Research Article

Asymptotic Normality of a Nonparametric Conditional Quantile Estimator for Random Fields

Sidi Ali Ould Abdi,[1] **Sophie Dabo-Niang,**[2]
Aliou Diop,[1] **and Ahmedoune Ould Abdi**[1]

[1] *Laboratoire LERSTAD, Université Gaston Berger, BP 234, Saint-Louis, Senegal*
[2] *Laboratoire EQUIPPE, Université Charles De Gaulle, Lille 3, Maison de la Recherche,
  Domaine du Pont de Bois, BP 60149, 59653 Villeneuve d'Ascq Cedex, France*

Correspondence should be addressed to Sophie Dabo-Niang, sophie.dabo@univ-lille3.fr

Given a stationary multidimensional spatial process $(Z_{\mathbf{i}} = (X_{\mathbf{i}}, Y_{\mathbf{i}}) \in \mathbb{R}^d \times \mathbb{R}, \mathbf{i} \in \mathbb{Z}^N)$, we investigate a kernel estimate of the spatial conditional quantile function of the response variable $Y_{\mathbf{i}}$ given the explicative variable $X_{\mathbf{i}}$. Asymptotic normality of the kernel estimate is obtained when the sample considered is an α-mixing sequence.

1. Introduction

In this paper, we are interested in nonparametric conditional quantile estimation for spatial data. Spatial data are modeled as finite realizations of random fields that is stochastic processes indexed in $\mathbb{Z}^N$, the integer lattice points in the N-dimensional Euclidean space $(N \geq 1)$. Such data are collected from different locations on the earth and arise in a variety of fields, including soil science, geology, oceanography, econometrics, epidemiology, environmental science, forestry, and many others; see Chilès and Delfiner [1], Guyon [2], Anselin and Florax [3], Cressie [4], or Ripley [5].
In the context of spatial data, the analysis of the influence of some covariates on a response variable is particularly difficult, due to the possibly highly complex spatial dependence among the various sites. This dependence cannot be typically modeled in any adequate way.

Conditional quantile analysis is of wide interest in modeling of spatial dependence and in the construction of confidence (predictive) intervals. There exist an extensive literature and various nonparametric approaches in conditional quantile estimation in the nonspatial case $(N = 1)$ for independent samples and time-dependent observations; see, for example,

Stute [6], Samanta [7], Portnoy [8], Koul and Mukherjee [9], Honda [10], Cai [11], Gannoun et al. [12], and Yu et al. [13]. Extending classical nonparametric conditional quantile estimation for dependent random variables to spatial quantile regression is far from being trivial. This is due to the absence of any canonical ordering in the space and of obvious definition of tail sigma-fields.

Although potential applications of conditional spatial quantile regressions are without number, only the papers of Koenker and Mizera [14], Hallin et al. [15], Abdi et al. [16], and Dabo-Niang and Thiam [17] have paid attention to study these regression methods. Hallin et al. [15] gave a Bahadur representation and an asymptotic normality results of a local linear conditional quantile estimator. The method of Koenker and Mizera [14] is a spatial smoothing technique rather than a spatial (auto) regression one and they do not take into account the spatial dependency structure of the data. The work of Abdi et al. [16] deals with $2r$—mean ($r \in \mathbb{N}^*$) and almost complete consistencies of a kernel estimate of conditional quantiles. The paper of Dabo-Niang and Thiam [17] gives the L_1 norm consistency and asymptotic normality of a kernel estimate of the spatial conditional quantile, but this estimate is less general than the one considered here.

However, conditional mean regression estimation for spatial data has been considered in several papers; some key references are Carbon et al. ([18, 19]), Biau and Cadre [20], Lu and Chen ([21, 22]), Hallin et al. ([23, 24]), Lahiri and Zhu [25], Carbon et al. [26], and Dabo-Niang and Yao [27].

The rest of the paper is organized as follows. In Section 2, we provided the notations and the kernel quantile estimate. Section 3 is devoted to assumptions. The asymptotic normality of the kernel estimate is stated in Section 4. Section 5 contains a prediction application based on quantile regression and applied to simulated data. Proofs and preliminary lemmas are given in the last section.

2. Kernel Conditional Quantile Estimator

Let $(Z_\mathbf{i} = (X_\mathbf{i}, Y_\mathbf{i}), \mathbf{i} \in \mathbb{Z}^N)$ ($N \geq 1$) be an $\mathbb{R}^d \times \mathbb{R}$-valued measurable strictly stationary spatial process ($d \geq 1$), with same distribution as the vector of variables (X, Y) and defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. A point $\mathbf{i} = (i_1, \dots, i_N)$ in $\mathbb{Z}^N$ will be referred to as a *site* and may also include a time component.

We assume that the process under study $(Z_\mathbf{i})$ is observed over a rectangular domain $\mathcal{O}_\mathbf{n} = \{\mathbf{i} = (i_1, \dots, i_N) \in \mathbb{Z}^N, 1 \leq i_k \leq n_k, k = 1, \dots, N\}$, $\mathbf{n} = (n_1, \dots, n_N) \in \mathbb{Z}^N$. We will write $\mathbf{n} \to \infty$ if $\min\{n_k\} \to \infty$ and $|n_j/n_k| < C$ for a constant C such that $0 < C < \infty$ for all j, k such that $1 \leq j$, $k \leq N$. In the sequel, all the limits are considered when $\mathbf{n} \to \infty$. For $\mathbf{n} = (n_1, \dots, n_N) \in \mathbb{Z}^N$, we set $\hat{\mathbf{n}} = n_1, \dots, n_N$.

Let S be a set of sites. $\mathcal{B}(S)$ will denote in what follows, the Borel σ-field generated by $Z_\mathbf{i}$, $\mathbf{i} \in S$. We assume that the regular version of the conditional probability of Y given X exists and has a bounded density with respect to Lebesgue's measure over $\mathbb{R}$. For all $x \in \mathbb{R}^d$, we denote by F^x (resp., f^x) the conditional distribution function (resp., the conditional density) of Y given $X = x$. In the following, x is a fixed point in $\mathbb{R}^d$ and we denote by V_x a neighborhood of this point. For $x \in \mathbb{R}^d$, we denote by $\|x\|$ the Euclidian norm of x. We suppose that the marginal and joint densities f_X and $f_{X,Y}$ of, respectively, X and (X, Y) exist with respect to Lebesgue measures on $\mathbb{R}^d$ and $\mathbb{R}^{d+1}$.

For $\alpha \in]0, 1[$, the conditional quantile of order α of F^x, denoted by $q_\alpha(x)$, can be written as a solution of the equation $F^x(q_\alpha(x)) = \alpha$.

To insure that $q_\alpha(x)$ exists and is unique, we assume that F^x is strictly increasing. Let $K_i = K_1((x - X_i)/h_1)$ and $H_i(y) = K_2((y - Y_i)/h_2)$. The conditional distribution F^x and the corresponding density f^x can be estimated by the following respective estimators:

$$
\begin{aligned}
\widehat{F}^x(y) &= \begin{cases} \dfrac{\sum_{i \in \mathcal{I}_n} K_i H_i(y)}{\sum_{i \in \mathcal{I}_n} K_i} & \text{if } \sum_{i \in \mathcal{I}_n} K_i \neq 0, \\ 0 & \text{else,} \end{cases} \\
\widehat{f}^x(y) &= \begin{cases} \dfrac{(1/h_2) \sum_{i \in \mathcal{I}_n} K_i K((y - Y_i)/h_2)}{\sum_{i \in \mathcal{I}_n} K_i} & \text{if } \sum_{i \in \mathcal{I}_n} K_i \neq 0, \\ 0 & \text{else,} \end{cases}
\end{aligned}
\tag{2.1}
$$

where K_1 is a kernel density, K_2 is a distribution function, K is the first derivative of K_2, and $h_1 = h_{1,n}$ (resp., $h_2 = h_{2,n}$) are sequences of positive real numbers tending to 0 when $n \to \infty$. Remark that we can write $\widehat{F}^x(y) = g_n(x, y)/\widehat{f}(x)$, where $g_n(x, y) = (1/\widehat{n}h_1^d) \sum_{i \in \mathcal{I}_n} K_i H_i(y)$ and $\widehat{f}(x) = (1/\widehat{n}h_1^d) \sum_{i \in \mathcal{I}_n} K_i$ are, respectively, the estimates of $g(x, y) = F^x(y) f_X(x)$ and $f_X(x)$. The kernel estimate of the conditional quantile is related to the conditional distribution estimator. A natural estimator $\widehat{q}_\alpha(x)$ of $q_\alpha(x)$ is defined such that

$$
\widehat{F}^x(\widehat{q}_\alpha(x)) = \alpha.
\tag{2.2}
$$

Remark 2.1. Another alternative characterization of the αth conditional quantile (see, e.g., Gannoun et al. [12]) is

$$
q_\alpha(x) = \underset{\theta \in \mathbb{R}}{\operatorname{argmin}} \, \mathbf{E}[(2\alpha - 1)(Y - \theta) + |Y - \theta| \mid X = x].
\tag{2.3}
$$

Then, one can consider the alternative local constant estimator (see Hallin et al. [15]) defined by

$$
\widetilde{q}_\alpha(x) = \underset{\theta \in \mathbb{R}}{\operatorname{argmin}} \sum_{i \in \mathcal{I}_n} (|Y_i - \theta| + (2\alpha - 1)(Y_i - \theta)) K_1\left(\frac{x - X_i}{h_1}\right).
\tag{2.4}
$$

Let us mention that it can be shown that (2.4) is equivalent to

$$
\widetilde{q}_\alpha(x) = \inf\left\{ t \in \mathbb{R}, \ \widetilde{F}^x(t) \geq \alpha \right\},
\tag{2.5}
$$

where $\widetilde{F}^x(\cdot)$ is the same estimator of $F^x(\cdot)$ as $\widehat{F}^x(\cdot)$ except that $H_i(y) = \mathbb{1}_{\{Y_i \leq y\}}$.

In this paper, we will focus on the study of the asymptotic behavior of $\widehat{q}_\alpha(x)$, since in practice some simulations permit to remark that the differences between this estimator and the local linear one are too small to affect any interpretations; see also Dabo-Niang and Thiam [17] and Gannoun et al. [12].

3. Assumptions

We denote by $g^{(j)}$ the derivative of order j of a function g. In what follows, C and C' will denote any positive constant.

3.1. General Assumptions

We will use the following four regularity conditions (see, e.g., Gannoun et al. [12] or Ferraty et al. [28]).

(H_1) The joint $f_{X,Y}$ and marginal f_X densities are, respectively, continuous on $\mathbb{R}^{d+1}$ and $\mathbb{R}^d$. Moreover $f_X(x) > 0$ and $f^x(q_\alpha(x)) > 0$.

If $f^x(q_\alpha(x)) = 0$, then one can use a condition like $F^x(\cdot)$ is of class $\mathcal{C}^j$, $F^{x^{(k)}}(q_\alpha(x)) = 0$ for $1 \le k < j$, and $0 < C < |F^{x^{(j)}}(y)| < C' < \infty$ as in Ferraty et al. [28].

(H_2) For all $(x_1, x_2) \in V_x \times V_x$, $\forall (y_1, y_2) \in \mathbb{R}^2$

$$\left| F^{x_1}(y_1) - F^{x_2}(y_2) \right| \le C\left(\|x_1 - x_2\|^{b_1} + |y_1 - y_2|^{b_2} \right), \quad b_1 > 0, \ b_2 > 0. \tag{3.1}$$

(H_3) We assume that the conditional density $f^{(X_i, X_j)}$ of (Y_i, Y_j) given (X_i, X_j) exists and is uniformly bounded in $\mathbf{i}, \mathbf{j}$.

For simplicity, we assume the following condition on the kernel K_1 (see, e.g., Devroye [29]).

(H_4) There exist C_1 and C_2, $0 < C_1 < C_2 < \infty$ such that

$$C_1 \mathbb{I}_{[0,1]}(\|t\|) < K_1(t) < C_2 \mathbb{I}_{[0,1]}(\|t\|). \tag{3.2}$$

(H_5) K_2 is of class $\mathcal{C}^1$ with a symmetric, Lipschitz, bounded, and compact support density K. In addition, we assume that the restriction of K_2 on $\{t \in \mathbb{R}, \ K_2(t) \in (0,1)\}$ is a strictly increasing function.

Assumption H_5 is classical in nonparametric estimation and is satisfied by usual kernels such as Epanechnikov and Biweight, whereas the Gaussian density K is also possible; it suffices to replace the compact support assumption by $\int_{\mathbb{R}^d} |t|^{b_2} K(t) dt < \infty$. Assumption H_5 ensures the existence and the uniqueness of the quantile estimate $\widehat{q}_\alpha(x)$.

(H_6) $\widehat{\mathbf{n}} h_1^{\theta_1} h_2^{\theta_2} \log \widehat{\mathbf{n}}^{\theta_3} \to +\infty$ with $\theta > 2N$.

(H_7) $\widehat{\mathbf{n}} h_1^{\theta_1'} h_2^{\theta_2'} \log \widehat{\mathbf{n}}^{\theta_3'} \to +\infty$ with $\theta > N(1 + 2\widetilde{\beta})$,

where θ and $\widetilde{\beta}$ are defined in Section 3.2.1, and

$$\theta_1 = \frac{d\theta}{\theta - 2N}, \qquad \theta_2 = \frac{\theta}{\theta - 2N}, \qquad \theta_3 = -1,$$
$$\theta_1' = \frac{d(N + \theta)}{\theta - N\left(1 + 2\widetilde{\beta}\right)}, \qquad \theta_2' = \frac{N + \theta}{\theta - N\left(1 + 2\widetilde{\beta}\right)}, \qquad \theta_3' = \frac{N - \theta}{\theta - N\left(1 + 2\widetilde{\beta}\right)}. \tag{3.3}$$

Hypotheses H_6 or H_7 on the bandwidths are similar to that of Carbon et al. [19] and imply the classical condition in nonparametric estimation, that is, $\widehat{\mathbf{n}} h_1^d h_2 / \log \widehat{\mathbf{n}} \to +\infty$.

(H_8) $\widehat{\mathbf{n}} h_1^{d+2b_1} \to 0$ and $\widehat{\mathbf{n}} h_1^d h_2^{2b_2} \to 0$.

(H_9) Let $r > 2$ be a given nonnegative real number. One assume that there exits a sequence of integers $q = q_{\mathbf{n}}$ tending to infinity and such that

(i) $[\widehat{\mathbf{n}}(h_1^d)^{1+(1-2/r)2N}]^{1/(2N)} \to \infty$,

(ii) $q = o([\widehat{\mathbf{n}}(h_1^d)^{1+(1-2/r)2N}]^{1/(2N)})$,

(iii) $\widehat{\mathbf{n}} q^{-\theta} \to 0$,

(iv) $(h_1^d)^{-(1-2/r)} q^{-\theta(1-2/r)+N} \to 0$,

with $\theta > N/(1 - 2/r)$.

3.2. Dependency Conditions

In spatial dependent data analysis, the dependence of the observations has to be measured. Here we will consider the following two dependence measures.

3.2.1. Mixing Condition

The spatial dependence of the process will be measured by means of α-mixing. Then, we consider the α-mixing coefficients of the field $(Z_{\mathbf{i}}, \mathbf{i} \in \mathbb{Z}^N)$, defined by the following: there exists a function $\varphi(t) \downarrow 0$ as $t \to \infty$, such that E and E' subsets of $\mathbb{Z}^N$ with finite cardinals are

$$\alpha\left(\mathcal{B}(E), \mathcal{B}(E')\right) = \sup_{B \in \mathcal{B}(E), \ C \in \mathcal{B}(E')} |\mathbb{P}(B \cap C) - \mathbb{P}(B)\mathbb{P}(C)|$$
$$\leq \psi\left(\mathrm{Card}(E), \mathrm{Card}(E')\right)\varphi\left(\mathrm{dist}(E, E')\right), \tag{3.4}$$

where $\mathrm{Card}(E)$ (resp., $\mathrm{Card}(E')$) is the cardinality of E (resp., E'), $\mathrm{dist}(E, E')$ the Euclidean distance between E and E', and $\psi : \mathbb{N}^2 \to \mathbb{R}^+$ is a symmetric positive function nondecreasing in each variable. Throughout the paper, it will be assumed that ψ satisfies either

$$\psi(n, m) \leq C \min(n, m), \quad \forall n, m \in \mathbb{N} \tag{3.5}$$

or

$$\psi(n, m) \leq C(n + m + 1)^{\widetilde{\beta}}, \quad \forall n, m \in \mathbb{N} \tag{3.6}$$

for some $\widetilde{\beta} \geq 1$ and some $C > 0$. We assume also that the process satisfies a polynomial mixing condition:

$$\varphi(t) \leq Ct^{-\theta}, \quad \theta > 0, \ t \in \mathbb{R}_+^*. \tag{3.7}$$

If $\psi \equiv 1$, then $Z_{\mathbf{i}}$ is called strongly mixing. Many stochastic processes, among them various useful time series models, satisfy strong mixing properties, which are relatively easy to check. Conditions (3.5)-(3.6) are used in Tran [30] and Carbon et al. [18, 19]. See Doukhan [31] for discussion on mixing and examples.

3.2.2. Local Dependence Condition

Since we aim to get the same rate of convergence as in the *i.i.d.* case, we need some local dependency assumptions. Then, we assume the following local dependency condition used in Tran [30].

