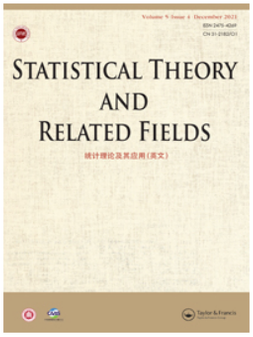

Statistical Theory and Related Fields

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tstf20

Asymptotic properties of a nonparametric conditional density estimator in the local linear estimation for functional data via a functional single-index model

Fadila Benaissa, Abdelmalek Gagui & Abdelhak Chouaf

To cite this article: Fadila Benaissa, Abdelmalek Gagui & Abdelhak Chouaf (2021): Asymptotic properties of a nonparametric conditional density estimator in the local linear estimation for functional data via a functional single-index model, Statistical Theory and Related Fields, DOI: 10.1080/24754269.2021.1965945

To link to this article: https://doi.org/10.1080/24754269.2021.1965945

Published online: 02 Sep 2021.

STATISTICAL THEORY AND RELATED FIELDS
https://doi.org/10.1080/24754269.2021.1965945

Taylor & Francis
Taylor & Francis Group

Asymptotic properties of a nonparametric conditional density estimator in the local linear estimation for functional data via a functional single-index model

Fadila Benaissa[a], Abdelmalek Gagui[b] and Abdelhak Chouaf[b]

[a]University Oran 1 Ahmed Ben Bella Oran, Oran, Algeria; [b]Laboratory of Statistics and Stochastic Processes, Djillali Liabes University, Sidi Bel Abbes, Algeria

ABSTRACT

This paper deals with the conditional density estimator of a real response variable given a functional random variable (i.e., takes values in an infinite-dimensional space). Specifically, we focus on the functional index model, this approach represents a good compromise between nonparametric and parametric models. Then we give under general conditions and when the variables are independent, the quadratic error and asymptotic normality of estimator by local linear method, based on the single-index structure. Finally, we complete these theoretical advances by some simulation studies showing both the practical result of the local linear method and the good behaviour for finite sample sizes of the estimator and of the Monte Carlo methods to create functional pseudo-confidence area.

ARTICLE HISTORY
Received 22 November 2020
Revised 14 January 2021
Accepted 5 July 2021

KEYWORDS
Mean squared error; single functional index; conditional density function; nonparametric estimation; local linear estimation; asymptotic normality; functional data

2000 MATHEMATICS SUBJECT CLASSIFICATIONS
62G05; 62G07; 62G20; 62G15

1. Introduction

The nonparametric estimation of the conditional density function plays a crucial role in statistical analysis. This subject can be approached from multiple perspectives depending on the complexity of the problem. Many techniques were studied in the literature to treat these various situations but all treat only real or multidimensional explanatory random variables.

Focusing on functional data for the kernel-type, the first results on the nonparametric estimate of this model, were got by Ferraty and Vieu (2006). They have studied the almost complete convergence the estimator of the conditional density and its derivates. Laksaci (2007) they studied quadratic error of this estimator, we return to Ferraty et al. (2010) established the uniform almost complete convergence of this model always.

Now, we show a few results on the local linear smoothing for functional data, actually this results considered by many authors recall them, Baìllo and Grané (2009) first proposed a local linear smoothing of the regression estimator in a Hilbert space, and came after them Barrientos-Marin et al. (2010) developed this method of local linear estimation of the regression in the semi-metric space for independent and identically distributed, Demongeot et al. (2013, 2014), has used this method to estimate conditional distribution and density function, in the case of spatial data (Laksaci et al., 2013) they established pointwise almost complete convergence with rate.

Furthermore, the functional index model plays a major role in statistics. The interest of this approach comes from its use to reduce the dimension of the data by projection in fractal space. The literature on this topic is closely limited, the first work which was interested in the single-index model on the nonparametric estimation is Ferraty et al. (2003) they stated for i.i.d. variables and obtained the almost complete convergence under some conditions. Based on the cross-validation procedure, Ait Saidi et al. (2008) proposed an estimator of this parameter, where the functional single-index is unknown. Recently, Attaoui et al. (2011) considered the nonparametric estimation of the conditional density in the single functional model. They established its pointwise and uniform almost complete convergence (a.co.) rates. In the same topic, Attaoui and Ling (2016) proved the asymptotic results of a nonparametric conditional cumulative distribution estimator for time series data. More recently, Tabti and Ait Saidi (2018) obtained the almost complete convergence and the uniform almost complete convergence of a kernel estimator of the hazard function with quasi-association condition when the observations are linked with functional single-index structure.

In this paper, we focus on the local linear estimation with the single-index structure to compute under some

CONTACT Fadila Benaissa fadila-@hotmail.com University Oran 1 Ahmed Ben Bella Oran, Oran 31000, Algeria

conditions, the quadratic error of the conditional density function estimator. In practice, this study has great importance, because, it permits to construct a prediction method based on the maximum risk estimation with a single functional index.

In Section 2, We introduce the estimator of our model in the single functional index. In Section 3, we introduce assumptions and asymptotic properties are given. Simulations are given in Section 4. Finally, Section 5 is devoted to the proofs of the results.

2. The model

Let $\{(X_i, Y_i), 1 \leq i \leq n\}$ be n random variables, independent and identically distributed as the random pair (X, Y) with values in $\mathcal{H} \times \mathbb{R}$, where $\mathcal{H}$ is a separable real Hilbert space with the norm $\| . \|$ generated by an inner product $\langle ., . \rangle$. We consider the semi-metric d_θ associated to the single-index $\theta \in \mathcal{H}$ defined by $\forall x_1, x_2 \in \mathcal{H} : d_\theta(x_1, x_2) :=| \langle x_1 - x_2, \theta \rangle |$. Assume that the explanation of Y given X is done through a fixed functional index θ in $\mathcal{H}$. In the sense that, there exists a θ in $\mathcal{H}$ (unique up to a scale normalization factor) such that: $\mathbb{E}[Y|X] = \mathbb{E}[Y|\langle \theta, X \rangle]$. The conditional density of Y given $X = x$ denoted by $f_\theta(.|x)$ exists and is given by $\forall y \in \mathbb{R}, f_\theta(y|x) := f(y|\langle x, \theta \rangle)$. In the following, we denote by $f(\theta, ., x)$, the conditional density of Y given $\langle x, \theta \rangle$ and we define the local linear estimator for single-index structure $\widehat{f}(\theta, ., x)$ of $f(\theta, ., x)$ by:

$$\widehat{f}(\theta, y, x) = \frac{\sum_{1 \leq i,j \leq n} W_{ij}(\theta, x) H(h_H{}^{-1}(y - Y_j))}{h_H \sum_{1 \leq i,j \leq n} W_{ij}(\theta, x)} = \frac{\sum_{1 \leq j \leq n} \Omega_j K_j H_j}{h_H \sum_{1 \leq j \leq n} \Omega_j K_j},$$

with

$$W_{ij}(\theta, x) = \beta_\theta(X_i, x)\Big(\beta_\theta(X_i, x) - \beta_\theta(X_j, x)\Big) K(h_K^{-1} d_\theta(x, X_i)) K(h_K^{-1} d_\theta(x, X_j)),$$

and $\Omega_j K_j = \sum_{i=1}^n W_{ij}$ with $\beta_\theta(X_i, x)$ is a known bi-functional operator from $\mathcal{H}^2$ into $\mathbb{R}$ where K and H are kernel functions and $h_K := h_{n,K}$ (resp. $h_H := h_{n,H}$) is a sequence that decrease to zero as n goes to infinity.

3. Assumptions and mains results

All along the paper, we will denote by C, C' and $C_{\theta,x}$ some strictly positive generic constants and by $K_i(\theta, x) := K(h_K^{-1} d_\theta(x, X_i))$, $\forall x \in \mathcal{H}, i = 1, \ldots, n$, $H_j := H(h_H{}^{-1}(y - Y_j)), \forall y \in \mathbb{R}, j = 1, \ldots, n.$, $\beta_{\theta,i} := \beta_\theta(X_i, x)$, $W_{ij}(\theta, x) := W_{\theta,ij}$ and we will use the notation $B_\theta(x, h_K) := \{x_1 \in \mathcal{H} : \ 0 < | < x - x_1, \theta > | < h_K\}$, the ball centred at x with radius h_K. Moreover, for find the results in our paper we denote: for any $l \in \{0, 2\} \psi_l(., y) := \frac{\partial^l f(., y, .)}{\partial y^l}$

$\Phi_l(s) = \mathbb{E}[\psi_l(X, y) - \psi_l(x, y)|d_\theta(x, X) = s]$

In order to study our asymptotic results, we need the following assumptions:

(H1) $\mathbb{P}(X \in B_\theta(x, h_K)) =: \phi_{\theta,x}(h_K) > 0$, and assume that there exists a function $\chi_{\theta,x}(\cdot)$ such that

$$\forall s \in [-1, 1] \lim_{h_K \to 0} \frac{\phi_{\theta,x}(s h_K, h_K)}{\phi_{\theta,x}(h_K)} = \chi_{\theta,x}(s).$$

(H2) For any $l \in \{0, 2\}$, the quantities $\Phi_l'(0)$ and $\Phi_l''(0)$ exist, where Φ_l' (resp.Φ_l'') denotes the first (resp. the second) derivative of Φ_l

(H3) The bi-functional $\beta_\theta(., .)$ satisfies:

$$\begin{cases} \forall x' \in \mathcal{F}, \ C_1 \, d_\theta(x, x') \leq |\beta_\theta(x, x')| \\ \leq C_2 \, d_\theta(x, x'), \text{ where } C_1 > 0, \ C_2 > 0, \\ \sup_{u \in B(x,r)} |\beta_\theta(u, x) - d_\theta(x, u)| = o(r) \\ h_K \int_{B(x,h_K)} \beta_\theta(u, x) dP(u) \\ = o\left(\int_{B(x,h_K)} \beta_\theta^2(u, x) \, \mathrm{d}P(u)\right). \end{cases}$$

Where $B_\theta(x, r) = \{x' \in \mathcal{H} / |d_\theta(x, x') \leq r\}$ and $dP(x)$ is the cumulative distribution of X.

