Mathematical Modelling and Analysis
Volume 27, Issue 3, 360–382, 2022
https://doi.org/10.3846/mma.2022.14909

# Local Linear Modelling of the Conditional Distribution Function for Functional Ergodic Data

**Somia Ayad[a], Ali Laksaci[b], Saâdia Rahmani[a] and Rachida Rouane[a]**

[a] *Laboratory of Stochastic Models, Statistics and Applications*
University of Saida–Dr. Moulay Tahar, 20000 Saida, Algeria
[b] *Department of Mathematics, College of Science*
King Khalid University, 61413 Abha, Saudi Arabia

E-mail: somia.ayad@univ-saida.dz

E-mail: alilak@yahoo.fr

E-mail(*corresp.*): saadia.rahmani@gmail.com

E-mail: rouane09@yahoo.fr

Received May 10, 2021; revised April 15, 2022; accepted April 24, 2022

**Abstract.** The focus of functional data analysis has been mostly on independent functional observations. It is therefore hoped that the present contribution will provide an informative account of a useful approach that merges the ideas of the ergodic theory and the functional data analysis by using the local linear approach. More precisely, we aim, in this paper, to estimate the conditional distribution function (CDF) of a scalar response variable given a random variable taking values in a semi-metric space. Under the ergodicity assumption, we study the uniform almost complete convergence (with a rate), as well as the asymptotic normality of the constructed estimator. The relevance of the proposed estimator is verified through a simulation study.

**Keywords:** ergodic sata, functional data, local linear estimator, conditional distribution function, nonparametric estimation, asymptotic properties.

**AMS Subject Classification:** 62G05; 62G08; 62G20; 62G35; 62H12.

## 1 Introduction and motivations

Over the almost last two decades, functional data analysis (FDA) has established itself as a dynamic and important field of statistical research. It has become very broad, with many specialized directions of research. This statistic area offers effective new tools and has stimulated novel methodological developments. With the availability of large amounts of data as well as the development of the computer instruments, (FDA) swept across various fields of applied sciences (for instance biometrics, geophysics and econometrics). There are many nonparametrc problems for functional data which have attracted a growing interest; one may refer to the famous work of Ferraty and Romain [13], the monograph of Ferraty and Vieu [14] and the pioneer book of Kokoszka and Reimherr [16] as well as the references therein.

Despite the simplicity of the classical kernel estimate and its availability in many statistical software packages, like R and Matlab make that it easy to understand and implement, its simplicity leads to some weaknesses; the most obvious of which is boundary bias effect. Moreover, it is well known that among the smoothing procedures, the local polynomial approach has various advantages over the classical kernel method. In particular, this method has better properties concerning the bias terms on the other one (cf. Fan and Gijbels [10] for an extensive discussion). In the context of the finite dimensional space, the local linear method is well established, frequently used and it has been the subject of considerable studies, and key references on this topic are Chen et al. [6], Fan and Yao [11] and references therein. However, before the pioneer work of Barrientos-Marin et al. [3], only few results are available for the local linear modeling in the functional statistics setup. Indeed, the first results, in this direction, were established by Baìllo and Grané [2]. This paper focuses on the local linear estimation of the regression operator when the explanatory variable takes values in a Hilbert space. The general case, where the regressors do not belong to a Hilbert space but to a semi-metric space, has been considered not only by Barrientos-Marin et al. [3], but also by El Methni and Rachdi [9], Demongeot et al. [8] and Laksaci et al. [19].

Recently, the paper of Bouanani et al. [4] has completed the theoretical advances presented by Laksaci et al. [19] by establishing the asymptotic normality of the local linear estimates for several conditional models.

Weak dependencies have been considered by many authors in the context of both discrete and continuous-time processes. We consider, in this paper, the ergodic framework which is more general than the weak dependencies. More precisely, we examine the local linear estimator's properties of the (CDF) when the data of our constructed estimator are ergodic.

In the literature, several real examples have been studied in order to emphasize the usefulness of such dependency. For instance, the ergodicity assumption models several phenomena in physics like the thermodynamic properties of gases, atoms or plasma. In a more general way, the ergodic theory becomes crucial because there are many phenomena which are neither independent nor $\alpha$-mixing either. In the past three decades, the study of statistical models adapted to such kind of dependency has been impressively large but mostly restricted to the standard multivariate situation where both the response and the explanatory variables are real or multivariate (see, Delecroix and Rosa [7] and Laïb and Ould-Saïd [18]). However, there are very few advances in this direction when the regressor is functional. One may refer to the work of Laib and Louani [17]. The authors studied under ergodicity assumption the asymptotic properties of an estimator of the regression operator. Related works can be found in the paper of Laib and Louani [17] when the data are completely observed and Chaouch et al. [5] for the right censored ones.

As it is mentioned above, the main aim of this paper is to construct and study under general conditions, the uniform almost complete convergence rate as well as the asymptotic normality of a local linear estimator of the CDF. For this purpose, it is assumed that the covariate takes its values in an infinite dimensional space and the data are sampled from a stationary ergodic process. Recall that uniform consistency results have been successfully used in the standard nonparametric setting (see for instance, Ferraty et al. [12], Ling et al. [20] and Kara-Zaitri et al. [15]). Each of these papers considers the case of the local constant method. However, in this contribution, we consider a more efficient estimate of the CDF by the local linear method. To make this paper as much self-contained as possible, the nonparametric model and its associated local linear estimator are constructed in Section 2. In the same section, we report some notations required for this contribution. The assumptions, under which the main results are valid, are stated and discussed in Section 3. Then, we derive theoretical results by giving a deep asymptotic study of the behaviour of the estimate, including the almost complete convergence of the CDF uniformly in the functional argument $x$ as well as the asymptotic gaussian distribution. The relevance of the proposed estimator is verified through a simulation study in Section 4. Finally, the paper is ended with a technical appendix.

## 2 Local linear estimator construction

Let $(X_i, Y_i)_{i=1,\ldots,n}$ be a strictly stationary (in an ergodic sense) process of $\mathcal{F} \times \mathbb{R}$-valued random elements, where $\mathcal{F}$ is a semi-metric space with semi-metric $d$. We assume that there exists a regular version of the conditional distribution of $Y$ given $X$, which is absolutely continuous with respect to the *Lebesgue* measure on $\mathbb{R}$.

Interest centers on the conditional behavior of $Y$ given $X$. To this end it is convenient to consider

$$F^x(y) = \mathbb{P}(Y_i \leq y | X_i = x).$$

Since the local linear approach requires a smoothing assumption that allows us to approximate locally the nonparametric CDF, we estimate the function $F^x(.)$ by assuming that it is smoothed enough to be locally approximated by a linear function. For this aim, we introduce two locating functions $\delta$ and $\rho$ (see Barrientos et al. [3] for more discussion on these bilinear continuous operators) and we consider a subset $\mathcal{C}_\mathcal{F}$ of $\mathcal{F}$ such that for $x_k \in \mathcal{C}_\mathcal{F}, \mathcal{C}_\mathcal{F} \subset \bigcup_{k=1}^{d_n} B(x_k, r_n)$ where $r_n$ (resp. $d_n$) is a sequence of positive real (resp. integer) numbers and

$B(x_k, r_n) = \{x'_k \in \mathcal{F}/|\delta(x'_k, x_k)| < r_n\}$. Such approximation can be expressed, for any $z \in \mathcal{C}_\mathcal{F}$ in the neighborhood of $x$ by:

$$F^z(y) = \alpha + \beta\rho(z, x) + o(\rho(z, x)). \tag{2.1}$$

We assume that the underlying process $(X_i, Y_i)$ is functional stationary ergodic. Then the estimator $\widehat{F}^x$ of $F^x$ can be seen as the solution of the following minimization problem

$$\min_{(\alpha,\beta)\in\mathbb{R}^2} \sum_{i=1}^{n} \left( J\left(\frac{y - Y_i}{h_J}\right) - \alpha - \beta\rho(X_i, x) \right)^2 K\left(\frac{\delta(x, X_i)}{h_K}\right), \tag{2.2}$$

where the bi-functional $\delta(.,.)$ is lied with the topological structure of the functional space $\mathcal{F}$, that means $|\delta(x, z)| = d(x, z)$, whereas, $\rho$ controls the local sharp of the model (see formula (2.1)). $K$ is a kernel, $J$ is a distribution function and $h_K = h_{K,n}$ (respectively $h_J = h_{J,n}$) is a sequence of positive real numbers. More precisely, the functional local linear estimator $\widehat{F}^x(y)$ of $F^x(y)$ is then $\widehat{\alpha}$ which is the first component of the pair $(\alpha, \beta)$ solution of the minimization problem (2.2). However, if the bi-functional operator $\rho$ is such that $\rho(z, z) = 0, \forall z \in \mathcal{F}$, then the quantity $\widehat{F}^x(y)$ is explicitly defined by:

$$\widehat{F}^x(y) = \frac{\displaystyle\sum_{j=1}^{n} \Gamma_j(x) K\left(\frac{\delta(x, X_j)}{h_K}\right) J\left(\frac{y - Y_j}{h_J}\right)}{\displaystyle\sum_{j=1}^{n} \Gamma_j(x) K\left(\delta(x, X_j)/h_K\right)}, \tag{2.3}$$

with

$$\Gamma_j(x) = \sum_{i=1}^{n} \rho_i^2(x) K_i(x) - \left(\sum_{i=1}^{n} \rho_i(x) K_i(x)\right) \rho_j,$$

where $\rho_i(x) = \rho(X_i, x)$, and $K_i(x) = K\left(\delta(x, X_i)/h_K\right)$.