The joint probability density $f_{(X_{\mathbf{i}}, X_{\mathbf{j}})}$ of $(X_{\mathbf{i}}, X_{\mathbf{j}})$ exists and satisfies

$$\left| f_{(X_{\mathbf{i}}, X_{\mathbf{j}})}(u, v) - f_{X_{\mathbf{i}}}(u) f_{X_{\mathbf{j}}}(v) \right| \leq C \tag{3.8}$$

for some constant C and for all $u, v, \mathbf{i}, \mathbf{j}$.

In addition, let the density $f_{(Z_{\mathbf{i}}, Z_{\mathbf{j}})}$ of $(Z_{\mathbf{i}}, Z_{\mathbf{j}})$ exist and

$$\left| f_{(Z_{\mathbf{i}}, Z_{\mathbf{j}})}(z, t) - f_{Z_{\mathbf{i}}}(z) f_{Z_{\mathbf{j}}}(t) \right| \leq C; \quad \forall z, t \in \mathbb{R}^{d+1}, \mathbf{i}, \mathbf{j}, \tag{3.9}$$

where C is a positive constant.

In the following, the notations $\xrightarrow{\mathcal{D}}$ and $\xrightarrow{P}$ mean, respectively, convergences in distribution and in probability.

4. Consistency Results

This section contains results on asymptotic normality of the conditional quantile estimate. The main result of this paper is given by the following theorem.

Theorem 4.1. *Under assumptions H_1–H_5 and H_8-H_9, (3.4), (3.7), (3.8), (3.9), and (3.5) combined with H_6 or (3.6) combined with H_7, one has*

$$\left(\frac{(f_{X,Y}(x, q_\alpha(x)))^2 \widehat{\mathbf{n}} h_1^d}{V(x)} \right)^{1/2} (\widehat{q}_\alpha(x) - q_\alpha(x) - C_{\mathbf{n}}(x)) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1), \tag{4.1}$$

where $V(x) = \alpha(1 - \alpha) f_X(x) \int_{\mathbb{R}^d} (K_1(t))^2 dt$ and $C_{\mathbf{n}}(x) = (1/f_{X,Y}(x, q_\alpha(x))) E[\alpha \widehat{f}(x) - g_{\mathbf{n}}(x, q_\alpha(x))]$, with $g_{\mathbf{n}}$ and $\widehat{f}$ being the estimates defined in Section 2.

The proof of this theorem is based on the following three lemmas.

Lemma 4.2. *Under assumptions H_2–H_5 and H_9, (3.4), (3.7), (3.8), (3.9), and (3.5) or (3.6), one has*

$$\left[\frac{\widehat{n}h_1^d}{V(x)}\right]^{1/2}\left(\left[\alpha\widehat{f}(x) - g_n(x, q_\alpha(x))\right] - E\left[\alpha\widehat{f}(x) - g_n(x, q_\alpha(x))\right]\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1). \tag{4.2}$$

Lemma 4.3. *Under assumptions H_2, H_4, and H_5, one has:*

$$E\left[\alpha\widehat{f}(x) - g_n(x, q_\alpha(x))\right] = O\left(h_1^{b_1} + h_2^{b_2}\right). \tag{4.3}$$

Lemma 4.4. *Under assumptions of Theorem 4.1, one has:*

$$g_n^{(1)}(x, q_\alpha^*(x)) \longrightarrow f_{X,Y}(x, q_\alpha(x)) \text{ in probability,} \tag{4.4}$$

where $q_\alpha^*(x)$ is an element of the interval of extremities $q_\alpha(x)$ and $\widehat{q}_\alpha(x)$.

Proof of Theorem 4.1. By assumption H_5, $g_n(x, \cdot)$ is of class $\mathcal{C}^1$. Then a Taylor expansion on a neighborhood of $q_\alpha(x)$ gives

$$g_n(x, \widehat{q}_\alpha(x)) = g_n(x, q_\alpha(x)) + g_n^{(1)}(x, q_\alpha^*(x))(\widehat{q}_\alpha(x) - q_\alpha(x)), \tag{4.5}$$

where $q_\alpha^*(x)$ is an element of the interval of extremities $q_\alpha(x)$ and $\widehat{q}_\alpha(x)$.
 It follows that

$$\begin{aligned}
\widehat{q}_\alpha(x) - q_\alpha(x) &= \frac{1}{g_n^{(1)}(x, q_\alpha^*(x))}(g_n(x, \widehat{q}_\alpha(x)) - g_n(x, q_\alpha(x))) \\
&= \frac{1}{g_n^{(1)}(x, q_\alpha^*(x))}\left(\alpha\widehat{f}(x) - g_n(x, q_\alpha(x))\right).
\end{aligned} \tag{4.6}$$

Then, we have

$$\begin{aligned}
&\left(\frac{(f_{X,Y}(x, q_\alpha(x)))^2\widehat{n}h_1^d}{V(x)}\right)^{1/2}(\widehat{q}_\alpha(x) - q_\alpha(x) - C_n(x)) \\
&= \left(\frac{(f_{X,Y}(x, q_\alpha(x)))^2\widehat{n}h_1^d}{V(x)}\right)^{1/2} \\
&\quad \times \left(\frac{1}{g_n^{(1)}(x, q_\alpha^*(x))}\left(\alpha\widehat{f}(x) - g_n(x, q_\alpha(x))\right) - \frac{1}{f_{X,Y}(x, q_\alpha(x))}E\left[\alpha\widehat{f}(x) - g_n(x, q_\alpha(x))\right]\right)
\end{aligned}$$

$$= \left[\frac{\widehat{\mathbf{n}} h_1^d}{V(x)} \right]^{1/2} \times \frac{f_{X,Y}(x, q_\alpha(x))}{g_{\mathbf{n}}^{(1)}(x, q_\alpha^*(x))}$$

$$\times \left(\left(\alpha \widehat{f}(x) - g_{\mathbf{n}}(x, q_\alpha(x)) \right) - \frac{g_{\mathbf{n}}^{(1)}(x, q_\alpha^*(x))}{f_{X,Y}(x, q_\alpha(x))} E \left(\alpha \widehat{f}(x) - g_{\mathbf{n}}(x, q_\alpha(x)) \right) \right)$$

$$= \frac{f_{X,Y}(x, q_\alpha(x))}{g_{\mathbf{n}}^{(1)}(x, q_\alpha^*(x))} \left[\frac{\widehat{\mathbf{n}} h_1^d}{V(x)} \right]^{1/2}$$

$$\times \left(\left(\alpha \widehat{f}(x) - g_{\mathbf{n}}(x, q_\alpha(x)) \right) - E \left(\alpha \widehat{f}(x) - g_{\mathbf{n}}(x, q_\alpha(x)) \right) \right)$$

$$+ \frac{f_{X,Y}(x, q_\alpha(x))}{g_{\mathbf{n}}^{(1)}(x, q_\alpha^*(x))} \left[\frac{\widehat{\mathbf{n}} h_1^d}{V(x)} \right]^{1/2} \times \left(1 - \frac{g_{\mathbf{n}}^{(1)}(x, q_\alpha^*(x))}{f_{X,Y}(x, q_\alpha(x))} \right) E \left(\alpha \widehat{f}(x) - g_{\mathbf{n}}(x, q_\alpha(x)) \right).$$

$$(4.7)$$

Lemmas 4.2 and 4.4 and Slutsky's theorem imply that the first term of the right-hand side of the last equality above tends in distribution to $N(0, 1)$. In addition, Lemma 4.3 permits to write

$$\left[\frac{\widehat{\mathbf{n}} h_1^d}{V(x)} \right]^{1/2} E \left(\alpha \widehat{f}(x) - g_{\mathbf{n}}(x, q_\alpha(x)) \right) \leq C \sqrt{\widehat{\mathbf{n}} h_1^d} \left(h_1^{b_1} + h_2^{b_2} \right). \qquad (4.8)$$

This last tends to 0 by H_8. Thus, the second term goes to zero in probability. This yields the proof. $\qquad \square$

Before going further, it should be interesting to give examples where all our conditions on the bandwidths are satisfied. That is done through the two following remarks.

Remark 4.5. Let us, for example, choose $b_2 = b_1 = 1$ and $h_2 = h_1 = h = \widehat{\mathbf{n}}^{-\nu}$, where $\nu > 0$ is a real number such that H_8 is verified. That is, $\widehat{\mathbf{n}} h^{d+2} = \widehat{\mathbf{n}}^{-\nu(d+2)+1} \to 0$. Thus, it suffices to choose $\nu = (1/(d + 2)) + \eta$, where $\eta > 0$ is a real number arbitrarily small.

The hypothesis H_6 holds if $\widehat{\mathbf{n}} h^{\theta_1 + \theta_2} \log n^{\theta_3} = \widehat{\mathbf{n}} h^{d + ((\theta + 2Nd)/(\theta - 2N))} \log n^{\theta_3} \to \infty$. Then, it suffices to have $(\theta + 2Nd)/(\theta - 2N) < 2$. That is, $\theta > 2N(d + 2)$. Similarly H_7 holds if $\widehat{\mathbf{n}} h^{\theta_1' + \theta_2'} \log n^{\theta_3'} = \widehat{\mathbf{n}} h^{d + (2Nd(1+\widetilde{\beta}) + N + \theta)/(\theta - N(1+2\widetilde{\beta}))} \log n^{\theta_3'} \to \infty$. That is, $\theta > 2Nd(1+\widetilde{\beta}) + N(3+4\widetilde{\beta})$.

Remark 4.6. Let $\gamma_r = d(1 - 2/r)2N$ and $h_1 = h_2 = h = \widehat{\mathbf{n}}^{-(1/(d+2))-\eta}$ be defined in the previous remark. Condition (i) of H_9 is equivalent to $\widehat{\mathbf{n}}^{((2-\gamma_r)/(2N(d+2)))-\eta(d+\gamma_r)/2N} \to \infty$. Then, it suffices to have $((2-\gamma_r)/(2N(d+2))) - \eta(d+\gamma_r)/2N > 0$. As η is arbitrarily small, it is enough to have $\gamma_r < 2$. That is, $2 < r < 2Nd/(Nd - 1)$. In this case, for any positive real number τ such that $((2-\gamma_r)/(2N(d+2))) - \eta((d+\gamma_r+\tau)/2N) > 0$, the choice $q = \lfloor \widehat{\mathbf{n}}^{((2-\gamma_r)/(2N(d+2)))-\eta((d+\gamma_r+\tau)/2N)} \rfloor + 1$ (where $\lfloor \cdot \rfloor$ denotes the integer part) gives an example where H_9(ii) is satisfied.

Let $q_* = \widehat{\mathbf{n}}^{((2-\gamma_r)/(2N(d+2)))-\eta(d+\gamma_r+\tau)/2N}$. We have $\widehat{\mathbf{n}} q^{-\theta} \leq \widehat{\mathbf{n}} q_*^{-\theta}$. Thus, to ensure that condition H_9(iii) holds, it suffices to have $\widehat{\mathbf{n}} q_*^{-\theta} \to 0$. That is, $\theta > 2N(d + 2)/((2 - \gamma_r) - \eta(d + 2)(d + \gamma_r + \tau))$.

Lastly, the condition H_9(iv) is $h^{-d(1-2/r)} q^{-\theta(1-2/r)+N} \to 0$. It can be written equivalently as $h^{-\gamma_r/2N} q^{-\theta(\gamma_r/2Nd)+N} \to 0$ or $h^{-\gamma_r d} q^{-\theta\gamma_r+2N^2d} \to 0$. Under the condition $\theta > 2N^2 d/\gamma_r$, we have $h^{-\gamma_r d} q^{-\theta\gamma_r+2N^2d} \leq h^{-\gamma_r d} q_*^{-\theta\gamma_r+2N^2d}$. Then, to satisfy the condition H_9(iv), it is enough to

have $h^{-\gamma_r d} q_*^{-\theta \gamma_r + 2N^2 d} \to 0$. That is, $\theta > (2N^2 d/\gamma_r) + ((2N(d + \eta d(d + 2))(d + 2))/((2 - \gamma_r) - \eta(d + 2)(d + \gamma_r + \tau)))$.

Thus, H_6, H_8, and H_9 (resp., H_7–H_9) are satisfied when $2 < r < 2Nd/(Nd - 1)$ and

$$
\begin{aligned}
&\theta > \max\left(\frac{N}{1 - 2/r}, \; 2N(d + 2), \; \frac{2N(d + 2)}{(2 - \gamma_r) - \eta(d + 2)(d + \gamma_r + \tau)}, \right. \\
&\qquad\qquad\left. \frac{2N^2 d}{\gamma_r} + \frac{2N(d + \eta d(d + 2))}{(2 - \gamma_r) - \eta(d + 2)(d + \gamma_r + \tau)} \right), \\
&\left(\text{resp. } \theta > \max\left(\frac{N}{1 - 2/r}, \; 2Nd\left(1 + \widetilde{\beta}\right) + N\left(3 + 4\widetilde{\beta}\right), \; \frac{2N(d + 2)}{(2 - \gamma_r) - \eta(d + 2)(d + \gamma_r + \tau)}, \right.\right. \\
&\qquad\qquad\left.\left. \frac{2N^2 d}{\gamma_r} + \frac{2N(d + \eta d(d + 2))}{(2 - \gamma_r) - \eta(d + 2)(d + \gamma_r + \tau)} \right)\right).
\end{aligned}
\tag{4.9}
$$

5. Application

In this section, we present a quantile prediction procedure and then apply it to simulated data.

5.1. Prediction Procedure

An application where a multidimensional stationary spatial process may be observed is the case of prediction of a strictly stationary $\mathbb{R}$-valued random field $(\xi_{\mathbf{i}}, \; \mathbf{i} \in \mathbb{Z}^N)$ at a given fixed point $\mathbf{i}_0$ when observations are taken from a subset of $\mathcal{O}_{\mathbf{n}}$, not containing $\mathbf{i}_0$, see Biau and Cadre [20] or Dabo-Niang and Yao [27].

Assume that the value of the field at a given location depends on the values taken by the field in a vicinity $\mathcal{U}_{\mathbf{i}_0}$ of $\mathbf{i}_0$ ($\mathbf{i}_0 \notin \mathcal{U}_{\mathbf{i}_0}$); then the random variables whose components are the $\{\xi_{\mathbf{i}}, \; \mathbf{i} \in \mathcal{U}_{\mathbf{i}_0}\}$ are an $\mathbb{R}^d$-valued spatial process, where d is the cardinal of $\mathcal{U}_{\mathbf{i}_0}$. In other words, we expect that the process $(\xi_{\mathbf{i}})$ satisfies a Markov property; see, for example, Biau and Cadre [20] or Dabo-Niang and Yao [27]. Moreover, we assume that $\mathcal{U}_{\mathbf{i}_0} = \mathcal{U} + \mathbf{i}_0$, where $\mathcal{U}$ is a fixed bounded set of sites that does not contain $\mathbf{0}$.

Suppose that $(\xi_{\mathbf{i}})$ is bounded and observed over a subset $\mathcal{O}_{\mathbf{n}}$ of $\mathcal{O}_{\mathbf{n}}$. The aim of this section is to predict $\xi_{\mathbf{i}_0}$, at a given fixed point $\mathbf{i}_0$ not in $\mathcal{O}_{\mathbf{n}} \subset \mathbb{N}^N$. It is well known that the best predictor of $\xi_{\mathbf{i}_0}$ given the data in $\mathcal{U}_{\mathbf{i}_0}$ in the sense of *mean-square error* is

$$
E\left(\widehat{\xi}_{\mathbf{i}_0} \mid \xi_{\mathbf{i}}, \; \mathbf{i} \in \mathcal{U}_{\mathbf{i}_0} \right).
\tag{5.1}
$$

Let $\mathcal{U}_{\mathbf{i}} = \mathcal{U} + \mathbf{i} = \{\mathbf{u} + \mathbf{i}, \; \mathbf{u} \in \mathcal{U}\}$ for each $\mathbf{i} \in \mathbb{N}^N$. To define a predictor of $\xi_{\mathbf{i}_0}$, let us consider the $\mathbb{R}^d$-valued random variables $\widetilde{\xi}_{\mathbf{i}} = \{\xi_{\mathbf{u}}, \; \mathbf{u} \in \mathcal{U}_{\mathbf{i}} \subset \mathcal{O}_{\mathbf{n}}\}$. The notation of the previous sections is used by setting $X_{\mathbf{i}} = \widetilde{\xi}_{\mathbf{i}}, \; Y_{\mathbf{i}} = \xi_{\mathbf{i}}, \; \mathbf{i} \in \mathbb{N}^N$.

As a predictor of $\xi_{\mathbf{i}_0}$, one can take the conditional median estimate $\widehat{\xi}_{\mathbf{i}_0} = \widehat{q}_{0.5}(\widetilde{\xi}_{\mathbf{i}_0})$.

We deduce from the previous consistency results the following corollary that gives the convergence of the predictor $\widehat{\xi}_{\mathbf{i}_0}$.

Corollary 5.1. *Under the conditions of Theorem 4.1*

$$
\left(\frac{\left(\widehat{f}^{\widetilde{\xi}_{\mathbf{i}_0}} \left(q_{0.5} \left(\widetilde{\xi}_{\mathbf{i}_0} \right) \right) \right)^2 \widehat{f} \left(\widetilde{\xi}_{\mathbf{i}_0} \right) \widehat{\mathbf{n}} h_1^d}{\widehat{V} \left(\widetilde{\xi}_{\mathbf{i}_0} \right)} \right)^{1/2} \left(\widehat{\xi}_{\mathbf{i}_0} - q_{0.5} \left(\widetilde{\xi}_{\mathbf{i}_0} \right) \right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1),
\tag{5.2}
$$

where

$$
\widehat{V}(x) = \alpha(1-\alpha)\widehat{f}(x) \int_{\mathbb{R}^d} \left(K_1(t) \right)^2 dt \quad (\alpha = 0.5).
\tag{5.3}
$$

This consistency result permits to have an approximation of a confidence interval and predictive intervals that consists of the $(\alpha_2 - \alpha_1)100\%$ confidence intervals with bounds $\widehat{q}_{\alpha_1}(\widetilde{\xi}_{\mathbf{i}_0})$ and $\widehat{q}_{\alpha_2}(\widetilde{\xi}_{\mathbf{i}_0})$, $(\alpha_1 < \alpha_2)$.

