

Remark 3.18.

It is clear that the results of Corollary 3.16 and Corollary 3.17 allow us to write

$$Var\left(\widehat{F}_D(x) - \widehat{F}_N(y|x)\right) = o\left(\frac{1}{\widehat{\mathbf{n}}h_H\phi_x(h_K)}\right).$$

Theorem 3.19.

Under assumptions (H1)-(H2), (H5) and (H8)-(H10) we have

$$\mathbb{E}\left[\widehat{h'}^{x}(y)-h'^{x}(y)\right]^{2}=B_{\mathbf{n}}^{2}(x,y)+\frac{\sigma_{h'}^{2}(x,y)}{\widehat{\mathbf{n}}h_{H}^{3}\phi_{x}(h_{K})}+o(h_{H}^{4}+h_{K})+o\left(\frac{1}{\widehat{\mathbf{n}}\,h_{H}^{3}\phi_{x}(h_{K})}\right),$$

where

$$B_{\mathbf{n}}(x,y)=\frac{(B_{H}^{f'}-h'^{x}(y)B_{H}^{F})h_{H}^{2}+(B_{K}^{f'}-h'^{x}(y)B_{K}^{F})h_{K}}{1-F^{x}(y)},$$

with

$$\begin{aligned}
B_{H}^{f'}(x,y)&=\frac{1}{2}\frac{\partial^{2}f^{x}(y)}{\partial y^{2}}\int t^{2}H''(t)dt,\\
B_{K}^{f'}(x,y)&=h_{K}\Phi_{0}'(0)\frac{\left(K(1)-\int_{0}^{1}(sK'(s))'\beta_{x}(s)ds\right)}{\left(K(1)-\int_{0}^{1}K''s)\beta_{x}(s)ds\right)},\\
B_{H}^{F}(x,y)&=\frac{1}{2}\frac{\partial^{2}F^{x}(y)}{\partial y^{2}}\int t^{2}H'(t)dt,\\
B_{K}^{F}(x,y)&=h_{K}\Psi_{0}'(0)\frac{\left(K(1)-\int_{0}^{1}(sK(s))'\beta_{x}(s)ds\right)}{\left(K(1)-\int_{0}^{1}K'(s)\beta_{x}(s)ds\right)},
\end{aligned}$$

and

$$\sigma_{h'}^{2}(x,y)=\frac{\beta_{2}h^{x}(y)}{(\beta_{1}^{2}(1-F^{x}(y)))}\int(H''(t))^{2}dt\ \ (\beta_{j}=K^{j}(1)-\int_{0}^{1}(K^{j})''(s)\beta_{x}(s)ds,\ \text{for},\ j=1,\,2).$$

Proof:

By using the same decomposition used in (Theorem 3.1 Rabhi et al. (2013); pp. 408), we show that the proof of Theorem 3.19 can be deduced from the following intermediate results:

Lemma 3.20.

Under the hypotheses of Theorem 3.19, we have

$$\mathbb{E}\left[\widehat{f'}_{N}^{x}(y)\right]-f'^{x}(y)=B_{H}^{f'}(x,y)h_{H}^{2}+B_{K}^{f'}(x,y)h_{K}+o(h_{H}^{4})+o(h_{K}),$$

and

$$\mathbb{E}\left[\widehat{F}_{N}^{x}(y)\right]-F^{x}(y)=B_{H}^{F}(x,y)h_{H}^{2}+B_{K}^{F}(x,y)h_{K}+o(h_{H}^{2})+o(h_{K}).$$

Remark 3.21.

Observe that, the result of this lemma permits to write

$$\left[\mathbb{E}\widehat{f'}^x_N(y) - f'^x(y)\right] = \mathcal{O}(h_H^4 + h_K).$$

Lemma 3.22.

Under the hypotheses of Theorem 3.19, we have

$$Var\left[\widehat{f'}^x_N(y)\right] = \frac{\sigma^2_{f'}(x,y)}{\widehat{\mathbf{n}}h_H^3\phi_x(h_K)} + o\left(\frac{1}{\widehat{\mathbf{n}}h_H^3\phi_x(h_K)}\right),$$

$$Var\left[\widehat{F}^x_N(y)\right] = o\left(\frac{1}{\widehat{\mathbf{n}}h_H\phi_x(h_K)}\right),$$

and

$$Var\left[\widehat{F}^x_D\right] = o\left(\frac{1}{\widehat{\mathbf{n}}h_H\phi_x(h_K)}\right),$$

where $\sigma^2_{f'}(x,y) := f^x(y)\int(H''(t))^2dt$.

Lemma 3.23.

Under the hypotheses of Theorem 3.19, we have

$$Cov(\widehat{f'}^x_N(y), \widehat{F}^x_D)o\left(\frac{1}{\widehat{\mathbf{n}}h_H^3\phi_x(h_K)}\right),$$

$$Cov(\widehat{f'}^x_N(y), \widehat{F}^x_N(y)) = o\left(\frac{1}{\widehat{\mathbf{n}}h_H^3\phi_x(h_K)}\right),$$

and

$$Cov(\widehat{F}^x_D, \widehat{F}^x_N(y)) = o\left(\frac{1}{\widehat{\mathbf{n}}h_H\phi_x(h_K)}\right).$$

□

Remark 3.24.

1. ***Notes on non-parametric model.*** In this paper, we chose a condition of derivability as our goal is to find an expression for the rate of convergence explicitly, asymptotically exact and keeps the usual form of the squared error (see Vieu (1991)). However, if one proceeds by a Lipschitz condition for example the conditional density of type:

$$\forall(y_1, y_2) \in N_y \times N_y, \forall(x_1, x_2) \in N_x \times N_x,$$

$$|f^{x_1}(y_1) - f^{x_2}(y_2)| \leq A_x((d(x_1, x_2)^2) + |y_1 - y_2|^2),$$

which is less restrictive, we obtain a result for the conditional distribution and conditional density respectively for example of type

$$\mathbb{E}\left[(\widehat{F}_Y^X(x,y) - F_Y^X(x,y))^2\right] = O(h_H^4 + h_K^4) + O\left(\frac{1}{n\phi(h_k)}\right),$$
$$\mathbb{E}\left[(\widehat{f}_Y^X(x,y) - f_Y^X(x,y))^2\right] = O(h_H^4 + h_K^4) + O\left(\frac{1}{nh_H\phi(h_k)}\right).$$

But for such an expression (implicitly) the rate of convergence will not allow us to properly determine the smoothing parameter. In other words, this condition of differentiability is a good compromise to obtain an explicit expression for the rate of convergence. Note that this condition is often taken in the case of finite dimension.

2. ***Notes on the squared error.*** The "dimensionality" of the observations (resp. model) is used in the expression of the rate of convergence (13). We find the "dimensionality" of the model in the way, while the "dimensionality" of the variable in the functional dispersion bias the property of concentration of the probability measure of the functional variable which is closely related to the topological structure of the functional space of the explanatory variable. Our asymptotic results highlight the importance of the concentration properties on small balls of the probability measure of the underlying functional variable. This highlights the role of semi-metric the quality of our estimate. A suitable choice of this parameter allows us to an interesting solution to the problem of the curse of dimensionality (see Rabhi et al. (2013). Another argument has a dramatic effect in our estimation. This is the smoothing parameter h_K (resp. h_H). The term of our rate of convergence, decomposed into two main parts, part bias proportional to h_K (resp. h_H), and part dispersion inversely proportional to h_K (resp. h_H)(ϕ is an increasing function depending on the h_K), makes this relatively easy choice minimizing the main part of this expression to determine this parameter.

4. Applications

In this section we emphasize the potential impact of our work by studying its practical interest in some important statistical problems. Moreover, in order to show the easy implementation of our approach on a concrete case, we discuss in the second part of this section the practical utilization of our model in risk analysis.

The choice of the smoothing parameter has paramount importance in estimating the kernel method. The rate of convergence given in the previous theorem allows us to make an optimal choice of this parameter. Indeed, just choose a smoothing parameter that minimizes the error mean square given by the above theorem (see Youndjé (1993) for the real case in mean of order p). We can also use another method of selection such as cross-validation.

Remember that the choice of the smoothing parameter for estimating the hazard function in the functional framework remains an open question, thus the use of optimal parameters of the conditional density and without theoretical validity, but it is justified by the close relationship between the conditional density, conditional distribution function and the conditional hazard function. Contrary to the statistical vector, the performance of estimation of the hazard function in functional statistics also depends on another additional argument that is the metric of the functional space of the explanatory variable. In practice the appropriate choice of this object plays a crucial role for the effectiveness of the model. Of the same as the smoothing parameter, the optimal choice of this factor is still an open question in this context of the statistic nonparametric functional. However, the choice from the family of semi-metric are used to sense if the curves are very smooth. The semi-metrics defined by the derivatives can be used. On the other hand, if the curves present a discontinuity, we make a call to the metric of the functional ACP. We refer to the package NFDA in R for codes of these semi-metrics. In our case, it is clear that the second family from the semi-metric is the most appropriate.

The bounds obtained here allow us to derive the same optimal rates of convergence as in the i.i.d case only if the space $\mathcal{F}$ is of finite dimension (for example $\mathbb{R}^d$). In the case of infinite dimensional space, as in the spatial case, finding an optimal rate is far from being proved (we refer for example Ferraty and Vieu (2006) for more discussion about this question).