# 3 Main results

## 3.1 Uniform almost complete convergence

### 3.1.1 Assumptions and notations

First we need to introduce some further notations. For $i = 1, \ldots, n$, let $\mathfrak{F}_i$ and $\mathcal{G}_i$ denote, respectively, the $\sigma$-field generated by $((X_1, Y_1), \ldots, (X_i, Y_i))$, and $((X_1, Y_1), \ldots, (X_i, Y_i), X_{i+1})$. For any fixed $y$ in $\mathbb{R}$, $\mathcal{N}_y$ denotes a fixed neighborhood of $y$. In the sequel, we will also need to define the small ball probability function by $\phi_x(h_1, h_2) = \mathbb{P}(h_2 \le \delta(X, x) \le h_1)$ and we will denote by $C$ and $C'$ some strictly positive generic constants. Finally, with some abuse of notations, we write $J_j(y)$ for $J\left((y - Y_j)/h_J\right)$ and we shall write $\phi$ instead of $\phi_x$.

Our consistency results are summarized in Theorem 1 and rely on the following seven assumptions:

**Structural hypotheses:**

**On the ergodic functional variables:**
We suppose that the strictly stationary ergodic process $(X_i, Y_i)_{i\in\mathbb{N}^*}$ satisfies: For all $r > 0$

(H1) $\forall x \in \mathcal{C}_{\mathcal{F}}$, $0 < C\phi(r) \leq \mathbb{P}(X \in B(x,r)) \leq C'\phi(r)$. Furthermore, $\exists\, \eta_0 > 0, \forall \eta < \eta_0, \phi'(\eta) < C$, where $\phi'$ denotes the first order derivative of $\phi$.

(H2) For all $i = 1, \dots, n$, there exist a determinist function $\phi_i$ such that:

**i)** $0 < C\phi_i(r) < \mathbb{P}(X_i \in B(x,r)|\mathfrak{F}_{i-1}) \leq C'\phi_i(r)$,

**ii)** $\frac{1}{n\phi(r)} \sum_{i=1}^{n} \phi_i(r) \longrightarrow 1\, a.co.$

**Technical and regularity conditions:**

(H3) **On the regularity of the model:**
There exist some positive constants $b_1$ and $b_2$ such that:
$\forall (x_1, x_2) \in \mathcal{C}_{\mathcal{F}} \times B(x_1, h_K)$ and $\forall (y_1, y_2) \in \mathcal{N}_y \times \mathcal{N}_y$:

$$|F^{x_1}(y_1) - F^{x_2}(y_2)| \leq C\left(|\delta(x_1,x_2)|^{b_1} + |y_1 - y_2|^{b_2}\right),$$

(H4) **On the bi-functional operators $\rho$ and $\delta$:**

(i) $\forall z \in \mathcal{F}$, $C|\delta\,(x,z)\,| \leq |\rho\,(x,z)\,| \leq C'|\delta\,(x,z)\,|$,

(ii) $\forall (x_1,x_2) \in \mathcal{C}_{\mathcal{F}} \times \mathcal{C}_{\mathcal{F}}, |\rho\,(x_1,x) - \rho\,(x_2,x)\,| \leq C'|\delta(x_1,x_2)|$.

(H5) **On the kernel $K$ and the distribution function $J$:**

(i) $K$ is a nonnegative bounded and Lipschitz kernel on its support $[-1;1]$.

(ii) $J$ is a differentiable function such that $\int |t|^{b_2} J^{(1)}(t)dt < \infty$.

(iii) $\mathbb{E}(J_j(y)|\mathcal{G}_{j-1}) = \mathbb{E}(J_j(y)|X_j)$.

(H6) Taking $r_n = O\,(\log n/n)$, the sequence $d_n$ satisfies:

$$\frac{(\log n)^2}{n\phi(h_K)} < \log d_n < \frac{n\phi(h_K)}{\log n} \text{ and } \sum_{n=1}^{\infty} d_n^{(1-\varrho)} < \infty \quad \text{for some} \quad \varrho > 1.$$

(H7) **On the bandwidth $h_K$ with respect to $\rho$ and $\phi$ :**

(i) There is a positive integer $n_0$, such that, $\forall n > n_0$:

$$-\frac{1}{\phi(h_K)}\int_{-1}^{1} \phi(zh_K, h_K)\frac{d}{dz}\left(z^2K(z)\right)dz > C,$$

(ii) $\lim_{n\to\infty} h_K = 0$, $\lim_{n\to\infty} h_H = 0$, and $\lim_{n\to\infty} \frac{\log n}{n\phi(h_K)} = 0$,

(iii) $h_K \int_{B(x,h_K)} \rho\,(u,x)\,dP\,(u) = o\left(\int_{B(x,h_K)} \rho^2\,(u,x)\,dP\,(u)\right)$, where $dP(x)$ is the cumulative distribution of $X$.

### 3.1.2 Comments on the hypotheses

(H1) involves the small ball techniques. It is clearly unrestrictive, since it is the same as that frequently used in the FDA context. We precise that the ergodic nature of the data is exploited by (H2) which is a very mild condition in comparison of that imposed by Laib and Louani [17].
Concerning (H3), this condition is necessary to evaluate the bias term in our asymptotic result. The constants $b_1$ and $b_2$ control the model's regularity and the only condition imposed is their positivity. In other words, this assumption guarantees slower variance of the operator compared to $X$. More the parameters $b_1$ and $b_2$ are small, more the curves of the operator's evolution as a function of $X$ are smooth, and less the estimate obtained is biased.

Condition (H4)(i) is unrestrictive condition and it is verified if $\rho(.,.) = \delta(.,.)$ (in this special case, (H7) (iii) means that the local expectation of $\rho$ is small enough with respect to its moment of second order); or if

$$\lim_{\delta(z,x)\to 0} |\rho(z,x)/\delta(z,x) - 1| = 0.$$

Indeed, $\forall z \in B(x,h_K)$, we have:

$$\left|\frac{\rho(z,x)-\delta(z,x)}{h_K}\right| \leq \left|\frac{\rho(z,x)}{\delta(z,x)} - 1\right| \longrightarrow 0 \text{ as } \delta(z,x) \to 0.$$

The Lipschitz condition (H4) (ii) on the locating function $\rho$ is the same used by Barrientos-Marin et al. [3] and it is typical in the context of local polynomial smoothing.

(H5)(i) could be replaced by another assumption such as the boundness of the kernel $K$. The slightly stronger assumption (H5) (i) just makes the proof of uniform convergence simpler. (H5)(ii) and (iii) are technical conditions imposed for brevity of proofs.

In (H6), the covering hypothesis on the subset $\mathcal{C}_{\mathcal{F}}$ is linked to the topological structure of our functional space $\mathcal{F}$. It controls Kolmogorov's entropy of the set $\mathcal{C}_{\mathcal{F}}$. Such consideration has been discussed and commented by Ferraty et al. [12]. The authors give several examples for which this condition is satisfied.

The assumption (H7) (i) precise the behaviour of the smoothing parameter $h_K$ in relation with the small ball probabilities and the kernel function $K$. The local behaviour of $\rho$ which models the local shape of our model is controlled by (H7)(iii).

We are now ready to state our first result which is the uniform almost complete convergence of the estimator $\widehat{F}^x(y)$ on the subset $\mathcal{C}_{\mathcal{F}}$.

**Theorem 1.** *As soon as assumptions (H1)–(H7) are fulfilled, we have*

$$\sup_{x\in\mathcal{C}_{\mathcal{F}}} |\widehat{F}^x(y) - F^x(y)| = O\left(h_K^{b_1} + h_J^{b_2}\right) + O\left(\sqrt{\frac{\log d_n}{n\phi(h_K)}}\right), \quad a.co.$$

Before starting the proof of this theorem, we introduce the following further notations:

$$\widehat{F}_N^x(y) := \frac{1}{n\mathbb{E}\left(\Gamma_1 K_1\right)} \sum_{j=1}^{n} \Gamma_j K_j J_j, \quad \bar{F}_N^x\left(y\right) := \frac{1}{n\mathbb{E}\left(\Gamma_1 K_1\right)} \sum_{j=1}^{n} \mathbb{E}\left(\Gamma_j K_j J_j | \mathfrak{F}_{j-1}\right),$$

$$\widehat{F}_D\left(x\right) := \frac{1}{n\mathbb{E}\left(\Gamma_1 K_1\right)} \sum_{j=1}^{n} \Gamma_j K_j, \quad \bar{F}_D\left(x\right) := \frac{1}{n\mathbb{E}\left(\Gamma_1 K_1\right)} \sum_{j=1}^{n} \mathbb{E}\left(\Gamma_j K_j | \mathfrak{F}_{j-1}\right).$$

Then, the proof of Theorem 1 is based on the following decomposition:

$$\widehat{F}^x(y) - F^x(y) = B_n(x,y) + \frac{1}{\widehat{F}_D\left(x\right)}\left[\left(B_n(x,y) + F^x\left(y\right)\right) A_n(x,y) + R_n(x,y)\right], \tag{3.1}$$

where

$$B_n(x,y) = \frac{\bar{F}_N^x\left(y\right)}{\bar{F}_D\left(x\right)} - F^x\left(y\right), \quad A_n(x,y) = \bar{F}_D\left(x\right) - \widehat{F}_D\left(x\right),$$

$$R_n(x,y) = \widehat{F}_N^x\left(y\right) - \bar{F}_N^x\left(y\right).$$

As immediate consequence of the decomposition (3.1), we need to prove the following lemmas:

**Lemma 1.** *Under assumptions (H1)–(H5) and (H7), we have that*

$$\sup_{x \in \mathcal{C}_{\mathcal{F}}} |B_n(x,y)| = O\left(h_K^{b_1} + h_J^{b_2}\right).$$

**Lemma 2.** *Under the hypotheses of Theorem* 1*, we obtain*

$$\sup_{x \in \mathcal{C}_{\mathcal{F}}} |R_n(x,y)| = O_{a.co}\left(\sqrt{\log d_n / \left(n\phi\left(h_K\right)\right)}\right).$$

**Lemma 3.** *Under assumptions (H1)–(H5) (i), (H6) and (H7), we get*

$$(i) \sup_{x \in \mathcal{C}_{\mathcal{F}}} |A_n(x,y)| = O_{a.co}\left(\sqrt{\frac{\log d_n}{n\phi(h_K)}}\right), \; (ii) \sum_{n=1}^{\infty} \mathbb{P}\left(\inf_{x \in \mathcal{C}_{\mathcal{F}}} \widehat{F}_D(x) < \frac{1}{2}\right) < \infty.$$

Thus, the proof of our main result is based on the previous lemmas combined with Lemma 5 of Ayad et al. [1] and the technical lemma 1 of Laib and Louani [17].

### 3.2 Asymptotic normality

Let us first focus on the supplementary assumptions we need to derive the asymptotic normality of our estimator.

(B1) The hypothesis (H1) holds and there exists a function $\Psi(\cdot)$ such that:

$$\forall t \in [-1, 1], \; \lim_{h_K \to 0} \frac{\phi(-h_K, th_K)}{\phi(h_K)} = \Psi(t).$$

(B2) The hypothesis (H3) holds and for all $(x_1, x_2, y_1, y_2) \in \mathcal{C}_\mathcal{F} \times \mathcal{C}_\mathcal{F} \times \mathcal{N}_y \times \mathcal{N}_y$:

$$\begin{cases} F : \mathcal{F} \times \mathbb{R} \longrightarrow \mathbb{R}, \lim_{|\delta(x_1,x_2)| \to 0} F^{x_1}(y) = F^{x_2}(y), \\ \lim_{|y_1 - y_2| \to 0} F^x(y_1) = F^x(y_2). \end{cases}$$

(B3) The locating functions $\rho$ and $\delta$ satisfy (H4), and

$$\sup_{u \in B(x,r)} |\rho(u,x) - \delta(x,u)| = o(r).$$

(B4) The hypothesis (H5) holds and the first derivative $K'$ of the kernel $K$ satisfies:

$$K^2(1) - \int_{-1}^{1} (K^2(u))' \Psi(u) du > 0.$$

(B5) Assumption (H7) holds and $\lim_{n\to\infty}(n-1)^k h_K^l \phi(h_K) = 0$, for $k = 1, 2$ and $l = 4, 5$.

In addition, we need to introduce the quantities $M_c$ and $N(a,b)$ which will appear in the computation of $\mathbb{E}\left(K_j^c | \mathfrak{F}_{j-1}\right)$.

$$M_c = K^c(1) - \int_{-1}^{1} (K^c(u))' \Psi(u) du, \qquad \text{where } c = 1, 2,$$

and for all $a > 0$ and $b = (2,4)$, $\quad N(a,b) = K^a(1) - \int_{-1}^{1} (u^b K^a(u))' \Psi(u) du$.

**Theorem 2.** *Under assumptions (B1)–(B5), (H1) and (H6) and if the smoothing parameters $h_K$ and $h_J$ satisfy $\sqrt{n\phi(h_K)}\left(h_K^{b_1} + h_J^{b_2}\right) \to 0$ as $n \to \infty$. We have*

$$\sqrt{n\phi(h_K)}\left(\widehat{F}^x(y) - F^x(y)\right) \xrightarrow{\mathcal{D}} N\left(0, V_{JK}(x,y)\right),$$

*where $V_{JK}(x,y) = \dfrac{M_2}{M_1^2} F^x(y)\left(1 - F^x(y)\right)$, and $\xrightarrow{\mathcal{D}}$ means the convergence in distribution.*

The first step of the proof consists in rewriting the decomposition (3.1) in the following way:

$$\widehat{F}^x(y) - F^x(y) = B_n(x,y) + \frac{C_n(x,y) + Q_n(x,y)}{\widehat{F}_D(x)},$$

where

$$C_n(x,y) = B_n(x,y) A_n(x,y), \quad Q_n(x,y) = R_n(x,y) + F^x(y) A_n(x,y).$$

Then, to state asymptotic normality, we remark that the hypothesis (B2) ensures the asymptotic negligence of $B_n(x,y)$. Moreover, according to Lemma 3 (i), $C_n$ converges almost completely to zero when $n$ goes to infinity. Consequently, the proof of Theorem 2 can be deduced from the following lemmas for which the proofs are relegated to the Appendix.

**Lemma 4.** *Under the assumptions of Theorem 2, we have*

$$\sqrt{n\phi\left(h_K\right)/V_{JK}(x,y)}\,Q_n(x,y) \xrightarrow{\mathcal{D}} N(0,1).$$

**Lemma 5.** *Under assumptions (H1)–(H5) (i), (H6) and (H7), we have*

$$\widehat{F}_D\left(x\right) - 1 = o_p(1).$$

## 4 On simulated data

We now conduct a simulation study in which the finite sample performance of the local linear estimator given in formula (2.3) is compared to the following local constant estimator:

$$\widetilde{F}^x(y) = \sum_{j=1}^{n} K_j(x)J_j(y) \Big/ \sum_{j=1}^{n} K_j(x).$$

Let us use the following regression model:

$$Y = r(X) + \varepsilon,$$

where the random variable $\varepsilon$ is normally distributed with a variance equal to 0.075. The explanatory functional variables are constructed by:

$$X_i(t) = 2w_i t^2 + 0.5\cos(\pi z_i t), \quad i = 1, \dots 200, \;\; t \in [0,1],$$

where $w_i$ are $n$ independent real random variables uniformly distributed over $[0,1]$ and $z_i = \frac{1}{3}z_{i-1} + \zeta_i$. Here $\zeta_i$ are i.i.d. realizations of $N(0,1)$ and are independent from $w_i$ and $z_i$, which is generated independently by $z_0 \sim N(0,1)$. All the curves $X_i$'s were discretized on the same grid generated from 200 equispaced measurements in (0,1) and are plotted in Figure 1.

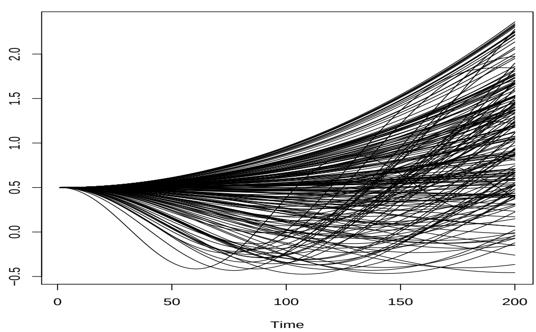


**Figure 1.** The curves $X_i$.

On the other hand, the building of the scalar response $Y$ is obtained by considering the following regression operator:

$$\left(\int_0^1 X'(t)dt\right)^2.$$

Recall that, the conditional distribution of $Y$ given $X = x$ corresponding to this model is explicitly given by the law of $\varepsilon_i$ shifted by $\left(\int_0^1 X'(t)dt\right)^2$.

When dealing with smooth curves such as those introduced herein, it is necessary to measure the proximity by means of a semi-metric based on the $L_2$ norm of some derivative of the curves. The smoothing parameters $h_K$ and $h_J$ are selected through automated cross validation, choosing a value that minimizes the average error on the withheld data. The behavior of our estimator is linked to the good choice of the functions $\delta$ and $\rho$. Because of the smoothness of the curves we take

$$\rho(x, x') = \int_0^1 \theta(t)(x^{(1)}(t) - x'^{(1)}(t))dt,$$

$$\delta(x, x') = \left(\int_0^1 (x^{(1)}(t) - x'^{(1)}(t))^2 dt\right)^{1/2}$$

with the functional index $\theta$ is selected among the eigenfunction of the empirical covariance operator $\frac{1}{n}\sum_{i=1}^n (X_i^{(1)} - \overline{X^{(1)}})^t \left((X_i^{(1)} - \overline{X^{(1)}}\right))$ corresponding to the biggest eigenvalues, where $\overline{X} = \frac{1}{n}\sum_{i=1}^n X_i$.

For both competitors, the kernel $K(u) = (1 - u^2)\mathbb{1}_{[0,1]}$ is used and the distribution function $J$ is defined by:

$$J(u) = \frac{3u}{4}\left(1 - \frac{u^2}{3}\right)\mathbb{1}_{[-1,1]} + \frac{1}{2}.$$

In this illustration, we have followed the following steps:

- *Step 1:* We generate $m$ replications of $(X_i, Y_i)_{i=1,\ldots,n}$.
- *Step 2:* We estimate the conditional local linear distribution (respectively the conditional kernel distribution).
- *Step 3:* We compare these estimators to the Gaussian distribution.