5.2. Numerical Properties

In this section, we study the performance of the conditional quantile predictor introduced in the previous section towards some simulations. Let us denote by $\mathrm{GRF}(m, \sigma^2, s)$ a Gaussian random field with mean m and covariance function defined by

$$
\vartheta(h) = \sigma^2 \exp\left\{ -\left(\frac{\|h\|}{s} \right)^2 \right\}, \quad h \in \mathbb{R}^2.
\tag{5.4}
$$

We consider a random field $(\xi_{\mathbf{i}})_{\mathbf{i} \in \mathbb{N}^2}$ from the three following models.

Model 1. One has

$$
\xi_{\mathbf{i}} = U_{\mathbf{i}} * \left(\sin(2X_{\mathbf{i}}) + 2\exp\left\{ -16X_{\mathbf{i}}^2 \right\} \right) + Z_{\mathbf{i}}, \quad \mathbf{i} \in \mathbb{N}^2.
\tag{5.5}
$$

Model 2. One has

$$
\xi_{\mathbf{i}} = U_{\mathbf{i}} * \left(2 + 2\cos(2X_{\mathbf{i}}) + \exp\left\{ -4X_{\mathbf{i}}^2 \right\} \right) + Z_{\mathbf{i}}, \quad \mathbf{i} \in \mathbb{N}^2.
\tag{5.6}
$$

Model 3. One has

$$
\xi_{\mathbf{i}} = U_{\mathbf{i}} * \left(1 + \left(\frac{2}{3} \right) \exp\{X_{\mathbf{i}}\} \right) + Z_{\mathbf{i}}, \quad \mathbf{i} \in \mathbb{N}^2,
\tag{5.7}
$$

where $X = (X_{\mathbf{i}}, \mathbf{i} \in \mathbb{N}^2)$ is a $\mathrm{GRF}(0, 5, 3)$, $Z = (Z_{\mathbf{i}}, \mathbf{i} \in \mathbb{N}^2)$ is a $\mathrm{GRF}(0, 0.1, 5)$ independent of X, and $U_{\mathbf{i}} = (1/\widehat{\mathbf{n}}) \sum_{\mathbf{j} \in \mathcal{I}_{\mathbf{n}}} \exp(-\|\mathbf{i} - \mathbf{j}\|/2)$. The choice of $U_{\mathbf{i}}$ in the models is motivated by a reinforcement of the spatial local dependency. Set

$$
\mathcal{I}_{\mathbf{n}} = \left\{ \mathbf{i} = (i,j) \in (\mathbb{N}^*)^2, \ 1 \le i \le 61, \ 1 \le j \le 61 \right\}.
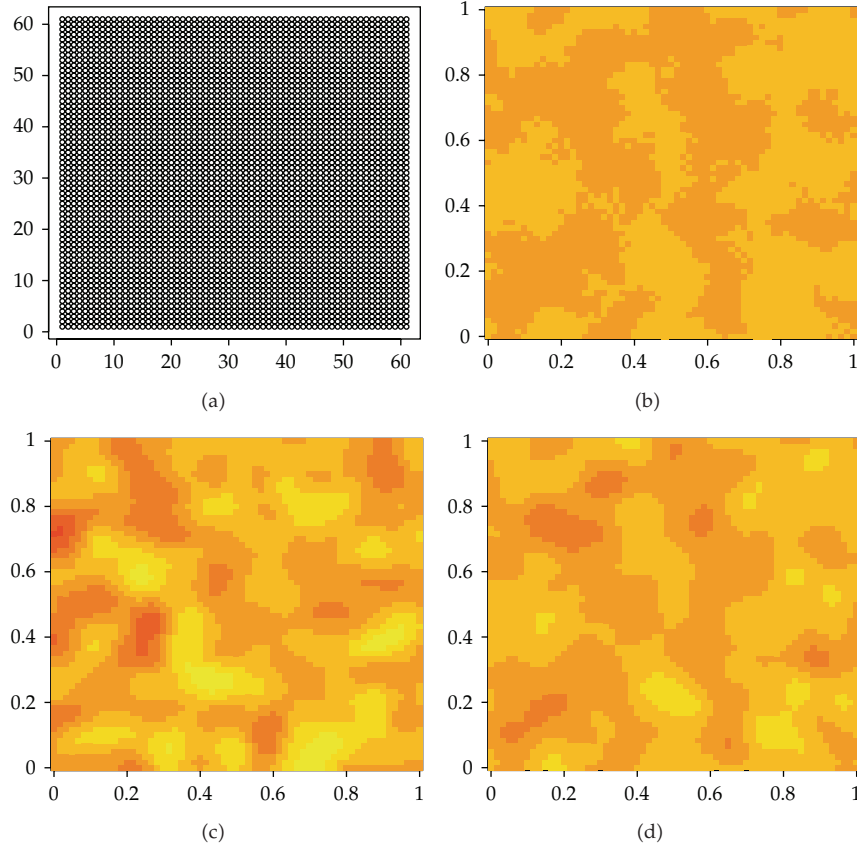\tag{5.8}
$$

Figure 1: The observation sites and the random fields ($\xi_\mathbf{i}$) of Models 1, 2, and 3 in, respectively, Panels (a), (b), (c), and (d).

We supposed that the field ($\xi_\mathbf{i}, \mathbf{i} \in \mathbb{N}^2$) is observable over the rectangular region $\mathcal{I}_\mathbf{n}$, observed over a subset $\mathcal{O}_\mathbf{n}$ and nonobserved in a subset $V_\mathbf{n}$. A sample of size $\widehat{\mathbf{n}} = 61 * 61 = 3721$ obtained from each model is plotted in Figure 1.

For the prediction purpose, subsets $\mathcal{O}_\mathbf{n}$ of size $M = 441$ and $V_\mathbf{n}$ (different with respect to the model) and the quantiles of order $0.5, 0.025, 0.975$ have been considered.

We want to predict the values $\xi_{\mathbf{i}_1}, \ldots, \xi_{\mathbf{i}_m}$ at given fixed sites $\mathbf{i}_1, \ldots, \mathbf{i}_m$ in $V_\mathbf{n}$, with $m = 10$. We provide the plots of the kernel densities estimators of the field of each model in Figure 2. The distributions of the models look asymmetric and highly heteroskedastic in Model 3. These graphics exhibit a strongly bimodal profile of Model 3. That means that a simple study of conditional mean or conditional median can miss some of the essential features of the dataset.

As explained above, for any $k \in \{1, \ldots, m\}$, we take the conditional median estimate $\widehat{\xi}_{\mathbf{i}_k}$, as a predictor of $\xi_{\mathbf{i}_k}$. We compute these predictors with the vicinity $\mathcal{U} = \{-1, 0, 1\} \times \{-1, 0, 1\} \setminus \{(0, 0)\}$ and select the standard normal density as kernel K_1 and the Epanechnikov kernel as K. For the bandwidth selection, we use two bandwidth choices.
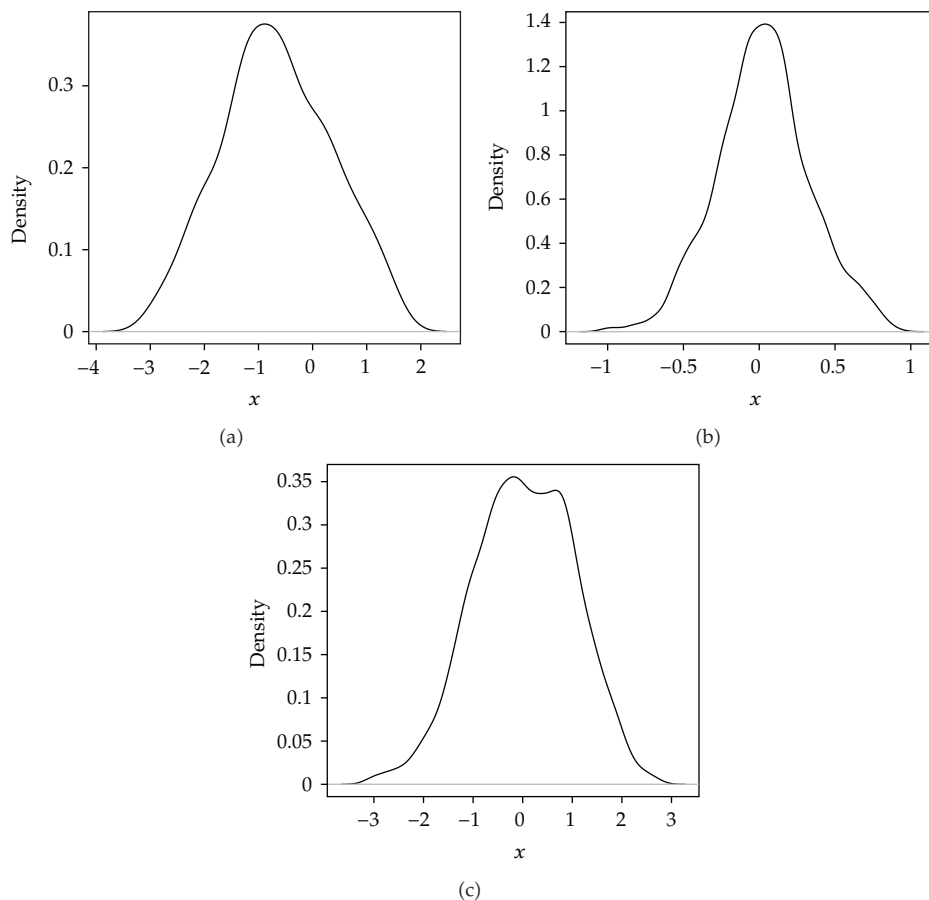
(a)

(b)

(c)

Figure 2: The densities of the random fields (ξ_i) of Models 1–3 in, respectively, Panels (a), (b), and (c).

(1) We consider first the rule developed in Yu and Jones [32] and take $h_1 = h_2 = h_{\mathbf{n}}$:

$$h_{\mathbf{n}} = h_{\text{mean}} \left(\frac{\alpha(1-\alpha)}{\phi\left(\Phi^{-1}(\alpha)\right)^2} \right)^{1/(4+d)}, \tag{5.9}$$

where h_{mean} is the bandwidth for kernel smoothing estimation of the regression mean obtained by cross-validation, and ϕ and Φ are, respectively, the standard normal density and distribution function.

(2) The second rule is to choose first $h_1 = h_{\mathbf{n}}$, and once h_1 is computed, use it and compute h_2 by cross-validation using the conditional distribution function estimate $\widehat{F}^x(\cdot)$.

Remark 5.2. The selection of the appropriate bandwidths in double-kernel smoothing conditional quantile is an important task as in classical smoothing techniques.

As we see above, one can choose either $h_1 = h_2$ or different bandwidths. The first choice can be motivated by the paper of Gannoun et al. [12] who dealt with a double kernel

estimator of the conditional quantile function with a single bandwidth. They compute the common value of the bandwidth by the use of the rule developed in Yu and Jones [32] for local linear quantile regression. One can also consider the rule where (see Gannoun et al., 2003) $h_1 = h_2 = h_n$ where h_{mean}, instead of being the bandwidth obtained by cross-validation of the regression mean estimate, is

$$h_{mean} = \left(C \frac{R_1(K_1)\sigma(x)}{\widehat{n}R_2(K_1)^2(m''(x))^2 g(x)} \right)^{1/(4+d)}, \tag{5.10}$$

where

$$R_1(K_1) = \int_{-\infty}^{\infty} K_1^2(u)du, \qquad R_2(K_1) = \int_{-\infty}^{\infty} u^2 K_1(u)du, \tag{5.11}$$

$m(x)$ and $\sigma(x)$ are, respectively, the conditional mean and conditional variance, and C is a constant. One can also choose the bandwidths using the following mean square error result of the conditional quantile estimate:

$$\text{MSE}(\widehat{q}_\alpha(x)) \approx \frac{R_1(K_1)\alpha(1-\alpha)}{\widehat{n}h_1^d g(x) f^x(q_\alpha(x))^2} + o\left(h_1^{2b_1} + h_2^{2b_2} + h_1^{b_1} h_2^{b_2} \right). \tag{5.12}$$

In the case of different bandwidths h_1 and h_2, we may still use the same rule as above for h_1 ($h_1 = h_n$). Therefore, only the choice of h_2 seems to deserve more future theoretical investigation since the simulations (see Tables 4 and 5) suggest that different choices of these two bandwidths are appreciable. This theoretical choice of h_2 (by, e.g., cross-validation) is beyond the scope of this paper and deserves future investigations.

To evaluate the performance of the quantile predictor $\widehat{\breve{\xi}}_{i_k}$ (with $h_1 = h_2$) or $\widehat{\breve{\xi}}_{i_k}^2$ (with $h_1 \neq h_2$) and compare it to the mean regression predictor (see Biau and Cadre [20]):

$$\breve{\xi}_{i_k} = \frac{\sum_{i \in \mathcal{O}_n} \breve{\xi}_i K\left(\left(\widetilde{\breve{\xi}}_{i_k} - \widetilde{\breve{\xi}}_i \right)/h_K \right)}{\sum_{i \in \mathcal{O}_n} K\left(\left(\widetilde{\breve{\xi}}_{i_k} - \widetilde{\breve{\xi}}_i \right)/h_K \right)}, \tag{5.13}$$

we compute the mean absolute errors:

$$\text{MAE} = \frac{1}{10}\sum_{k=1}^{10} \left| \breve{\xi}_{i_k} - \xi_{i_k}^n \right|, \quad \text{where } \xi_{i_k}^n = \widehat{\breve{\xi}}_{i_k}, \ \widehat{\breve{\xi}}_{i_k}^2, \text{ or } \breve{\xi}_{i_k}. \tag{5.14}$$

Tables 1–6 give the quantile estimates for $\breve{\xi}_{i_k}$, $k = 1, \ldots, m$, $p \in \{0.025, 0.5, 0.975\}$, $\mathcal{U} = \{-1, 0, 1\} \times \{-1, 0, 1\} \setminus \{(0, 0)\}$, and the quantile and mean regression prediction errors of the predictors $\widehat{\breve{\xi}}_{i_k}$ and $\widetilde{\breve{\xi}}_{i_k}$. In each table, Inf (resp., Sup) is the lower (resp., upper) bound of the confidence interval (see above) of the median estimator. The bandwidth choice considered in these first three tables is the first choice $h_1 = h_2 = h_n$. Clearly, the predictors give good results.

Table 1: Quantile estimates ($\hat{q}_\alpha(\widetilde{\xi}_{i_0})$, $\alpha \in \{0.025, 0.5, 0.975\}$) and mean regression predictor of Model 1 and confidence interval for $\mathcal{U}$ and $\hat{n} = 441$.

$\hat{\mathbf{n}} = 441$	$\alpha = 0.025$	Inf	True	$\hat{\widehat{\xi}}_{\mathbf{i}_k}$	Sup	$\alpha = 0.975$	$\check{\xi}_{\mathbf{i}_k}$
	−0.341	−0.267	**−0.204**	−0.170	−0.074	0.056	−0.160
	−0.701	−0.556	**−0.560**	−0.483	−0.410	−0.150	−0.481
	−0.558	− 0.317	**−0.232**	−0.246	−0.175	0.017	−0.278
	−0.369	−0.118	**−0.107**	−0.072	−0.026	0.196	−0.056
	−0.649	−0.452	**−0.225**	−0.319	−0.185	−0.002	−0.481
	0.008	0.192	**0.292**	0.271	0.349	0.513	0.270
	−0.402	−0.273	**−0.235**	−0.185	−0.097	0.068	−0.140
	−0.701	−0.603	**−0.489**	−0.483	−0.363	−0.088	−0.507
	−0.613	−0.486	**−0.353**	−0.382	−0.278	−0.130	−0.457
	−0.613	−0.529	**−0.463**	−0.437	−0.345	−0.193	−0.444
MAE				0.038			0.073

Table 2: Quantile estimates ($\hat{q}_\alpha(\widetilde{\xi}_{i_0})$, $\alpha \in \{0.025, 0.5, 0.975\}$) and mean regression predictor of Model 1 and confidence interval for $\mathcal{U}$ and $\hat{n} = 441$.

$\hat{\mathbf{n}} = 441$	$\alpha = 0.025$	Inf	True	$\hat{\widehat{\xi}}_{\mathbf{i}_k}$	Sup	$\alpha = 0.975$	$\check{\xi}_{\mathbf{i}_k}$
	−0.486	−0.311	**−0.228**	−0.248	−0.185	0.033	−0.248
	−0.190	−0.165	**−0.107**	−0.100	−0.036	0.080	−0.095
	−0.138	−0.314	**−0.077**	−0.073	0.168	−0.006	−0.079
	−0.508	−0.615	**−0.366**	−0.367	−0.119	−0.285	−0.368
	−0.266	−0.349	**−0.162**	−0.171	0.008	−0.011	−0.176
	0.154	0.196	**0.251**	0.258	0.320	0.372	0.258
	−0.285	−0.267	**−0.118**	−0.136	−0.004	−0.057	−0.144
	−0.432	−0.494	**−0.338**	−0.336	−0.177	−0.032	−0.316
	0.046	−0.351	**0.212**	0.215	0.781	0.358	0.213
	−0.015	0.043	**0.162**	0.154	0.266	0.372	0.148
MAE				0.008			0.012

The median prediction errors are rather smaller than those of mean regression for the first two models (less than 0.038 (resp., 0.012) for Model 1 (resp., Model 2)). Notice that the quantile and mean regression prediction results of Model 3 are similar and rather worse than those of Models 1 and 2. This can be explained by the fact that multimodality of the field of Model 3 cannot be captured by conditional median and the conditional mean.