(H4) The kernel K is a positive, differentiable function and its derivative K' exists and is such that there exist two constants C and C' with $-\infty < C < K'(t) < C' < 0$, for $t \in [-1, 1]$ and $K(1) > 0$.

(H5) The kernel H is a differentiable function and bounded, such that:

$$\int H(t) \mathrm{d}t = 1, \quad \int t^2 H(t) \mathrm{d}t < \infty \quad \text{and} \quad \int H^2(t) \mathrm{d}t < \infty.$$

Where $b_2 > 0$.

(H6) The bandwidths h_K,h_H satisfies:

(i) $\lim_{n \to \infty} h_K = 0$, $\lim_{n \to \infty} h_H = 0$ and $\lim_{n \to \infty} n h_H \phi_{\theta,x}(h_K) = \infty$,

(ii) $\lim_{n \longrightarrow \infty} n h_H^5 \phi_{\theta,x}(h_K) = 0$ and $\lim_{n \longrightarrow \infty} n h_H h_K^2 \phi_{\theta,x}(h_K) = 0$.

Comments on assumptions: Notice that, (H1) and (H2) are a simple adaptation of the conditions in Ferraty et al. (2007) on the regression operator, when we replace the semi-metric by some bi-functional d_θ. The second part of the condition (H3) is unrestrictive and is verified, for instance, if $d_\theta(\cdot, \cdot) = \beta_\theta(\cdot, \cdot)$, moreover if

$$\lim_{d_\theta(x,u) \to 0} \left| \frac{\beta_\theta(u, x)}{d_\theta(x, u)} - 1 \right| = 0.$$

Assumptions (H4)–(H6) are classical in this context of quadratic errors and asymptotic normality in functional statistic.

3.1. Mean square convergence

In this part, we are going to show the asymptotic results of quadratic-mean convergence

Theorem 3.1: *Under assumptions* (H1)–*(H6), we obtain:*

$$\begin{aligned}&\mathbb{E}\left[\widehat{f}(\theta,y,x)-f(\theta,y,x)\right]^2\\&=B_H^2(\theta,x,y)h_H^4+B_K^2(\theta,x,y)h_K^2+\frac{V_{HK}(\theta,x,y)}{nh_H\phi_{\theta,x}(h_K)}\\&\quad+o(h_H^4)+o(h_K^2)+o\left(\frac{1}{nh_H\phi_{\theta,x}(h_K)}\right)\end{aligned}$$

where

$$B_H(\theta,x,y)=\frac{1}{2}\frac{\partial^2 f(\theta,y,x)}{\partial y^2}\int t^2H(t)\mathrm{d}t,$$

$$B_K(\theta,x,y)=\Phi_0'(0)\frac{M_0}{M_1}h_K,$$

and

$$V_{HK}(\theta,x,y)=\frac{M_2 f(\theta,y,x)}{M_1^2}\left(\int H^2(t)\mathrm{d}t\right),$$

with

$$M_0=K(1)-\int_0^1 sK'(s)\chi_{\theta,x}(s)\mathrm{d}s\quad\textit{and}$$

$$M_j=K^j(1)-\int_0^1 (K^j)'(s)\chi_{\theta,x}(s)\mathrm{d}s\ \textit{for}\ j=1,2.$$

we set

$$\widehat{f}(\theta,y,x)=\frac{\widehat{f}_N(\theta,y,x)}{\widehat{f}_D(\theta,x)}.$$

where

$$\widehat{f}_N(\theta,y,x)=\frac{1}{n(n-1)h_H\mathbb{E}[W_{12}(\theta,x)]}\sum_{1\le i\ne j\le n}W_{ij}(\theta,x)H(h_H^{-1}(y-Y_j)),$$

and

$$\widehat{f}_D(\theta,x)=\frac{1}{n(n-1)\mathbb{E}[W_{12}(\theta,x)]}\sum_{1\le i\ne j\le n}W_{ij}(\theta,x),$$

The following lemmas will be useful for proof of Theorem 3.1.

Lemma 3.2: *Under the assumptions of Theorem* 3.1*, we obtain:*

$$\begin{aligned}&\mathbb{E}\left[\widehat{f}_N(\theta,y,x)\right]-f(\theta,y,x)\\&=B_H(\theta,x,y)h_H^2+B_K(\theta,x,y)h_K+o(h_H^2)+o(h_K).\end{aligned}$$

Lemma 3.3: *Under the assumptions of Theorem* 3.1*, we obtain:*

$$Var\left[\widehat{f}_N(\theta,y,x)\right]=\frac{V_{HK}(\theta,x,y)}{nh_H\phi_{\theta,x}(h_K)}+o\left(\frac{1}{nh_H\phi_{\theta,x}(h_k)}\right).$$

Lemma 3.4: *Under the assumptions of Theorem* 3.1*, we get:*

$$Cov(\widehat{f}_N(\theta,y,x),\widehat{f}_D(\theta,x))=O\left(\frac{1}{n\phi_{\theta,x}(h_K)}\right).$$

Lemma 3.5: *Under the assumptions of Theorem* 3.1*, we get:*

$$Var\left[\widehat{f}_D(\theta,x)\right]=O\left(\frac{1}{n\phi_{\theta,x}(h_K)}\right).$$

3.2. Asymptotic normality

This section contains results on the asymptotic normality of $\widehat{f}(\theta,y,x)$. Before announcing our main results, we introduce the quantity $N(a,b)$, which will appear in the bias and variance dominant terms:

$$N(a,b)=K^a(1)-\int_{-1}^1 (u^bK^a(u))'\chi_x(u)\,\mathrm{d}u$$
$$\text{for all } a>0 \text{ and } b=2,4$$

Then, we have the following theorem:

Theorem 3.6: *Under assumptions (H1)–(H6), we obtain:*

$$\begin{aligned}&\sqrt{nh_H\phi_{\theta,x}(h_K)}(\widehat{f}(\theta,y,x)-f(\theta,y,x)-\mathbb{B}_n(\theta,x,y))\\&\quad\xrightarrow{D}\mathcal{N}(0,V_{HK}(\theta,x,y))\end{aligned}\tag{1}$$

where,

$$V_{HK}(\theta,x,y)=\frac{M_2}{M_1^2}f(\theta,y,x)\left(\int H^2(t)\mathrm{d}t\right)\tag{2}$$

and

$$\mathbb{B}_n(\theta,x,y)=\frac{\mathbb{E}(\widehat{f}_N(\theta,y,x)(y))}{\mathbb{E}(\widehat{f}_D(\theta,x))}-f(\theta,y,x)\tag{3}$$

with $\xrightarrow{D}$ *denoting the convergence in distribution.*

Proof of Theorem 3.6: Inspired by the decomposition given in Masry (2005) , we set.

$$\begin{aligned}&\widehat{f}(\theta,y,x)-f(\theta,y,x)-\mathbb{B}_n(\theta,x,y)\\&=\frac{\widehat{f}_N(\theta,y,x)-f(\theta,y,x)\widehat{f}_D(\theta,x)-\widehat{f}_D(\theta,x)\mathbb{B}_n(\theta,x,y)}{\widehat{f}_D(\theta,x)}\end{aligned}$$

If we denote by

$$\begin{aligned}&Q_n(\theta,x,y)\\&=\widehat{f}_N(\theta,y,x)-f(\theta,y,x)\widehat{f}_D(\theta,x)\end{aligned}$$

$$- \mathbb{E}(\widehat{f}_N(\theta,y,x) - f(\theta,y,x)\widehat{f}_D(\theta,x)$$
$$= \widehat{f}_N(\theta,y,x) - f(\theta,y,x)\widehat{f}_D(\theta,x) - \mathbb{B}_n(\theta,x,y) \quad (4)$$

since

$$\widehat{f}_N(\theta,y,x) - f(\theta,y,x)\widehat{f}_D(\theta,x) = Q_n(\theta,x,y) + \mathbb{B}_n(\theta,x,y)$$

then the proof of this theorem will be completed from the following expression

$$\widehat{f}(\theta,y,x) - f(\theta,y,x) - \mathbb{B}_n(\theta,x,y) = \frac{Q_n(\theta,x,y) - \mathbb{B}_n(\theta,x,y)(\widehat{f}_D(\theta,x) - \mathbb{E}(\widehat{f}_D(\theta,x)))}{\widehat{f}_D(\theta,x)} \quad (5)$$

and the following auxiliary results which play a main role and for which proofs are given in the appendix. ■

Lemma 3.7: *Under assumptions* (H1)–(H5)*, we have*

$$\widehat{f}_D(\theta,x) \xrightarrow{P} \mathbb{E}(\widehat{f}_D(\theta,x)) = 1$$

where $\xrightarrow{P}$ *denotes the convergence in probability.*

So, Lemma 3.7, implies that $\widehat{f}_D(\theta,x) \to 1$. Moreover, $\mathbb{B}_n(\theta,x,y) = o(1)$ as $n \to \infty$ because of the continuity of $f(\theta,y,x)$. Then, we obtain that

$$\widehat{f}(\theta,y,x) - f(\theta,y,x) - \mathbb{B}_n(\theta,x,y) = \frac{Q_n(\theta,x,y)}{\widehat{f}_D(\theta,x)}(1 + o_p(1))$$

Lemma 3.8: *Under assumptions (H1)–(H5), we have*

$$\sqrt{nh_H\phi_{\theta,x}(h_K)}Q_n(\theta,x,y) \xrightarrow{D} \mathcal{N}(0, V_{HK}(\theta,x,y)), \quad (6)$$

where $V_{HK}(\theta,x,y)$ is defined by (2).