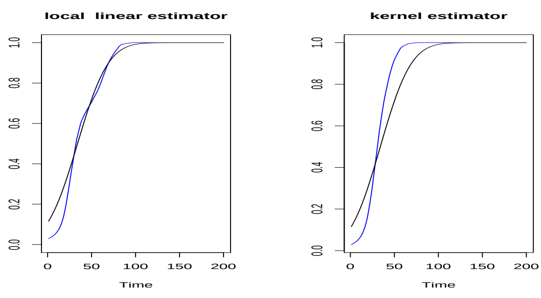


**Figure 2.** Comparison between the both estimators.

In order to eliminate the zero weight, we have removed the negative weighing. The obtained results are plotting in Figure 2. It is clear that the local linear estimator of the CDF operator convincingly outperforms the local constant one.

**Acknowledgements**

The authors would like to thank the Editor and the two anonymous reviewers for their careful reading of the first version of this paper, and for their insightful and constructive comments which improved substantially the quality of the manuscript.

## References

[1] S. Ayad, A. Laksaci, S. Rahmani and R. Rouane. On the local linear modelization of the conditional density for functional and ergodic data. *METRON*, **78**(2):237–254, 2020. https://doi.org/10.1007/s40300-020-00174-6.

[2] A. Baíllo and A. Grané. Local linear regression for functional predictor and scalar response. *Journal of Multivariate Analysis*, **100**(1):102–111, 2009. https://doi.org/10.1016/j.jmva.2008.03.008.

[3] J. Barrientos-Marin, F. Ferraty and P. Vieu. Locally modelled regression and functional data. *Journal of Nonparametric Statistics*, **22**(5):617–632, 2010. https://doi.org/10.1080/10485250903089930.

[4] O. Bouanani, A. Laksaci, M. Rachdi and S. Rahmani. Asymptotic normality of some conditional nonparametric functional parameters in high-dimensional statistics. *Behaviormetrika*, **46**(1):199–233, 2019. https://doi.org/10.1007/s41237-018-0057-9.

[5] M. Chaouch, N. Laïb and E. Ould-Saïd. Nonparametric M-estimation for right censored regression model with stationary ergodic data. *Statistical Methodology*, **33**:234–255, 2016.

[6] M-Y. Cheng, J. Fan and J.S Marron. On automatic boundary corrections. *The Annals of Statistics*, **25**(4):1691–1708, 1997. https://doi.org/10.1214/aos/1031594737.

[7] M. Delecroix and A.C. Rosa. Nonparametric estimation of a regression function and its derivatives under an ergodic hypothesis. *Journal of Nonparametric Statistics*, **6**(4):367–382, 1996. https://doi.org/10.1080/10485259608832682.

[8] J. Demongeot, A. Naceri, A. Laksaci and M. Rachdi. Local linear regression modelization when all variables are curves. *Statistics & Probability Letters*, **121**:37–44, 2017. https://doi.org/10.1016/j.spl.2016.09.021.

[9] M. El Methni and M. Rachdi. Local weighted average estimation of the regression operator for functional data. *Communications in Statistics-Theory and Methods*, **40**(17):3141–3153, 2011. https://doi.org/10.1080/03610921003778209.

[10] J. Fan and I. Gijbels. *Local Polynomial Modelling and Its Applications*. Chapman & Hall/CRC, 1996.

[11] J. Fan and Q. Yao. *Nonlinear time series: nonparametric and parametric methods*. Springer, New York, 2003.

[12] F. Ferraty, A. Laksaci, A. Tadj and P. Vieu. Rate of uniform consistency for nonparametric estimates with functional variables. *Journal of Statistical Planning and Inference*, **140**(2):335–352, 2010. https://doi.org/10.1016/j.jspi.2009.07.019.

[13] F. Ferraty and Y. Romain. *The Oxford handbook of functional data analaysis*. Oxford University Press, 2011.

[14] F. Ferraty and P. Vieu. *Nonparametric functional data analysis: theory and practice*. Springer Science & Business Media, 2006. https://doi.org/10.1007/0-387-36620-2.

[15] L. Kara-Zaitri, A. Laksaci, M. Rachdi and P. Vieu. Uniform in bandwidth consistency for various kernel estimators involving functional data. *Journal of Nonparametric Statistics*, **29**(1):85–107, 2017.

[16] P. Kokoszka and M. Reimherr. *Introduction to functional data analysis.* CRC press, 2017. https://doi.org/10.1201/9781315117416.

[17] N. Laib and D. Louani. Rates of strong consistencies of the regression function estimator for functional stationary ergodic data. *Journal of Statistical Planning and Inference*, **141**(1):359–372, 2011. https://doi.org/10.1016/j.jspi.2010.06.009.

[18] N. Laïb and E. Ould-Saïd. A robust nonparametric estimation of the autoregression function under an ergodic hypothesis. *Canadian Journal of Statistics*, **28**(4):817–828, 2000. https://doi.org/10.2307/3315918.

[19] A. Laksaci, M. Rachdi and S. Rahmani. Spatial modelization: local linear estimation of the conditional distribution for functional data. *Spatial Statistics*, **6**:1–23, 2013. https://doi.org/10.1016/j.spasta.2013.04.004.

[20] N. Ling, S. Meng and P. Vieu. Uniform consistency rate of kNN regression estimation for functional time series data. *Journal of Nonparametric Statistics*, **31**(2):451–468, 2019. https://doi.org/10.1080/10485252.2019.1583338.

[21] Z. Zhou and Z. Lin. Asymptotic normality of locally modelled regression estimator for functional data. *Journal of Nonparametric Statistics*, **28**(1):116–131, 2016. https://doi.org/10.1080/10485252.2015.1114112.

# Appendix

## 4.1 Preliminary technical lemmas

Firstly, we state the following lemmas which are needed to establish our asymptotic results.

***Lemma. A.1*** *Under assumptions (H1), (H2), (H4)(i), (H5) and (H7), we obtain*

$$(i) \sup_{x\in\mathcal{C}_\mathcal{F}} \bar{F}_D(x) = O(1), \quad (ii) \inf_{x\in\mathcal{C}_\mathcal{F}} \bar{F}_D(x) = O(1).$$

***Lemma. A.2*** *Under assumptions (B1), (H2), (B2)–(B5), we have that*

$$(i)\ h_K\mathbb{E}\left(\rho_j K_j^a|\mathfrak{F}_{j-1}\right) = o(h_K^2\phi_j(h_K)) \quad \forall a>0,$$

$$(ii)\ \frac{1}{n\phi(h_K)}\sum_{j=1}^n \mathbb{E}\left(K_j^c|\mathfrak{F}_{j-1}\right) = M_c + o(1)\ \ for\ \ c=1,2, \quad (iii)\ \frac{1}{n\phi(h_K)}$$

$$\times\sum_{j=1}^n \mathbb{E}\left(\Gamma_j^2K_j^2|\mathfrak{F}_{j-1}\right) = (n-1)^2(N(1,2))^2h_K^4\phi^2(h_K)M_2 + o\left(h_K^4\phi^2(h_K)\right).$$

*Proof of Lemma A.1.* Before we start the proof of i), it is clear that by using Lemma A.1 of [3], we obtain

$$nCh_K^2\phi\left(h_K\right) \le \mathbb{E}\left(\Gamma_1(x)K_1(x)\right) \le nC'h_K^2\phi\left(h_K\right). \tag{4.1}$$

Then, by considering Lemma 5 of Ayad et al. [1] and (4.1) we get

$$\sup_{x\in\mathcal{C}_{\mathcal{F}}} \bar{F}_D(x) = O(1) \sup_{x\in\mathcal{C}_{\mathcal{F}}} \frac{1}{n\phi\left(h_K\right)} \sum_{j=1}^{n} \phi_j\left(h_K\right).$$

So, the claimed result i) of this lemma is a direct consequence of the assumption (H2)(ii). The proof of (ii) is similar to that of (i), and is therefore omitted. □

*Proof of Lemma A.2.* Firstly, the proof of (i) and (ii) are similar to the proof of (b) and (a) of Lemma A.1 in [21]. Secondly, in order to prove (iii), we use the definition of the conditional variance. Indeed,

$$\begin{aligned} &\frac{1}{n\phi(h_K)} \sum_{j=1}^{n} \mathbb{E}\left(\Gamma_j^2 K_j^2 | \mathfrak{F}_{j-1}\right) \\ &= \frac{1}{n\phi(h_K)} \sum_{j=1}^{n} \left(\operatorname{Var}\left(\Gamma_j K_j | \mathfrak{F}_{j-1}\right) + \left(\mathbb{E}\left(\Gamma_j K_j | \mathfrak{F}_{j-1}\right)\right)^2\right). \end{aligned} \tag{4.2}$$

It remains to study each term of (4.2). For the first term on the right hand side of this equation, we have

$$\begin{aligned} &\operatorname{Var}\left(\Gamma_j K_j | \mathfrak{F}_{j-1}\right) \\ &= (n-1)\left(\operatorname{Var}\left(\rho_1^2(x)K_1(x)K_j|\mathfrak{F}_{j-1}\right) + \operatorname{Var}\left(\rho_1(x)K_1(x)\rho_j K_j|\mathfrak{F}_{j-1}\right)\right) \\ &= (n-1)\Big(\underbrace{\mathbb{E}\left(\rho_1^4(x)K_1^2(x)\right)\mathbb{E}\left(K_j^2|\mathfrak{F}_{j-1}\right)}_{T_1} - \underbrace{\left(\mathbb{E}\left(\rho_1^2(x)K_1(x)\right)\mathbb{E}\left(K_j|\mathfrak{F}_{j-1}\right)\right)^2}_{T_2} \\ &\quad + \underbrace{\mathbb{E}\left(\rho_1^2(x)K_1^2(x)\right)\mathbb{E}\left(\rho_j^2 K_j^2|\mathfrak{F}_{j-1}\right)}_{T_3} - \underbrace{\left(\mathbb{E}\left(\rho_1(x)K_1(x)\right)\mathbb{E}\left(\rho_j K_j|\mathfrak{F}_{j-1}\right)\right)^2}_{T_4}\Big). \end{aligned}$$