We derive from the results of the tables 95% confidence and predictive intervals where the extremities are the 2.5% and 97.5% quantiles estimates, for each of the 10 prediction sites. Note that the confidence interval is generally more precise than the predictive one.

Tables 4–6 give the same estimates (for Models 1–3) as the first three tables but using both the same and the different choices of h_1 and h_2 described the aftermentioned rules. As proved by the numerical experiments (all the results are not presented here for seek of simplicity), the prediction results are improved when the sample size increases but notice that the improvement depends on the number of sites which are dependent on the prediction site in interest. Depending on the position of the sites, the conditional mean regression gives sometimes better prediction than the conditional median and vice versa. A next step would be to apply the predictor to a spatial real data that deserves future investigations.

Table 3: Quantile estimates ($\widehat{q}_\alpha(\widetilde{\xi}_{i_0})$, $\alpha \in \{0.025, 0.5, 0.975\}$) and mean regression predictor of Model 2 and confidence interval for $\mathcal{U}$ and $\widehat{n} = 441$.

$\widehat{n} = 441$	$\alpha = 0.025$	Inf	True	$\widehat{\xi}_{\mathbf{i}_k}$	Sup	$\alpha = 0.975$	$\widecheck{\xi}_{\mathbf{i}_k}$
	−0.508	−0.331	**−0.269**	−0.276	−0.222	−0.140	−0.303
	−0.057	0.069	**0.123**	0.120	0.172	0.291	0.140
	−0.487	−0.428	**−0.357**	−0.367	−0.306	−0.175	−0.399
	−0.550	−0.484	**−0.397**	−0.402	−0.321	−0.267	−0.373
	−0.011	0.136	**0.177**	0.177	0.218	0.373	0.153
	0.011	0.093	**0.132**	0.132	0.171	0.291	0.125
	−0.069	0.111	**0.150**	0.143	0.175	0.324	0.134
	0.154	0.266	**0.314**	0.316	0.365	0.410	0.309
	0.068	0.160	**0.227**	0.227	0.294	0.366	0.237
	−0.402	−0.328	**−0.276**	−0.278	−0.228	−0.128	−0.296
MAE				0.004			0.02

Table 4: Quantile estimates ($\widehat{q}_\alpha(\widetilde{\xi}_{i_0})$, $\alpha \in \{0.025, 0.5, 0.975\}$) for different bandwidth choices and the conditional median $\widehat{\xi}_{i_k}$ obtained with the same bandwidths h_1 and h_2 of Model 1 and confidence interval for $\mathcal{U}$ and $\widehat{n} = 441$.

$\widehat{n} = 441$	$\alpha = 0.025$	Inf	True	$\widehat{\xi}_{\mathbf{i}_k}^2$	Sup	$\alpha = 0.975$	$\widehat{\xi}_{\mathbf{i}_k}$
	−1.13	−1.27	**−0.71**	−0.73	−0.2	−0.19	−1.13
	−0.67	−3.78	**0.02**	−0.2	3.37	0.18	−0.67
	−1.13	−2.73	**−0.44**	−0.71	1.32	−0.03	−0.92
	0.33	0.74	**0.93**	0.83	0.93	1.22	0.33
	−0.44	−0.47	**0.09**	−0.06	0.36	0.18	−0.16
	0.18	0.41	**0.28**	0.56	1.35	0.97	0.18
	0.35	0.68	**0.89**	0.84	1.01	1.41	0.46
	0.11	0.54	**0.62**	0.69	0.84	1.38	0.33
	0.18	0.28	**0.24**	0.53	0.79	0.79	0.18
	0.64	0.82	**0.95**	0.95	1.09	1.25	0.64
MAE				0.15			0.37

Table 5: Quantile estimates ($\widehat{q}_\alpha(\widetilde{\xi}_{i_0})$, $\alpha \in \{0.025, 0.5, 0.975\}$) for different bandwidth choices and the conditional median $\widehat{\xi}_{i_k}$ obtained with the same bandwidths h_1 and h_2 of Model 2 and confidence interval for $\mathcal{U}$ and $\widehat{n} = 441$.

$\widehat{n} = 441$	$\alpha = 0.025$	Inf	True	$\widehat{\xi}_{\mathbf{i}_k}^2$	Sup	$\alpha = 0.975$	$\widehat{\xi}_{\mathbf{i}_k}$
	−2.18	−1.97	**−1.4**	−1.69	−1.4	−1.17	−1.97
	−1.02	−0.48	**−0.96**	−0.39	−0.31	0.38	−1.02
	−0.98	−0.88	**−0.81**	−0.76	−0.63	−0.39	−1.12
	−0.88	−0.84	**−0.44**	−0.59	−0.35	−0.04	−0.88
	−0.16	0.3	**0.28**	0.41	0.51	0.91	−0.25
	−0.88	−0.55	**−0.52**	−0.43	−0.3	0.23	−0.88
	0.64	1.21	**1.52**	1.29	1.37	1.94	0.57
	−0.76	−0.41	**−0.09**	−0.14	0.14	0.54	−0.59
	0.03	0.42	**0.73**	0.54	0.67	1.29	−0.04
	−0.76	−0.55	**−0.45**	−0.39	−0.22	−0.04	−0.76
MAE				0.18			0.48

Table 6: Quantile estimates $(\hat{q}_\alpha(\tilde{\xi}_{i_0}))$, $\alpha \in \{0.025, 0.5, 0.975\})$ for different bandwidth choices and the conditional median $\hat{\xi}_{i_k}$ obtained with the same bandwidths h_1 and h_2 of Model 3 and confidence interval for $\mathcal{U}$ and $\hat{n} = 441$.

$\hat{n} = 441$	$\alpha = 0.025$	Inf	True	$\hat{\xi}^2_{i_k}$	Sup	$\alpha = 0.975$	$\hat{\xi}_{i_k}$
	0.19	0.7	**0.5**	0.79	0.88	1.33	0.79
	−0.18	0.24	**0.07**	0.35	0.47	1.19	0.35
	−0.48	−1.7	**−0.48**	−0.14	1.42	0.29	−0.14
	−2.35	−3.02	**−2.54**	−2.17	−1.31	−1.68	−2.17
	−2.35	−2.29	**−2.1**	−2.17	−2.05	−1.41	−2.17
	0.28	0.65	**0.77**	0.75	0.85	1.65	0.77
	−2.35	−6.14	**−1.44**	−1.68	6.11	−0.47	−1.68
	−0.66	−0.06	**0.23**	0.21	0.49	1.12	0.21
	−0.04	0.32	**0.62**	0.42	0.53	1.19	0.42
	−0.66	−0.63	**−0.3**	−0.1	0.43	0.69	−0.1
MAE				0.2			0.2

Appendix

This section is devoted to proofs of lemmas used to establish the main result of Section 4. To this end, some needed additional lemmas and results will be stated and proved in this section.

Before giving the proof of Lemma 4.2, we introduce the following notation and establish the following preliminary lemma.

Let $L_{\mathbf{i}}$ be the random variable defined by

$$L_{\mathbf{i}} = \frac{1}{h_1^{d/2}} \left[\alpha K_{\mathbf{i}} - K_{\mathbf{i}} H_{\mathbf{i}}(q_\alpha(x)) - E(\alpha K_{\mathbf{i}} - K_{\mathbf{i}} H_{\mathbf{i}}(q_\alpha(x))) \right]. \tag{A.1}$$

Lemma A.1. *Under assumptions H_2, H_4, and H_5, (3.4), (3.7) (with $\theta > N + 1$), and (3.5) or (3.6), one has for all $\mathbf{i}$ and $\mathbf{j}$*

(i) $\text{var}(L_{\mathbf{i}}) \to V(x)$,

(ii) $\sum_{\mathbf{i},\mathbf{j} \in \mathcal{O}_n, \mathbf{i} \neq \mathbf{j}} \text{Cov}(L_{\mathbf{i}}, L_{\mathbf{j}}) = o(\hat{n})$,

(iii) $(1/\hat{n}) \, \text{var}(\sum_{\mathbf{i} \in \mathcal{O}_n} L_{\mathbf{i}}) \to V(x)$,

where $V(x)$ is defined in Theorem 4.1.

Remark A.2. If one takes $2 < r \leq 4$ in H_9, then $\theta > N + 1$.

Proof of Lemma A.1. Note that we can deduce easily from assumption H_4 the existence of two positive constants C and C' such that

$$C h_1^d \leq E K_{\mathbf{i}}^s \leq C' h_1^d, \tag{A.2}$$

for all $\mathbf{i}$ and any given integer s.

Now, let us calculate the variance term. We have

$$\operatorname{var}(L_i) = \frac{1}{h_1^d}\left[EK_i^2(\alpha - H_i(q_\alpha(x)))^2 - (EK_i(\alpha - H_i(q_\alpha(x))))^2 \right] = A_1 - A_2. \tag{A.3}$$

Let us first consider A_2. By taking the conditional expectation with respect to X_i, we get

$$\begin{aligned} \left| EK_i(H_i(q_\alpha(x)) - \alpha) \right| &= \left| EK_i[E(H_i(q_\alpha(x)) \mid X_i) - \alpha] \right| \\ &\leq EK_i \left| E(H_i(q_\alpha(x)) \mid X_i) - \alpha \right|. \end{aligned} \tag{A.4}$$

From an integration by parts and an usual change of variables, we obtain

$$\begin{aligned} \left| E(H_i(q_\alpha(x)) \mid X_i) - \alpha \right| &= \left| \int_{\mathbb{R}} K_2\left(\frac{q_\alpha(x) - z}{h_2} \right) f^{X_i}(z) dz - F^x(q_\alpha(x)) \right| \\ &= \left| \int_{\mathbb{R}} (h_2)^{-1} K\left(\frac{q_\alpha(x) - z}{h_2} \right) F^{X_i}(z) dz - F^x(q_\alpha(x)) \right| \\ &\leq \int_{\mathbb{R}} K(t) \left| F^{X_i}(q_\alpha(x) - th_2) - F^x(q_\alpha(x)) \right| dt. \end{aligned} \tag{A.5}$$

Then, using, respectively, H_4, H_2, and H_5, we have

$$\begin{aligned} \mathbb{1}_{\|X_i - x\| \leq h_1} \left| E(H_i(q_\alpha(x)) \mid X_i) - \alpha \right| &\leq C\left(h_1^{b_1} + h_2^{b_2} \int_{\mathbb{R}} |t|^{b_2} K(t) dt \right) \\ &\leq C\left(h_1^{b_1} + h_2^{b_2} \right). \end{aligned} \tag{A.6}$$

Thus, we can write

$$\begin{aligned} A_2 &= \frac{1}{h_1^d} \left| EK_i(\alpha - H_i(q_\alpha(x))) \right|^2 \leq \frac{1}{h_1^d} \left(EK_i \left| \alpha - E(H_i(q_\alpha(x)) \mid X_i) \right| \right)^2 \\ &\leq C\left(h_1^{b_1} + h_2^{b_2} \right)^2 \frac{1}{h_1^d} (EK_i)^2. \end{aligned} \tag{A.7}$$

Using (A.2) and the fact that the bandwidths tend to 0, we get $A_2 = O(h_1^{b_1} + h_2^{b_2})$.

Concerning A_1, we have

$$(\alpha - H_i(q_\alpha(x)))^2 = \left(H_i^2(q_\alpha(x)) - \alpha \right) - 2\alpha \cdot (H_i(q_\alpha(x)) - \alpha) + \alpha - \alpha^2. \tag{A.8}$$

Then, we can write

$$A_1 = \frac{1}{h_1^d} \left[EK_i^2\left(H_i^2(q_\alpha(x)) - \alpha \right) - 2\alpha EK_i^2(H_i(q_\alpha(x)) - \alpha) \right] + \alpha(1 - \alpha)\frac{EK_i^2}{h_1^d}. \tag{A.9}$$

The conditional expectation with respect to X_i permits to write

$$
A_1 = \frac{1}{h_1^d} EK_i^2 \left[E\left(H_i^2(q_\alpha(x)) \mid X_i \right) - \alpha \right] - \frac{2\alpha}{h_1^d} EK_i^2 \left[E(H_i(q_\alpha(x)) \mid X_i) - \alpha \right]
$$
$$
+ \alpha(1-\alpha)\frac{EK_i^2}{h_1^d}.
$$
(A.10)

For the second term of the right-hand side of (A.10), the same argument as that used above and (A.2) permit to obtain

$$
\frac{2\alpha}{h_1^d} EK_i^2 \left| \left[E(H_i(q_\alpha(x)) \mid X_i) - \alpha \right] \right| = O\left(h_1^{b_1} + h_2^{b_2} \right).
$$
(A.11)

For the first term, an integration by parts, assumptions H_2, H_4, and H_5 lead to

$$
\mathbb{1}_{\|X_i - x\| \le h_1} \left| E\left(H_i^2(q_\alpha(x)) \mid X_i \right) - \alpha \right|
$$
$$
= \mathbb{1}_{\|X_i - x\| \le h_1} \left| \int_{\mathbb{R}} K_2^2\left(\frac{q_\alpha(x) - z}{h_2} \right) f^{X_i}(z) dz - F^x(q_\alpha(x)) \right|
$$
$$
= \mathbb{1}_{\|X_i - x\| \le h_1} \left| \int_{\mathbb{R}} 2K_2(t)K(t)\left(F^{X_i}(q_\alpha(x) - h_2 t) - F^x(q_\alpha(x)) \right) dt \right|
$$
$$
\le Ch_1^{b_1} \int_{\mathbb{R}} 2K_2(t)K(t) dt + Ch_2^{b_2} \int_{\mathbb{R}} 2K_2(t)|t|^{b_2} K(t) dt
$$
$$
\le Ch_1^{b_1} + Ch_2^{b_2} \int_{\mathbb{R}} 2|t|^{b_2} K(t) dt \le C\left(h_1^{b_1} + h_2^{b_2} \right).
$$
(A.12)

Thus, we have

$$
\frac{1}{h_1^d} \left| EK_i^2 \left[E\left(H_i^2(q_\alpha(x)) \mid X_i \right) - \alpha \right] \right| \le C\left(h_1^{b_1} + h_2^{b_2} \right) \frac{1}{h_1^d} EK_i^2.
$$
(A.13)

From (A.2), we have $(1/h_1^d)EK_i^2 \le C$. It is then clear that

$$
\frac{1}{h_1^d} \left| EK_i^2 \left[E\left(H_i^2(q_\alpha(x)) \mid X_i \right) - \alpha \right] \right| \le C\left(h_1^{b_1} + h_2^{b_2} \right).
$$
(A.14)

Therefore, the term A_1 has the same limit as the last term of the right-hand side of (A.10) which goes directly (by Bochner's lemma) to $V(x) = \alpha(1-\alpha) f_X(x) \int_{\mathbb{R}^d} (K_1(t))^2 dt$.

Consequently, we have $\mathrm{var}(L_i) \to V(x)$. This gives the part (i) of the lemma.