Remark 3.9: As mentioned in Demongeot et al. (2013), the function $\phi_{\theta,x}(t)$ can be empirically estimated by

$$\hat{\phi}_{\theta,x}(t) = \frac{\sharp\{i : |d(X_i,x)| \le t\}}{n},$$

where $\sharp(A)$ denote the cardinality of the set A. So, if we take advantage of the following assumptions, (H6) $\lim_{n\to+\infty} \sqrt{nh_H\phi_{\theta,x}(h_K)}\mathbb{B}_n(\theta,x,y) = 0$, we can cancel the bias term and obtain the following corollary.

Corollary 3.10: *Under the assumptions of Theorem* 3.6*, we get*

$$\sqrt{\frac{nh_H\hat{\phi}_{\theta,x}(h_K)}{V_{HK}(\theta,x,y)}}(\widehat{f}(\theta,y,x) - f(\theta,y,x)) \to \mathcal{N}(0,1)$$

4. Simulation study

We first construct the simulation of the explanatory functional variables. In the second part, we focus on the ability of the nonparametric functional regression to predict responses variable from functional predictors. Finally we illustrate the MONTE-CARLO methodology and we will appropriate to test the efficiency of the asymptotic normality results parallel the practical experiment and we build functional pseudo-confidence area.

For this purpose, we consider the following process explanatory functional variables for $n = 350$:

$$X_i(t) = \sum_{j=1}^{3} V_{ij}\cos((3+j)t) + W_i(t-\pi)^2, \quad \forall t \in [0,100]$$

where V_{ij} and W_i are n independent real random variables (r.r.v.) uniformly distributed over [0.3; 2] (resp. [1; 3]), t is assumed that these curves are observed on a discretization grid of 100 points in the interval. These functional variables are represented in Figure 1.

For response variables Y_i, we consider the following model for all $i = 1,\dots n$ and $j = 1,\dots 100$:

$$Y = r(\langle\theta_k, X\rangle) + \epsilon$$

where $r(U_i) = \int_0^{100} \frac{1}{1-U_i^2(\nu)}\,d\nu$ and ϵ is a centred normal variable and assumed to be independent of $(X_i)_i$. Then, we can get the corresponding conditional density, which is explicitly defined by

$$f(\theta_k,y,x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(y-r(\langle\theta_k,x\rangle))^2}$$

Our goal in this illustration is to show the usefulness of conditional density in a context of forecasting. Thus the use of optimal parameters of the conditional density and without theoretical validity.

Now, we precise the different parameters of our estimators. Indeed, first of all, it is clear that the shape of the curves allows us to use

$$d(x_1,x_2) = \sqrt{\int_0^1 (x_1(t) - x_2(t))^2};$$
$$\forall x_1, x_2 \in \mathcal{H} \text{ where } \mathcal{H} \text{ is a semi-metric space}$$

We choose particularly the quadratic kernels defined by

$$\frac{3}{2}(1-x^2)\mathbf{1}_{[-1,1)} \quad \text{and} \quad \mathbf{K(1)} > \mathbf{0}$$

In this illustration, we select the functional index θ_k on the set of eigenvectors of the empirical covariance

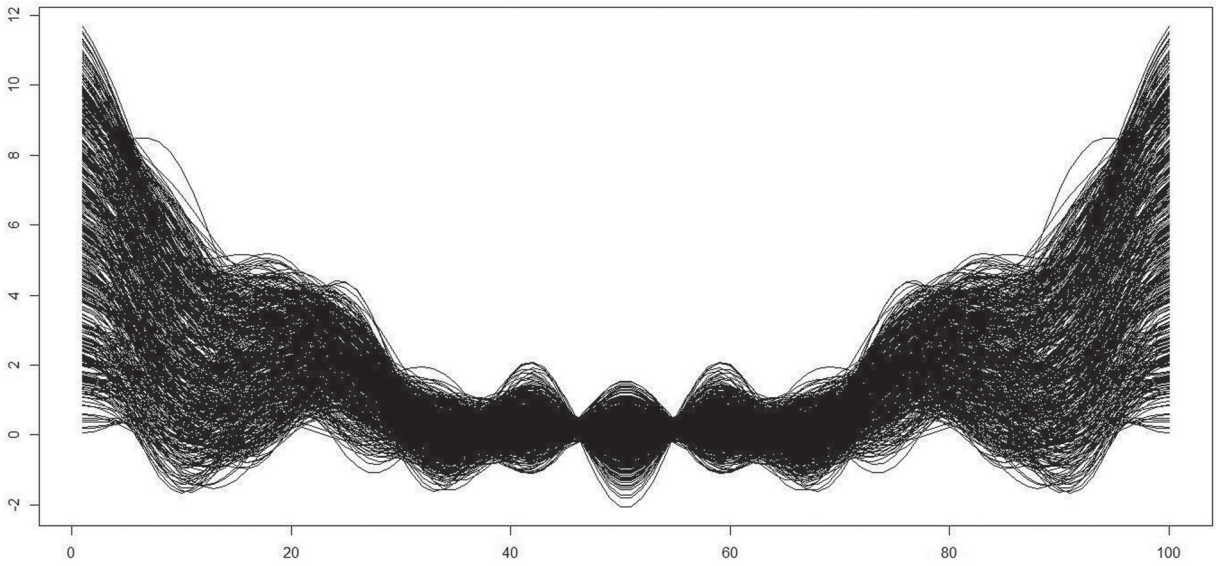

Figure 1. The curves $X_i, i = 1, \ldots 200$.

operator.

$$\Gamma_n(X) = \frac{1}{200}\sum_{i=1}^{200}(X_i - \bar{X})^t((X_i - \bar{X})).$$

Indeed, We recall that the ideas of Ait Saidi et al. (2008) can be adapted to find a method of practical selection for θ_k. However, this adaptation in the case of the conditional density requires tools and additional preliminary results (See the discussion Attaoui et al. (2011) and Attaoui (2014)).

For this purpose, we divide our observations on two packets learning sample $(X_i, Y_i)_{i=1,\ldots 200}$ and test sample $(X_i, Y_i)_{i=201,\ldots 250}$. For the choice of smoothing parameters h_K and h_H, we will adopt the selection criterion used by Ferraty et al. (2006) in the case of the kernel method for which h_K and h_H are obtained. by minimizing the next criterion

$$\frac{1}{n}\sum_{i=1}^{n} W_1(X_i)\int \widehat{f}^{-i^2}_{(h_K,h_H)}(X_i, y)\, W_2(y)\mathrm{d}y - \frac{2}{n}\sum_{i=1}^{n}\widehat{f}^{-i}_{(h_K,h_H)}(X_i, Y_i)\, W_1(X_i)\, W_2(Y_i) \qquad (7)$$

where

$$\widehat{f}^{-k}_{(h_K,h_H)}(X_k, y) = \frac{\sum_{k\neq i,j=1}^{n} W_{ij}(X_k) H(h_H^{-1}(y - Y_i))}{h_H \sum_{i,j=1}^{n} W_{ij}(X_k)}$$

A first way of assessing the quality of prediction is to compare predicted functional responses $(\widehat{f}(\theta, y, x)$ for any X in the testing sample) versus the true of conditional density operator (i.e., $f(\theta, y, x)$) as in Figure 2.

For the next simulation algorithm, we used:

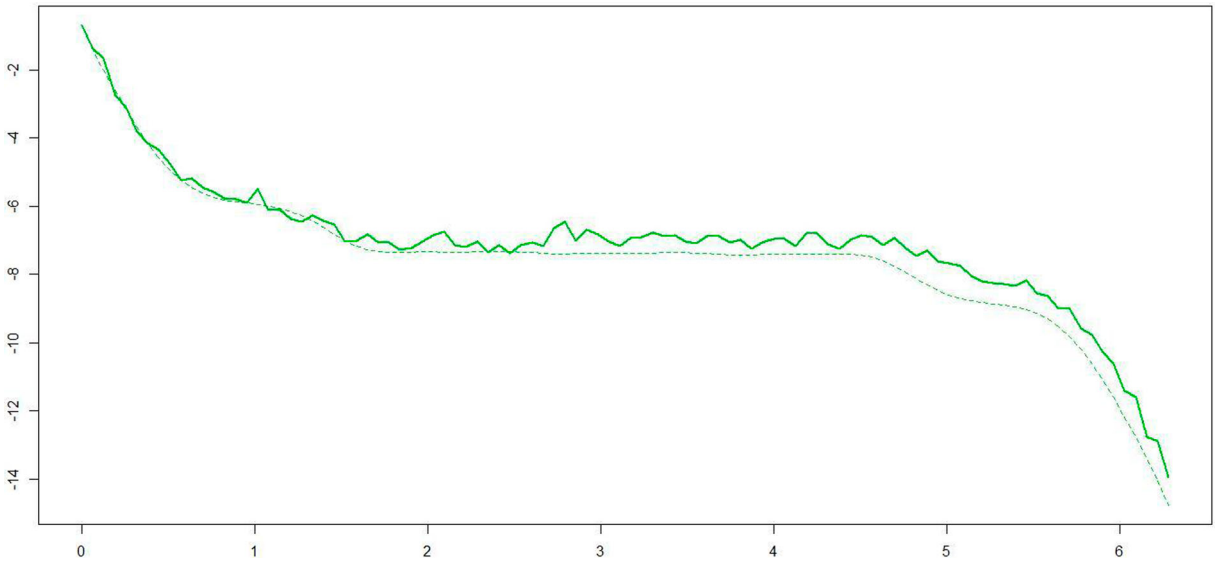

Figure 2. Predicted functional responses (solid lines); observed functional responses (dashed lines).