A. Some derivatives

• *On the choices of the bandwidths parameters.* As all smoothing by a kernel method, the choice of bandwidths parameters has a crucial role in determining the performance of the estimators. The mean quadratic error given in Theorem 3.19 is a basic ingredient to solve this problem. Usually, the ideal theoretical choices are obtained by minimizing this error. Here, we have explicated its leading term which is

$$B_{\mathbf{n}}^2(x,y) + \frac{\sigma_{h'}^2(x,y)}{\widehat{\mathbf{n}} h_H^3 \phi_x(h_K)}.$$

Then the smoothing parameters minimizing this leading term are asymptotically optimal with respect the L^2-error. However the practical utilization of this criterium requires some additional computational efforts. More precisely, it requires the estimation of the unknown quantities Ψ_0', Φ_0', $f'^x(y)$ and $F^x(y)$. Clearly, all these estimations can be obtained by using pilots estimators of the conditional distribution function $F^x(y)$ and of the conditional density $f'^x(y)$. Such estimations are possible by using the kernel methods, with a separate choice of the bandwidths parameters between both models. More preciously, for the conditional density, we propose to adopt, to the functional case, the bandwidths selectors studied by Bouraine et al. (2010) by considering the following criterion

$$CVPDF = \frac{1}{\widehat{\mathbf{n}}} \sum_{\mathbf{i} \in I_{\mathbf{n}}} W_1(X_{\mathbf{i}}) \int \widehat{f'}^{X_{\mathbf{i}}^{-\mathbf{i}^2}}(y) W_2(y) dy - \frac{2}{\widehat{\mathbf{n}}} \sum_{\mathbf{i} \in I_{\mathbf{n}}} \widehat{f'}^{X_{\mathbf{i}}^{-\mathbf{i}}}(Y_{\mathbf{i}}) W_1(X_{\mathbf{i}}) W_2(Y_{\mathbf{i}}), \qquad (14)$$

while for the the conditional distribution function we can use the cross-validation rule proposed by De Gooijer and Gannoun (2000) (in vectorial case)

$$CVCDF = \frac{1}{\widehat{\mathbf{n}}} \sum_{\mathbf{k},\mathbf{l} \in I_{\mathbf{n}}} \left[\mathbb{I}_{Y_{\mathbf{k}} \leq Y_{\mathbf{l}}} - \widehat{F}^{X_{\mathbf{k}}^{-\mathbf{k}}}(Y_{\mathbf{l}}) \right]^2 W(X_{\mathbf{k}}),$$

where W_1, W_2 and W are some suitable trimming functions and

$$\widehat{F}^{X_{\mathbf{k}}^{-\mathbf{k}}}(Y_{\mathbf{l}}) = \frac{\sum_{\mathbf{i} \in I_{\mathbf{n},\varsigma_{\mathbf{n}}}^{\mathbf{k},\mathbf{l}}} K(h_K^{-1} d(X_{\mathbf{k}}, X_{\mathbf{i}})) H(h_H^{-1}(Y_{\mathbf{l}} - Y_{\mathbf{i}}))}{\sum_{\mathbf{i} \in I_{\mathbf{n},\varsigma_{\mathbf{n}}}^{\mathbf{k},\mathbf{l}}} K(h_K^{-1} d(X_{\mathbf{k}}, X_{\mathbf{i}}))},$$

and

$$\widehat{f'}^{X_{\mathbf{i}}^{-\mathbf{i}}}(y) = \frac{h_H^{-2} \sum_{\mathbf{j} \in I_{\mathbf{n},\varsigma_{\mathbf{n}}}^{\mathbf{i}}} K(h_K^{-1} d(X_{\mathbf{i}}, X_{\mathbf{j}})) H''(h_H^{-1}(y - Y_{\mathbf{j}}))}{\sum_{\mathbf{j} \in I_{\mathbf{n},\varsigma_{\mathbf{n}}}^{\mathbf{i}}} K(h_K^{-1} d(X_{\mathbf{i}}, X_{\mathbf{j}}))},$$

with

$$I_{\mathbf{n},\varsigma_{\mathbf{n}}}^{\mathbf{k},\mathbf{l}} = \{\mathbf{i} \text{ such that } \; \|\mathbf{i} - \mathbf{k}\| \geq \varsigma_{\mathbf{n}} \; \text{ and } \; \|\mathbf{i} - \mathbf{l}\| \geq \varsigma_{\mathbf{n}}\} \text{ and } I_{\mathbf{n},\varsigma_{\mathbf{n}}}^{\mathbf{i}} = \{\mathbf{j} \text{ such that } \; \|\mathbf{j} - \mathbf{i}\| \geq \varsigma_{\mathbf{n}} \; \}.$$

Of course, we can also adopt another selection method such that the parametric bootstrap method, proposed by Hall et al. (1999) and Hyndman et al. (1996), respectively, for the conditional cumulative distribution function and the conditional density in the finite dimensional case. Nevertheless, a data-driven method allows to overcome this additional computation is very important in practice and is one of the natural prospects of the present work.

• *Confidence intervals.* The main application from the result of asymptotic normality is to build confidence band for the true value of $h'^x(y)$.

Similar to the previous application, the practical utilization of our result in this topic requires the estimation of the quantity $\sigma_{h'}^2(x, y)$. A plug-in estimate for the asymptotic standard deviation $\sigma_{h'}^2(x, y)$ can be obtained by using the estimators $\widehat{f'}^x(y)$ and $\widehat{F}^x(y)$ of $f'^x(y)$ and $F^x(y)$. Then we get

$$\widehat{\sigma}_{h'}^2(x, y) := \frac{\widehat{\beta}_2 \widehat{f'}^x(y)}{\left(\widehat{\beta}_1^{\,2} (1 - \widehat{F}^x(y))^2 \right)},$$

where

$$\begin{cases} \widehat{\beta}_1 = \dfrac{1}{\widehat{\mathbf{n}}\phi_x(h_K)} \displaystyle\sum_{\mathbf{i}\in I_{\mathbf{n}}} K(h_K^{-1}d(x, X_{\mathbf{i}})) \\ \text{and} \quad \widehat{\beta}_2 = \dfrac{1}{\widehat{\mathbf{n}}\phi_x(h_K)} \displaystyle\sum_{\mathbf{i}\in I_{\mathbf{n}}} K^2(h_K^{-1}d(x, X_{\mathbf{i}})). \end{cases}$$

Clearly the function $\phi_x(\cdot)$ does not appear in the calculation of the confidence interval by simplification. More precisely, we obtain the following approximate $(1-\zeta)$ confidence band for $h'^x(y)$

$$\widehat{h}'^x(y) \pm t_{1-\zeta/2} \times \left(\frac{\widehat{\sigma}^2_{h'}(x,y)}{\widehat{\mathbf{n}} h_H^3 \phi_x(h_K)} \right)^{1/2},$$

where $t_{1-\zeta/2}$ denotes the $1-\zeta/2$ quantile of the standard normal distribution.

B. Application to continuously indexed random fields

Clearly, a continuously indexed random filed $(Z_t, t \in \mathbb{R}^N)$ is one of the important examples of functional spatial data.

Indeed, let $(Z_t, t \in \mathbb{R}^N)$ be a $\mathbb{R}$-valued strictly stationary random spatial process assumed to be bounded and observed over some subset $I \subset \mathbb{R}^N$. Then our approach can be used to predict the value Z_{s_0} at an unobserved location $s_0 \in I$ by taking into account the observed part of the process $(Z_t, t \in \mathbb{R}^N)$ in its continuous form. For this, we suppose that the value of Z_{s_0} depends only on the values of the process (Z_t) in a bounded neighborhood $\mathcal{V}_{s_0} \subset I$ of s_0. From Z_t we may construct m functional spatial random variables as follows: Consider some grid $\mathcal{G}_n = \{t_i = (t_{i,1}, \ldots, t_{i,N}) \in I,\ 1 \le t_{i,j} \le n_j,\ j = 1, \ldots, N,\ i = 1, \ldots, m\}$ such that

$$\forall i = 1, \ldots, m, \ \min_{1\le j\le N-1} (t_{i,j+1} - t_{i,j}) \ge C > 0 \quad \text{for some constant } C,$$

and we define

$$\forall i = 1, \ldots, m,\ X_{t_i} = (Z_t,\ t \in \mathcal{V}_{t_i}),$$

where $\mathcal{V}_{t_i} = \mathcal{V} + t_{t_i}$, $\mathcal{V} = \mathcal{V}_{s_0} - s_0$, which does not contain 0. So the predictor that we proposed (see Biau and Cadre (2004) and Dabo-Niang and Yao (2007) for the finite dimension mean regression case) aims to evaluate a real characteristic denoted $Y_{s_0} = Z_{s_0}$, at a site s_0, given $X_{s_0} = (Z_t,\ t \in \mathcal{V}_{s_0})$.

5. Proofs of Technical Lemmas

In the following, we will denote $\forall \mathbf{i}$

$$K_{\mathbf{i}} = K(h_H^{-1}d(x, X_{\mathbf{i}})), \quad H_{\mathbf{i}} = H(h_H^{-1}(y-Y_{\mathbf{i}}), \quad H'_{\mathbf{i}} = H'(h_H^{-1}(y-Y_{\mathbf{i}}) \quad \text{and} \quad H''_{\mathbf{i}} = H''(h_H^{-1}(y-Y_{\mathbf{i}}).$$

First of all, we state the following lemmas which are due to Carbon et al. (1997). They are needed for the convergence of our estimates. There proofs will then be omitted.

Lemma 6.1.

Suppose $E_1, \ldots, E_r$ are sets containing m sites each with $dist(E_i,\ E_j) \geq \gamma$ for all $i \neq j$ where $1 \leq i \leq r$ and $1 \leq j \leq r$. Suppose $Z_1, \ldots, Z_r$ is a sequence of real-valued r.v.'s measurable with respect to $\mathcal{B}(E_1), \ldots, \mathcal{B}(E_r)$, respectively, and Z_i takes values in $[a, b]$. Then there exists a sequence of independent r.v.'s $Z_1^*, \ldots, Z_r^*$ independent of $Z_1, \ldots, Z_r$ such that Z_i^* has the same distribution as Z_i and satisfies

$$\sum_{i=1}^{r} E|Z_i - Z_i^*| \leq 2r(b-a)\psi((r-1)m, m)\varphi(\gamma).$$

Lemma 6.2.