Let us now focus on the covariance term. Let

$$
S_1 = \{ i, j \in \mathcal{O}_n : 0 < \|i - j\| \le c_n \}, \qquad S_2 = \{ i, j \in \mathcal{O}_n : \|i - j\| > c_n \},
$$
(A.15)

where $c_{\mathbf{n}}$ is a sequence of integers that converges to infinity and will be given later. We have

$$
\begin{aligned}
|\mathrm{Cov}(L_{\mathbf{i}}, L_{\mathbf{j}})| &= |EL_{\mathbf{i}}L_{\mathbf{j}}| \\
&= \left| \frac{1}{h_1^d} \left[EK_{\mathbf{i}}K_{\mathbf{j}}(\alpha - H_{\mathbf{i}}(q_\alpha(x)))(\alpha - H_{\mathbf{j}}(q_\alpha(x))) - (EK_{\mathbf{i}}(\alpha - H_{\mathbf{i}}(q_\alpha(x))))^2 \right] \right| \\
&\le \frac{1}{h_1^d} EK_{\mathbf{i}}K_{\mathbf{j}} |(\alpha - H_{\mathbf{i}}(q_\alpha(x)))(\alpha - H_{\mathbf{j}}(q_\alpha(x)))| + \frac{1}{h_1^d} \left[EK_{\mathbf{i}}(\alpha - H_{\mathbf{i}}(q_\alpha(x))) \right]^2 .
\end{aligned}
\tag{A.16}
$$

Taking the conditional expectation with respect to $X_{\mathbf{i}}$ and using the convexity of the function $x \mapsto |x|$, we have

$$
\begin{aligned}
\frac{1}{h_1^d} \left[EK_{\mathbf{i}}(\alpha - H_{\mathbf{i}}(q_\alpha(x))) \right]^2 &= \frac{1}{h_1^d} |EK_{\mathbf{i}}(\alpha - E(H_{\mathbf{i}}(q_\alpha(x)) \mid X_{\mathbf{i}}))|^2 \\
&\le \frac{1}{h_1^d} \left[EK_{\mathbf{i}} |E(H_{\mathbf{i}}(q_\alpha(x)) \mid X_{\mathbf{i}}) - \alpha| \right]^2 .
\end{aligned}
\tag{A.17}
$$

Since K_2 is a distribution function, we have clearly $|E(H_{\mathbf{i}}(q_\alpha(x)) \mid X) - \alpha| \le C$. Thus, by (A.2) we get $(1/h_1^d)[EK_{\mathbf{i}}(\alpha - H_{\mathbf{i}}(q_\alpha(x)))]^2 \le Ch_1^d$. As $|(\alpha - H_{\mathbf{i}}(q_\alpha(x)))(\alpha - H_{\mathbf{j}}(q_\alpha(x)))| \le C$, we have

$$
\begin{aligned}
\frac{1}{h_1^d} &\left[EK_{\mathbf{i}}K_{\mathbf{j}} |(\alpha - H_{\mathbf{i}}(q_\alpha(x)))(\alpha - H_{\mathbf{j}}(q_\alpha(x)))| \right] \\
&\le C \frac{1}{h_1^d} \left[EK_{\mathbf{i}}K_{\mathbf{j}} \right] \\
&= C \frac{1}{h_1^d} \int_{\mathbb{R}^{2d}} K_1 \left(\frac{x-t}{h_1} \right) K_1 \left(\frac{x-s}{h_1} \right) f_{(X_{\mathbf{i}}, X_{\mathbf{j}})}(t, s) dt \, ds .
\end{aligned}
\tag{A.18}
$$

Remark that by condition (3.8), the density $f_{(X_{\mathbf{i}}, X_{\mathbf{j}})}$ is bounded. Then, we have

$$
\begin{aligned}
\frac{1}{h_1^d} &\left[EK_{\mathbf{i}}K_{\mathbf{j}} |(\alpha - H_{\mathbf{i}}(q_\alpha(x)))(\alpha - H_{\mathbf{j}}(q_\alpha(x)))| \right] \\
&\le C \frac{1}{h_1^d} \int_{\mathbb{R}^{2d}} K_1 \left(\frac{x-t}{h_1} \right) K_1 \left(\frac{x-s}{h_1} \right) ds \, dt \\
&\le Ch_1^d \int_{\mathbb{R}^{2d}} K_1(u) K_1(v) du \, dv \le Ch_1^d .
\end{aligned}
\tag{A.19}
$$

Hence,

$$
|\mathrm{Cov}(L_{\mathbf{i}}, L_{\mathbf{j}})| \le Ch_1^d, \qquad \sum_{S_1} |\mathrm{Cov}(L_{\mathbf{i}}, L_{\mathbf{j}})| \le C\widehat{\mathbf{n}} c_{\mathbf{n}}^N h_1^d .
\tag{A.20}
$$

Lemma 4.1(ii) of Carbon et al. [19] and $|L_\mathbf{i}| \leq Ch_1^{-d/2}$ permit to write

$$\left|\mathrm{Cov}(L_\mathbf{i}, L_\mathbf{j})\right| \leq Ch_1^{-d}\varphi(\|\mathbf{i}-\mathbf{j}\|),$$

$$\sum_{S_2}\left|\mathrm{Cov}(L_\mathbf{i}, L_\mathbf{j})\right| \leq Ch_1^{-d}\sum_{(\mathbf{i},\mathbf{j})\in S_2}\varphi(\|\mathbf{i}-\mathbf{j}\|) \leq C\widehat{\mathbf{n}}h_1^{-d}\sum_{\mathbf{i}:\,\|\mathbf{i}\|>c_\mathbf{n}}\varphi(\|\mathbf{i}\|) \tag{A.21}$$

$$\leq C\widehat{\mathbf{n}}h_1^{-d}c_\mathbf{n}^{-(N+\eta)}\sum_{\mathbf{i}:\,\|\mathbf{i}\|>c_\mathbf{n}}\|\mathbf{i}\|^{N+\eta}\varphi(\|\mathbf{i}\|),$$

where $\eta > 0$ is a given sufficiently small real number. Using (A.20) and (A.21), we have

$$\sum_{\substack{\mathbf{i},\mathbf{j}\in\mathcal{O}_\mathbf{n}\\\mathbf{i}\neq\mathbf{j}}}\left|\mathrm{Cov}(L_\mathbf{i}, L_\mathbf{j})\right| \leq \left(C\widehat{\mathbf{n}}c_\mathbf{n}^N h_1^d + C\widehat{\mathbf{n}}h_1^{-d}c_\mathbf{n}^{-(N+\eta)}\sum_{\mathbf{i}:\,\|\mathbf{i}\|>c_\mathbf{n}}\|\mathbf{i}\|^{N+\eta}\varphi(\|\mathbf{i}\|)\right). \tag{A.22}$$

Let $c_\mathbf{n} = h_1^{-d/(N+\eta)}$; then we get

$$\sum_{\substack{\mathbf{i},\mathbf{j}\in\mathcal{O}_\mathbf{n}\\\mathbf{i}\neq\mathbf{j}}}\left|\mathrm{Cov}(L_\mathbf{i}, L_\mathbf{j})\right| \leq \left(C\widehat{\mathbf{n}}h_1^{d(1-N/(N+\eta))} + C\widehat{\mathbf{n}}\sum_{\mathbf{i}:\,\|\mathbf{i}\|>c_\mathbf{n}}\|\mathbf{i}\|^{N+\eta}\varphi(\|\mathbf{i}\|)\right)$$

$$\leq \left(C\widehat{\mathbf{n}}h_1^{d(1-N/(N+\eta))} + C\widehat{\mathbf{n}}\sum_{\mathbf{i}:\,\|\mathbf{i}\|>c_\mathbf{n}}\|\mathbf{i}\|^{N+\eta-\theta}\right). \tag{A.23}$$

Since $\theta > N + 1$ and $c_\mathbf{n} \to \infty$, if $\eta < \theta - (N+1)$, we obtain $\sum_{\mathbf{i},\mathbf{j}\in\mathcal{O}_\mathbf{n},\mathbf{i}\neq\mathbf{j}}\left|\mathrm{Cov}(L_\mathbf{i}, L_\mathbf{j})\right| = o(\widehat{\mathbf{n}})$. This implies $\sum_{\mathbf{i},\mathbf{j}\in\mathcal{O}_\mathbf{n},\mathbf{i}\neq\mathbf{j}}\mathrm{Cov}(L_\mathbf{i}, L_\mathbf{j}) = o(\widehat{\mathbf{n}})$ and then gives the part (ii). For (iii), we have $(1/\widehat{\mathbf{n}})\mathrm{var}(\sum_{\mathbf{i}\in\mathcal{O}_\mathbf{n}}L_\mathbf{i}) = (\mathrm{var}(L_\mathbf{i}) + (1/\widehat{\mathbf{n}})\sum_{\mathbf{i},\mathbf{j}\in\mathcal{O}_\mathbf{n},\mathbf{i}\neq\mathbf{j}}\mathrm{Cov}(L_\mathbf{i}, L_\mathbf{j})) \to V(x)$. This completes the proof. $\qquad\square$

Proof of Lemma 4.2. The proof is similar to that of Lemma 3.2 of Tran [30]. For the sake of brevity, we detail the main parts. We have

$$\alpha\widehat{f}(x) - g_\mathbf{n}(x, q_\alpha(x)) = \frac{1}{\widehat{\mathbf{n}}h_1^d}\sum_{\mathbf{i}\in\mathcal{O}_\mathbf{n}}(\alpha K_\mathbf{i} - K_\mathbf{i}H_\mathbf{i}(q_\alpha(x))). \tag{A.24}$$

Let $S_\mathbf{n} = \sum_{i_k=1,\ k=1,\dots,N}^{n_k}L_\mathbf{i}$, with $L_\mathbf{i} = (1/h_1^{d/2})[\alpha K_\mathbf{i} - K_\mathbf{i}H_\mathbf{i}(q_\alpha(x)) - E(\alpha K_\mathbf{i} - K_\mathbf{i}H_\mathbf{i}(q_\alpha(x)))]$. A simple computation gives

$$\left[\frac{\widehat{\mathbf{n}}h_1^d}{V(x)}\right]^{1/2}\left(\left[\alpha\widehat{f}(x) - g_\mathbf{n}(x, q_\alpha(x))\right] - E\left[\alpha\widehat{f}(x) - g_\mathbf{n}(x, q_\alpha(x))\right]\right) = (\widehat{\mathbf{n}}V(x))^{-1/2}S_\mathbf{n}. \tag{A.25}$$

By assumptions H_9(i)-(ii), there exists a sequence $s_\mathbf{n} \to \infty$ of positive integers such that

$$s_\mathbf{n} q_\mathbf{n} = o\left(\left[\hat{\mathbf{n}} \left(h_1^d \right)^{1+(1-2/r)2N} \right]^{1/(2N)} \right). \tag{A.26}$$

It suffices, for example, to take $s_\mathbf{n} = ([\hat{\mathbf{n}}(h_1^d)^{1+(1-2/r)2N}]^{1/(2N)}/q_\mathbf{n})^{1/2}$.

Let $p = p_\mathbf{n} = [(\hat{\mathbf{n}} h_1^d)^{1/(2N)}/s_\mathbf{n}]$; then we have

$$
\begin{aligned}
qp^{-1} &= qs_\mathbf{n} \left[\left(\hat{\mathbf{n}} h_1^d \right)^{-1/(2N)} \right] \\
&\leq C\left(\left[\hat{\mathbf{n}} h_1^{d(1+(1-2/r)2N)} \right]^{1/(2N)} \left[(\hat{\mathbf{n}} h_1^d) \right]^{-1/(2N)} \right) \leq Ch_1^{d(1-2/r)}.
\end{aligned} \tag{A.27}
$$

This last term tends to 0 when $\mathbf{n} \to +\infty$. Thus, it can be assumed without loss of generality that $q < p$. Suppose that there exist integers $r_1, \ldots, r_N$ such that $n_1 = r_1(p + q), \ldots, n_N = r_N(p + q)$.

As in Tran [30], the random variables $L_\mathbf{i}$ are set into large blocks and small blocks. Let

$$
\begin{aligned}
U(1, \mathbf{n}, x, \mathbf{j}) &= \sum_{\substack{i_k = j_k(p+q)+1 \\ k=1,\ldots,N}}^{J_k(p+q)+p} L_\mathbf{i}, \\[2ex]
U(2, \mathbf{n}, x, \mathbf{j}) &= \sum_{\substack{i_k = j_k(p+q)+1 \\ k=1,\ldots,N-1}}^{j_k(p+q)+p} \sum_{i_N = j_N(p+q)+p+1}^{(j_N+1)(p+q)} L_\mathbf{i}, \\[2ex]
U(3, \mathbf{n}, x, \mathbf{j}) &= \sum_{\substack{i_k = j_k(p+q)+1 \\ k=1,\ldots,N-2}}^{j_k(p+q)+p} \sum_{i_{N-1} = j_{N-1}(p+q)+p+1}^{(j_{N-1}+1)(p+q)} \sum_{i_N = j_N(p+q)+1}^{j_N(p+q)+p} L_\mathbf{i}, \\[2ex]
U(4, \mathbf{n}, x, \mathbf{j}) &= \sum_{\substack{i_k = j_k(p+q)+1 \\ k=1,\ldots,N-2}}^{j_k(p+q)+p} \sum_{i_{N-1} = j_{N-1}(p+q)+p+1}^{(j_{N-1}+1)(p+q)} \sum_{i_N = j_N(p+q)+p+1}^{(j_N+1)(p+q)} L_\mathbf{i},
\end{aligned} \tag{A.28}
$$

and so on. Finally

$$
\begin{aligned}
U\left(2^{N-1}, \mathbf{n}, x, \mathbf{j} \right) &= \sum_{\substack{i_k = j_k(p+q)+p+1 \\ k=1,\ldots,N-1}}^{(j_k+1)(p+q)} \sum_{i_N = j_N(p+q)+1}^{j_N(p+q)+p} L_\mathbf{i}, \\[2ex]
U\left(2^N, \mathbf{n}, x, \mathbf{j} \right) &= \sum_{\substack{i_k = j_k(p+q)+p+1 \\ k=1,\ldots,N}}^{(J_k+1)(p+q)} L_\mathbf{i}.
\end{aligned} \tag{A.29}
$$

Setting $\mathcal{Q} = \{0, \ldots, r_1 - 1\} \times \cdots \times \{0, \ldots, r_N - 1\}$, we define for each integer $i = 1, \ldots, 2^N$,

$$T(\mathbf{n}, x, i) = \sum_{\mathbf{j} \in \mathcal{Q}} U(i, \mathbf{n}, x, \mathbf{j}). \tag{A.30}$$

Then, we have $S_{\mathbf{n}} = \sum_{i=1}^{2^N} T(\mathbf{n}, x, i)$. Note that $T(\mathbf{n}, x, 1)$ is the sum of variables $L_{\mathbf{i}}$ in big blocks. The variables $T(\mathbf{n}, x, i)$, $2 \leq i \leq 2^N$ are variables in small blocks.

As raised by Biau and Cadre [20], if one does not have the equalities $n_i = r_i(p + q)$, the term say $T(\mathbf{n}, x, 2^N + 1)$ (which contains the $L_{\mathbf{i}}$'s at the ends not included in the blocks above) can be added. This will not change the proof much.

Thus, to prove the lemma it suffices to prove that

$$\frac{S_{\mathbf{n}}}{(\hat{\mathbf{n}}V(x))^{1/2}} = \frac{T(\mathbf{n}, x, 1)}{(\hat{\mathbf{n}}V(x))^{1/2}} + \frac{\sum_{i=2}^{2^N} T(\mathbf{n}, x, i)}{(\hat{\mathbf{n}}V(x))^{1/2}} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1). \tag{A.31}$$

The general approach is to show that as $\mathbf{n} \to \infty$,

$$
\begin{aligned}
Q_1 &\equiv \left| E \exp \left[iuT(\mathbf{n}, x, 1) \right] - \prod_{\substack{j_k = 0 \\ k = 1, \ldots, N}}^{r_k - 1} E \exp \left[iuU(1, \mathbf{n}, x, \mathbf{j}) \right] \right| \longrightarrow 0, \\
Q_2 &\equiv \hat{\mathbf{n}}^{-1} E \left(\sum_{i=2}^{2^N} T(\mathbf{n}, x, i) \right)^2 \longrightarrow 0, \\
Q_3 &\equiv \hat{\mathbf{n}}^{-1} \sum_{\mathbf{j} \in \mathcal{Q}} E[U(1, \mathbf{n}, x, \mathbf{j})]^2 \longrightarrow V(x), \\
Q_4 &\equiv \hat{\mathbf{n}}^{-1} \sum_{\mathbf{j} \in \mathcal{Q}} E \left[(U(1, \mathbf{n}, x, \mathbf{j}))^2 \mathbf{1}_{\{|U(1, \mathbf{n}, x, \mathbf{j})| > \varepsilon(V(x)\hat{\mathbf{n}})^{1/2}\}} \right] \longrightarrow 0
\end{aligned}
\tag{A.32}
$$

for all $\varepsilon > 0$.

This can be easily done as in Tran [30].   □

The following two lemmas will be used to prove Lemma 4.4.