Then, by using (i) of Lemma A.1 in [3] and (i) of Lemma 5 in [1], we find

$$\frac{n-1}{n\phi(h_K)} \sum_{j=1}^{n} T_i = O\left((n-1)h_K^4\phi(h_K)\right) \quad \text{for } i = 1,2,3,4.$$

It follows that

$$\frac{1}{n\phi(h_K)} \sum_{j=1}^{n} \operatorname{Var}\left(\Gamma_j K_j | \mathfrak{F}_{j-1}\right) \longrightarrow 0, \qquad \text{as } n \to \infty.$$

On the other side, to complete the proof of (iii), we have to study the first term on the right hand side of (4.2). For that, we write:

$$\begin{aligned} \frac{1}{n\phi(h_K)} \sum_{j=1}^{n} \left(\mathbb{E}\left(\Gamma_j K_j|\mathfrak{F}_{j-1}\right)\right)^2 &= \frac{1}{n\phi(h_K)} \sum_{j=1}^{n} \Big(\mathbb{E}\Big(\sum_{i=1}^{n} \rho_i^2(x)K_i(x)K_j \\ &\quad - \sum_{i=1}^{n} \rho_i(x)K_i(x)\rho_j K_j|\mathfrak{F}_{j-1}\Big)\Big)^2 = \gamma_{n1} + \gamma_{n2} + \gamma_{n3}, \end{aligned}$$

where

$$\gamma_{n1} = \frac{(n-1)^2}{n\phi(h_K)} \left(\mathbb{E}\left(\rho_1^2(x)K_1(x)\right)\right)^2 \sum_{j=1}^{n} \left(\mathbb{E}\left(K_j|\mathfrak{F}_{j-1}\right)\right)^2,$$

$$\gamma_{n2} = \frac{(n-1)^2}{n\phi(h_K)} \left(\mathbb{E}\left(\rho_1(x)K_1(x)\right)\right)^2 \sum_{j=1}^{n} \left(\mathbb{E}\left(\rho_j K_j|\mathfrak{F}_{j-1}\right)\right)^2, \quad \gamma_{n3} = -\frac{2\,(n-1)^2}{n\phi(h_K)}$$

$$\times\, \mathbb{E}\left(\rho_1^2(x)K_1(x)\right) \mathbb{E}\left(\rho_1(x)K_1(x)\right) \sum_{j=1}^{n} \mathbb{E}\left(K_j|\mathfrak{F}_{j-1}\right) \mathbb{E}\left(\rho_j K_j|\mathfrak{F}_{j-1}\right).$$

Concerning the term $\gamma_{n1}$, by applying Jensen's inequality, we have

$$\gamma_{n1} \leq \frac{(n-1)^2}{n\phi(h_K)} \left(\mathbb{E}\left(\rho_1^2(x)K_1(x)\right)\right)^2 \sum_{j=1}^{n} \mathbb{E}\left(K_j^2|\mathfrak{F}_{j-1}\right),$$

then, we apply (c) of Lemma A.1 in [21] and (ii) of Lemma A.2 to obtain:

$$\gamma_{n1} = (n-1)^2 \left((N(1,2))^2 h_K^4 \phi^2(h_K) M_2 + o\left(h_K^4 \phi^2(h_K)\right)\right). \tag{4.3}$$

Concerning $\gamma_{n2}$, we apply (b) of Lemma A.1 in [21] and (i) of Lemma A.2 to get:

$$\gamma_{n2} = o\left((n-1)^2 h_K^4 \phi(h_K)\right). \tag{4.4}$$

For the last term, we apply (i) of Lemma A.1 in [3], (i) of Lemma 5 in [1] and (i) of Lemma A.2 to get:

$$\gamma_{n3} = o\left((n-1)^2 h_K^5 \phi(h_K)\right). \tag{4.5}$$

Combining (4.3), (4.4) and (4.5) permits to obtain the claimed result. □

### 4.2 Proofs of main results

*Proof of Lemma* 1. We start by writing

$$\sup_{x\in\mathcal{C}_{\mathcal{F}}} |B_n(x,y)| = \sup_{x\in\mathcal{C}_{\mathcal{F}}} |\widetilde{B}_n(x,y)| \big/ \inf_{x\in\mathcal{C}_{\mathcal{F}}} |\bar{F}_D(x)|,$$

where $\widetilde{B}_n(x,y) = \bar{F}_N^x(y) - F^x(y)\,\bar{F}_D(x)$.

First, observe that $\widetilde{B}_n(x,y)$ can be written as

$$\widetilde{B}_n(x,y) = \frac{1}{n\mathbb{E}\left(\varGamma_1 K_1\right)} \sum_{j=1}^{n} \left\{\mathbb{E}\left(\varGamma_j K_j J_j|\mathfrak{F}_{j-1}\right) - F^x(y)\,\mathbb{E}\left(\varGamma_j K_j|\mathfrak{F}_{j-1}\right)\right\}$$

$$= \frac{1}{n\mathbb{E}\left(\varGamma_1 K_1\right)} \sum_{j=1}^{n} \left\{\mathbb{E}\left(\varGamma_j K_j \mathbb{E}\left(J_j|\mathcal{G}_{j-1}\right)|\mathfrak{F}_{j-1}\right) - F^x(y)\,\mathbb{E}\left(\varGamma_j K_j|\mathfrak{F}_{j-1}\right)\right\}$$

$$\leq \frac{1}{n\mathbb{E}\left(\varGamma_1 K_1\right)} \sum_{j=1}^{n} \left\{\mathbb{E}\left(\varGamma_j K_j \left|\mathbb{E}\left[J_j|X_j\right] - F^x(y)\right| |\mathfrak{F}_{j-1}\right)\right\}. \tag{4.6}$$

The last inequality is obtained by using (H5) (iii).

Next, we have directly after integrating by parts, and changing of variables

$$|\mathbb{E}\left[J_j|X_j\right] - F^x\left(y\right)| \leq \int_{\mathbb{R}} J^{(1)}\left(t\right)|F^x\left(y-h_J t\right) - F^x\left(y\right)| dt.$$

Thus, from assumptions (H3) and (H5)(i) we get:

$$1\!\!1_{B(x,h_k)}\left(X_j\right)|\mathbb{E}\left[J_j|X_j\right] - F^x\left(y\right)| \leq \int_{\mathbb{R}} J^{(1)}\left(t\right)\left(h_K^{b_1} + |t|^{b_2} h_J^{b_2}\right) dt.$$

Since, $J^{(1)}$ is probability density function, and under the assumption (H5) (ii), we find that:

$$1\!\!1_{B(x,h_k)}\left(X_j\right)|\mathbb{E}\left[J_j|X_j\right] - F^x\left(y\right)| \leq C\left(h_K^{b_1} + h_J^{b_2}\right). \tag{4.7}$$

Hence, by combining (4.6) together with (i) of Lemma A.1, we obtain

$$\sup_{x\in\mathcal{C}_{\mathcal{F}}} |\widetilde{B}_n\left(x,y\right)| \quad = \quad O\left(h_K^{b_1} + h_J^{b_2}\right) \sup_{x\in\mathcal{C}_{\mathcal{F}}} \bar{F}_D\left(x\right),$$

and the claimed result of this lemma is now checked. □

*Proof of Lemma* 2. First, by using (H4) (i) and because the kernel $K$ is bounded on $[-1,1]$, it can be easily seen that

$$|\Gamma_j(x)| \quad \leq \quad nCh_K^2 + nCh_K|\rho_j(x)|. \tag{4.8}$$

Second, for all $x \in \mathcal{C}_{\mathcal{F}}$, we denote: $k(x) = arg \min_{k\in\{1,2,\ldots,d_n\}} |\delta(x,x_k)|$

$$\sup_{x\in\mathcal{C}_{\mathcal{F}}} |R_n(x,y)| \leq \underbrace{\sup_{x\in\mathcal{C}_{\mathcal{F}}} |\widehat{F}_N^x(y) - \widehat{F}_N^{x_{k(x)}}(y)|}_{\mathcal{Q}_1} + \underbrace{\sup_{x\in\mathcal{C}_{\mathcal{F}}} |\widehat{F}_N^{x_{k(x)}}(y) - \bar{F}_N^{x_{k(x)}}(y)|}_{\mathcal{Q}_2}$$
$$+ \underbrace{\sup_{x\in\mathcal{C}_{\mathcal{F}}} |\bar{F}_N^{x_{k(x)}}(y) - \bar{F}_N^x(y)|}_{\mathcal{Q}_3}.$$