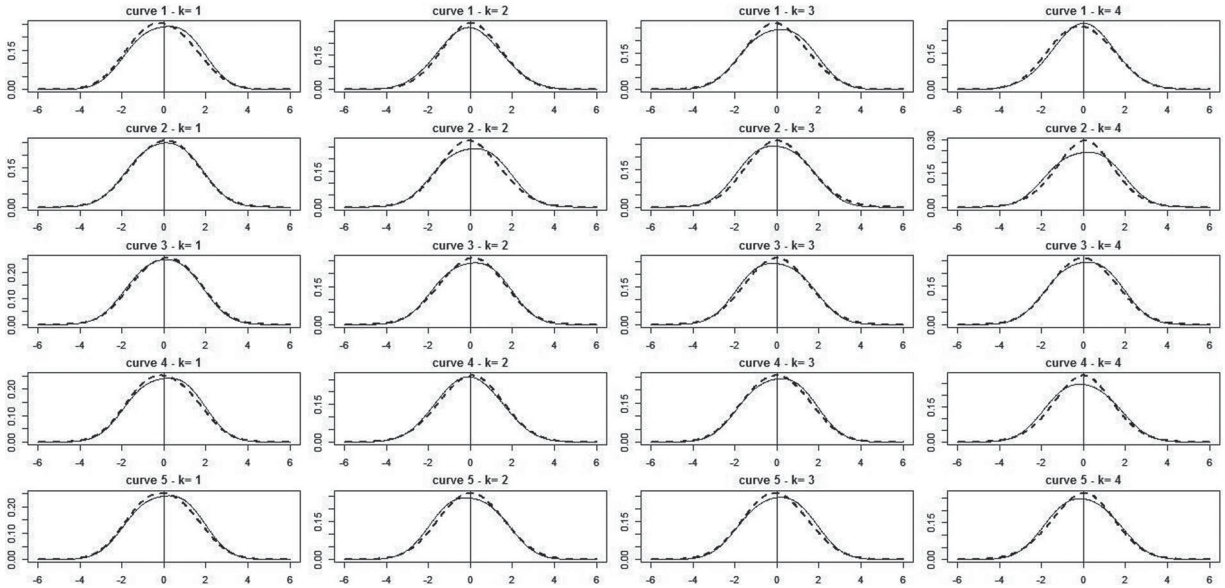

Figure 3. Representation of the estimated density for $k = 1, 2, 3, 4$.

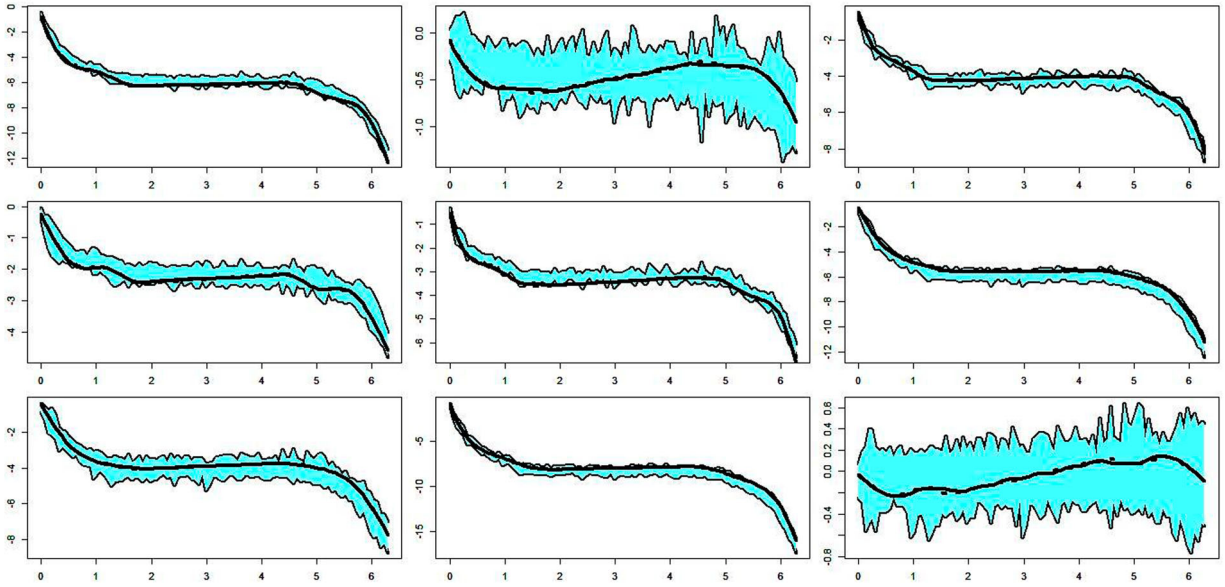

Figure 4. Functional pseudo-confidence areas.

- Simulate a sample of size n.
- Calculate the smoothing parameters h_K and h_H that are varied over an interval [0,1] and which minimizes by (7).
- We compute for $k = 1, 2, 3, 4$ the quantities

$$(nh_H\phi_{\theta_k,x})^{1/2}(\widehat{f}(\theta_k, y, x) - f(\theta_k, y, x))$$

  where $\widehat{f}(\theta_k, y, x)$ is the functional kernel estimator from the sample $(X_i, Y_i)_{i=1,\ldots,200,}$ and $k = 1, 2, 3, 4$.
- compute a standard density estimator by local linear method.
- compare the estimated $\widehat{f}(\theta_k, y, x)$ with the corresponding estimated $f(\theta_k, y, x)$).

The obtained results are shown in Figure 3. It can be seen that, both densities are very well approximated and have good behaviours with respect to the standard normal distribution.

An application of results of Theorem 3.6 is to build the functional pseudo-confidence areas. To this aim, let us set for any component k, $(k = 1, 2, \ldots, K)$ and $\eta_k = \eta/K$ with $\eta \in [0, 1]$, confidence intervals $E_k^{\eta_k}$ such that

$$P\left(\bigcap_{k=1}^{K} r_k(U) \in \widehat{E}_K^{\eta_k}\right) \geq 1 - \eta$$

where $U = \langle X, \theta_k \rangle$ with $\theta_1, \ldots, \theta_K$ is a data-driven orthonormal basis the K eigenfunctions associated to the K largest eigenvalues of Γ.

The results from the asymptotic normality of the conditional density is expressed in Corollary 3.10 and we can approximate $(1 - \eta)$ confidence interval of

$f(\theta, y, x)$ by:

$$\widehat{f}(\theta, y, x) \pm t_{\eta/2} \times \left(\frac{\widehat{V_{HK}}(\theta, x, y)}{nh_H \hat{\phi}_{\theta,x}(h_K)} \right)^{1/2}.$$

where $t_{\eta/2}$ denotes the $\eta/2$ quantile of the standard normal $N(0, 1)$.

Figure 4 represents a functional pseudo-confidence zone for 9 different fixed curves with $\eta = 0.05$ and $K = 4$. We can remark that $r(\langle x, \theta_k \rangle)$ and its K-dimensional projection onto $\widehat{\theta}_1, \ldots, \widehat{\theta}_K$ are very close. This conclusion shows the good performance of our asymptotic normality, indeed when one replaces the data-driven basis with the eigenfunctions of Γ, one gets very similar functional pseudo-confidence areas.

5. Conclusion

In this paper, we are mainly interested in the nonparametric estimation of the conditional density function by the local linear method for a variable explanatory functionally conditioned to an actual response variable via a functional single-index model. We show that the estimator provides good predictions under this model. One of the main contributions of this work is the choice of semi-metric. Indeed, it is well known that, in nonparametric functional statistics, the semi-metric of the projection type is very important for increasing the concentration property. The functional index model is a special case of this family of semi-metrics because it is based on the projection on a functional direction which is important for the implementation of our method in practice. Therefore, we can draw zones of functional pseudo-confidence, which is a very interesting tool for assessing the quality of the prediction.

Acknowledgments

The authors are very grateful to the Editor and the anonymous reviewers for their comments which improved the quality of this paper. The authors wish to thank two anonymous referees for their helpful comments and suggestions, which greatly improved the quality of this paper.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

Ait Saidi, A., Ferraty, F., Kassa, P., & Vieu, P. (2008). Cross-validated estimations in the single functional index model. *Statistics*, *42*(6), 475–494. https://doi.org/10.1080/02331880801980377

Attaoui, S. (2014). Strong uniform consistency rates and asymptotic normality of conditional density estimator in the single functional index modeling for time series data. AStA Advances in Statistical Analysis., *98*(3), 257–286. https://doi.org/10.1007/s10182-014-0227-3

Attaoui, S., Laksaci, A., & Ould Said, F. (2011). A note on the conditional density estimate in the single functional index model. *Statistics & Probability Letters*, *81*(1), 45–53. https://doi.org/10.1016/j.spl.2010.09.017

Attaoui, S., & Ling, N. (2016). Asymptotic results of a nonparametric conditional cumulative distribution estimator in the single functional index modeling for time series data with applications. *Metrika*, *79*(3), 485–511. https://doi.org/10.1007/s00184-015-0564-6

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

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

Bosq, D., & Lecoutre, J. P. (1987). *Théorie de l'estimation fonctionnelle*. Ed. Economica.