Lemma A.3. *Under conditions* H_1, H_3, *and* H_4, (3.4), (3.7), (3.8), (3.9), *and*

$$
\begin{aligned}
&(i) \text{ assumption } (3.5) \text{ and } H_6 \\
&\qquad\qquad\qquad or \\
&(ii) \text{ assumption } (3.6) \text{ and } H_7,
\end{aligned}
\tag{A.33}
$$

one has

$$\left| \hat{q}_\alpha(x) - q_\alpha(x) \right| \xrightarrow{P} 0. \tag{A.34}$$

Lemma A.4. *Under conditions of Lemma A.3, one has*

$$\left| \widehat{f}^x(q_\alpha(x)) - f^x(q_\alpha(x)) \right| \xrightarrow{P} 0. \tag{A.35}$$

Proof of Lemma 4.4. We have

$$\left| g_n^{(1)}(x, q_\alpha^*(x)) - f_{X,Y}(x, q_\alpha(x)) \right| \leq \widehat{f}(x) \left| \widehat{f}^x(q_\alpha^*(x)) - f^x(q_\alpha^*(x)) \right|$$
$$+ f^x(q_\alpha^*(x)) \left| \widehat{f}(x) - f_X(x) \right| \tag{A.36}$$
$$+ f_X(x) \left| f^x(q_\alpha^*(x)) - f^x(q_\alpha(x)) \right|.$$

Lemma A.3 implies the convergence in probability of $q_\alpha^*(x)$ to $q_\alpha(x)$. Then, by H_1, the third term of the right-hand side of the previous inequality tends to zero in probability. Lemma A.7 with Lemma A.3 gives the convergence in probability to zero of the second term. To deal with the first term, we can notice that

$$\left| \widehat{f}^x(q_\alpha^*(x)) - f^x(q_\alpha^*(x)) \right| \leq \left| \widehat{f}^x(q_\alpha^*(x)) - \widehat{f}^x(q_\alpha(x)) \right|$$
$$+ \left| \widehat{f}^x(q_\alpha(x)) - f^x(q_\alpha(x)) \right| + \left| f^x(q_\alpha(x)) - f^x(q_\alpha^*(x)) \right|. \tag{A.37}$$

The third term of the right-hand side (of the previous inequality) is the same as the third term of the previous inequality. The second one is treated by Lemma A.4. For the first term, we have

$$\left| \widehat{f}^x(q_\alpha^*(x)) - \widehat{f}^x(q_\alpha(x)) \right| \leq \frac{1}{h_2} \sum_{i \in \mathcal{I}_n} W_i \left| K\left(\frac{q_\alpha^*(x) - Y_i}{h_2} \right) - K\left(\frac{q_\alpha(x) - Y_i}{h_2} \right) \right|, \tag{A.38}$$

where $W_{ni} = K_i / \sum_{i \in \mathcal{I}_n} K_i$. Therefore, using the fact that K is Lipschitz (assumption H_5), we get

$$\left| \widehat{f}^x(q_\alpha^*(x)) - \widehat{f}^x(q_\alpha(x)) \right| \leq \frac{1}{h_2^2} \left| q_\alpha^*(x) - q_\alpha(x) \right|. \tag{A.39}$$

Then, for any $\varepsilon > 0$, we have

$$P\left(\left| \widehat{f}^x(q_\alpha^*(x)) - \widehat{f}^x(q_\alpha(x)) \right| > \varepsilon \right) \leq P\left(\left| q_\alpha^*(x) - q_\alpha(x) \right| > \varepsilon h_2^2 \right). \tag{A.40}$$

This permits to conclude that the weak convergence of $q_\alpha^*(x)$ to $q_\alpha(x)$ (Lemma A.3) implies that of $\widehat{f}^x(q_\alpha^*(x))$ to $\widehat{f}^x(q_\alpha(x))$. This yields the proof. $\qquad\square$

Proof of Lemma A.3. By hypothesis H_5, the function $\widehat{F}^x(\cdot)$ is continuous and strictly increasing. Thus, the inverse function $\widehat{F}^x(\cdot)^{(-1)}$ is continuous and strictly increasing. In particular, the

continuity of $\widehat{F}^x(\cdot)^{(-1)}$ at $\widehat{F}^x(q_\alpha(x))$ gives (see also the proof of Theorem 3.1 of Ferraty et al. [28], page 383)

$$\forall \varepsilon > 0, \ \exists \eta = \eta(\varepsilon) > 0, \ \forall y, \quad \left| \widehat{F}^x(y) - \widehat{F}^x(q_\alpha(x)) \right| \leq \eta \Longrightarrow |y - q_\alpha(x)| \leq \varepsilon, \qquad (A.41)$$

and thus

$$\begin{aligned}
\forall \varepsilon > 0, \ \exists \eta > 0, \quad P\big(|\widehat{q}_\alpha(x) - q_\alpha(x)| > \varepsilon\big) &\leq P\Big(\big|\widehat{F}^x(\widehat{q}_\alpha(x)) - \widehat{F}^x(q_\alpha(x))\big| > \eta\Big) \\
&= P\Big(\big|F^x(q_\alpha(x)) - \widehat{F}^x(q_\alpha(x))\big| > \eta\Big).
\end{aligned} \qquad (A.42)$$

So, to prove the weak convergence of the quantile estimate $\widehat{q}_\alpha(x)$ to $q_\alpha(x)$ it suffices to prove that the conditional distribution estimate $\widehat{F}^x(q_\alpha(x))$ converges in probability to $F^x(q_\alpha(x))$. That is done in the following lemma. $\qquad \square$

Lemma A.5. *Under conditions of Lemma A.3, one has*

$$\widehat{F}^x(q_\alpha(x)) - F^x(q_\alpha(x)) \xrightarrow{P} 0. \qquad (A.43)$$

In order to establish Lemma A.4, we introduce the following notations and state the following three technical lemmas. For $y \in \mathbb{R}$, let

$$\begin{aligned}
\Gamma_{\mathbf{i}}(x, y) &= \frac{1}{\widehat{\mathbf{n}} h_1^d h_2} K_1\left(\frac{x - X_{\mathbf{i}}}{h_1}\right) K\left(\frac{y - Y_{\mathbf{i}}}{h_2}\right), \qquad \Delta_{\mathbf{i}}(x, y) = \Gamma_{\mathbf{i}}(x, y) - E\Gamma_{\mathbf{i}}(x, y), \\
f_{\mathbf{n}}(x, y) &= \frac{1}{\widehat{\mathbf{n}} h_1^d h_2} \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} K_1\left(\frac{x - X_{\mathbf{i}}}{h_1}\right) K\left(\frac{y - Y_{\mathbf{i}}}{h_2}\right),
\end{aligned} \qquad (A.44)$$

and $S_{\mathbf{n}}(x, y) = \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} \Delta_{\mathbf{i}}(x, y) = f_{\mathbf{n}}(x, y) - E f_{\mathbf{n}}(x, y)$,

$$I_{\mathbf{n}}(x, y) = \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} E(\Delta_{\mathbf{i}}(x, y))^2, \qquad R_{\mathbf{n}}(x, y) = \sum_{\substack{\mathbf{i} \neq \mathbf{j} \\ \mathbf{i}, \mathbf{j} \in \mathcal{I}_{\mathbf{n}}}} |E\Delta_{\mathbf{i}}(x, y) \Delta_{\mathbf{j}}(x, y)|. \qquad (A.45)$$

Lemma A.6. *Under the conditions of Lemma A.3, one has for $y \in \mathbb{R}$*

$$|f_{\mathbf{n}}(x, y) - f_{X,Y}(x, y)| \xrightarrow{P} 0. \qquad (A.46)$$

Lemma A.7. *If the conditions of Lemma A.3 are satisfied, then $\widehat{f}(x) \xrightarrow{P} f_X(x)$.*

Lemma A.8. *Under the conditions of Lemma A.3, one has for $y \in \mathbb{R}$*

$$I_{\mathbf{n}}(x, y) + R_{\mathbf{n}}(x, y) = O\left(\frac{1}{\widehat{\mathbf{n}} h_K^d h_H}\right). \qquad (A.47)$$

Proof of Lemma A.4. Let $y = q_\alpha(x)$. We have

$$\left| \widehat{f^x}(y) - f^x(y) \right| \leq \frac{1}{\widehat{f}(x)} \left| f_{\mathbf{n}}(x, y) - f_{X,Y}(x, y) \right| + \frac{1}{\widehat{f}(x)} f^x(y) \left| \widehat{f}(x) - f_X(x) \right|. \tag{A.48}$$

Lemmas A.7 and A.6 give, respectively, the weak convergence of $\widehat{f}(x)$ to $f_X(x)$ and $f_{\mathbf{n}}(x, y)$ to $f_{X,Y}(x, y)$. This finishes the proof since, by H_1, $f_X(x) > 0$ and $f^x(y)$ is bounded. □

In order to prove Lemma A.5, we introduce the following notations and state the following three lemmas. For $y \in \mathbb{R}$, let

$$\widetilde{\Gamma}_{\mathbf{i}}(x, y) = \frac{1}{\widehat{\mathbf{n}} h_1^d} K_1 \left(\frac{x - X_{\mathbf{i}}}{h_1} \right) K_2 \left(\frac{y - Y_{\mathbf{i}}}{h_2} \right), \qquad \widetilde{\Delta}_{\mathbf{i}}(x, y) = \widetilde{\Gamma}_{\mathbf{i}}(x, y) - E\widetilde{\Gamma}_{\mathbf{i}}(x, y),$$

$$\widetilde{I}_{\mathbf{n}}(x, y) = \sum_{\mathbf{i} \in \mathcal{O}_{\mathbf{n}}} E\left(\widetilde{\Delta}_{\mathbf{i}}(x, y) \right)^2, \qquad \widetilde{R}_{\mathbf{n}}(x, y) = \sum_{\substack{\mathbf{i} \neq \mathbf{j} \\ \mathbf{i}, \mathbf{j} \in \mathcal{O}_{\mathbf{n}}}} \left| E\widetilde{\Delta}_{\mathbf{i}}(x, y) \widetilde{\Delta}_{\mathbf{j}}(x, y) \right|, \tag{A.49}$$

$$g_{\mathbf{n}}(x, y) = \sum_{\mathbf{i} \in \mathcal{O}_{\mathbf{n}}} \widetilde{\Gamma}_{\mathbf{i}}(x, y), \qquad g(x, y) = F^x(y) f_X(x).$$

Lemma A.9. *Under conditions of Lemma A.5, one has*

$$g_{\mathbf{n}}(x, q_\alpha(x)) \xrightarrow{P} g(x, q_\alpha(x)). \tag{A.50}$$

Lemma A.10. *Under conditions of Lemma A.5, one has*

$$\widetilde{I}_{\mathbf{n}}(x, q_\alpha(x)) + \widetilde{R}_{\mathbf{n}}(x, q_\alpha(x)) = O\left(\frac{1}{\widehat{\mathbf{n}} h_1^d} \right). \tag{A.51}$$

Proof of Lemma A.5. Let $y = q_\alpha(x)$; it is easy to see that

$$\left| \widehat{F}^x(y) - F^x(y) \right| \leq \frac{1}{\widehat{f}(x)} \left\{ \left| g_{\mathbf{n}}(x, y) - g(x, y) \right| + \left(F^x(y) \left| f_X(x) - \widehat{f}(x) \right| \right) \right\}. \tag{A.52}$$

Lemmas A.7 and A.9 give, respectively, the weak convergence of $\widehat{f}(x)$ to $f_X(x)$ and $g_{\mathbf{n}}(x, y)$ to $g(x, y)$ and yield the proof. □

Proof of Lemma A.6. Remark that

$$\left| f_{\mathbf{n}}(x, y) - f_{X,Y}(x, y) \right| \leq \left| f_{\mathbf{n}}(x, y) - E f_{\mathbf{n}}(x, y) \right| + \left| E f_{\mathbf{n}}(x, y) - f_{X,Y}(x, y) \right|. \tag{A.53}$$

The asymptotic behavior of the bias term is standard, in the sense that it is not affected by the dependence structure of the data. We have

$$
\begin{aligned}
&\left| Ef_{\mathbf{n}}(x,y) - f_{X,Y}(x,y) \right| \\
&= \left| \frac{1}{h_K^d h_H} \int_{\mathbb{R}^{d+1}} K\left(\frac{x-u}{h_K} \right) H\left(\frac{y-v}{h_H} \right) f_{X,Y}(u,v)\, du\, dv - f_{X,Y}(x,y) \right| \\
&\leq \int_{\mathbb{R}^{d+1}} K(t) H(s) \left| f_{X,Y}(x - h_K t, y - h_H s) - f_{X,Y}(x,y) \right| dt\, ds.
\end{aligned}
\tag{A.54}
$$

This last term goes to zero by H_1 and the Lebesgue dominated Theorem.

The proof of the weak convergence of $S_{\mathbf{n}}(x) = f_{\mathbf{n}}(x,y) - Ef_{\mathbf{n}}(x,y)$ to zero is similar to that of Theorem 3.3 of Carbon et al. [19] or Lemma 3.2 of Dabo-Niang and Yao [27]. For the sake of completeness, we present it entirely. Let us now introduce a spatial block decomposition that has been used by Carbon et al. [19].

Without loss of generality, assume that $n_i = 2pq_i$ for $1 \leq i \leq N$. The random variables $\Delta_{\mathbf{i}}(x,y)$ can be grouped into $2^N q_1 \ldots q_N$ cubic blocks $U(1, \mathbf{n}, \mathbf{j})$ of side p (as previously). Then, for each integer $1 \leq i \leq 2^N$, define $T(\mathbf{n}, i) = \sum_{j_k=0,\ k=1,\ldots,N}^{q_k-1} U(i, \mathbf{n}, \mathbf{j})$ and get the following decomposition $S_{\mathbf{n}}(x,y) = \sum_{i=1}^{2^N} T(\mathbf{n}, i)$.

Observe that, for any $\varepsilon > 0$,

$$
P\left(\left| S_{\mathbf{n}}(x,y) \right| > \varepsilon \right) = P\left(\left| \sum_{i=1}^{2^N} T(\mathbf{n}, i) \right| > \varepsilon \right) \leq 2^N P\left(|T(\mathbf{n}, 1)| > \frac{\varepsilon}{2^N} \right).
\tag{A.55}
$$

We enumerate in an arbitrary way the $\hat{q} = q_1 \ldots q_N$ terms $U(1, \mathbf{n}, \mathbf{j})$ of the sum $T(\mathbf{n}, 1)$ that we call $W_1, \ldots, W_{\hat{q}}$. Note that $U(1, \mathbf{n}, \mathbf{j})$ is measurable with respect to the σ-field generated by $V_{\mathbf{i}}(x,y)$, with $\mathbf{i}$ such that $2j_k p + 1 \leq i_k \leq (2j_k + 1)p,\ k = 1, \ldots, N$. These sets of sites are separated by a distance at least p, and since K_1 and K_2 are bounded, then we have for all $i = 1, \ldots, \hat{q}$,

$$
|W_i| \leq C \left(\hat{\mathbf{n}} h_1^d h_2 \right)^{-1} p^N \|K_1\|_\infty \|K_2\|_\infty.
\tag{A.56}
$$

Lemma 4.4 in Carbon et al. [19] ensures that there exist independent random variables $W_1^*, \ldots, W_{\hat{q}}^*$ such that

$$
\sum_{i=1}^{\hat{q}} E \left| W_i - W_i^* \right| \leq C \hat{q} \left(\hat{\mathbf{n}} h_1^d h_2 \right)^{-1} p^N \|K_1\|_\infty \|K_2\|_\infty \psi\left(\hat{\mathbf{n}}, p^N \right) \varphi(p).
\tag{A.57}
$$

Markov's inequality leads to

$$
P\left(\sum_{i=1}^{\hat{q}} \left| W_i - W_i^* \right| > \frac{\varepsilon}{2^{N+1}} \right) \leq C 2^{N+1} \left(\hat{\mathbf{n}} h_1^d h_2 \right)^{-1} p^N \hat{q} \psi\left(\hat{\mathbf{n}}, p^N \right) \varepsilon^{-1} \varphi(p).
\tag{A.58}
$$

By Bernstein's inequality, we have

$$P\left(\left| \sum_{i=1}^{\hat{q}} W_i^* \right| > \frac{\varepsilon}{2^{N+1}} \right) \le 2 \exp\left\{ \frac{-\varepsilon^2/(2^{N+1})^2}{4\sum_{i=1}^{\hat{q}} EW_i^{*2} + 2C\left(\hat{\mathbf{n}} h_1^d h_2\right)^{-1} p^N \varepsilon / 2^{N+1}} \right\}. \tag{A.59}$$

Combining (A.55), (A.58), and (A.59), we get

$$\begin{aligned}
P(|S_{\mathbf{n}}(x,y)| > \varepsilon) &\le 2^N P\left(\sum_{i=1}^{\hat{q}} |W_i - W_i^*| > \frac{\varepsilon}{2^{N+1}} \right) + 2^N P\left(\left| \sum_{i=1}^{\hat{q}} W_i^* \right| > \frac{\varepsilon}{2^{N+1}} \right) \\
&\le 2^{N+1} \exp\left\{ \frac{-\varepsilon^2/(2^{N+1})^2}{4\sum_{i=1}^{\hat{q}} EW_i^{*2} + 2C\left(\hat{\mathbf{n}} h_1^d h_2\right)^{-1} p^N \varepsilon / 2^{N+1}} \right\} \\
&\quad + C 2^{2N+1} \psi\left(\hat{\mathbf{n}}, p^N\right) \left(\hat{\mathbf{n}} h_1^d h_2\right)^{-1} p^N \hat{q} \varepsilon^{-1} \varphi(p).
\end{aligned} \tag{A.60}$$

Let $\lambda > 0$ and set

$$\varepsilon = \varepsilon_{\mathbf{n}} = \left(\frac{\log \hat{\mathbf{n}}}{\hat{\mathbf{n}} h_1^d h_2} \right)^{1/2}, \qquad p = p_{\mathbf{n}} = \left(\frac{\hat{\mathbf{n}} h_1^d h_2}{\log \hat{\mathbf{n}}} \right)^{1/2N}. \tag{A.61}$$

By the fact that W_i^* and W_i have the same distribution, we have

$$\sum_{i=1}^{\hat{q}} EW_i^{*2} = \sum_{i=1}^{\hat{q}} EW_i^2 \le I_{\mathbf{n}}(x, q_\alpha(x)) + R_{\mathbf{n}}(x, q_\alpha(x)). \tag{A.62}$$

Then, by Lemma A.8, we get $\sum_{i=1}^{\hat{q}} EW_i^{*2} = O(1/\hat{\mathbf{n}} h_1^d h_2)$. Thus, for the case (i) of the theorem a simple computation shows that for sufficiently large $\mathbf{n}$,

$$\begin{aligned}
P(|S_{\mathbf{n}}(x,y)| > \lambda \varepsilon_{\mathbf{n}}) &\le 2^{N+1} \exp\left\{ \frac{-\lambda^2 \log \hat{\mathbf{n}}}{2^{2N+4}C + 2^{N+2}C\lambda} \right\} \\
&\quad + C 2^{N+1} p^N h_1^{-d} h_2^{-1} \lambda^{-1} \varepsilon_{\mathbf{n}}^{-1} \varphi(p) \\
&\le C \hat{\mathbf{n}}^{-b} + C 2^{N+1} p^N h_1^{-d} h_2^{-1} \lambda^{-1} \varepsilon_{\mathbf{n}}^{-1} \varphi(p),
\end{aligned} \tag{A.63}$$

where $b > 0$ and depends on λ. For case (ii), we obtain

$$\begin{aligned}
P(|S_{\mathbf{n}}(x,y)| > \lambda \varepsilon_{\mathbf{n}}) &\le 2^{N+1} \exp\left\{ \frac{-\lambda^2 \log \hat{\mathbf{n}}}{2^{2N+4}C + 2^{N+2}C\lambda} \right\} \\
&\quad + C 2^{N+1} \hat{\mathbf{n}}^{\tilde{\beta}} h_1^{-d} h_2^{-1} \lambda^{-1} \varepsilon_{\mathbf{n}}^{-1} \varphi(p) \\
&\le C \hat{\mathbf{n}}^{-b} + C 2^{N+1} \hat{\mathbf{n}}^{\tilde{\beta}} h_1^{-d} h_2^{-1} \lambda^{-1} \varepsilon_{\mathbf{n}}^{-1} \varphi(p).
\end{aligned} \tag{A.64}$$

Then, (A.63) and (A.64) can be condensed in

$$P\left(\left|f_{\mathbf{n}}(x,y) - Ef_{\mathbf{n}}(x,y)\right| > \lambda\varepsilon_{\mathbf{n}}\right) \leq \begin{cases} C\widehat{\mathbf{n}}^{-b} + C\lambda^{-1}h_1^{-d}h_2^{-1}\varepsilon_{\mathbf{n}}^{(\theta-2N)/N} & \text{under (i),} \\ C\widehat{\mathbf{n}}^{-b} + C\lambda^{-1}h_1^{-d}h_2^{-1}\widehat{\mathbf{n}}^{\widetilde{\beta}}\varepsilon_{\mathbf{n}}^{(\theta-N)/N} & \text{under (ii).} \end{cases} \tag{A.65}$$

Then, to prove the convergence in probability of $U_{1\mathbf{n}}(x)$, it suffices to show that for, respectively, (i) and (ii)

$$C\widehat{\mathbf{n}}^{-b} \longrightarrow 0, \qquad \lambda^{-1}h_1^{-d}h_2^{-1}\varepsilon_{\mathbf{n}}^{(\theta-2N)/N} \longrightarrow 0,$$

$$C\widehat{\mathbf{n}}^{-b} \longrightarrow 0, \qquad \lambda^{-1}h_1^{-d}h_2^{-1}\widehat{\mathbf{n}}^{\widetilde{\beta}}\varepsilon_{\mathbf{n}}^{(\theta-N)/N} \longrightarrow 0. \tag{A.66}$$

Clearly, $\widehat{\mathbf{n}}^{-b} \to 0$ if $b > 0$. On the one hand, we have

$$\lambda^{-1}h_1^{-d}h_2^{-1}\varepsilon_{\mathbf{n}}^{(\theta-2N)/N} \leq C\left[\widehat{\mathbf{n}}h_1^{d\theta/(\theta-2N)}h_2^{\theta/(\theta-2N)}(\log\widehat{\mathbf{n}})^{-1}\right]^{(2N-\theta)/2N}. \tag{A.67}$$

This last goes to 0 by H_6.