We will now treat each of the three terms involved in this decomposition. We start by the consistency of the term $\mathcal{Q}_1$. By using (4.1) and the boundeness on $K$ and $J$, one can write:

$$\mathcal{Q}_1 \leq \sup_{x\in\mathcal{C}_{\mathcal{F}}} \frac{1}{n}\sum_{j=1}^n |J_j(y)|\,|\frac{1}{\mathbb{E}\left(\Gamma_1(x)K_1(x)\right)}\Gamma_j(x)K_j(x)1\!\!1_{B(x,h_K)}(X_j)$$
$$- \frac{1}{\mathbb{E}\left(\Gamma_1(x_{k(x)})K_1(x_{k(x)})\right)}\Gamma_j(x_{k(x)})K_j(x_{k(x)})1\!\!1_{B(x_{k(x)},h_K)}(X_j)|$$
$$\leq \Big(\frac{C}{n^2h_K^2\phi(h_K)}\sup_{x\in\mathcal{C}_{\mathcal{F}}}\sum_{j=1}^n |\Gamma_j(x)1\!\!1_{B(x,h_K)}(X_j)|$$

$$\times |K_j(x) - K_j(x_{k(x)})1\!\!1_{B(x_{k(x)},h_K)}(X_j)|\Big)$$
$$+ \Big(\frac{C}{n^2h_K^2\phi(h_K)} \sup_{x\in\mathcal{C}_\mathcal{F}} \sum_{j=1}^{n} K_j(x_{k(x)})1\!\!1_{B(x_{k(x)},h_K)}(X_j)$$
$$\times |\Gamma_j(x)1\!\!1_{B(x,h_K)}(X_j) - \Gamma_j(x_{k(x)})|\Big) := F_1 + F_2.$$

Let us first deal with the term $F_1$. Because the kernel $K$ satisfy the Lipschitz condition, and by using the inequality (4.8), we have

$$|\Gamma_j(x)|1\!\!1_{B(x,h_K)}(X_j)|K_j(x) - K_j(x_{k(x)})1\!\!1_{B(x_{k(x)},h_K)}(X_j)|$$
$$\leq nCh_K^2 \left(\frac{r_n}{h_K} 1\!\!1_{B(x,h_K)\cap B(x_{k(x)},h_K)}(X_j) + 1\!\!1_{B(x,h_K)\cap \overline{B(x_{k(x)},h_K)}}(X_j)\right),$$

which implies that:

$$F_1 \leq \frac{Cr_n}{nh_K\phi(h_K)} \sup_{x\in\mathcal{C}_\mathcal{F}} \sum_{j=1}^{n} 1\!\!1_{B(x,h_K)\cap B(x_{k(x)},h_K)}(X_j)$$
$$+ \frac{C}{n\phi(h_K)} \sup_{x\in\mathcal{C}_\mathcal{F}} \sum_{j=1}^{n} 1\!\!1_{B(x,h_K)\cap \overline{B(x_{k(x)},h_K)}}(X_j).$$

Concerning the term $F_2$, we have that:

$$1\!\!1_{B(x_{k(x)},h_K)}(X_j)|\Gamma_j(x)1\!\!1_{B(x,h_K)}(X_j) - \Gamma_j(x_{k(x)})|$$
$$\leq \underbrace{1\!\!1_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j)|\Gamma_j(x) - \Gamma_j(x_{k(x)})|}_{A}$$
$$+ \underbrace{nCh_K^2 1\!\!1_{B(x_{k(x)},h_K)\cap \overline{B(x,h_K)}}(X_j)}_{B}.$$

Now, we calculate the first part of the right side of this inequality

$$A = 1\!\!1_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j)\Big|\Big(\sum_{i=1}^{n} \rho_i^2(x)K_i(x) - \rho_i^2(x_{k(x)})K_i(x_{k(x)})\Big)$$
$$- \Big(\big(\sum_{i=1}^{n} \rho_i(x)K_i(x)\big)\rho_j(x)\big) - \Big(\sum_{i=1}^{n} \rho_i(x_{k(x)})K_i(x_{k(x)})\rho_j(x_{k(x)})\Big)\Big| \leq A_1 + A_2,$$

where

$$A_1 = 1\!\!1_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j)\Big|\sum_{i=1}^{n} \rho_i^2(x)K_i(x) - \rho_i^2(x_{k(x)})K_i(x_{k(x)})\Big|,$$
$$A_2 = 11_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j)\Big|\Big(\sum_{i=1}^{n} \rho_i(x)K_i(x)\Big)\rho_j(x)$$
$$- \Big(\sum_{i=1}^{n} \rho_i(x_{k(x)})K_i(x_{k(x)})\Big)\rho_j(x_{k(x)})\Big|$$

Let us now examine the terms $A_1$ and $A_2$ by putting

$$T^{k,l} = 1\!\!1_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j)\Big|\Big(\sum_{i=1}^{n}\rho_i^k(x)K_i(x)\Big)\rho_j^l(x) - \Big(\sum_{i=1}^{n}\rho_i^k(x_{k(x)})K_i(x_{k(x)})\Big)\rho_j^l(x_{k(x)})\Big| \quad \text{with } k=1,2 \text{ and } l=0,1.$$

Therefore,

$$T^{k,l} \leq T_1^{k,l} + T_2^{k,l}$$

with

$$T_1^{k,l} = 1\!\!1_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j)\Big(\sum_{i=1}^{n}|\rho_i^k(x)|K_i(x)\times|\rho_j^l(x)-\rho_j^l(x_{k(x)})|\Big),$$

$$T_2^{k,l} = 1\!\!1_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j)\Big(|\rho_j^l(x_{k(x)})| \times |\sum_{i=1}^{n}(\rho_i^k(x)K_i(x)-\rho_i^k(x_{k(x)})K_i(x_{k(x)})|\Big).$$

By the assumption (H4)(ii) for $l=1$, we can write:

$$1\!\!1_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j)|\rho_j(x)-\rho_j(x_{k(x)})|\leq Cr_n 1\!\!1_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j).$$

So, for $l=0, k=2$

$$T_1^{k,l} = 0, \tag{4.9}$$

and for $l=1, k=1$

$$T_1^{k,l} \leq nCr_nh_K 1\!\!1_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j). \tag{4.10}$$

We now turn to the term $T_2^{k,l}$

$$T_2^{k,l} \leq 1\!\!1_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j)\left(\sum_{i=1}^{n}|\rho_j^l(x_{k(x)})|K_i(x)\times|\rho_i^k(x)-\rho_i^k(x_{k(x)})|\right) + 1\!\!1_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j)\left(\sum_{i=1}^{n}|\rho_j^l(x_{k(x)})||\rho_i^k(x_{k(x)})|\,\big|K_i(x)-K_i(x_{k(x)})\big|\right).$$

Observe that:

$$1\!\!1_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j)|\rho_i^2(x)-\rho_i^2(x_{k(x)})| \leq Cr_nh_K 1\!\!1_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j),$$

which implies that for $k=1,2$

$$1\!\!1_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j)|\rho_i^k(x)-\rho_i^k(x_{k(x)})| \leq Cr_nh_K^{k-1} 1\!\!1_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j).$$

Therefore, for $l = 0$, and $k = 2$

$$T_2^{k,l} \leq nCr_n h_K 1\!\!1_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j), \tag{4.11}$$

and for $l = 1$, and $k = 1$

$$T_2^{k,l} \leq nCr_n h_K 1\!\!1_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j). \tag{4.12}$$

Then, by combining (4.9) with (4.11), we find that

$$A_1 \leq nCr_n h_K 1\!\!1_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j).$$

In addition, by combining (4.10) with (4.12), allows us to find

$$A_2 \leq nCr_n h_K 1\!\!1_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j),$$

which implies that

$$A \leq nCr_n h_K 1\!\!1_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j).$$

Thus,

$$F_2 \leq \frac{Cr_n}{nh_K\phi(h_K)} \sup_{x\in\mathcal{C}_\mathcal{F}} \sum_{j=1}^{n} 1\!\!1_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j) + \frac{C}{n\phi(h_K)} \sup_{x\in\mathcal{C}_\mathcal{F}} \sum_{j=1}^{n} 1\!\!1_{B(x_{k(x)},h_K)\cap \overline{B(x,h_K)}}(X_j).$$

Consequently, we obtain

$$\mathcal{Q}_1 \leq C \sup_{x\in\mathcal{C}_\mathcal{F}} \left(\mathcal{Q}_{1.1} + \mathcal{Q}_{1.2} + \mathcal{Q}_{1.3}\right),$$

where

$$\mathcal{Q}_{1.1} = \frac{C}{n\phi(h_K)} \sum_{j=1}^{n} 1\!\!1_{B(x_{k(x)},h_K)\cap \overline{B(x,h_K)}}(X_j),$$

$$\mathcal{Q}_{1.2} = \frac{Cr_n}{nh_K\phi(h_K)} \sum_{j=1}^{n} 1\!\!1_{B(x,h_K)\cap B(x_{k(x)},h_K)}(X_j),$$

$$\mathcal{Q}_{1.3} = \frac{C}{n\phi(h_K)} \sum_{j=1}^{n} 1\!\!1_{B(x,h_K)\cap \overline{B(x_{k(x)},h_K)}}(X_j).$$

Now, we evaluate those last terms by an application of the standard inequality for sums of bounded random variables with $Z_j$ identified such that:

$$Z_j = \begin{cases} \dfrac{1}{\phi(h_K)}[1\!\!1_{B(x_{k(x)},h_K)\cap \overline{B(x,h_K)}}(X_j)] & \text{for } \mathcal{Q}_{1.1}, \\ \dfrac{r_n}{h_K\phi(h_K)}[1\!\!1_{B(x,h_K)\cap B(x_{k(x)},h_K)}(X_j)] & \text{for } \mathcal{Q}_{1.2}, \\ \dfrac{1}{\phi(h_K)}[1\!\!1_{B(x,h_K)\cap \overline{B(x_{k(x)},h_K)}}(X_j)] & \text{for } \mathcal{Q}_{1.3}. \end{cases}$$