Demongeot, J., Laksaci, A., Madani, F., & Rachdi, M. (2013). Functional data: local linear estimation of the conditional density and its application. *Statistics: A Journal of Theoretical and Applied Statistics.*, *76*(2), 328–355. https://doi.org/10.1080/02331888.2011.568117

Demongeot, J., Laksaci, A., Rachdi, M., & Rahmani, S. (2014). On the local linear modelization of the conditional distribution for functional data. *Sankhya: The Indian Journal of Statistics.*, *76*(2), 328–355. https://doi.org/10.1007/s13171-013-0050-z

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

Ferraty, F, Laksaci, A, & Vieu, P. (2006). Estimating some characteristics of the conditional distribution in nonparametric functional model. *Statistical Inference for Stochastic* Processes , *9*(1), 47–76. https://doi.org/10.1007/s.11203-004-3561-3

Ferraty, F., Mas, A., Vieu, P., & Vieu, P. (2007). Advances in nonparametric regression for functional variables. *Australian & New Zealand Journal of Statistics*, *49*(1), 1–20. https://doi.org/10.1111/j.1467-842X.2006.00454.x

Ferraty, F., Peuch, A., & Vieu, P. (2003). Modéle à indice fonctionnel simple. *Comptes Rendus Mathématique de l'Académie des Sciences Paris*, *336*(12), 1025–1028. https://doi.org/10.1016/S1631-073X(03)00239-5(in French)

Ferraty, F., & Vieu, P. (2006). *Nonparametric functional data analysis. Theory and Practice*. Springer Series in Statistics.

Laksaci, A. (2007). Convergence en moyenne quadratique de l'estimateur a noyau de la densité conditionnelle avec variable explicative fonctionnelle. *Publications de l'Institut de statistique de l'Université de Paris*, *51*(3), 69–80.

Laksaci, A., Rachdi, M., & Rahmani, S. (2013). Spatial modelization: local linear estimation of the conditional distribution for functional data. *Spatial Statistics*, *6*(4), 1–23. https://doi.org/10.1016/j.spasta.2013.04.004

Masry, E. (2005). Nonparametric regression estimation for dependent functional data: asymptotic normality. Stochastic Processes and their Applications, *115*(1), 155–177. https://doi.org/10.1016/j.spa.2004.07.006

Sarda, P., & Vieu, P. (2000). *Kernel regression*. Wiley Series in Probability and Statistics (pp. 43–70).

Tabti, H., & Ait Saidi, A. (2018). Estimation and simulation of conditional hazard function in the quasi-associated framework when the observations are linked via a functional single-index structure. *Communications*

in Statistics – Theory and Methods, *47*(4), 816–838. https://doi.org/10.1080/03610926.2016.1213294

Appendix

Proof of Theorem 3.1: We know the theorem is a consequence of a separate computes two quantities (bias and variance) of $\widehat{f}(\theta,y,x)$, we have

$$\mathbb{E}\left[\widehat{f}(\theta,y,x)-f(\theta,y,x)\right]^2 = \left[\mathbb{E}\left(\widehat{f}(\theta,y,x)\right)-f(\theta,y,x)\right]^2+\mathrm{Var}\left[\widehat{f}(\theta,y,x)\right]$$

By classical calculations, we obtain

$$\begin{aligned}\widehat{f}(\theta,y,x)-f(\theta,y,x) &= \left(\widehat{f}_N(\theta,y,x)-f(\theta,y,x)\right)\\ &\quad-\left(\widehat{f}_N(\theta,y,x)-\mathbb{E}[\widehat{f}_N(\theta,y,x)]\right)\left(\widehat{f}_D(\theta,x)-1\right)\\ &\quad-\mathbb{E}[\widehat{f}_N(\theta,y,x)]\left(\widehat{f}_D(\theta,x)-1\right)\\ &\quad+\left(\widehat{f}_D(\theta,x)-1\right)^2\widehat{f}(\theta,y,x).\end{aligned}$$

which implies that:

$$\begin{aligned}\mathbb{E}\left[\widehat{f}(\theta,y,x)\right]-f(\theta,y,x) &= \left(\mathbb{E}[\widehat{f}_N(\theta,y,x)]-f(\theta,y,x)\right)\\ &\quad-Cov\left(\widehat{f}_N(\theta,y,x),\widehat{f}_D(\theta,x)\right)\\ &\quad+\mathbb{E}\left[\left(\widehat{f}_D(\theta,x)-\mathbb{E}[\widehat{f}_D(\theta,x)]\right)^2\widehat{f}(\theta,y,x)\right].\end{aligned}$$

Under the assumption (H5), we can bound $\widehat{f}(\theta,y,x)$ by a constant $C>0$ where $\widehat{f}(\theta,y,x)\le C/h_H$. Hence:

$$\begin{aligned}\mathbb{E}\left[\widehat{f}(\theta,y,x)\right]-f(\theta,y,x) &= \left(\mathbb{E}[\widehat{f}_N(\theta,y,x)]-f(\theta,y,x)\right)\\ &\quad-\mathrm{Cov}\left(\widehat{f}_N(\theta,y,x),\widehat{f}_D(\theta,x)\right)\\ &\quad+\mathrm{Var}\left[\widehat{f}_D(\theta,x)\right]O(h_H^{-1}).\end{aligned}$$

Now, by similar technics as those Sarda and Vieu (2000) and by Bosq and Lecoutre (1987), the variance term is

$$\begin{aligned}\mathrm{Var}\left[\widehat{f}(\theta,y,x)\right] &= \mathrm{Var}\left[\widehat{f}_N(\theta,y,x)\right]\\ &\quad-2\mathbb{E}[\widehat{f}_N(\theta,y,x)]Cov\left(\widehat{f}_N(\theta,y,x),\widehat{f}_D(\theta,x)\right)\\ &\quad+\left(\mathbb{E}[\widehat{f}_N(\theta,y,x)]\right)^2\mathrm{Var}\left(\widehat{f}_D(\theta,x)\right)\\ &\quad\times o\left(\frac{1}{nh_H\phi_{\theta,x}(h_K)}\right).\end{aligned}$$

■

Proof of Lemma 3.2: We have:

$$\begin{aligned}\mathbb{E}[\widehat{f}_N(\theta,y,x)] &= \mathbb{E}\left[\frac{1}{n(n-1)h_H\mathbb{E}[W_{12}(\theta,x)]}\right.\\ &\quad\left.\times\sum_{1\le i\ne j\le n}W_{ij}(\theta,x)H(h_H^{-1}(y-Y_j))\right]\\ &= \frac{1}{h_H\mathbb{E}[W_{\theta,12}]}\mathbb{E}\left[W_{\theta,12}\mathbb{E}[H_2|X_2]\right].\end{aligned}$$

By using a Taylor's expansion and under assumption (H5), we have

$$\mathbb{E}[H_2|X_2] = f(\theta,y,X_2)+\frac{h_H^2}{2}\left(\int t^2H(t)\mathrm{d}t\right)\frac{\partial^2 f(\theta,y,X_2)}{\partial y^2}+o(h_H^2).$$

Now, we can re-written as:

$$\mathbb{E}[H_2|X_2] = \psi_0(X_2,y)+\frac{h_H^2}{2}\left(\int t^2H(t)\,\mathrm{d}t\right)\psi_2(X_2,y)+o(h_H^2).$$

Thus, we obtain

$$\begin{aligned}\mathbb{E}\left[\widehat{f}_N(\theta,y,x)\right] &= \frac{1}{E[W_{\theta,12}]}\left(\mathbb{E}\left[W_{\theta,12}\psi_2(X_2,y)\right]\right.\\ &\quad\left.+\left(\int t^2H(t)\mathrm{d}t\right)\mathbb{E}\left[W_{\theta,12}\psi_2(X_2,y)\right]+o(h_H^2)\right).\end{aligned}$$

Accordingly to Ferraty et al. (2007), for $l\in\{0,2\}$, we show that

$$\begin{aligned}\mathbb{E}[W_{\theta,12}\psi_l(X_2,y)] &= \psi_l(x,y)\mathbb{E}[W_{\theta,12}]+\mathbb{E}[W_{\theta,12}(\psi_l(X_2,y)-\psi_l(x,y))]\\ &= \psi_l(x,y)\mathbb{E}[W_{\theta,12}]\\ &\quad+\mathbb{E}[W_{\theta,12}\mathbb{E}[\psi_l(X_2,y)-\psi_l(x,y)|d_\theta(X_2,x)]]\\ &= \psi_l(x,y)\mathbb{E}[W_{\theta,12}]+\mathbb{E}[W_{\theta,12}\Phi_l(d_\theta(X_2,x))].\end{aligned}$$

Since $\Phi_l(0)=0$, we obtain

$$\mathbb{E}[W_{\theta,12}\Phi_l(d_\theta(X_2,x))] = \Phi_l'(0)\mathbb{E}[d_\theta(X_2,x)W_{\theta,12}]+o(\mathbb{E}\left[d_\theta(X_2,x)W_{\theta,12}\right]).$$

Then,

$$\begin{aligned}\mathbb{E}\left[\widehat{f}_N(\theta,y,x)\right] &= f(\theta,y,x)+\frac{h_H^2}{2}\frac{\partial^2 f(\theta,y,x)}{\partial y^2}\int t^2H(t)\mathrm{d}t\\ &\quad+o\left(h_H^2\frac{\mathbb{E}\left[d_\theta(X_2,x)W_{\theta,12}\right]}{\mathbb{E}[W_{\theta,12}]}\right)\\ &\quad+\Phi_0'(0)\frac{\mathbb{E}\left[d_\theta(X_2,x)W_{\theta,12}\right]}{\mathbb{E}[W_{\theta,12}]}\\ &\quad+o\left(\frac{\mathbb{E}\left[d_\theta(X_2,x)W_{\theta,12}\right]}{\mathbb{E}[W_{\theta,12}]}\right).\end{aligned}$$