On the other hand, we have

$$\lambda^{-1}h_1^{-d}h_2^{-1}\widehat{\mathbf{n}}^{\widetilde{\beta}}\varepsilon_{\mathbf{n}}^{(\theta-N)/N}$$

$$\leq C\left[\widehat{\mathbf{n}}h_1^{d(\theta+N)/(\theta-N(1+2\widetilde{\beta}))}h_2^{(\theta+N)/(\theta-N(1+2\widetilde{\beta}))}(\log\widehat{\mathbf{n}})^{(N-\theta)/(\theta-N(1+2\widetilde{\beta}))}\right]^{(N(1+2\widetilde{\beta})-\theta)/2N}, \tag{A.68}$$

which goes to 0 by H_7. This yields the proof. □

Proof of Lemma A.7. By using the same arguments as in **Lemma A.6**, we get

$$\left|E\widehat{f}(x) - f_X(x)\right| = \left|\int_{\mathbb{R}^d} K_1(s)\left[f_X(x-sh) - f_X(x)\right]ds\right|. \tag{A.69}$$

This last term tends to zero by Lebesgue dominated Theorem. Let

$$\Gamma_{\mathbf{i}}(x) = \frac{1}{\widehat{\mathbf{n}}h_1^d}K_1\left(\frac{x-X_{\mathbf{i}}}{h_1}\right), \qquad \Delta_{\mathbf{i}}(x) = \Gamma_{\mathbf{i}}(x) - E\Gamma_{\mathbf{i}}(x). \tag{A.70}$$

Then we have $\widehat{f}(x) - E\widehat{f}(x) = \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\Delta_{\mathbf{i}}(x) = S_{\mathbf{n}}(x)$.

Let $I_{\mathbf{n}}(x) = \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}E(\Delta_{\mathbf{i}}(x))^2$ and $R_{\mathbf{n}}(x) = \sum_{\mathbf{i}\neq\mathbf{j},\ \mathbf{i},\mathbf{j}\in\mathcal{I}_{\mathbf{n}}}|E\Delta_{\mathbf{i}}(x)\Delta_{\mathbf{j}}(x)|$. Lemma 2.2 of Tran [30] gives that

$$\mathrm{var}\left(\widehat{f}(x)\right) \leq I_{\mathbf{n}}(x) + R_{\mathbf{n}}(x) = O\left(\frac{1}{\widehat{\mathbf{n}}h_1^d}\right). \tag{A.71}$$

Consider $\varepsilon_{\mathbf{n}} = (\log \hat{\mathbf{n}} / \hat{\mathbf{n}} h_1^d)^{1/2}$ and $p = (\hat{\mathbf{n}} h_1^d / \log \hat{\mathbf{n}})^{1/2N}$ and use the same arguments as in the proof of Lemma A.6, to get for sufficiently large $\mathbf{n}$

$$
P\left(|S_{\mathbf{n}}(x)| > \lambda \sqrt{\frac{\log \hat{\mathbf{n}}}{\hat{\mathbf{n}} h_1^d}} \right) \leq
\begin{cases}
C\hat{\mathbf{n}}^{-b} + C2^{N+1} p^N h_1^{-d} \varepsilon_{\mathbf{n}}^{-1} \varphi(p) & \text{under (i),} \\
C\hat{\mathbf{n}}^{-b} + C2^{N+1} \hat{\mathbf{n}}^{\tilde{\beta}} h_1^{-d} \varepsilon_{\mathbf{n}}^{-1} \varphi(p) & \text{under (ii),}
\end{cases}
\tag{A.72}
$$

with $b > 0$. It suffices to show that for the case (i) (resp., (ii)) $p^N h_1^{-d} \varepsilon_{\mathbf{n}}^{-1} \varphi(p) \to 0$ (resp., $n^{\tilde{\beta}} h_1^{-d} \varepsilon_{\mathbf{n}}^{-1} \varphi(p) \to 0$). A simple computation shows for, respectively, (i) and (ii)

$$
p^N h_1^{-d} \varepsilon_{\mathbf{n}}^{-1} \varphi(p) \leq C\left[\hat{\mathbf{n}} h_1^{d\theta/(\theta-2N)} \log \hat{\mathbf{n}}^{-1} \right]^{(2N-\theta)/2N},
$$
$$
\hat{\mathbf{n}}^{\tilde{\beta}} h_1^{-d} \varepsilon_{\mathbf{n}}^{-1} \varphi(p) \leq C\left[\hat{\mathbf{n}} h_1^{d(N+\theta)/(\theta-N(1+2\tilde{\beta}))} (\log \hat{\mathbf{n}})^{(N-\theta)/(\theta-N(1+2\tilde{\beta}))} \right]^{(N(1+2\tilde{\beta})-\theta)/2N},
\tag{A.73}
$$

and these last go to 0 by, respectively, H_6 and H_7, which yields the proof. $\qquad\square$

Proof of Lemma A.8. We have $\hat{\mathbf{n}} h_1^d h_2 I_{\mathbf{n}}(x, y) = \hat{\mathbf{n}} h_1^d h_2 \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} (E\Gamma_{\mathbf{i}}^2(x, y) - E^2\Gamma_{\mathbf{i}}(x, y))$. First, remark that

$$
\hat{\mathbf{n}} h_1^d h_2 \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} E\Gamma_{\mathbf{i}}^2(x, y) = h_1^{-d} h_2^{-1} \int_{\mathbb{R}^{d+1}} K_1^2\left(\frac{x-z}{h_1}\right) K^2\left(\frac{y-v}{h_2}\right) f_{X,Y}(z, v) \, dz \, dv
$$
$$
= \int_{\mathbb{R}^{d+1}} K_1^2(z) K^2(v) f_{X,Y}(x - h_1 z, y - h_2 v) \, dz \, dv.
\tag{A.74}
$$

By assumption H_1 and Lebesgue dominated Theorem, this last integral converges to

$$
f_{X,Y}(x, y) \int_{\mathbb{R}^{d+1}} K_1^2(z) K^2(v) \, dz \, dv.
\tag{A.75}
$$

Next, notice that

$$
\hat{\mathbf{n}} h_1^d h_2 \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} E^2\Gamma_{\mathbf{i}}(x, y) = h_1^{-d} h_2^{-1} \left(\int_{\mathbb{R}^{d+1}} K_1\left(\frac{x-z}{h_1}\right) K\left(\frac{y-v}{h_2}\right) f_{X,Y}(z, v) \, dz \, dv \right)^2.
\tag{A.76}
$$

By an usual change of variables, we obtain

$$
\hat{\mathbf{n}} h_1^d h_2 \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} E^2\Gamma_{\mathbf{i}}(x, y) = h_1^d h_2 \left(\int_{\mathbb{R}^{d+1}} K_1(z) K(v) f_{X,Y}(x - h_1 z, y - h_2 v) \, dz \, dv \right)^2.
\tag{A.77}
$$

This last term tends to 0 by H_1 and Lebesgue dominated Theorem.

Let us now prove that for $\mathbf{n}$ being large enough, there exists C such that $\hat{\mathbf{n}} h_1^d h_2 R_{\mathbf{n}}(x, y) < C$. Let $S = \{\mathbf{i}, \mathbf{j}, \text{dist}(\mathbf{i}, \mathbf{j}) \le s_{\mathbf{n}}\}$, where $s_{\mathbf{n}}$ is a real sequence that converges to infinity and will be specified later. Then $R_{\mathbf{n}}(x, y) = R_{\mathbf{n}}^1(x, y) + R_{\mathbf{n}}^2(x, y)$, with

$$R_{\mathbf{n}}^1(x, y) = \sum_{\mathbf{i}, \mathbf{j} \in S} \left| E \Delta_{\mathbf{i}}(x, y) \Delta_{\mathbf{j}}(x, y) \right|, \qquad R_{\mathbf{n}}^2(x, y) = \sum_{\mathbf{i}, \mathbf{j} \in S^c} \left| E \Delta_{\mathbf{i}}(x, y) \Delta_{\mathbf{j}}(x, y) \right|, \qquad \text{(A.78)}$$

where S^c stands for the complement of S. Now, by change of variables, H_3, and Lebesgue dominated Theorem, we get

$$\begin{aligned} E \left[\left| K\left(\frac{y - Y_{\mathbf{i}}}{h_2} \right) \right| \left| K\left(\frac{y - Y_{\mathbf{j}}}{h_2} \right) \right| \mid (X_{\mathbf{i}}, X_{\mathbf{j}}) \right] \\ = h_H^2 \int_{\mathbb{R}^2} K(t) K(s) f^{(X_{\mathbf{i}}, X_{\mathbf{j}})}(y - h_2 t, y - h_2 s) dt\, ds = O\left(h_2^2 \right). \end{aligned} \qquad \text{(A.79)}$$

Similarly, we have

$$E \left[\left| K\left(\frac{y - Y_{\mathbf{i}}}{h_2} \right) \right| \mid X_{\mathbf{i}} \right] = h_2 \int_{\mathbb{R}} K(t) f^{X_{\mathbf{i}}}(y - h_2 t) dt = O(h_2). \qquad \text{(A.80)}$$

In addition, by (3.8), we get $E K_{\mathbf{i}} K_{\mathbf{j}} = O(h_1^{2d})$ and $E K_{\mathbf{i}} = O(h_1^d)$. Let us consider $R_{\mathbf{n}}^1(x, y)$. We have

$$\begin{aligned} \left| E \Delta_{\mathbf{i}}(x, y) \Delta_{\mathbf{j}}(x, y) \right| &= \left| E \Gamma_{\mathbf{i}}(x, y) \Gamma_{\mathbf{j}}(x, y) - E \Gamma_{\mathbf{i}}(x, y) E \Gamma_{\mathbf{j}}(x, y) \right| \\ &\le E \left[E \left| \Gamma_{\mathbf{i}}(x, y) \Gamma_{\mathbf{j}}(x, y) \right| \mid (X_{\mathbf{i}}, X_{\mathbf{j}}) \right] + \left(E \left[E \left| \Gamma_{\mathbf{i}}(x, y) \right| \mid X_{\mathbf{i}} \right] \right)^2 \\ &\le \hat{\mathbf{n}}^{-2} h_1^{-2d} h_2^{-2} E K_{\mathbf{i}} K_{\mathbf{j}} E \left[\left| K\left(\frac{y - Y_{\mathbf{i}}}{h_2} \right) \right| \left| K\left(\frac{y - Y_{\mathbf{j}}}{h_2} \right) \right| \mid (X_{\mathbf{i}}, X_{\mathbf{j}}) \right] \\ &\quad + \hat{\mathbf{n}}^{-2} h_1^{-2d} h_2^{-2} \left(E K_{\mathbf{i}} E \left[\left| K\left(\frac{y - Y_{\mathbf{i}}}{h_2} \right) \right| \mid X_{\mathbf{i}} \right] \right)^2 \le C \hat{\mathbf{n}}^{-2}. \end{aligned} \qquad \text{(A.81)}$$

Then $\hat{\mathbf{n}} h_1^d h_2 R_{\mathbf{n}}^1(x, y) \le \hat{\mathbf{n}}^{-1} h_1^d h_2 \sum_{\mathbf{i}, \mathbf{j} \in S} 1 \le C h_1^d h_2 s_{\mathbf{n}}^N$. Let us now compute $R_{\mathbf{n}}^2(x, y)$. Since K_1 and K are bounded, by applying Lemma 4.1(ii) of Carbon et al. [19] we get

$$\left| E \Delta_{\mathbf{i}}(x, y) \Delta_{\mathbf{j}}(x, y) \right| \le C \hat{\mathbf{n}}^{-2} h_1^{-2d} h_2^{-2} \psi(1, 1) \varphi(\|\mathbf{i} - \mathbf{j}\|). \qquad \text{(A.82)}$$

Then, we obtain that

$$\begin{aligned} \hat{\mathbf{n}} h_1^d h_2 R_{\mathbf{n}}^2(x, y) &\le C \hat{\mathbf{n}}^{-1} h_1^{-d} h_2^{-1} \sum_{\mathbf{i}, \mathbf{j} \in S^c} \psi(1, 1) \varphi(\|\mathbf{i} - \mathbf{j}\|) \\ &\le C h_1^{-d} h_2^{-1} s_{\mathbf{n}}^{-N} \sum_{\|\mathbf{i}\| > s_{\mathbf{n}}} \|\mathbf{i}\|^N \varphi(\|\mathbf{i}\|) \\ &\le C h_1^{-d} h_2^{-1} s_{\mathbf{n}}^{-N} \sum_{\|\mathbf{i}\| > s_{\mathbf{n}}} \|\mathbf{i}\|^{N - \theta}. \end{aligned} \qquad \text{(A.83)}$$

As $\theta > N + 1$, the choice $s_{\mathbf{n}} = \left(h_1^d h_2 \right)^{-1/N}$ gives the desired result and yields the proof. $\qquad \square$

Proof of Lemma A.9. We have

$$
\begin{aligned}
&\left| E g_{\mathbf{n}}(x,y) - g(x,y) \right| \\
&= \left| \frac{1}{\hat{\mathbf{n}} h_1^d} \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} \int_{\mathbb{R}^{d+1}} K_1\left(\frac{x-z}{h_1}\right) K_2\left(\frac{y-v}{h_2}\right) f_{X,Y}(z,v) dz\, dv - \int_{-\infty}^{y} f_{X,Y}(x,u) du \right| \\
&= \left| \int_{-\infty}^{y} \int_{\mathbb{R}^{d+1}} K_1(s) K(t) \left[f_{X,Y}(x - sh_1, u - th_2) - f_{X,Y}(x,u) \right] du\, ds\, dt \right|,
\end{aligned}
\tag{A.84}
$$

and this last term tends to zero by H_1 and Lebesque dominated Theorem. We have

$$
g_{\mathbf{n}}(x,y) - E g_{\mathbf{n}}(x,y) = \sum_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}} \widetilde{\Delta}_{\mathbf{i}}(x,y) = \widetilde{S}_{\mathbf{n}}(x,y).
\tag{A.85}
$$

The sequel of the proof uses the same decomposition and similar lines as in the proof of Lemma A.6. Taking $\varepsilon_{\mathbf{n}} = \lambda (\log \hat{\mathbf{n}} / \hat{\mathbf{n}} h_1^d)^{1/2}$, $p = (\hat{\mathbf{n}} h_1^d / \log \hat{\mathbf{n}})^{1/2N}$ and making use of Lemma A.10, we obtain, for $\lambda > 0$, the existence of $b > 0$ such that, for $\mathbf{n}$ being large enough,

$$
P\left(\left| g_{\mathbf{n}}(x,y) - E g_{\mathbf{n}}(x,y) \right| > \varepsilon_{\mathbf{n}} \right) \leq
\begin{cases}
C\hat{\mathbf{n}}^{-b} + C2^{N+1} p^N h_1^{-d} \varepsilon_{\mathbf{n}}^{-1} \varphi(p) & \text{under (i)}, \\
C\hat{\mathbf{n}}^{-b} + C2^{N+1} \hat{\mathbf{n}}^{\widetilde{\beta}} h_1^{-d} \varepsilon_{\mathbf{n}}^{-1} \varphi(p) & \text{under (ii)}.
\end{cases}
\tag{A.86}
$$