It is clear that for $\mathcal{Q}_{1.1}$ and $\mathcal{Q}_{1.3}$, we have under the second part of (H1):

$$Z_j = O\left(\frac{1}{\phi(h_K)}\right), \quad \mathbb{E}[Z_j] = O\left(\frac{r_n}{\phi(h_K)}\right), \quad \mathbb{E}\left(Z_j^2\right) = O\left(\frac{r_n}{\phi(h_K)^2}\right).$$

Therefore,

$$\mathcal{Q}_{1.1} = O\left(\frac{r_n}{\phi(h_K)}\right) + O_{a.co}\left(\sqrt{r_n \log n / n\phi(h_K)^2}\right).$$

With the same manner, the assumption (H6) allows to get, for $\mathcal{Q}_{1.2}$:

$$Z_j = O\left(\frac{r_n}{h_K\phi(h_K)}\right), \quad \mathbb{E}[Z_j] = O\left(\frac{r_n}{h_K}\right), \quad \mathbb{E}\left(Z_j^2\right) = O\left(\frac{r_n^2}{h_K^2\phi(h_K}\right),$$

which implies that:

$$\mathcal{Q}_{1.2} = O_{a.co}\left(\sqrt{\log d_n / n\phi(h_K)}\right).$$

To finish the study of the term $\mathcal{Q}_1$, we need to put together all the intermediate result and to employ the second part of (H6) to obtain

$$\mathcal{Q}_1 = O_{a.co}\left(\sqrt{\log d_n / n\phi(h_K)}\right).$$

Concerning the term $\mathcal{Q}_2$, we have for all $\varepsilon > 0$,

$$\begin{aligned}
\mathbb{P}\left(\mathcal{Q}_2 > \varepsilon\sqrt{\log d_n / n\phi(h_K)}\right) &= \mathbb{P}\left(\max_{k\in 1,\dots,d_n} |\widehat{F}_N^{x_{k(x)}}(y) - \bar{F}_N^{x_{k(x)}}(y)| > \varepsilon\right) \\
&\leq d_n \max_{k\in 1,\dots,d_n} \mathbb{P}\left(|\widehat{F}_N^{x_{k(x)}}(y) - \bar{F}_N^{x_{k(x)}}(y)| > \varepsilon\sqrt{\log d_n / n\phi(h_K)}\right).
\end{aligned}$$

Let

$$\widehat{F}_N^{x_{k(x)}}(y) - \bar{F}_N^{x_{k(x)}}(y) = \frac{1}{\mathbb{E}(\Gamma_1 K_1)} \sum_{j=1}^{n} S_j$$

with

$$S_j = \Gamma_j(x_{k(x)})K_j(x_{k(x)})J_j(y) - \mathbb{E}\left(\Gamma_j(x_{k(x)})K_j(x_{k(x)})J_j(y)|\mathfrak{F}_{j-1}\right),$$

where $S_j$ is a triangular array of bounded martingale differences with respect to the sequence of $\sigma$-fields $\left(\mathfrak{F}_{j-1}\right)_{j\geq 1}$. So, we have

$$\begin{aligned}
\mathbb{E}\left(S_j^2|\mathfrak{F}_{j-1}\right) &= \mathbb{E}\left(\left(\Gamma_j K_j\right)^2 J_j^2|\mathfrak{F}_{j-1}\right) - \mathbb{E}\left(\left(\Gamma_j K_j J_j|\mathfrak{F}_{j-1}\right)\right)^2 \\
&\leq \mathbb{E}\left(\left(\Gamma_j K_j\right)^2 J_j^2|\mathfrak{F}_{j-1}\right).
\end{aligned}$$

As $J_j \leq 1$, we deduce that

$$\mathbb{E}\left(S_j^2|\mathfrak{F}_{j-1}\right) \leq \mathbb{E}\left(\Gamma_j^2 K_j^2|\mathfrak{F}_{j-1}\right).$$

By using Equation (4.8), (H4)(i) and (H5)(i), we obtain that:

$$\mathbb{E}\left(S_j^2|\mathfrak{F}_{j-1}\right) \le 2Cn^2h_K^4\phi_j\left(h_K\right).$$

Now, we use the exponential inequality of Lemma 1 of [17] (with $d_j^2 = Cn^2h_K^4\phi_j(h_K)$) to obtain for all $\varepsilon > 0$,

$$\begin{aligned}&\mathbb{P}\left(\left|\widehat{F}_N^{x_{k(x)}}(y) - \bar{F}_N^{x_{k(x)}}(y)\right| > \varepsilon\sqrt{\log d_n/n\phi(h_K)}\right)\\ &\quad = \mathbb{P}\Big(\Big|\frac{1}{n\mathbb{E}\left(\Gamma_1K_1\right)}\sum_{j=1}^n S_j\Big| > \varepsilon\sqrt{\log d_n/n\phi(h_K)}\Big) \le 2\exp\left\{-C\varepsilon^2\log d_n\right\}.\end{aligned}$$

Thus, by choosing $\varepsilon$ such that $C\varepsilon^2 = \varsigma$, we get

$$d_n\max_{k\in\{1,\ldots,d_n\}}\mathbb{P}\left(|\widehat{F}_N^{x_{k(x)}}(y) - \bar{F}_N^{x_{k(x)}}(y)| > \varepsilon\sqrt{\log d_n/n\phi(h_K)}\right) \le Cd_n^{1-\varsigma}.$$

Since $\sum_{n=1}^\infty d_n^{1-\varsigma} < \infty$, we obtain that:

$$\mathcal{Q}_2 = O_{a.co}\left(\sqrt{\log d_n/n\phi(h_K)}\right).$$

For the term $\mathcal{Q}_3$, clearly we have

$$\mathcal{Q}_3 \le \mathbb{E}\left(\sup_{x\in\mathcal{C}_\mathcal{F}}|\widehat{F}_N^{x_{k(x)}}(y) - \bar{F}_N^{x_{k(x)}}(y)||\mathfrak{F}_{j-1}\right).$$

Subsequently, we follow the same steps used in studying the term $\mathcal{Q}_1$ to find

$$\mathcal{Q}_3 = O_{a.co}\left(\sqrt{\log d_n/n\phi(h_K)}\right).$$

This is enough to complete the proof of Lemma 2. □

*Proof of Lemma* 3.

i) This result can be deduced from Lemma 2 by taking $J_j = 1$. In this case, (H5) (ii) and (iii) are not necessary.

ii) It is easy to see that $\inf_{x\in\mathcal{C}_\mathcal{F}}|\widehat{F}_D(x)| \le \frac{1}{2}$ implies that there exist $x \in \mathcal{C}_\mathcal{F}$ such that

$$1 - \widehat{F}_D(x) \ge \frac{1}{2} \Longrightarrow \sup_{x\in\mathcal{C}_\mathcal{F}}|1 - \widehat{F}_D(x)| \ge \frac{1}{2}.$$

According to (i) of this lemma, we have:

$$\mathbb{P}\left(\inf_{x\in\mathcal{C}_\mathcal{F}}|\widehat{F}_D(x)| \le 0.5\right) \le \mathbb{P}\left(\sup_{x\in\mathcal{C}_\mathcal{F}}|1 - \widehat{F}_D(x)| \ge 0.5\right).$$

Consequently,

$$\sum_{n=1}^\infty\mathbb{P}\left(\inf_{x\in\mathcal{C}_\mathcal{F}}|\widehat{F}_D(x)| \le \frac{1}{2}\right) < \infty,$$

which ends the proof. □

*Proof of Lemma* 4. For all $j = 1, \ldots, n$, let us denote

$$\eta_{n,j} = \frac{\sqrt{n\phi\left(h_K\right)}}{n\mathbb{E}\left(\Gamma_1 K_1\right)}\left(J_j - F^x(y)\right)\Gamma_j K_j,$$

and define $\xi_{n,j} = \eta_{n,j} - \mathbb{E}\left(\eta_{n,j}|\mathfrak{F}_{j-1}\right)$. It is clear that

$$\sqrt{n\phi\left(h_K\right)}Q_n(x,y) = \sum_{j=1}^{n}\xi_{n,j}. \tag{4.13}$$

The summands in Equation (4.13) form a triangular array of stationary martingale differences with respect to the $\sigma$-fields $\mathfrak{F}_{j-1}$. Accordingly, the asymptotic normality of $Q_n(x,y)$ can be established by applying the central limit theorem for discrete time arrays of real-valued martingales. Therefore, to show Lemma 4, it suffices to prove the following two claims:

$$\sum_{j=1}^{n}\mathbb{E}\left(\xi_{n,j}^2|\mathfrak{F}_{j-1}\right) \xrightarrow{\mathbb{P}} V_{JK}(x,y), \tag{4.14}$$

$$\forall\varepsilon > 0 \;\; n\mathbb{E}\left(\xi_{n,j}^2 1\!\!1_{[|\xi_{n,j}|>\varepsilon]}\right) = o(1) \quad \text{(Lindeberg condition).} \tag{4.15}$$