Therefore, it remains to determine the quantities $E[d_\theta(X_2,x)W_{\theta,12}]$ and $E[W_{\theta,12}]$. According to the definition of $W_{\theta,12}$, the two quantities $E[d_\theta(X_2,x)W_{\theta,12}]$ and $E[W_{\theta,12}]$ are based on the evaluation asymptotic of $E[K_1^a\beta_1^b]$. To do that, we treat firstly, the case $b=1$. For this case, we use the assumptions (H3) and (H4) to get

$$\begin{aligned}h_K\mathbb{E}[K_1^a\beta_{\theta,1}] &= o\left(\int_{B(x,h_K)}\beta_\theta^2(u,x)\mathrm{d}P(u)\right)\\ &= o(h_K^2\phi_{\theta,x}(h_K)).\end{aligned}$$

So, we obtain that,

$$\mathbb{E}[K_1^a\beta_{\theta,1}] = o(h_K\phi_{\theta,x}(h_K)). \tag{A1}$$

Moreover, for all $b>1$, and after simplifications of the expressions, permits to write that

$$\mathbb{E}[K_1^a\beta_{\theta,1}^b] = \mathbb{E}[K_1^a d_\theta^b(x,X)]+o(h_K^b\phi_{\theta,x}(h_K)).$$

Concerning the first term, we write

$$\begin{aligned}
&h_K^{-b}\mathbb{E}[K_1^a d_\theta^b] \\
&= \int v^b K^a(v)\mathrm{d}P^{h_K^{-1}d_\theta(x,X)}(v) \\
&= \int_{-1}^{1}\left[K^a(1) - \int_v^1 \left((s^bK^a(s))'\right)\mathrm{d}u\right]\mathrm{d}P^{h_K^{-1}d_\theta(x,X)}(v) \\
&= \left(K(1)\phi_{\theta,x}(h_K) - \int_{-1}^{1}(s^bK^a(s))^{(1)}\phi_{\theta,x}(sh_K,h_K)\mathrm{d}s\right) \\
&= \phi_{\theta,x}(h_K)\left(K(1) - \int_{-1}^{1}(s^bK^a(s))'\frac{\phi_{\theta,x}(sh_K,h_K)}{\phi_{\theta,x}(h_K)}\mathrm{d}s\right).
\end{aligned}$$

Finally, under assumptions (H1), we get

$$\begin{aligned}
\mathbb{E}[K_1^a\beta_{\theta 1}^b] &= h_K^b\phi_{\theta,x}(h_K)\left(K(1) - \int_{-1}^{1}(s^bK^a(u))'\chi_{\theta,x}(s)\mathrm{d}s\right) \\
&\quad + o(h_K^b\phi_{\theta,x}(h_K)). \qquad \text{(A2)}
\end{aligned}$$

So,

$$\begin{aligned}
&\frac{\mathbb{E}[d_\theta(X_2,x)W_{\theta,12}]}{\mathbb{E}[W_{\theta,12}]} \\
&= h_K\left(\frac{K(1) - \int_{-1}^{1}(sK(s))'\chi_{\theta,x}(s)\mathrm{d}s}{K(1) - \int_{-1}^{1}(K'(u)\chi_{\theta,x}(s)\mathrm{d}s}\right) + o(h_K).
\end{aligned}$$

Hence,

$$\begin{aligned}
&\mathbb{E}\left[\widehat{f}_N(\theta,y,x)\right] \\
&= f(\theta,y,x) + \frac{h_H^2}{2}\frac{\partial^2 f(\theta,y,x)}{\partial y^2}\int t^2H(t)\mathrm{d}t + o(h_H^2) \\
&\quad + h_K\Phi_0'(0)\frac{\left(K(1) - \int_{-1}^{1}(sK(s))'\chi_{\theta,x}(s)\mathrm{d}s\right)}{\left(K(1) - \int_{-1}^{1}K'(s)\chi_{\theta,x}(s)\mathrm{d}s\right)} \\
&\quad + o(h_K). \qquad \blacksquare
\end{aligned}$$

Proof of Lemma 3.3.: We know

$$\begin{aligned}
&Var\left(\widehat{f}_N(\theta,y,x)\right) \\
&= \frac{1}{\left(n(n-1)h_H(\mathbb{E}[W_{\theta,12}])\right)^2}Var\left(\sum_{1\le i\ne j\le n}W_{\theta,ij}H_j\right) \\
&= \frac{1}{\left(n(n-1)h_H(\mathbb{E}[W_{\theta,12}])\right)^2} \\
&\quad \times\Big[n(n-1)\mathbb{E}[W_{\theta,12}^2(H_2^2] \\
&\quad + n(n-1)\mathbb{E}[W_{\theta,12}W_{\theta,21}H_2H_1] \\
&\quad + n(n-1)(n-2)\mathbb{E}[W_{\theta,12}W_{\theta,13}H_2H_3] \\
&\quad + n(n-1)(n-2)\mathbb{E}[W_{\theta,12}W_{\theta,23}H_2H_3] \\
&\quad + n(n-1)(n-2)\mathbb{E}[W_{\theta,12}W_{\theta,31}H_2H_1] \\
&\quad + n(n-1)(n-2)\mathbb{E}[W_{\theta,12}W_{\theta,32}H_2^2] \\
&\quad - n(n-1)(4n-6)\mathbb{E}[W_{\theta,12}H_2]^2\Big]. \qquad \text{(A3)}
\end{aligned}$$

By direct calculations, we get

$$\begin{cases}
\mathbb{E}[W_{\theta,12}^2H_2] = O(h_K^4h_H\phi_{\theta,x}^2(h_K)), \quad \mathbb{E}[W_{\theta,12}W_{\theta,21}H_2H_1] \\
\quad = O(h_K^4h_H^2\phi_{\theta,x}^2(h_K)), \\
\mathbb{E}[W_{\theta,12}W_{\theta,13}H_2H_3] = \mathbb{E}[W_{\theta,12}W_{\theta,31}H_2H_1] \\
\quad = \mathbb{E}[W_{\theta,12}W_{\theta,23}H_2H_3] = O(h_K^4h_H^2\phi_{\theta,x}^3(h_K)), \\
\mathbb{E}[W_{\theta,12}W_{\theta,32}H_2^2] = \mathbb{E}^2[\beta_1^2K_1]\mathbb{E}[K_1^2H_1^2] \\
\quad +o(h_K^4h_H\phi_{\theta,x}^3(h_K)).
\end{cases}$$

Clearly, the latter term in the last cases is the leading one, and can be evaluated in (A3) by using

$$\frac{(n-2)}{n(n-1)(h_H\mathbb{E}[W_{\theta,12}])^2}\mathbb{E}^2[\beta_{\theta,1}^2K_1]\mathbb{E}[K_1^2H_1^2]$$

by the same arguments used in Lemma 3.2 proofs, we obtain

$$Var\left(\widehat{f}_N(\theta,y,x)\right) = \frac{\mathbb{E}[K_1^2H_1^2]}{n(h_H\mathbb{E}[K_1])^2} + o\left(\frac{1}{nh_H\phi_{\theta,x}(h_K)}\right). \qquad \text{(A4)}$$

Observe that

$$\begin{aligned}
\mathbb{E}[K_1^2H_1^2] &= \mathbb{E}[K_1^2\mathbb{E}((H_1^{(1)})^2|\langle\theta,X_1\rangle)] \\
&= \mathbb{E}\left[K_1^2\int H^2(h_H^{-1}(y-z))f(\theta,z,X_1)\mathrm{d}z\right].
\end{aligned}$$

Thus, by the change of variables $t = h_H^{-1}(y-z)$, we get

$$\mathbb{E}[K_1^2H_1^2] = h_H\mathbb{E}\left[K_1^2\int H^2(t)f(\theta,y-h_Ht,X_1)\mathrm{d}t\right].$$

By using Taylor's expansion of order 1 of $f(\theta,\cdot,y)$ we get

$$\begin{aligned}
f(\theta,y-h_Ht,X_1) &= f(\theta,y,X_1) + O(h_H) \\
&= f(\theta,y,X_1) + o(1).
\end{aligned}$$

Then

$$\begin{aligned}
\mathbb{E}[K_1^2H_1^2] &= h_H\left(\int H^2(t)\mathrm{d}t\right)\mathbb{E}\left[K_1^2f(\theta,y,X_1)\right] \\
&\quad + o(h_H\mathbb{E}[K_1^2]).
\end{aligned}$$

Also, by the same steps in proof of Lemma 3.2, we obtain

$$\mathbb{E}\left[K_1^2f(\theta,y,X_1)\right] = f(\theta,y,x)\mathbb{E}[K_1^2] + o(\mathbb{E}[K_1^2])$$

which give that:

$$\mathbb{E}[K_1^2H_1^2] = h_Hf(\theta,y,x)\mathbb{E}[K_1^2]\int H^2(t)\mathrm{d}t + o(h_H\mathbb{E}[K_1^2]). \qquad \text{(A5)}$$

Finally, we obtain from (A2), (A4) and (A5), that

$$\begin{aligned}
&Var\left(\widehat{f}_N(\theta,y,x)\right) \\
&= \frac{f(\theta,y,x)}{nh_H\phi_{\theta,x}(h_K)}\left(\int H^2(t)\mathrm{d}t\right) \\
&\quad \times\left[\frac{\left(K^2(1) - \int_{-1}^{1}(K^2(s))'\chi_{\theta,x}(s)\mathrm{d}s\right)}{\left(K(1) - \int_{-1}^{1}(K(s))'\chi_{\theta,x}(s)\mathrm{d}s\right)^2}\right] \\
&\quad + o\left(\frac{1}{nh_H\phi_{\theta,x}(h_K)}\right) \\
&= \frac{M_2f(\theta,y,x)}{M_1^2nh_H\phi_{\theta,x}(h_K)}\left(\int H^2(t)\mathrm{d}t\right) \\
&\quad + o\left(\frac{1}{nh_H\phi_{\theta,x}(h_K)}\right). \qquad \blacksquare
\end{aligned}$$