To prove the convergence of the lemma, it suffices to show, respectively, for (i) and (ii) that

$$
\begin{aligned}
p^N h_1^{-d} \varepsilon_{\mathbf{n}}^{-1} \varphi(p) &\longrightarrow 0, \\
\hat{n}^{\widetilde{\beta}} h_1^{-d} \varepsilon_{\mathbf{n}}^{-1} \varphi(p) &\longrightarrow 0.
\end{aligned}
\tag{A.87}
$$

A simple computation gives

$$
p^N h_1^{-d} \varepsilon_{\mathbf{n}}^{-1} \varphi(p) \leq C \left[\hat{\mathbf{n}} h_1^{d\theta/(\theta-2N)} \log \hat{\mathbf{n}}^{-1} \right]^{(2N-\theta)/2N},
\tag{A.88}
$$

and this last goes to 0 by Assumption H_6. Analogously, we have

$$
h_1^{-d} \varepsilon_{\mathbf{n}}^{-1} \varphi(p) \leq C \left[\hat{\mathbf{n}} h_1^{d(N+\theta)/(\theta - N(1+2\widetilde{\beta}))} (\log \hat{\mathbf{n}})^{(N-\theta)/(\theta - N(1+2\widetilde{\beta}))} \right]^{(N(1+2\widetilde{\beta})-\theta)/2N},
\tag{A.89}
$$

which goes to 0 by Assumption H_7. This yields the proof. $\square$

Proof of Lemma A.10. Our conditions and the proof of Lemma 2.2 in Tran [30] yield easily the result. More precisely, let $y = q_\alpha(x)$ and write

$$\hat{\mathbf{n}}h_1^d \tilde{I}_{\mathbf{n}}(x,y) = \hat{\mathbf{n}}h_1^d \sum_{\mathbf{i} \in \mathbf{i}} \left(E\left(\left(\tilde{\Gamma}_{\mathbf{i}}(x,y) \right)^2 \right) - E^2\left(\tilde{\Gamma}_{\mathbf{i}}(x,y) \right) \right),$$

$$\hat{\mathbf{n}}h_1^d \sum_{\mathbf{i} \in \mathbf{i}} E\left(\left(\tilde{\Gamma}_{\mathbf{i}}(x,y) \right)^2 \right) = h_1^{-d} \int_{\mathbb{R}^{d+1}} K_1^2\left(\frac{x-z}{h_1} \right) K_2^2\left(\frac{y-v}{h_2} \right) f_{X,Y}(z,v) dz\, dv, \qquad \text{(A.90)}$$

$$\hat{\mathbf{n}}h_1^d \sum_{\mathbf{i} \in \mathbf{i}} \left(E^2\left(\tilde{\Gamma}_{\mathbf{i}}(x,y) \right) \right) = h_1^{-d} \left(\int_{\mathbb{R}^{d+1}} K_1\left(\frac{x-z}{h_1} \right) K_2\left(\frac{y-v}{h_2} \right) f_{X,Y}(z,v) dz\, dv \right)^2.$$

Since $K_2(t) \leq 1$ for any $t \in \mathbb{R}$, we have

$$\hat{\mathbf{n}}h_1^d \sum_{\mathbf{i} \in \mathbf{i}} \left(E^2\left(\tilde{\Gamma}_{\mathbf{i}}(x,y) \right) \right) \leq h_1^{-d} \left(\int_{\mathbb{R}^d} K_1\left(\frac{x-z}{h_1} \right) \left[\iint_{\mathbb{R}} f_{X,Y}(z,v) dv \right] dz \right)^2$$

$$= h_1^d \left(\int_{\mathbb{R}^d} K_1(t) f_X(x-th_1) dt \right)^2, \qquad \text{(A.91)}$$

which goes to 0 by Assumption H_1 and Lebesgue dominated Theorem. By a change of variables, we have

$$\hat{\mathbf{n}}h_1^d \sum_{\mathbf{i} \in \mathbf{i}} E\left(\left(\tilde{\Gamma}_{\mathbf{i}}(x,y) \right)^2 \right) = \int_{\mathbb{R}^{d+1}} K_1^2(s) K_2^2\left(\frac{y-v}{h_2} \right) f_{X,Y}(x-sh_1,v) ds\, dv$$

$$= \int_{\mathbb{R}^{d+1}} K_1^2(s) K_2^2\left(\frac{y-v}{h_2} \right) \qquad \text{(A.92)}$$

$$\times \left(\mathbb{1}_{]-\infty,y[}(v) + \mathbb{1}_{]y,\infty[}(v) \right) f_{X,Y}(x-sh_1,v) ds\, dv.$$

Remark that if $v \in \,]y,\infty[$, then $K_2^2((y-v)/h_2) \to 0$, and if $v \in \,]-\infty,y[$, then $K_2^2((y-v)/h_2) \to 1$. Thus

$$\lim_{\mathbf{n} \to \infty} K_1^2(s) K_2^2\left(\frac{y-v}{h_2} \right) \mathbb{1}_{]-\infty,y[}(v) f_{X,Y}(x-sh,v) = K_1^2(s) \mathbb{1}_{]-\infty,y[}(v) f_{X,Y}(x,v). \qquad \text{(A.93)}$$

Then, we deduce from the dominated convergence Theorem that

$$\lim_{\mathbf{n} \to \infty} \hat{\mathbf{n}}h_1^d \sum_{\mathbf{i} \in \mathbf{i}} E\left(\left(\tilde{\Gamma}_{\mathbf{i}}(x,y) \right)^2 \right) = \int_{\mathbb{R}^{d+1}} K_1^2(s) \mathbb{1}_{]-\infty,y[}(v) f_{X,Y}(x,v) ds\, dv$$

$$= \int_{-\infty}^{y} f_{X,Y}(x,v) dv \int_{\mathbb{R}^d} K_1^2(s) ds. \qquad \text{(A.94)}$$

Let us now prove that there exists C such that $\hat{\mathbf{n}}h_1^d \tilde{R}_{\mathbf{n}} < C$, for $\mathbf{n}$ being large enough. As K_1 and K_2 are bounded, by applying Lemma 4.1(ii) of Carbon et al. [19], we obtain

$$\left| E\tilde{\Gamma}_{\mathbf{i}}(x,y)\tilde{\Gamma}_{\mathbf{j}}(x,y) \right| \leq C\hat{\mathbf{n}}^{-2}h_1^{-2d}\psi(1,1)\varphi(\|\mathbf{i}-\mathbf{j}\|). \qquad \text{(A.95)}$$

Let $D_{\mathbf{n}}$ be a sequence of real numbers tending to 0 as $\mathbf{n} \to \infty$ and $S = \{\mathbf{i}, \mathbf{j} \in \mathcal{I}_{\mathbf{n}}, \text{dist}(\mathbf{i}, \mathbf{j}) \leq D_{\mathbf{n}}\}$. Then, $\widetilde{R}_{\mathbf{n}} \leq \widetilde{R}_{\mathbf{n}}^1 + \widetilde{R}_{\mathbf{n}}^2$, with

$$\widetilde{R}_{\mathbf{n}}^1 = \sum_{\mathbf{i}, \mathbf{j} \in S} \left| E \widetilde{\Gamma}_{\mathbf{i}}(x, y) \widetilde{\Gamma}_{\mathbf{j}}(x, y) \right|, \qquad \widetilde{R}_{\mathbf{n}}^2 = \sum_{\mathbf{i}, \mathbf{j} \in S^c} \left| E \widetilde{\Gamma}_{\mathbf{i}}(x, y) \widetilde{\Gamma}_{\mathbf{j}}(x, y) \right|, \tag{A.96}$$

where S^c stands for the complement of S. Now,

$$
\begin{aligned}
\widetilde{R}_{\mathbf{n}}^1 = \frac{1}{\widehat{\mathbf{n}}^2 h_1^{2d}} \sum_{\mathbf{i}, \mathbf{j} \in S} & \left| \int_{\mathbb{R}^{2d+2}} K_1\left(\frac{x - u_1}{h_1}\right) K_2\left(\frac{y - v_1}{h_2}\right) K_1\left(\frac{x - u_2}{h_2}\right) \right. \\
& \times K_2\left(\frac{y - v_2}{h_2}\right) f_{Z_{\mathbf{i}}, Z_{\mathbf{j}}}(u_1, v_1, u_2, v_2) du_1\, dv_1\, du_2\, dv_2 \\
& - \int_{\mathbb{R}^{d+1}} K_1\left(\frac{x - u_1}{h_1}\right) K_2\left(\frac{y - v_1}{h_2}\right) f_{X,Y}(u_1, v_1) du_1 dv_1 \\
& \left. \times \int_{\mathbb{R}^{d+1}} K_1\left(\frac{x - u_2}{h_1}\right) K_2\left(\frac{y - v_2}{h_2}\right) f_{X,Y}(u_2, v_2) du_2 dv_2 \right| \\
\leq \frac{1}{\widehat{\mathbf{n}}^2 h_1^{2d}} \sum_{\mathbf{i}, \mathbf{j} \in S} & \left| \int_{\mathbb{R}^{2d}} K_1\left(\frac{x - u_1}{h_1}\right) K_1\left(\frac{x - u_2}{h_1}\right) \right. \\
& \left. \times \int_{\mathbb{R}^2} \left[f_{Z_{\mathbf{i}}, Z_{\mathbf{j}}}(u_1, v_1, u_2, v_2) - f_{X,Y}(u_1, v_1) f_{X,Y}(u_2, v_2) \right] dv_1\, dv_2 \right| du_1\, du_2 \\
\leq \frac{1}{\widehat{\mathbf{n}}^2 h_1^{2d}} \sum_{\mathbf{i}, \mathbf{j} \in S} & \int_{\mathbb{R}^{2d}} \left(K_1\left(\frac{x - u_1}{h}\right) K_1\left(\frac{x - u_2}{h_1}\right) \right) \times \left| f_{X_{\mathbf{i}}, X_{\mathbf{j}}}(u_1, u_2) - f_X(u_1) f_X(u_2) \right| du_1 du_2.
\end{aligned}
\tag{A.97}
$$

Under H_3 and (3.8), we get

$$\widehat{\mathbf{n}} h_1^d \widetilde{R}_{\mathbf{n}}^1 \leq C \left(\int_{\mathbb{R}^d} |K_1(u)| du \right)^2 \widehat{\mathbf{n}}^{-1} h_1^d \sum_{\mathbf{i}, \mathbf{j} \in S_0} 1 \leq C \widehat{\mathbf{n}}^{-1} h_1^d D_n^N \widehat{\mathbf{n}} = C h_1^d D_{\mathbf{n}}^N. \tag{A.98}$$

Now, using (A.95), one can write

$$
\begin{aligned}
\widehat{\mathbf{n}} h_1^d \widetilde{R}_{\mathbf{n}}^2 = C \widehat{\mathbf{n}}^{-1} h_1^{-d} \sum_{\mathbf{i}, \mathbf{j} \in S^c} \{ \varphi(\|\mathbf{i} - \mathbf{j}\|) \} & \leq C h_1^{-d} \sum_{\|\mathbf{i}\| > D_{\mathbf{n}}} \{ \varphi(\|\mathbf{i}\|) \} \\
& \leq C h_1^{-d} D_{\mathbf{n}}^{-N} \sum_{\|\mathbf{i}\| > D_{\mathbf{n}}} \|\mathbf{i}\|^{N - \theta}.
\end{aligned}
\tag{A.99}
$$

As $\theta > N + 1$, the choice $D_{\mathbf{n}} = h_1^{-d/N}$ gives the desired result. This ends the proof. $\qquad \square$

Acknowledgments

References

[1] J. P. Chilès and P. Delfiner, *Statistics for Spatial Data*, John Wiley & Sons, New York, NY, USA, 1999.

[2] X. Guyon, *Random Fields on a Network—Modeling, Statistics, and Applications*, Probability and Its Applications, Springer, New York, NY, USA, 1995.

[3] L. Anselin and R. J. G. M. Florax, *New Directions in Spatial Econometrics*, Springer, Berlin, Germany, 1995.

[4] N. A. C. Cressie, *Statistics for Spatial Data*, Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics, John Wiley & Sons, New York, NY, USA, 1991.

[5] B. D. Ripley, *Spatial Statistics*, John Wiley & Sons, New York, NY, USA, 1981, Wiley Series in Probability and Mathematical Statistics.

[6] W. Stute, "Conditional empirical processes," *The Annals of Statistics*, vol. 14, no. 2, pp. 638–647, 1986.

[7] M. Samanta, "Nonparametric estimation of conditional quantiles," *Statistics & Probability Letters*, vol. 7, no. 5, pp. 407–412, 1989.

[8] S. Portnoy, "Asymptotic behavior of regression quantiles in nonstationary, dependent cases," *Journal of Multivariate Analysis*, vol. 38, no. 1, pp. 100–113, 1991.

[9] H. L. Koul and K. Mukherjee, "Regression quantiles and related processes under long range dependent errors," *Journal of Multivariate Analysis*, vol. 51, no. 2, pp. 318–337, 1994.

[10] T. Honda, "Nonparametric estimation of a conditional quantile for α-mixing processes," *Annals of the Institute of Statistical Mathematics*, vol. 52, no. 3, pp. 459–470, 2000.

[11] Z. Cai, "Regression quantiles for time series," *Econometric Theory*, vol. 18, no. 1, pp. 169–192, 2002.

[12] A. Gannoun, J. Saracco, and K. Yu, "Nonparametric prediction by conditional median and quantiles," *Journal of Statistical Planning and Inference*, vol. 117, no. 2, pp. 207–223, 2003.

[13] K. Yu, Z. Lu, and J. Stander, "Quantile regression: applications and current research areas," *Journal of the Royal Statistical Society Series D*, vol. 52, no. 3, pp. 331–350, 2003.

[14] R. Koenker and I. Mizera, "Penalized triograms: total variation regularization for bivariate smoothing," *Journal of the Royal Statistical Society Series B*, vol. 66, no. 1, pp. 145–164, 2004.

[15] M. Hallin, Z. Lu, and K. Yu, "Local linear spatial quantile regression," *Bernoulli*, vol. 15, no. 3, pp. 659–686, 2009.

[16] S. A. O. Abdi, S. Dabo-Niang, A. Diop, and A. O. Abdi, "Consistency of a nonparametric conditional quantile estimator for random fields," *Mathematical Methods of Statistics*, vol. 19, no. 1, pp. 1–21, 2010.

[17] S. Dabo-Niang and B. Thiam, "Robust quantile estimation and prediction for spatial processes," *Statistics & Probability Letters*, vol. 80, no. 17-18, pp. 1447–1458, 2010.

[18] M. Carbon, M. Hallin, and L. T. Tran, "Kernel density estimation for random fields: the L_1 theory," *Journal of Nonparametric Statistics*, vol. 6, no. 2-3, pp. 157–170, 1996.

[19] M. Carbon, L. T. Tran, and B. Wu, "Kernel density estimation for random fields (density estimation for random fields)," *Statistics & Probability Letters*, vol. 36, no. 2, pp. 115–125, 1997.

[20] G. Biau and B. Cadre, "Nonparametric spatial prediction," *Statistical Inference for Stochastic Processes*, vol. 7, no. 3, pp. 327–349, 2004.

[21] Z.-D. Lu and X. Chen, "Spatial nonparametric regression estimation: non-isotropic case," *Acta Mathematicae Applicatae Sinica English Series*, vol. 18, no. 4, pp. 641–656, 2002.

[22] Z. Lu and X. Chen, "Spatial kernel regression estimation: weak consistency," *Statistics & Probability Letters*, vol. 68, no. 2, pp. 125–136, 2004.

[23] M. Hallin, Z. Lu, and L. T. Tran, "Kernel density estimation for spatial processes: the L_1 theory," *Journal of Multivariate Analysis*, vol. 88, no. 1, pp. 61–75, 2004.

[24] M. Hallin, Z. Lu, and L. T. Tran, "Local linear spatial regression," *The Annals of Statistics*, vol. 32, no. 6, pp. 2469–2500, 2004.

[25] S. N. Lahiri and J. Zhu, "Resampling methods for spatial regression models under a class of stochastic designs," *The Annals of Statistics*, vol. 34, no. 4, pp. 1774–1813, 2006.

[26] M. Carbon, C. Francq, and L. T. Tran, "Kernel regression estimation for random fields," *Journal of Statistical Planning and Inference*, vol. 137, no. 3, pp. 778–798, 2007.

[27] S. Dabo-Niang and A.-F. Yao, "Kernel regression estimation for continuous spatial processes," *Mathematical Methods of Statistics*, vol. 16, no. 4, pp. 298–317, 2007.

[28] F. Ferraty, A. Rabhi, and P. Vieu, "Conditional quantiles for dependent functional data with application to the climatic El Niño phenomenon," *Sankhyā*, vol. 67, no. 2, pp. 378–399, 2005.

[29] L. Devroye, "On the almost everywhere convergence of nonparametric regression function estimates," *The Annals of Statistics*, vol. 9, no. 6, pp. 1310–1319, 1981.

[30] L. T. Tran, "Kernel density estimation on random fields," *Journal of Multivariate Analysis*, vol. 34, no. 1, pp. 37–53, 1990.

[31] P. Doukhan, *Mixing Properties and Examples*, vol. 85 of *Lecture Notes in Statistics*, Springer-Verlag, New York, 1994.

[32] K. Yu and M. C. Jones, "Local linear quantile regression," *Journal of the American Statistical Association*, vol. 93, no. 441, pp. 228–237, 1998.