(i) Suppose that (1) holds. Denote by $\mathcal{L}_r(\mathcal{F})$ the class of $\mathcal{F}-$measurable r.v.'s X satisfying $\|X\|_r = (E|X|^r)^{1/r} < \infty$. Suppose $X \in \mathcal{L}_r(\mathcal{B}(E))$ and $Y \in \mathcal{L}_s(\mathcal{B}(E'))$. Assume also that $1 \leq r,\ s,\ t < \infty$ and $r^{-1} + s^{-1} + t^{-1} = 1$. Then,

$$|\mathbb{E}XY - \mathbb{E}X\mathbb{E}Y| \leq C\|X\|_r\|Y\|_s\{\psi(Card(E), Card(E'))\varphi(dist(E, E'))\}^{1/t}. \tag{15}$$

(ii) For r.v.'s bounded with probability 1, the right-hand side of (15) can be replaced by

$$C\psi(Card(E), Card(E'))\varphi(dist(E, E')).$$

We also need the following lemma due to Nakhapeytyan (1987).

Lemma 6.3.

Let, $Z_1 \ldots Z_n$ be a random vector such that $\left|\mathbb{E}\prod_{s=i}^{n} Z_s\right| < \infty,\ i = 1, \ldots, n-1,\ |Z_i| \leq C$, $i = 1, \ldots, n$. Then,

$$\left|\mathbb{E}\prod_{s=1}^{n} Z_s - \prod_{s=1}^{n}\mathbb{E}Z_s\right| \leq \sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\left|\mathbb{E}(Z_i - 1)(Z_j - 1)\prod_{s=j+1}^{n} Z_s - \mathbb{E}(Z_i - 1)\mathbb{E}(Z_j - 1)\prod_{s=j+1}^{n} Z_s\right|.$$

Proof of Lemma 3.3:

Let $H_{\mathbf{i}}^{(j)}(y) = H^{(j)}\left(h_H^{-1}(y - Y_{\mathbf{i}})\right)$ and $\forall\ y \in \mathcal{S}_{\mathbb{R}}$, we have

$$\begin{aligned} \left|\mathbb{E}\widehat{F}_N^{(j)}(y|x) - F^{(j)}(y|x)\right| &= \Big|\frac{h_H^{-j}}{\mathbb{E}\left(K_{\mathbf{1}}(x)\right)}\mathbb{E}\left\{K\left(h_K^{-1}d(x, X_{\mathbf{1}})\right)\mathbb{E}\left(H^{(j)}\left(h_H^{-1}(y - Y_{\mathbf{1}})|X_{\mathbf{1}}\right)\right)\right\} \\ &\qquad -F^{(j)}(y|x)\Big| \\ &= \frac{1}{h_H^j\mathbb{E}\left(K_{\mathbf{1}}(x)\right)}\Big|\mathbb{E}\Big\{K_{\mathbf{1}}(x)\mathbb{I}_{B(x,h_K)}(X)\Big(\mathbb{E}\left(H_1^{(j)}(y)|X\right) \\ &\qquad -F^{(j)}(y|x)\Big)\Big\}\Big|. \end{aligned} \tag{16}$$

Moreover, we have

$$\mathbb{E}\left(H_1^{(j)}(y)|X_{\mathbf{1}}\right) = \int_{\mathbb{R}} H^{(j)}\left(h_H^{-1}(y - v)\right) f(v|X_{\mathbf{1}})dv,$$

and integrating by parts, we obtain that

$$\begin{aligned} \mathbb{E}\left(H_{\mathbf{1}}^{(j)}(y)|X_{\mathbf{1}}\right) &= -\sum_{l=1}^{j} h_H^l\left[H^{(j-l)}\left(h_H^{-1}(y - v)\right)F^{(l)}(v|X_{\mathbf{1}})\right]_{-\infty}^{+\infty} \\ &\quad + h_H^{j-1}\int_{\mathbb{R}} H^{(1)}\left(h_H^{-1}(y - v)\right)F^{(j)}(v|X_{\mathbf{1}})dv. \end{aligned} \tag{17}$$

Condition (H6) allows us to cancel the first term in the right side of (17) and considering the usual change of variable $t = \frac{y-v}{h_H}$, we can write:

$$\left|\mathbb{E}\left(H_{\mathbf{1}}^{(j)}(y)|X_{\mathbf{1}}\right) - h_H^j F^{(j)}(y|x)\right| \leq h_H^j \int_{\mathbb{R}} H^{(1)}(t)\left|F^{(j)}(y - h_H t|X_{\mathbf{1}}) - F^{(j)}(y|x)\right| dt.$$

Finally, (H3) allows to write $\forall y \in \mathcal{S}_{\mathbb{R}}$

$$\mathbb{I}_{B(x,h_K)}(X)\left|\mathbb{E}\left(H_{\mathbf{1}}^{(j)}(y)|X\right) - h_H^j F^{(j)}(y|x)\right| \leq C h_H^j \int_{\mathbb{R}} H^{(1)}(t)\left(h_K^{b_1} + |t|^{b_2}h_H^{b_2}\right)dt. \tag{18}$$

The use of (H5), (16) and Lemma 3.2 achieve the proof of Lemma 3.3. □

Proof of Lemma 3.4:

Using the compactness of $\mathcal{S}_{\mathbb{R}}$, we can write that $\mathcal{S}_{\mathbb{R}} \subset \bigcup_{k=1}^{z_{\mathbf{n}}}(t_k - l_{\mathbf{n}}, t_k + l_{\mathbf{n}})$. Take $k(y) = \arg\min_{k\in\{1,2,\ldots,z_{\mathbf{n}}\}} |y - t_k|$, $z_{\mathbf{n}} \leq \widehat{\mathbf{n}}^{\frac{3}{2}\alpha+\frac{1}{2}}$. Thus, we obtain this decomposition

$$\begin{aligned}\left|\widehat{F}_N^{(j)}(y|x)-\mathbb{E}\widehat{F}_N^{(j)}(y|x)\right| &\leq \underbrace{\left|\widehat{F}_N^{(j)}(y|x)-\widehat{F}_N^{(j)}(t_{k(y)}|x)\right|}_{T_1}+\underbrace{\left|\widehat{F}_N^{(j)}(t_{k(y)}|x)-\mathbb{E}\widehat{F}_N^{(j)}(t_{k(y)}|x)\right|}_{T_2}\\ &+ \underbrace{\left|\mathbb{E}\widehat{F}_N^{(j)}(t_{k(y)}|x)-\mathbb{E}\widehat{F}_N^{(j)}(y|x)\right|}_{T_3}. \end{aligned} \tag{19}$$

• Concerning (T_1):

$$\begin{aligned}\left|\widehat{F}_N^{(j)}(y|x)-\widehat{F}_N^{(j)}(t_{k(y)}|x)\right| &\leq \frac{1}{\widehat{\mathbf{n}}h_H^j\mathbb{E}\left(K_{\mathbf{1}}(x)\right)}\sum_{\mathbf{i}\in I_{\mathbf{n}}}K_{\mathbf{i}}(x)\left|H_{\mathbf{i}}^{(j)}(y)-H_{\mathbf{i}}^{(j)}(t_{k(y)})\right|\\ &\leq C\frac{h_H^{-j}}{\widehat{\mathbf{n}}\mathbb{E}\left(K_{\mathbf{1}}(x)\right)}\sum_{\mathbf{i}\in I_{\mathbf{n}}}K_{\mathbf{i}}(x)\left|H_{\mathbf{i}}^{(j)}(y)-H_{\mathbf{i}}^{(j)}(t_{k(y)})\right|\\ &\leq \frac{Ch_H^{-j}|y-t_{k(y)}|}{h_H}\widehat{\mathbf{n}}\mathbb{E}\left[K_{\mathbf{1}}(x)\right]\sum_{\mathbf{i}\in I_{\mathbf{n}}}K_{\mathbf{i}}(x)\\ &\leq C\frac{h_H^{-j}l_{\mathbf{n}}}{h_H}\widehat{F}_D(x)\\ &\leq C\frac{h_H^{-j}l_{\mathbf{n}}}{h_H}. \end{aligned} \tag{20}$$

The second inequality is obtained by considering a Lipschitz argument whereas the last one comes from the definition of $\widehat{F}_D(x)$.

Take now $l_{\mathbf{n}}=\widehat{\mathbf{n}}^{-\frac{3\alpha}{2}-\frac{1}{2}}$ and note that, because of (H7), we have $\dfrac{l_{\mathbf{n}}}{h_H^{j+1}}=o\left(\sqrt{\dfrac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,h_H^{2j-1}\phi(h_K)}}\right)$.

Thus the almost complete convergence of $\widehat{F}_D(x)$, we can write

$$\left|\widehat{F}_N^{(j)}(y|x)-\widehat{F}_N^{(j)}(t_{k(y)}|x)\right|=o_{a.co.}\left(\sqrt{\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,h_H^{2j-1}\phi(h_K)}}\right). \tag{21}$$

• Concerning (T_2): The proof is based on ideas similar to those used by Carbon et al. (1997).