Proof of Lemma 3.4: The proof of this Lemma is similar to Lemma 3.3 proof, it permits to write

$$\begin{aligned}
&Cov\left(\widehat{f}_N(\theta,y,x),\widehat{f}_D(\theta,x)\right) \\
&= \frac{1}{\left(n(n-1)h_H\mathbb{E}[W_{\theta,12}]\right)^2} \\
&\quad \times Cov\left(\sum_{1\le i\ne j\le n}W_{\theta,ij}H_j, \sum_{1\le i'\ne j'\le n}W_{\theta,i'j'}\right)
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{\big(n(n-1)h_H\mathbb{E}[W_{\theta,12}]\big)^2}\Big[n(n-1)\mathbb{E}[W_{\theta 12}^2H_1]\\
&\quad + n(n-1)\mathbb{E}[W_{\theta,12}W_{\theta,21}H_2]\\
&\quad + n(n-1)(n-2)\mathbb{E}[W_{\theta,12}W_{\theta,13}H_2]\\
&\quad + n(n-1)(n-2)\mathbb{E}[W_{\theta,12}W_{\theta,23}H_2]\\
&\quad + n(n-1)(n-2)\mathbb{E}[W_{\theta,12}W_{\theta,31}H_2]\\
&\quad + n(n-1)(n-2)\mathbb{E}[W_{\theta,12}W_{\theta,32}H_2]\\
&\quad - n(n-1)(4n-6)(\mathbb{E}[W_{\theta,12}H_2]\mathbb{E}[W_{\theta,12}]\Big].
\end{aligned}
$$

By direct calculations, we get

$$
\begin{cases}
\mathbb{E}[W_{\theta,12}^2H_2] = \mathbb{E}[W_{\theta,12}W_{\theta,21}H_2] = O(h_K^4h_H\phi_{\theta,x}^2(h_K)),\\
\mathbb{E}[W_{\theta,12}W_{\theta,13}H_2] = \mathbb{E}[W_{\theta,12}W_{\theta,31}H_2]\\
\quad = O(h_K^4h_H\phi_{\theta,x}^3(h_K)),\\
\mathbb{E}[W_{\theta,12}W_{\theta,23}H_2] = \mathbb{E}[W_{\theta,12}W_{\theta,32}H_2]\\
\quad = O(h_K^4h_H\phi_{\theta,x}^3(h_K)).
\end{cases}
$$

Since $\mathbb{E}[W_{\theta,12}] = O(h_K^2\phi_{\theta,x}^2(h_K))$, we obtain

$$
\mathrm{Cov}\left(\widehat{f}_N(\theta,y,x),\widehat{f}_D(\theta,x)\right) = O\left(\frac{1}{n\phi_{\theta,x}(h_K)}\right). \qquad \blacksquare
$$

Proof of Lemma 3.5: We have that

$$
\mathrm{Var}(\widehat{f}_D(\theta,x)) = \frac{1}{(n(n-1)\mathbb{E}[W_{\theta,12}])^2}\mathrm{Var}\left(\sum_{1\le i\ne j\le n} W_{\theta,ij}\right).
$$

Similarly to Lemma 3.3 proofs, we get

$$
\mathrm{Var}\left(\widehat{f}_D(\theta,x)\right) = \frac{\mathbb{E}[K_1^2]}{n(\mathbb{E}[K_1])^2} + o\left(\frac{1}{n\phi_{\theta,x}(h_K)}\right).
$$

We have as $n\longrightarrow\infty$ $\frac{1}{\phi_{\theta,x}(h_K)}\mathbb{E}[K_1^j]\longrightarrow M_j, j=1,2$ (see Ferraty et al., 2007) then, we can write finally

$$
\begin{aligned}
\mathrm{Var}\left(\widehat{f}_D(\theta,x)\right) &= \frac{M_2\phi_{\theta,x}(h_K)}{n(M_1\phi_{\theta,x}(h_K))^2} + o\left(\frac{1}{n\phi_{\theta,x}(h_K)}\right)\\
&= O\left(\frac{1}{n\phi_{\theta,x}(h_K)}\right). \qquad \blacksquare
\end{aligned}
$$

Proof of Lemma 3.8: We have

$$
\begin{aligned}
&\sqrt{nh_H\phi_{\theta,x}(h_K)}Q_n(\theta,x,y)\\
&\quad = \frac{\sqrt{nh_H\phi_{\theta,x}(h_K)}}{n\mathbb{E}(\Omega_1K_1)}\left(\sum_{j=1}^n \Omega_jK_j(H_j - f(\theta,y,x)\right)\\
&\qquad - \mathbb{E}\left(\sum_{j=1}^n \Omega_jK_j(H_j - f(\theta,y,x))\right)
\end{aligned}
$$

then, combined with (4) implies that

$$
\begin{aligned}
&\sqrt{nh_H\phi_{\theta,x}(h_K)}Q_n(\theta,x,y)\\
&\quad = \frac{1}{n\mathbb{E}(\beta_1^2K_1)}\sum_{i=1}^n \beta_i^2K_i\frac{\sqrt{nh_H\phi_{\theta,x}(h_K)}\mathbb{E}(\beta_1^2K_1)}{\mathbb{E}(\Omega_1K_1)}\\
&\qquad \times \sum_{j=1}^n K_j(H_j - f(\theta,y,x))\\
&\qquad - \frac{1}{n\mathbb{E}(\beta_1K_1)}\sum_{i=1}^n \beta_iK_i\frac{\sqrt{nh_H\phi_{\theta,x}(h_K)}\mathbb{E}(\beta_1K_1)}{\mathbb{E}(\Omega_1K_1)}\\
&\qquad \times \sum_{j=1}^n \beta_jK_j(H_j - f(\theta,y,x))\\
&\qquad - \mathbb{E}\left(\frac{1}{n\mathbb{E}(\beta_1^2K_1)}\sum_{i=1}^n \beta_i^2K_i\frac{\sqrt{nh_H\phi_{\theta,x}(h_K)}\mathbb{E}(\beta_1^2K_1)}{\mathbb{E}(\Omega_1K_1)}\right.\\
&\qquad \left.\times \sum_{j=1}^n K_j(H_j - f(\theta,y,x))\right)\\
&\qquad + \mathbb{E}\left(\frac{1}{n\mathbb{E}(\beta_1K_1)}\sum_{i=1}^n \beta_iK_i\frac{\sqrt{nh_H\phi_{\theta,x}(h_K)}\mathbb{E}(\beta_1K_1)}{\mathbb{E}(\Omega_1K_1)}\right.\\
&\qquad \left.\times \sum_{j=1}^n \beta_jK_j(H_j - f(\theta,y,x))\right)
\end{aligned}
$$

Denote by

$$
\begin{aligned}
S_1 &= \frac{1}{n\mathbb{E}(\beta_1^2K_1)}\sum_{i=1}^n \beta_i^2K_i,\\
S_2 &= \frac{\sqrt{nh_H\phi_{\theta,x}(h_K)}\mathbb{E}(\beta_1^2K_1)}{\mathbb{E}(\Omega_1K_1)}\sum_{j=1}^n K_j(H_j - f(\theta,y,x))\\
S_3 &= \frac{1}{n\mathbb{E}(\beta_1K_1)}\sum_{i=1}^n \beta_iK_i \quad \text{and}\\
S_4 &= \frac{\sqrt{nh_H\phi_{\theta,x}(h_K)}\mathbb{E}(\beta_1K_1)}{\mathbb{E}(\Omega_1K_1)}\sum_{j=1}^n \beta_jK_j(H_j - f(\theta,y,x))
\end{aligned}
$$

It remains to show that,

$$
\begin{aligned}
&\sqrt{nh_H\phi_{\theta,x}(h_K)}Q_n(\theta,x,y)\\
&\quad = S_1S_2 - S_3S_4 - \mathbb{E}(S_1S_2 - S_3S_4)\\
&\quad = (S_1S_2 - \mathbb{E}(S_1S_2)) - (S_3S_4 - \mathbb{E}(S_3S_4))
\end{aligned}
$$

Hence by Slutsky's theorem, to show (A3), it suffices to prove the following two claims:

$$
S_1S_2 - \mathbb{E}(S_1S_2) \xrightarrow{D} \mathcal{N}(0, V_{HK}(\theta,x,y)) \tag{A6}
$$

$$
S_3S_4 - \mathbb{E}(S_3S_4) \xrightarrow{P} 0, \tag{A7}
$$

Proof of (A6) We can write that

$$
\begin{aligned}
&S_1S_2 - \mathbb{E}(S_1S_2)\\
&\quad = S_2 - \mathbb{E}(S_2) + (S_1-1)S_2 - \mathbb{E}((S_1-1)S_2).
\end{aligned}
$$

by Slutsky's theorem, we get the following intermediate results:

$$
(S_1-1)S_2 - \mathbb{E}((S_1-1)S_2) \xrightarrow{P} 0 \tag{A8}
$$

and

$$
S_2 - \mathbb{E}(S_2) \xrightarrow{D} \mathcal{N}(0, V_{HK}(\theta,x,y)) \tag{A9}
$$

Concerning the proof of (A8), by applying the Bienaymé–Tchebychv's inequality, we obtain for all $\epsilon > 0$

$$
\begin{aligned}
&\mathbb{P}(|(S_1-1)S_2 - \mathbb{E}((S_1-1)S_2)| > \epsilon)\\
&\quad \le \frac{\mathbb{E}(|(S_1-1)S_2 - \mathbb{E}(S_1-1)S_2)|)}{\epsilon}.
\end{aligned}
$$