Let us start the proof of (4.14) by remarking that

$$\mathbb{E}\left(\xi_{n,j}^2|\mathfrak{F}_{j-1}\right) = \mathbb{E}\left(\eta_{n,j}^2|\mathfrak{F}_{j-1}\right) - \left(\mathbb{E}\left(\eta_{n,j}|\mathfrak{F}_{j-1}\right)\right)^2.$$

Thus, it remains to check that

$$\lim_{n\to\infty}\sum_{j=1}^{n}\Big(\mathbb{E}\left(\eta_{n,j}|\mathfrak{F}_{j-1}\right)\Big)^2 = 0 \quad \text{in probability,} \tag{4.16}$$

$$\lim_{n\to\infty}\sum_{j=1}^{n}\mathbb{E}\left(\eta_{n,j}^2|\mathfrak{F}_{j-1}\right) = V_{JK}(x,y) \quad \text{in probability.} \tag{4.17}$$

Concerning the proof of (4.16), by applying Equation (4.7), Equation (4.1) and Lemma 5 of Ayad et *al.* [1], we get that

$$\begin{aligned}
\left|\mathbb{E}\left(\eta_{n,j}|\mathfrak{F}_{j-1}\right)\right| &= \frac{\sqrt{n\phi\left(h_K\right)}}{n\mathbb{E}\left(\Gamma_1 K_1\right)}\left|\mathbb{E}\Big(\left(J_j - F^x(y)\right)\Gamma_j K_j|\mathfrak{F}_{j-1}\Big)\right| \\
&\leq C\sqrt{n\phi\left(h_K\right)}\left(h_K^{b_1} + h_J^{b_2}\right)\frac{1}{n\phi\left(h_K\right)}\phi_j\left(h_K\right).
\end{aligned}$$

Thus, by using (H2) (ii), we find

$$\sum_{j=1}^{n}\Big(\mathbb{E}\left(\eta_{n,j}|\mathfrak{F}_{j-1}\right)\Big)^2 = O_{a.co}\Big(n\phi\left(h_K\right)\left(h_K^{b_1} + h_J^{b_2}\right)^2\Big).$$

For the proof of Equation (4.17), we use (H5)(iii) to obtain

$$\sum_{j=1}^{n}\mathbb{E}\left(\eta_{n,j}^2|\mathfrak{F}_{j-1}\right)=\frac{\phi\left(h_K\right)}{n\left(\mathbb{E}\left(\Gamma_1K_1\right)\right)^2}\sum_{j=1}^{n}\mathbb{E}\left(\Gamma_j^2K_j^2\left(J_j-F^x(y)\right)^2|\mathfrak{F}_{j-1}\right)$$
$$=\frac{\phi\left(h_K\right)}{n\left(\mathbb{E}\left(\Gamma_1K_1\right)\right)^2}\sum_{j=1}^{n}\mathbb{E}\left(\Gamma_j^2K_j^2\mathbb{E}\left[\left(J_j-F^x(y)\right)^2|X_j\right]|\mathfrak{F}_{j-1}\right).$$

Next, by using the definition of the conditional variance, we find

$$\mathbb{E}\left[\left(J_j-F^x(y)\right)^2|X_j\right]=\mathrm{Var}\left[J_j|X_j\right]+\left[\mathbb{E}\left(J_j|X_j\right)-F^x(y)\right]^2:=\beta_{n1}+\beta_{n2}.$$

Concerning the term $\beta_{n1}$, we have

$$\mathrm{Var}\left[J_j|X_j\right]=\mathbb{E}\left(J_j^2|X_j\right)-\left(\mathbb{E}\left(J_j|X_j\right)\right)^2. \tag{4.18}$$

An integration by part followed by a change of variable with the assumption (B2) permit us to deduce

$$\mathbb{E}\left[J_j|X_j\right]=\int_{\mathbb{R}}J^{(1)}\left(t\right)\left[F^x\left(y-h_Jt\right)-F^x\left(y\right)\right]dt+F^x\left(y\right)=F^x\left(y\right). \tag{4.19}$$

Similarly, the first term on the right hand side of Equation (4.18) is treated directly by using again (B2) combined with an integration by part and a change of variable. It follows that

$$\mathbb{E}\left[J_j^2|X_j\right]=\int_{\mathbb{R}}J^2\left((y-z)/h_J\right)f^x(z)dz$$
$$=\int_{\mathbb{R}}2J\left(t\right)J^{(1)}\left(t\right)\left[F^x\left(y-h_Jt\right)-F^x\left(y\right)\right]dt+\int_{\mathbb{R}}2J\left(t\right)J^{(1)}\left(t\right)F^x\left(y\right)dt.$$

Since $\int_{\mathbb{R}}2J\left(t\right)J^{(1)}\left(t\right)F^x\left(y\right)dt=F^x\left(y\right)$, we infer that:

$$\mathbb{E}\left[J_j^2|X_j\right]\longrightarrow F^x\left(y\right),\quad \text{as } n\to\infty. \tag{4.20}$$

Now, by combining the result (4.19) with (4.20), we arrive directly at the following result:

$$\mathrm{Var}\left[J_j|X_j\right]=F^x\left(y\right)\left(1-F^x\left(y\right)\right). \tag{4.21}$$

Concerning the term $\beta_{n2}$, we deduce by (4.19) that $\beta_{n2}\longrightarrow 0$, as $n\to\infty$. Therefore,

$$\sum_{j=1}^{n}\mathbb{E}\left(\eta_{n,j}^2|\mathfrak{F}_{j-1}\right)=\frac{\phi\left(h_K\right)}{n\left(\mathbb{E}\left(\Gamma_1K_1\right)\right)^2}\sum_{j=1}^{n}\mathbb{E}\left(\Gamma_j^2K_j^2\beta_{n1}|\mathfrak{F}_{j-1}\right).$$

Combining (d) of Lemma A.1 [21], (iii) of Lemma A.2 and Equation (4.21) allow to obtain

$$\sum_{j=1}^{n}\mathbb{E}\left(\eta_{n,j}^2|\mathfrak{F}_{j-1}\right)\longrightarrow\frac{M_2}{M_1^2}F^x\left(y\right)\left(1-F^x\left(y\right)\right)=V_{JK}(x,y),\qquad \text{as } n\to\infty,$$

which completes the proof of the claim (4.14).

Concerning the proof of (4.15), the Lindeberg condition implies that

$$n\mathbb{E}\left(\xi_{n,j}^2 1\!\!1_{[|\xi_{n,j}|>\varepsilon]}\right) \leq 4n\mathbb{E}\left(\eta_{n,j}^2 1\!\!1_{[|\eta_{n,j}|>\frac{\varepsilon}{2}]}\right).$$

By using Markov's and Hölder's inequalities, we can write for all $\varepsilon > 0$,

$$\mathbb{E}\left(\eta_{n,j}^2 1\!\!1_{[|\eta_{n,j}|>\frac{\varepsilon}{2}]}\right) \leq \frac{\mathbb{E}\left(|\eta_{n,j}|\right)^{2a}}{(\varepsilon/2)^{2a/b}}.$$

Taking $a = 1 + \frac{\delta}{2}$ for any $\delta > 0$, such that $\bar{G}_{2+\delta} = \mathbb{E}\left(|J_j - F^x(y)|^{2+\delta}|X_j\right)$ is a continuous function. It follows that

$$\begin{aligned}
&4n\mathbb{E}\left(\eta_{n,j}^2 1\!\!1_{[|\eta_{n,j}|>\frac{\varepsilon}{2}]}\right) \\
&\leq C\left(\frac{\phi(h_K)}{n}\right)^{\frac{2+\delta}{2}} \frac{n}{\left(\mathbb{E}\left(\Gamma_1 K_1\right)\right)^{2+\delta}}\mathbb{E}\left([|\left(J_j - F^x(y)\right)|\Gamma_j K_j]^{2+\delta}\right) \\
&\leq C\left(\frac{\phi(h_K)}{n}\right)^{\frac{2+\delta}{2}} \frac{n}{\left(\mathbb{E}\left(\Gamma_1 K_1\right)\right)^{2+\delta}}\mathbb{E}\left(|\Gamma_j K_j|^{2+\delta}\left[\mathbb{E}\left(|J_j - F^x(y))\right|^{2+\delta}|X_j\right]\right) \\
&\leq C\left(\frac{\phi(h_K}{n}\right)^{\frac{2+\delta}{2}} \frac{n}{\left(\mathbb{E}\left(\Gamma_1 K_1\right)\right)^{2+\delta}}\mathbb{E}\left(|\Gamma_j K_j|^{2+\delta}\right)\bar{G}_{2+\delta} \\
&= O\left((n\phi(h_K))^{\frac{-\delta}{2}}\right) \quad \text{as } n \to \infty,
\end{aligned}$$

and the claimed result is checked. □

*Proof of Lemma* 5. Observe that

$$\widehat{F}_D(x) - 1 = \underbrace{\widehat{F}_D(x) - \bar{F}_D(x)}_{I_1} + \underbrace{\bar{F}_D(x) - 1}_{I_2}.$$

Since $I_2 \longrightarrow 0$ almost completely as $n \to \infty$ in view of (H2) (ii), it suffices to show that $I_1 = o(1)$ as $n \to \infty$ Indeed, by using Lemma 3 (i) we obtain

$$\widehat{F}_D(x) - \bar{F}_D(x) = o(1)\,\text{almost completely as}\,n \to \infty.$$

Which completes the proof of Lemma 5. □