We have $\forall\eta>0$

$$
\begin{aligned}
&\mathbb{P}\left(\sup_{y\in\mathcal{S}_{\mathbb{R}}}\left|\widehat{F}_N^{(j)}(t_{k(y)}|x)-\mathbb{E}\widehat{F}_N^{(j)}(t_{k(y)}|x)\right|>\eta\sqrt{\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,h_H^{2j-1}\phi(h_K)}}\right)\\
&\quad= P\left(\max_{k\in\{1,2,\ldots,z_{\mathbf{n}}\}}\left|\widehat{F}_N^{(j)}(t_k|x_k)-\mathbb{E}\widehat{F}_N^{(j)}(t_k|x_k)\right|>\eta\sqrt{\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,h_H^{2j-1}\phi(h_K)}}\right)\\
&\quad\le z_{\mathbf{n}}\max_{k\in\{1,2,\ldots,z_{\mathbf{n}}\}}\mathbb{P}\left(\left|\widehat{F}_N^{(j)}(t_k|x_k)-\mathbb{E}\widehat{F}_N^{(j)}(t_k|x_k)\right|>\eta\sqrt{\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,h_H^{2j-1}\phi(h_K)}}\right).
\end{aligned}
$$

Let, $\Gamma_{\mathbf{i}}^{(j)}(x,t_k)=K_{\mathbf{i}}(x)H_{\mathbf{i}}^{(j)}(t_k)-\mathbb{E}\left[K_{\mathbf{i}}(x)H_{\mathbf{i}}^{(j)}(t_k)\right]$, then

$$
\widehat{F}_N^{(j)}(t_k|x)-\mathbb{E}\widehat{F}_N^{(j)}(t_k|x)=\frac{h_H^{-j}}{\widehat{\mathbf{n}}\mathbb{E}\left(K_{\mathbf{1}}(x)\right)}\sum_{\mathbf{i}\in I_{\mathbf{n}}}\Gamma_{\mathbf{i}}^{(j)}(x,t_k).
$$

Consider, the spatial block decomposition on the random variables $\Gamma_{\mathbf{i}}^{(j)}(x,t_k)=K_{\mathbf{i}}(x)H_{\mathbf{i}}^{(j)}(t_k)-\mathbb{E}\left[K_{\mathbf{i}}(x)H_{\mathbf{i}}^{(j)}(t_k)\right]$ for fixed integer $p_{\mathbf{n}}$, (depending on $\mathbf{n}$) as follows

$$
\begin{aligned}
U(1,\mathbf{n},x,t_k,\mathbf{m})&=\textstyle\sum_{\substack{i_k=2m_kp_{\mathbf{n}}+1\\k=1,\ldots,N}}^{2m_kp_{\mathbf{n}}+p_{\mathbf{n}}}\Gamma_{\mathbf{i}}^{(j)}(x,t_k),\\
U(2,\mathbf{n},x,t_k,\mathbf{m})&=\sum_{\substack{i_k=2m_kp_{\mathbf{n}}+1\\k=1,\ldots,N-1}}^{2m_kp_{\mathbf{n}}+p_{\mathbf{n}}}\ \sum_{i_N=2m_Np_{\mathbf{n}}+p_{\mathbf{n}}+1}^{2(m_N+1)p_{\mathbf{n}}}\Gamma_{\mathbf{i}}^{(j)}(x,t_k),\\
U(3,\mathbf{n},x,t_k,\mathbf{m})&=\sum_{\substack{i_k=2m_kp_{\mathbf{n}}+1\\k=1,\ldots,N-2}}^{2m_kp_{\mathbf{n}}+p_{\mathbf{n}}}\ \sum_{i_{N-1}=2m_{N-1}p_{\mathbf{n}}+p_{\mathbf{n}}+1}^{2(m_{N-1}+1)p_{\mathbf{n}}}\ \sum_{i_N=2m_Np_{\mathbf{n}}+1}^{2m_Np_{\mathbf{n}}+p_{\mathbf{n}}}\Gamma_{\mathbf{i}}^{(j)}(x,t_k),\\
U(4,\mathbf{n},x,t_k,\mathbf{m})&=\sum_{\substack{i_k=2m_kp_{\mathbf{n}}+1\\k=1,\ldots,N-2}}^{2m_kp_{\mathbf{n}}}\ \sum_{i_{N-1}=2m_{N-1}p_{\mathbf{n}}+p_{\mathbf{n}}+1}^{2(m_{N-1}+1)p_{\mathbf{n}}}\ \sum_{i_N=2m_Np_{\mathbf{n}}+p_{\mathbf{n}}+1}^{2(m_N+1)p_{\mathbf{n}}}\Gamma_{\mathbf{i}}^{(j)}(x,t_k),
\end{aligned}
$$

and so an. Finally,

$$
\begin{aligned}
U(2^{N-1},\mathbf{n},x,t_k,\mathbf{m})&=\sum_{\substack{i_k=2m_kp_{\mathbf{n}}+p_{\mathbf{n}}+1\\k=1,\ldots,N-1}}^{2(m_k+1)p_{\mathbf{n}}}\ \sum_{i_N=2m_Np_{\mathbf{n}}+1}^{2m_Np_{\mathbf{n}}+p_{\mathbf{n}}}\Gamma_{\mathbf{i}}^{(j)}(x,t_k),\\
U(2^{N},\mathbf{n},x,t_k,\mathbf{m})&=\sum_{\substack{i_k=2m_kp_{\mathbf{n}}+p_{\mathbf{n}}+1\\k=1,\ldots,N}}^{2(m_k+1)p_{\mathbf{n}}}\Gamma_{\mathbf{i}}^{(j)}(x,t_k).
\end{aligned}
$$

This blocking scheme is similar to that used in Tran (1990).

Now, for $\mathcal{M} = \{0, \ldots, r_1 - 1\} \times \ldots \times \{0, \ldots, r_N - 1\}$ where, $r_i = 2^{-1} n_i p_{\mathbf{n}}^{-1}, i = 1, \ldots, N$, we define

$$T(\mathbf{n}, x, t_k, i) = \sum_{\mathbf{m} \in \mathcal{M}} U(i, \mathbf{n}, x, t_k, \mathbf{j}),$$

and we write,

$$\widehat{F}_N^{(j)}(t_k|x_k) - \mathbb{E}\widehat{F}_N^{(j)}(t_k|x_k) = \frac{h_H^{-j}}{\widehat{\mathbf{n}}\mathbb{E}\left[K_{\mathbf{1}}(x)\right]} \sum_{i=1}^{2^N} T(\mathbf{n}, x, t_k, i). \tag{22}$$

Note that, as raised by Biau and Cadre (2004), if one does not have the equalities $n_i = 2r_i p_{\mathbf{n}}$, the term, say $T(\mathbf{n}, x, t_k, 2^N + 1)$ (which contains the $\Gamma_{\mathbf{i}}^{(j)}(x, t_k)$'s at the ends not included in the blocks above) can be added. This will not change the proof a lot.

Now, under the last equation (22), we can say that, for all $\eta > 0$

$$\mathbb{P}\left(\left|\widehat{F}_N^{(j)}(t_k|x_k) - \mathbb{E}\widehat{F}_N^{(j)}(t_k|x_k)\right| \geq \eta\right) \leq 2^N \max_{i=1,\ldots} \mathbb{P}\left(T(\mathbf{n}, x, t_k, i) \geq \eta\widehat{\mathbf{n}} h_H^j \mathbb{E}\left[K_{\mathbf{1}}(x)\right]\right).$$

Finally, the desired result follows from the evaluation of the following quantities:

$$\mathbb{P}\left(T(\mathbf{n}, x, t_k, i) \geq \eta\widehat{\mathbf{n}} h_H^j \mathbb{E}\left[K_{\mathbf{1}}(x)\right]\right), \qquad \text{for all } i = 1, \ldots, 2^N.$$

Without loss of generality, we will only consider the case $i = 1$. For this case, we enumerate the variable $(U(1, \mathbf{n}, x, t_k, \mathbf{m}); \mathbf{m} \in \mathcal{M})$ and we apply Lemma 4.5 of Carbon et al. (1997) on variables enumerated. The variables with the new enumeration will be noted $Z_1, \ldots Z_L$ where $L = \prod_{k=1}^N r_k = 2^{-N}\widehat{\mathbf{n}} p_{\mathbf{n}}^{-N} \leq \widehat{\mathbf{n}} p_{\mathbf{n}}^{-N}$. Thus for each Z_d there exists a certain $\mathbf{m}_m$ in $\mathcal{M}$ such that

$$Z_d = \sum_{\mathbf{i} \in I(1, \mathbf{n}, x, \mathbf{m}_m)} \Gamma_{\mathbf{i}}^{(j)}(x),$$

where $I(1, \mathbf{n}, x, \mathbf{m}_m) = \{\mathbf{i} : 2m_{k_m} p_{\mathbf{n}} + 1 \leq i_k \leq 2m_{k_m} p_{\mathbf{n}} + p_{\mathbf{n}} \quad ; k = 1, \ldots, N\}$. Clearly the sets $I(1, \mathbf{n}, x, \mathbf{m}_m)$ contain $p_{\mathbf{n}}^N$ sites and are separated by a distance of at least $p_{\mathbf{n}}$. So, according to Lemma 4.5 of Carbon et al. (1997), one can find independent random variables $Z_1^*, \ldots Z_L^*$ having the same law as $(Z_d)_{d=1,\ldots L}$, such that

$$\sum_{d=1}^L \mathbb{E}|Z_d - Z_d^*| \leq 2CLp_{\mathbf{n}}^N \psi\left((L-1)p_{\mathbf{n}}^N, p_{\mathbf{n}}^N\right)\varphi(p_{\mathbf{n}}).$$

Therefore, by the Bernstein and Markov inequalities we have:

$$\mathbb{P}\left(T(\mathbf{n},x,i)\geq \eta\widehat{\mathbf{n}}h_H^j\mathbb{E}\left[K_{\mathbf{1}}(x)\right]\right)\leq B_1+B_2,$$

where

$$B_1=\mathbb{P}\left(\left|\sum_{d=1}^{L}Z_d^*\right|\geq \frac{L\eta\widehat{\mathbf{n}}h_H^j\mathbb{E}\left[K_{\mathbf{1}}(x)\right]}{2L}\right)\leq 2\exp\left(-\frac{(\eta\widehat{\mathbf{n}}h_H^j\mathbb{E}\left[K_{\mathbf{1}}(x)\right])^2}{LVar\left[Z_1^*\right]+Cp_{\mathbf{n}}^N\eta\widehat{\mathbf{n}}h_H^j\mathbb{E}\left[K_{\mathbf{1}}(x)\right]}\right),$$

and

$$\begin{aligned}B_2 &= \mathbb{P}\left(\sum_{d=1}^{L}|Z_d-Z_d^*|\geq \frac{\eta\widehat{\mathbf{n}}h_H^j\mathbb{E}\left[K_{\mathbf{1}}(x)\right]}{2}\right)\\ &\leq \frac{1}{\eta\widehat{\mathbf{n}}h_H^j\mathbb{E}\left[K_{\mathbf{1}}(x)\right]}\sum_{d=1}^{L}\mathbb{E}|Z_d-Z_d^*|\\ &\leq 2Lp_{\mathbf{n}}^N(\eta\widehat{\mathbf{n}}h_H^j\mathbb{E}\left[K_{\mathbf{1}}(x)\right])^{-1}\psi\left((L-1)p_{\mathbf{n}}^N,p_{\mathbf{n}}^N\right)\varphi(p_{\mathbf{n}}).\end{aligned}$$

Since, $\widehat{\mathbf{n}}=2^NLp_{\mathbf{n}}^N$ and $\psi\left((L-1)p_{\mathbf{n}}^N,p_{\mathbf{n}}^N\right)\leq p_{\mathbf{n}}^N$ we get for $\eta=\eta_0\sqrt{\dfrac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,h_H^{2j-1}\phi_x(h_K)}}$,

$$B_2\leq \widehat{\mathbf{n}}p_{\mathbf{n}}^N\left(\log\widehat{\mathbf{n}}\right)^{-1/2}\left(\widehat{\mathbf{n}}h_H^{2j-1}\phi_x(h_K)\right)^{-1/2}\varphi(p_{\mathbf{n}}),$$

with, $p_{\mathbf{n}}=C\left(\frac{\widehat{\mathbf{n}}h_H^{2j-1}\phi_x(h_K)}{\log\widehat{\mathbf{n}}}\right)^{1/2N}$, we can write

$$B_2\leq \left(\log\widehat{\mathbf{n}}\right)^{-1}\widehat{\mathbf{n}}\,\varphi(p_{\mathbf{n}}). \tag{23}$$

Consequently, from (H7), we have

$$\sum_{\mathbf{n}\in I_{\mathbf{n}}}\widehat{\mathbf{n}}\,\varphi(p_{\mathbf{n}})<\infty.$$

Let us focus now on B_1. For this, let us evaluate asymptotically $Var\left[Z_1^*\right]$. Indeed,

$$Var\left[Z_1^*\right]=Var\left[\sum_{\mathbf{i}\in I(1,\mathbf{n},x,\mathbf{1})}\Gamma_{\mathbf{i}}^{(j)}(x)\right]=\sum_{\mathbf{i},\mathbf{i}'\in I(1,\mathbf{n},x,\mathbf{1})}\left|Cov(\Gamma_{\mathbf{i}}^{(j)}(x),\Gamma_{\mathbf{i}'}^{(j)}(x))\right|.$$

Let, $Q_{\mathbf{n}}=\displaystyle\sum_{\mathbf{i}\in I(1,\mathbf{n},x,\mathbf{1})}Var\left[\Gamma_{\mathbf{i}}^{(j)}(x)\right]$ and $R_{\mathbf{n}}=\displaystyle\sum_{\mathbf{i}\neq\mathbf{i}'\in I(1,\mathbf{n},x,\mathbf{1})}\left|Cov(\Gamma_{\mathbf{i}}^{(j)}(x),\Gamma_{\mathbf{i}'}^{(j)}(x))\right|$.

It is clear that under assumption (H6), we have

$$Var[\Gamma_{\mathbf{i}}^{(j)}(x)] \leq h_H^j Var[\Delta_{\mathbf{i}}(x)].$$

As well as by assumptions (H1) and (H2), we have

$$Var[\Gamma_{\mathbf{i}}^{(j)}(x)] = O\left(h_H^j \phi_x(h_K)\right),$$

therefore,

$$Q_{\mathbf{n}} = O\left(p_{\mathbf{n}}^N h_H^j \phi_x(h_K)\right).$$

Concerning $R_{\mathbf{n}}$, we introduce the followings sets:

$$S_1 = \{\mathbf{i}, \mathbf{j} \in I(1, \mathbf{n}, x, \mathbf{1}) : 0 < \|\mathbf{i} - \mathbf{j}\| \leq c_{\mathbf{n}}\}, \qquad S_2 = \{\mathbf{i}, \mathbf{j} \in I(1, \mathbf{n}, x, \mathbf{1}) : \|\mathbf{i} - \mathbf{j}\| > c_{\mathbf{n}}\},$$

where $c_{\mathbf{n}}$ is a real sequence that converges to $+\infty$ and will be made precise later. Split $R_{\mathbf{n}}$ into two separate summations over sites in S_1 and S_2:

$$\begin{aligned} R_{\mathbf{n}} &= \sum_{(\mathbf{i},\mathbf{j}')\in S_1} \left|Cov\left(\Gamma_{\mathbf{i}}^{(j)}(x), \Gamma_{\mathbf{i}'}^{(j)}(x)\right)\right| + \sum_{(\mathbf{i},\mathbf{j})\in S_2} \left|Cov\left(\Gamma_{\mathbf{i}}^{(j)}(x), \Gamma_{\mathbf{i}'}^{(j)}(x)\right)\right| \\ &= R_{\mathbf{n}}^1 + R_{\mathbf{n}}^2. \end{aligned}$$

Note that (H6) and for $j = 1$ the conditional density of $(Y_{\mathbf{i}}, Y_{\mathbf{i}'})$ given $(X_{\mathbf{i}}, X_{\mathbf{i}'})$ is continuous at (t_k, t_k) allow to show that

$$\mathbb{E}\left(H^{(j)}(h_H^{-1}(t_k - Y_{\mathbf{i}}))H^{(j)}(h_H^{-1}(t_k - Y_{\mathbf{i}'}))|(X_{\mathbf{i}}, X_{\mathbf{i}'})\right) = O(h_H^2),$$

while (H1) and (H3) imply that

$$\mathbb{E}\left(H^{(j)}(h_H^{-1}(t_k - Y_{\mathbf{i}}))|X_{\mathbf{i}}\right) = O(h_H).$$

Moreover, on one hand, we have:

$$\begin{aligned} R_{\mathbf{n}}^1 &= \sum_{(\mathbf{i},\mathbf{j})\in S_1} \left|\mathbb{E}\left[K_{\mathbf{i}}(x)H_{\mathbf{i}}^{(j)}(t_k)K_{\mathbf{j}'}(x)H_{\mathbf{i}'}^{(j)}(t_k)\right] - \mathbb{E}\left[K_{\mathbf{i}}(x)H_{\mathbf{i}}^{(j)}(t_k)\right]\mathbb{E}\left[K_{\mathbf{i}'}(x)H_{\mathbf{i}'}^{(j)}(t_k)\right]\right| \\ &\leq Cp_{\mathbf{n}}^N c_{\mathbf{n}}^N h_H^2 \phi_x(h_K)\left((\phi_x(h_K))^{1/a} + h_H\phi_x(h_K)\right) \\ &\leq Cp_{\mathbf{n}}^N c_{\mathbf{n}}^N h_H^2 \phi_x(h_K)^{(a+1)/a}. \end{aligned}$$

On the other hand, as the random variables $K_{\mathbf{i}}$ and $(x)H_{\mathbf{i}}^{(j)}$ are bounded, from Lemma 2.1(ii) of Tran (1990), we deduce

$$|Cov\left(\Delta_{\mathbf{i}}(x), \Delta_{\mathbf{i}'}(x)\right)| \leq C\varphi\left(\|\mathbf{i}-\mathbf{i}'\|\right),$$

then,

$$\begin{aligned} R_{\mathbf{n}}^2 &\leq C \sum_{(\mathbf{i},\mathbf{i}')\in S_2} \varphi\left(\|\mathbf{i}-\mathbf{i}'\|\right) \leq C p_{\mathbf{n}}^N \sum_{\mathbf{i}:\|\mathbf{i}\|\geq c_{\mathbf{n}}} \varphi\left(\|\mathbf{i}\|\right) \\ &\leq C p_{\mathbf{n}}^N c_{\mathbf{n}}^{-Na} \sum_{\mathbf{i}:\|\mathbf{i}\|\geq c_{\mathbf{n}}} \|\mathbf{i}\|^{Na} \varphi\left(\|\mathbf{i}\|\right). \end{aligned}$$

Let $c_{\mathbf{n}} = \left(h_H \phi_x(h_K)\right)^{-1/Na}$, then, we have

$$\begin{aligned} R_{\mathbf{n}}^2 &\leq C p_{\mathbf{n}}^N c_{\mathbf{n}}^{-Na} \sum_{\mathbf{i}:\|\mathbf{i}\|\geq c_{\mathbf{n}}} \|\mathbf{i}\|^{Na} \varphi\left(\|\mathbf{i}\|\right) \\ &\leq C p_{\mathbf{n}}^N h_H \phi_x(h_K) \sum_{\mathbf{i}:\|\mathbf{i}\|\geq c_{\mathbf{n}}} \|\mathbf{i}\|^{Na} \varphi\left(\|\mathbf{i}\|\right). \end{aligned}$$

Because of (3) and (H2), we get $R_{\mathbf{n}}^2 \leq C p_{\mathbf{n}}^N h_H \phi_x(h_K)$. Furthermore, under this choice of $c_{\mathbf{n}}$ we have $R_{\mathbf{n}}^1 \leq C p_{\mathbf{n}}^N h_H \phi_x(h_K)$. Hence

$$Var\left[Z_1^*\right] = O\left(p_{\mathbf{n}}^N h_H \phi_x(h_K)\right).$$

By using this last result, together with the definitions of $p_{\mathbf{n}}$, L and η, we get

$$B_1 \leq \exp\left(-C\eta_0 \log \widehat{\mathbf{n}}\right).$$

Consequently, an appropriate choice of η_0 completes the proof of the first part of this lemma.

• Concerning (T_3): because of (20) we have:

$$\sup_{y\in\mathcal{S}_{\mathbb{R}}} \left|\mathbb{E}\widehat{F}_N^{(j)}(y|x) - \mathbb{E}\widehat{F}_N^{(j)}(t_{k(y)}|x)\right| \leq C\frac{l_n}{h_H^{j+1}}.$$

Using analogous arguments as for T_1, we can show for n large enough:

$$\mathbb{P}\left(\sup_{y\in\mathcal{S}_{\mathbb{R}}} \left|\mathbb{E}\widehat{F}_N^{(j)}(y|x) - \mathbb{E}\widehat{F}_N^{(j)}(t_{k(y)}|x)\right| > \eta\sqrt{\frac{\log n}{n\, h_H^{2j-1}\phi_x(h_K)}}\right) = 0. \tag{24}$$

□

Proof of Lemma 3.20:

First, for $\mathbb{E}[\widehat{f'}^x_N(y)]$, we start by writing

$$\mathbb{E}[\widehat{f'}^x_N(y)] = \frac{\mathbb{E}\left[K_{\mathbf{1}}\mathbb{E}[h_H^{-2}H_1^{''}|X]\right]}{\mathbb{E}[K_1]} \quad \text{with,} \quad h_H^{-2}\mathbb{E}\left[H_1^{''}|X\right] = \int_{\mathbb{R}} H^{''}(t)f^X(y-h_Ht)dt.$$

The latter can be rewritten, by using a Taylor expansion under (H8), as follows

$$h_H^{-2}\mathbb{E}[H_1^{''}|X] = f^x(y) + \frac{h_H^2}{2}\left(\int t^2H^{''}(t)dt\right)\frac{\partial^2 f^x(y)}{\partial^2 y} + o(h_H^2).$$

Thus we get

$$\begin{aligned}\mathbb{E}\left[\widehat{f'}^x_N(y)\right] &= \frac{1}{\mathbb{E}[K_{\mathbf{1}}]}\left(\mathbb{E}\left[\frac{h_H^2}{2}K_{\mathbf{1}}\frac{\partial^2 f^x(y)}{\partial^2 y}\right]\int t^2H^{''}(t)dt\right)\\ &\quad+\frac{1}{\mathbb{E}[K_{\mathbf{1}}]}\left(\mathbb{E}\left[K_{\mathbf{1}}f^X(y)\right]+o(h_H^2)\right).\end{aligned}$$

Let, $\psi_l(\cdot,y) := \frac{\partial^l f^\cdot(y)}{\partial^l y}$ for $l\in\{0,2\}$. Since $\Phi_l(0)=0$, we have

$$\begin{aligned}\mathbb{E}\left[K_1\psi_l(X,y)\right] &= \psi_l(x,y)\mathbb{E}[K_{\mathbf{1}}] + \mathbb{E}\left[K_{\mathbf{1}}\left(\psi_l(X,y)-\psi_l(x,y)\right)\right]\\ &= \psi_l(x,y)\mathbb{E}[K_{\mathbf{1}}] + \mathbb{E}\left[K_{\mathbf{1}}\left(\Phi_l(d(x,X)\right)\right]\\ &= \psi_l(x,y)\mathbb{E}[K_{\mathbf{1}}] + \Phi_l'(0)\mathbb{E}\left[d(x,X)K_{\mathbf{1}}\right] + o(\mathbb{E}\left[d(x,X)K_{\mathbf{1}}\right]).\end{aligned}$$

So,

$$\begin{aligned}\mathbb{E}\left[\widehat{f'}^x_N(y)\right] &= f^x(y) + \frac{h_H^2}{2}\frac{\partial^2 f^x(y)}{\partial y^2}\int t^2H^{''}(t)dt + o\left(h_H^2\frac{\mathbb{E}\left[d(x,X)K_1\right]}{\mathbb{E}[K_{\mathbf{1}}]}\right)\\ &\quad+\Phi_0'(0)\frac{E\left[d(x,X)K_{\mathbf{1}}\right]}{\mathbb{E}[K_{\mathbf{1}}]} + o\left(\frac{\mathbb{E}\left[d(x,X)K_{\mathbf{1}}\right]}{\mathbb{E}[K_{\mathbf{1}}]}\right).\end{aligned}$$

Similarly to Ferraty et al. (2007); we show that

$$\frac{1}{\phi_x(h_K)}\mathbb{E}\left[d(x,X)K_{\mathbf{1}}\right] = h_K\left(K(1) - \int_0^1 (sK(s))'\beta_x(s)ds + o(1)\right),$$

and

$$\frac{1}{\phi_x(h_K)}\mathbb{E}\left[K_{\mathbf{1}}\right] = K(1) - \int_0^1 K'(s)\beta_x(s)ds + o(1).$$

Hence,

$$\begin{aligned}\mathbb{E}\left[\widehat{f'}^x_N(y)\right] &= f^x(y) + \frac{h_H^2}{2}\frac{\partial^2 f^x(y)}{\partial y^2}\int t^2 H''(t)dt \\ &+h_K\Phi_0'(0)\frac{\left(K(1) - \int_0^1 (sK(s))'\beta_x(s)ds\right)}{\left(K(1) - \int_0^1 K'(s)\beta_x(s)ds\right)} + o(h_H^2) + o(h_K).\end{aligned}$$

Second, concerning $\mathbb{E}[\widehat{F}^x_N(y)]$, we write by an integration by parts

$$\mathbb{E}[\widehat{F}^x_N(y)] = \frac{1}{\mathbb{E}[K_1]}\mathbb{E}\left[K_1 E[H_1|X]\right] \text{ with } \mathbb{E}\left[H_1|X\right] = \int_{\mathbb{R}} H'(t)F^X(y - h_H t)dt.$$

The same steps used to study $\mathbb{E}[\widehat{f}^x_N(y)]$ can be followed to prove that

$$\begin{aligned}\mathbb{E}\left[\widehat{F}^x_N(y)\right] &= F^x(y) + \frac{h_H^2}{2}\frac{\partial^2 F^x(y)}{\partial y^2}\int t^2 H'(t)dt \\ &+h_K\Psi_0'(0)\frac{\left(K(1) - \int_0^1 (sK(s))'\beta_x(s)ds\right)}{\left(K(1) - \int_0^1 K'(s)\beta_x(s)ds\right)} + o(h_H^2) + o(h_K).\end{aligned}$$

□

Proof of Lemma 3.22:

For the first quantity $Var[\widehat{f'}^x_N(y)]$, we have

$$s_n^2 = Var[\widehat{f'}^x_N(y)] = \frac{1}{(\widehat{\mathbf{n}}h_H^2\mathbb{E}\left[K_{\mathbf{1}}(x)\right])^2}Var\left[\sum_{\mathbf{i}\in I_{\mathbf{n}}}\Gamma_{\mathbf{i}}(x)\right],$$

where

$$\Gamma_{\mathbf{i}}(x) = K_{\mathbf{i}}(x)H_{\mathbf{i}}''(y) - \mathbb{E}\left[K_{\mathbf{i}}(x)H_{\mathbf{i}}''(y)\right].$$

Thus,

$$\begin{aligned}Var[\widehat{f}^x_N(y)] &= \frac{1}{(\widehat{\mathbf{n}}h_H^2\mathbb{E}\left[K_{\mathbf{1}}\right])^2}\underbrace{\sum_{\mathbf{i}\neq\mathbf{j}} Cov\left(\Gamma_{\mathbf{i}}(x), \Gamma_{\mathbf{j}}(x)\right)}_{s_n^{cov}} + \underbrace{\sum_{\mathbf{i}\in I_{\mathbf{n}}} Var\left(\Gamma_{\mathbf{i}}(x)\right)}_{s_n^{var}} \\ &= \frac{Var\left[\Gamma_{\mathbf{1}}\right]}{\widehat{\mathbf{n}}(h_H^2\mathbb{E}\left[K_{\mathbf{1}}\right])^2} + \frac{1}{(\widehat{\mathbf{n}}h_H^2\mathbb{E}\left[K_{\mathbf{1}}\right])^2}\sum_{\mathbf{i}\neq\mathbf{j}} Cov(\Gamma_{\mathbf{i}}, \Gamma_{\mathbf{j}}).\end{aligned}$$

Let us calculate the quantity $Var\left[\Gamma_{\mathbf{1}}(x)\right]$. We have

$$\begin{aligned} Var\left[\Gamma_{\mathbf{1}}(x)\right] &= \mathbb{E}\left[K_{\mathbf{1}}^2(x)H_{\mathbf{1}}^{''2}(y)\right] - \left(\mathbb{E}\left[K_{\mathbf{1}}(x)H_{\mathbf{1}}^{''}(y)\right]\right)^2 \\ &= \mathbb{E}\left[K_{\mathbf{1}}^2(x)\right]\frac{\mathbb{E}\left[K_{\mathbf{1}}^2(x)H_{\mathbf{1}}^{''2}(y)\right]}{\mathbb{E}\left[K_{\mathbf{1}}^2(x)\right]} \\ &\quad -(\mathbb{E}\left[K_{\mathbf{1}}(x)\right])^2\left(\frac{\mathbb{E}\left[K_{\mathbf{1}}(x)H_{\mathbf{1}}'(y)\right]}{\mathbb{E}\left[K_{\mathbf{1}}(x)\right]}\right)^2. \end{aligned}$$

So, by using the same arguments as those used in previous lemma, we get

$$\begin{aligned} \frac{1}{\phi_x(h_K)}\mathbb{E}\left[K_{\mathbf{1}}^2(x)\right] &= K^2(1) - \int_0^1 (K^2(s))'\beta_x(s)ds + o(1) \\ \frac{\mathbb{E}\left[K_{\mathbf{1}}^2(x)H_{\mathbf{1}}^{''2}(y)\right]}{\mathbb{E}\left[K_{\mathbf{1}}^2(x)\right]} &= h_H^2 f^x(y)\int H^{''2}(t)dt + o(h_H^2) \\ \frac{\mathbb{E}[K_{\mathbf{1}}(x)H_{\mathbf{1}}^{''}(y)]}{\mathbb{E}\left[K_{\mathbf{1}}(x)\right]} &= h_H^2 f^x(y) + o(h_H^2), \end{aligned}$$

which implies that

$$\begin{aligned} Var\left[\Gamma_{\mathbf{i}}(x)\right] &= h_H^2\phi_x(h_K)f^x(y)\int H^{''2}(t)dt\left(K^2(1) - \int_0^1 (K^2(s))'\beta_x(s))ds\right) \\ &\quad + o\left(h_H^2\phi_x(h_K)\right). \end{aligned} \tag{25}$$

Now let us focus on the covariance term. To do that, we define

$$E_1 = \{\mathbf{i},\mathbf{j} \in \mathcal{I}_{\mathbf{n}} : 0 < \|\mathbf{i}-\mathbf{j}\| \le c_{\mathbf{n}}\} \text{ and } E_2 = \{\mathbf{i},\mathbf{j} \in \mathcal{I}_{\mathbf{n}} : \|\mathbf{i}-\mathbf{j}\| > c_{\mathbf{n}}\}.$$

For all $(\mathbf{i},\mathbf{j}) \in E_1^2$, we write

$$Cov\left(\Gamma_{\mathbf{i}}(x),\Gamma_{\mathbf{j}}(x)\right) = \mathbb{E}\left[K_{\mathbf{i}}(x)K_{\mathbf{j}}(x)H_{\mathbf{i}}^{''}(y)H_{\mathbf{j}}^{''}(y)\right] - \left(\mathbb{E}\left[K_{\mathbf{i}}(x)H_{\mathbf{i}}^{''}(y)\right]\right)^2,$$

and we use the fact that

$\mathbb{E}\left[H_{\mathbf{i}}^{''}(y)H_{\mathbf{j}}^{''}(y)|(X_{\mathbf{i}},X_{\mathbf{j}})\right] = \mathcal{O}(h_H^4)$; $\forall i \neq j$, $\mathbb{E}\left[H_{\mathbf{i}}^{''}(y)|X_{\mathbf{i}}\right] = \mathcal{O}(h_H^2)$, $\forall\mathbf{i}$.

Under (H2) and (H10), we get

$$\mathbb{E}\left[K_{\mathbf{i}}K_{\mathbf{j}}H_{\mathbf{i}}^{''}H_{\mathbf{j}}^{''}\right] \le Ch_H^4\mathbb{P}\left[(X_{\mathbf{i}},X_{\mathbf{j}}) \in B(x,h_K)\times B(x,h_K)\right],$$

and

$$\mathbb{E}\left[K_{\mathbf{i}}(x)H_{\mathbf{i}}^{''}(y)\right] \leq Ch_H^2 \mathbb{P}\left(X_{\mathbf{i}} \in B(x, h_K)\right).$$

It follows that the hypotheses (H1), (H2), and (H10) imply that

$$Cov\left(\Gamma_{\mathbf{i}}(x), \Gamma_{\mathbf{j}}(x)\right) \leq Ch_H^2 \phi_x(h_K)\left(\phi_x(h_K) + (\phi_x(h_K))^{1/a}\right).$$

So,

$$\sum_{E_1} Cov\left(\Gamma_{\mathbf{i}}(x), \Gamma_{\mathbf{j}}(x)\right) \leq C\left(\widehat{\mathbf{n}} c_{\mathbf{n}}^N h_H^4 \left(\phi_x(h_K)\right)^{1+\frac{1}{a}}\right).$$

On the other hand, Lemma 6.2 and $|\Gamma_{\mathbf{i}}| \leq C$ permit us to write that $(\mathbf{i}, \mathbf{j}) \in E_2^2$

$$\left|Cov\left(\Gamma_{\mathbf{i}}(x), \Gamma_{\mathbf{j}}(x)\right)\right| \leq C\varphi\left(\|\mathbf{i} - \mathbf{j}\|\right),$$

and

$$\begin{aligned}
\sum_{E_2} Cov\left(\Gamma_{\mathbf{i}}(x), \Gamma_{\mathbf{j}}(x)\right) &\leq C \sum_{E_2} \varphi\left(\|\mathbf{i} - \mathbf{j}\|\right) \\
&\leq C\widehat{\mathbf{n}} \sum_{\mathbf{i}:\|\mathbf{i}\| > c_{\mathbf{n}}} \varphi\left(\|\mathbf{i}\|\right) \\
&\leq C\widehat{\mathbf{n}} c_{\mathbf{n}}^{-Na} \sum_{\mathbf{i}:\|\mathbf{i}\| > c_{\mathbf{n}}} \|\mathbf{i}\|^{Na} \varphi\left(\|\mathbf{i}\|\right).
\end{aligned}$$

Finally, we have:

$$\sum_{\mathbf{i} \neq \mathbf{j}} Cov(D_{\mathbf{i}}, D_{\mathbf{j}}) \leq \left(C\widehat{\mathbf{n}} c_{\mathbf{n}}^N h_H^4 \phi_x^{1+\frac{1}{a}}(h_K) + C\widehat{\mathbf{n}} c_{\mathbf{n}}^{-Na} \sum_{\mathbf{i}:\|\mathbf{i}\| > c_{\mathbf{n}}} \|\mathbf{i}\|^{Na} \varphi\left(\|\mathbf{i}\|\right)\right).$$

Let $c_{\mathbf{n}} = (h_H^{4/(a+1)} \phi_x^{1/a}(h_K))^{-1/N}$. Then we obtain that

$$\sum Cov\left(\Gamma_{\mathbf{i}}(x), \Gamma_{\mathbf{j}}(x)\right) = o\left(\widehat{\mathbf{n}} h_H^2 \phi_x(h_K)\right).$$

In conclusion, we have

$$\begin{aligned} Var[\widehat{f'}^x_N(y)] &= \frac{f^x(y)}{\widehat{\mathbf{n}} h_H^4 \phi_x(h_K)} \left(\int H''^2(t)dt\right) \left(\frac{\left(K^2(1) - \int_0^1 (K^2(s))'\beta_x(s)ds\right)}{\left(K(1) - \int_0^1 K'(s)\beta_x(s)ds\right)^2} \right) \\ &\quad + o\left(\frac{1}{\widehat{\mathbf{n}} h_H^2 \phi_x(h_K)}\right). \end{aligned} \tag{26}$$

Now, for $\widehat{F}^x_N(y)$, (resp. $\widehat{F}^x_D$) we replace $H''_{\mathbf{i}}(y)$ by $H_{\mathbf{i}}(y)$ (resp. by 1) and we follow the same ideas, under the fact that $H \leq 1$

$$\begin{aligned} Var[\widehat{F}^x_N(y)] &= \frac{F^x(y)}{\widehat{\mathbf{n}} \phi_x(h_K)} \left(\int H'^2(t)dt\right) \left(\frac{\left(K^2(1) - \int_0^1 (K^2(s))'\beta_x(s)ds\right)}{\left(K(1) - \int_0^1 K'(s)\beta_x(s)ds\right)^2} \right) \\ &\quad + o\left(\frac{1}{\widehat{\mathbf{n}} \phi_x(h_K)}\right), \end{aligned}$$

and

$$Var[\widehat{F}^x_D] = \frac{1}{\widehat{\mathbf{n}} \phi_x(h_K)} \left(\frac{\left(K^2(1) - \int_0^1 (K^2(s))'\beta_x(s)ds\right)}{\left(K(1) - \int_0^1 K'(s)\beta_x(s)ds\right)^2} \right) + o\left(\frac{1}{\widehat{\mathbf{n}} \phi_x(h_K)}\right).$$

This yields the proof. □

Proof of Lemma 3.23:

The proof of this lemma follows the same steps as the previous Lemma. For this, we keep the same notation and we write

$$\begin{aligned} Cov(\widehat{f'}^x_N(y), \widehat{F}^x_N(y)) &= \frac{1}{\widehat{\mathbf{n}} h_H^2 (\mathbb{E}\left[K_{\mathbf{1}}(x)\right])^2} Cov\left(\Gamma_{\mathbf{1}}(x), \Delta_{\mathbf{1}}(x)\right) \\ &\quad + \frac{1}{\widehat{\mathbf{n}}^2 h_H^2 (\mathbb{E}\left[K_{\mathbf{1}}(x)\right])^2} \sum_{\mathbf{i} \neq \mathbf{j}} Cov(\Gamma_{\mathbf{i}}(x), \Delta_{\mathbf{j}}(x)), \end{aligned}$$

where,

$$\Delta_{\mathbf{i}}(x) = K_{\mathbf{i}}(x) H_{\mathbf{i}}(y) - \mathbb{E}\left[K_{\mathbf{i}}(x) H_{\mathbf{i}}(y)\right].$$

For the first term, we have under (H9)

$$\begin{aligned} Cov\left(\Gamma_{\mathbf{1}}(x), \Delta_{\mathbf{1}}(x)\right) &= \mathbb{E}[K_{\mathbf{1}}^2(x)H_{\mathbf{1}}(y)H_{\mathbf{1}}^{''}(y)] - \mathbb{E}[K_{\mathbf{1}}(x)H_{\mathbf{1}}(y)]\mathbb{E}[K_{\mathbf{1}}(x)H_{\mathbf{1}}^{''}(y)] \\ &= \mathcal{O}(h_H^2\phi_x(h_K)) + \mathcal{O}(h_H^2\phi_x^2(h_K)) \\ &= \mathcal{O}(h_H^2\phi_x(h_K)). \end{aligned}$$

Therefore,

$$\begin{aligned} \frac{1}{\widehat{\mathbf{n}}h_H^2(\mathbb{E}\left[K_{\mathbf{1}}(x)\right])^2} Cov\left(\Gamma_{\mathbf{1}}(x), \Delta_{\mathbf{1}}(x)\right) &= \mathcal{O}\left(\frac{1}{\widehat{\mathbf{n}}\phi_x(h_K)}\right) \\ &= o\left(\frac{1}{\widehat{\mathbf{n}}h_H^2\phi_x(h_K)}\right). \end{aligned} \tag{27}$$

Furthermore, for the covariance term we split the sum over the two subsets E_1 and E_2 defined with, $c_{\mathbf{n}} = (h_H^{2/(a+1)}\phi_x^{1/a}(h_K))^{-1/N}$ and we use once again the boundedness of K and H, and the fact that (1) and (H5) imply that $\mathbb{E}\left(H_{\mathbf{i}}^{''}(y)|X_{\mathbf{i}}\right) = \mathcal{O}(h_H^2),\ \forall;\mathbf{i}$ to get over $\forall \mathbf{i}, \mathbf{j} \in E_1$ that

$$Cov\left(D_{\mathbf{i}}, \Delta_{\mathbf{j}}\right) \leq C h_H^2 \phi_x(h_K)(\phi_x(h_K) + \phi_x^{\frac{1}{a}}(h_K)).$$

While over E_2 we apply Lemma 6.2 to write that

$$\left|Cov\left(\Gamma_{\mathbf{i}}(x), \Delta_{\mathbf{j}}(x)\right)\right| \leq C\varphi\left(\|\mathbf{i} - \mathbf{j}\|\right).$$

Consequently, because of the definition of $c_{\mathbf{n}}$, we have

$$\sum_{\mathbf{i}\neq\mathbf{j}} Cov(\Gamma_{\mathbf{i}}(x), \Delta_{\mathbf{j}}(x)) = \sum_{E_1} Cov\left(\Gamma_{\mathbf{i}}(x), \Delta_{\mathbf{j}}(x)\right) + \sum_{E_2} Cov\left(\Gamma_{\mathbf{i}}(x), \Delta_{\mathbf{j}}(x)\right) = o\left(\widehat{\mathbf{n}}\phi_x(h_K)\right). \tag{28}$$

From (27) and (28), we deduce that

$$Cov(\widehat{f'}_N^x(y), \widehat{F}_N^x(y)) = o\left(\frac{1}{\widehat{\mathbf{n}}h_H^2\phi_x(h_K)}\right).$$

The same arguments can be used to shows that

$$Cov(\widehat{f'}_N^x(y), \widehat{F}_D^x) = o\left(\frac{1}{\widehat{\mathbf{n}}h_H^2\phi_x(h_K)}\right),$$

and

$$Cov(\widehat{F}_N^x(y), \widehat{F}_D^x) = o\left(\frac{1}{\widehat{\mathbf{n}}\phi_x(h_K)}\right).$$

□

Acknowledgments

The authors are thankful to anonymous referees and the Editor-in-Chief Professor Aliakbar Montazer Haghighi for useful comments and suggestions towards the improvement of this paper.

REFERENCES

Biau, G. and Cadre, B. (2004). Nonparametric Spatial Prediction. Stat. Inf. Stoch. Proc., Vol. 7, pp. 327 - 349.

Bouchentouf, A.A., Djebbouri, T., Rabhi, A. and Sabri, K. (2014). Strong uniform consistency rates of some characteristics of the conditional distribution estimator in the functional single index model. Appl. Math. (Warsaw), Vol. 41, No. 4, pp. 301 - 322.

Bouraine, M., Aït Saidi, A., Ferraty, F. and Vieu, P. (2010). Choix optimal de l'indice multi-fonctionnel: méthode de validation croisée. R. Roumaine Math. Pures et Appl., Vol. 50, No. 5, pp. 355 - 367.

Carbon, M., Hallin, M. and Tran, L.T. (1996). Kernel density estimation for random fields. J. Stat. Prob. Lett., Vol. 36, pp. 115 - 125.

Carbon, M., Tran, L.T. and Wu, B. (1997). Kernel density estimation for random fields: the L_1 theory. J. Nonparamet. Stat., Vol. 6, pp. 157 - 170.

Carbon, M., Francq, C. and Tran, L.T. (2007). Kernel regression estimation for random fields. J. Stat. Plann. Inf., Vol. 137, pp. 778 - 798.

Dabo-Niang, S., Rachdi, M. and Yao, A.F. (2011). Spatial kernel regression estimation prediction for functional rom fields. Far East J. Stat., Vol. 37, pp. 77 - 113.

Dabo-Niang, S. and Yao, A.F. (2007). Kernel regression estimation for continuous spatial processes. Math. Methods Statist., Vol. 16, pp. 298 - 317.

Djebbouri, T. Hamel, E.H and Rabhi, A. (2015). On conditional hazard function estimate for functional mixing data. New Trends in Mathematical Sciences (NTMSCI), Vol. 3, No. 2, pp. 79 - 95.

Doukhan, P. (1994). *Mixing: Properties examples*, Lecture Notes in Statistics 85, Springer Verlag, New York.

Ferraty, F., Mas, A. and Vieu, P. Advances in nonparametric regression for functional variables. Australian and New Zealand Journal of Statistics, Vol. 49, pp. 1 - 20.

Ferraty, F., Rabhi, A. and Vieu, P. (2005). Conditional Quantiles for Functionally Dependent Data with Application to the Climatic El Niño Phenomeno Sankhyā: The Indian Journal of Statistics, Special Issue on Quantile Regression and Related Methods, Vol. 67, No. 2, pp. 399 - 417.

Ferraty, F., Rabhi, A. and Vieu, P. (2008). Estimation nonparamétrique de la fonction de hasard avec variable explicative fonctionnelle. Rev. Roum. Math Pures & Appl, Vol. 53, pp. 1 - 18.

Ferraty, F. and Vieu, P. (2006). *Nonparametric Functional Data Analysis. Theory and Practice*. Springer Series in Statistics. Springer, New York.

Guyon, X. (1987). *Estimation d'un champ par pseudo-vraisemblance conditionnelle: Etude asym ptotique et application au cas Markovien*, Proceedings of the Sixth Franco-Belgian Meeting of Statisticians.

Hall, P., Wolef, R.C.L. and Yao, Q. (1999). Methods for estimating a conditional distribution function. Ann. Stat. Assoc., Vol. 94, pp. 154 - 163.

Hyndman, R.J., Bashtannyk, D.M. and Grunwald, G.K. (1996). Estimating and visualizing conditional densities. J. Comput. Graph. Stat., Vol. 5, pp. 315 - 336.

Laksaci, A. and Mechab, B. (2010). Estimation nonparamétrique de la fonction de hasard avec variable explicative fonctionnelle: cas des données spatiales. Rev. Roumaine Math Pures & Appl., Vol. 55, pp. 35 - 51.

Li, J. and Tran, L.T. (2009). Nonparametric estimation of conditional expectation. J. Stat. Plann. Inf., Vol. 139, pp. 164 - 175.

Li, J. and Tran, L.T. (2007). Hazard rate estimation on rom fields. J. Multivariate Anal., Vol. 98, pp. 1337 - 1355.

Lu, Z. and Chen, X. (2004). Spatial kernel regression estimation: weak concistency. J. Stat. Prob. Lett., Vol. 68, pp. 125 - 136.

NAC, C. (1991). *Statistics for spatial Data*, Wiley Series in Probability Mathematical Statistics, New York.

Quintela-del-Rio, A. (2008). Hazard function given a functional variable: Non-parametric estimation under strong mixing conditions. J. Nonparam. Stat., Vol. 20, pp. 413 - 430.

Rabhi, A., Benaissa, S., Hamel, E.H. and Mechab, B. (2013). Mean square error of the estimator of the conditional hazard function. Appl. Math. (Warsaw), Vol. 40, No. 4, pp. 405 - 420.

Ramsay, J. (2008). *Fda problems that I like to talk about,* Personal communication.

Ripley, B. (1981). *Spatial Statistics*, Wiley New-York.

Roussas, G. (1989). Hazard rate estimation under dependence conditions .J. Stat. Plann. Inf., Vol. 22, pp. 81 - 93.

Tran, L.T. (1990). Kernel density estimation on rom fields. J. Multivariate Anal., Vol. 34, pp. 37 - 53.

Vieu, P. (1991). Quadratic errors for nonparametric estimates under dependence. J. Multivariate Anal., Vol. 39, pp. 324 - 347.

Watson, G.S. and Leadbetter, M.R. (1964). Hazard analysis. I. Biometrika, Vol. 51, pp. 175-184.

Youndjé, E. (1993). *Estimation non paramétrique de la densité conditionnelle par la méthode du noyau*. Thèse de Doctorat, Université de Rouen.