Then, the Cauchy–Schwarz inequality implies that

$$
\begin{aligned}
&\mathbb{E}(|(S_1-1)S_2 - \mathbb{E}((S_1-1)S_2)|)\\
&\quad \le 2\mathbb{E}(|(S_1-1)S_2)|) \le 2\sqrt{\mathbb{E}((S_1-1)^2)}\sqrt{\mathbb{E}((S_2)^2)}
\end{aligned}
$$

On one side, by using (A1) and (A2), we obtain

$$\begin{aligned}\mathbb{E}((S_1-1)^2) = Var(S_1) &= \frac{1}{n^2\mathbb{E}^2(\beta_1^2K_1)}n\mathrm{Var}(\beta_1^2K_1)\\ &\le \frac{1}{nO(h^4\phi^2_{\theta,x}(h_K))}\mathbb{E}(\beta_1^4K_1^2)\\ &= O\left(\frac{1}{nh_H\phi_{\theta,x}(h_K)}\right).\end{aligned}$$

and on the other side, we obtain

$$\begin{aligned}\mathbb{E}((S_2)^2) &= \frac{nh_H\phi_{\theta,x}(h_K)\mathbb{E}^2(\beta_1^2K_1)}{\mathbb{E}^2(\Omega_1K_1)}\\ &\quad\times\mathbb{E}\left(\sum_{j=1}^n K_j(H_j-f(\theta,y,x))\right)^2\\ &= \frac{n}{(n-1)^2O(\phi_{\theta,x}(h_K))}(nO(\phi_{\theta,x}(h_K))\\ &\quad+n(n-1)o(\phi^2_{\theta,x}(h_K)))\\ &= O(1)+o(n\phi^2_{\theta,x}(h_K))\end{aligned}$$

Thus

$$\begin{aligned}&\mathbb{E}(|(S_1-1)S_2-\mathbb{E}((S_1-1)S_2)|)\\ &\le 2\sqrt{\mathbb{E}((S_1-1)^2)}\sqrt{\mathbb{E}((S_2)^2)}\\ &\le 2\sqrt{O\left(\frac{1}{nh_H\phi_{\theta,x}(h_K)}\right)(O(1)+o(nh_H\phi_{\theta,x}(h_K)))}\\ &= o(1),\end{aligned}$$

which implies that $(S_1-1)S_2-\mathbb{E}(S_1-1)S_2)=o_p(1)$. Then, as $n\to\infty$, we get

$$\begin{aligned}&\mathbb{P}(|(S_1-1)S_2-\mathbb{E}(S_1-1)S_2)|)>\epsilon)\\ &\le\frac{\mathbb{E}(|(S_1-1)S_2-\mathbb{E}(S_1-1)S_2)|)}{\epsilon}\to 0.\end{aligned}$$

Concerning the proof of (A9)

$$\begin{aligned}P_n &= S_2-\mathbb{E}(S_2)\\ &= \frac{\sqrt{nh_H\phi_{\theta,x}(h_K)}\mathbb{E}(\beta_1^2K_1)}{\mathbb{E}(\Omega_1K_1)}\sum_{j=1}^n K_j(H_j-f(\theta,y,x))\\ &\quad-\mathbb{E}(K_j(H_j-f(\theta,y,x)))\\ &= \frac{\sqrt{nh_H\phi_{\theta,x}(h_K)}\mathbb{E}(\beta_1^2K_1)}{\mathbb{E}(\Omega_1K_1)}\sum_{j=1}^n\mu_{nj}(x,y),\end{aligned}$$

where

$$\mu_{nj}(x,y)=K_j(H_j-f(\theta,y,x))-\mathbb{E}(K_j(H_j-f(\theta,y,x)))$$

By the fact that $\mu_{nj}(x,y)$ are i.i.d., it follows that

$$\begin{aligned}var(P_n(x,y)) &= \frac{n^2\phi_{\theta,x}(h_K)\mathbb{E}^2(\beta_1^2K_1)}{\mathbb{E}^2(\Omega_1K_1)}var(\mu_{n1}(x,y))\\ &= \frac{n^2\phi_{\theta,x}(h_K)\mathbb{E}^2(\beta_1^2K_1)}{\mathbb{E}^2(\Omega_1K_1)}\mathbb{E}(\mu_{n1}^2(x,y))\end{aligned}$$

Thus

$$\begin{aligned}&var(P_n(x,y))\\ &= \frac{n^2\phi_{\theta,x}(h_K)\mathbb{E}^2(\beta_1^2K_1)}{\mathbb{E}^2(\Omega_1K_1)}(\mathbb{E}(K_1^2(H_1-f(\theta,y,x))^2)\\ &\quad-(\mathbb{E}(K_1(H_1-f(\theta,y,x)))^2).\end{aligned}\tag{A10}$$

Concerning the second term on the right-hand side of (A10), we have

$$\begin{aligned}&(\mathbb{E}(K_1(H_1-f(\theta,y,x)))^2\\ &\quad=(\mathbb{E}(\mathbb{E}(K_1(H_1-f(\theta,y,x))|X_1))^2\\ &\quad=(\mathbb{E}(K_1\mathbb{E}((H_1|X_1)-f(\theta,y,x))))^2,\end{aligned}$$

where

$$\mathbb{E}((H_1|X_1)-f(\theta,y,x))\to 0\quad\text{as } n\to\infty\tag{A11}$$

Now let us return to the first term of the right hand of (A10). We have

$$\begin{aligned}&\frac{n^2\phi_{\theta,x}(h_K)\mathbb{E}^2(\beta_1^2K_1)}{\mathbb{E}^2(\Omega_1K_1)}(\mathbb{E}(K_1^2(H_1-f(\theta,y,x))^2)\\ &=\frac{n^2\phi_{\theta,x}(h_K)\mathbb{E}^2(\beta_1^2K_1)}{\mathbb{E}^2(\Omega_1K_1)}(\mathbb{E}(\mathbb{E}((H_1-f(\theta,y,x))^2|X_1)K_1^2)\\ &=\frac{n^2\phi_{\theta,x}(h_K)\mathbb{E}^2(\beta_1^2K_1)}{\mathbb{E}^2(\Omega_1K_1)}\mathbb{E}(var(H_1|X_1)K_1^2)\\ &\quad+\frac{n^2\phi_{\theta,x}(h_K)\mathbb{E}^2(\beta_1^2K_1)}{\mathbb{E}^2(\Omega_1K_1)}\\ &\quad\times(\mathbb{E}(\mathbb{E}((H_1|X_1)-f(\theta,y,x))^2)K_1^2)\end{aligned}$$

By using (A9), that allows to have, as $n\to\infty$

$$\frac{n^2\phi_{\theta,x}(h_K)\mathbb{E}^2(\beta_1^2K_1)}{\mathbb{E}^2(\Omega_1K_1)}(\mathbb{E}(\mathbb{E}((H_1|X_1)-f(\theta,y,x))^2)K_1^2)\to 0$$

Combining (5) and (6), we obtain as $n\to\infty$

$$\begin{aligned}\mathbb{E}(var(H_1|X_1)K_1^2)&\to\mathbb{E}(K_1^2)f(\theta,y,x)\left(\int H^2(t)\mathrm{d}t\right)\\ &= M_2f(\theta,y,x)\left(\int H^2(t)\mathrm{d}t\right)\phi_{\theta,x}(h_K).\end{aligned}$$

Therefore, by using (5) and (6), Equation (A10) becomes

$$\begin{aligned}&\mathrm{Var}(P_n(x,y))\\ &= \frac{n^2\phi_{\theta,x}(h_K)(N(1,2)h_K^2\phi_{\theta,x}(h_K))^2}{((n-1)N(1,2)M_1h_K^2\phi_{\theta,x}(h_K))^2}M_2f(\theta,y,x)\\ &\quad\times\left(\int H^2(t)\mathrm{d}t\right)\phi_{\theta,x}(h_K)\\ &= \frac{n^2M_2}{(n-1)^2M_1^2}f(\theta,y,x)\left(\int H^2(t)\mathrm{d}t\right)\\ &\to \frac{M_2}{M_1^2}f(\theta,y,x)\left(\int H^2(t)\mathrm{d}t\right)\\ &= V_{HK}(\theta,x,y)\quad\text{as } n\to\infty\end{aligned}$$

Now, in order to end the proof of (A6), we focus on the central limit theorem. So, the proof of (13) is completed if Lindberg's condition is verified. In fact, Lindberg's condition holds since, for any $\eta>0$

$$\begin{aligned}\sum_{j=1}^n\mathbb{E}(\mu_{nj}^2\mathbb{1}_{(|\mu_{nj}|>\eta)}) &= n\mathbb{E}(\mu_{n1}^2\mathbb{1}_{(|\mu_{n1}|>\eta)})\\ &= \mathbb{E}((\sqrt{n}\mu_{n1})^2\mathbb{1}_{(|\sqrt{n}\mu_{n1}|>\sqrt{n}\eta)})\end{aligned}$$

as

$$\mathbb{E}((\sqrt{n}\mu_{n1})^2)=n\mathbb{E}(\mu_{n1}^2)\to\frac{M_2}{M_1^2}f(\theta,y,x)\left(\int H^2(t)\mathrm{d}t\right).$$

Proofs of (A7). To use the same arguments as those invoked to prove (A6), let us write

$$\begin{aligned} &S_3S_4 - \mathbb{E}(S_3S_4) \\ &\quad = S_4 - \mathbb{E}(S_4) + (S_3 - 1)S_4 - \mathbb{E}(S_3 - 1)S_4)). \end{aligned}$$

By applying Bienaymé–Tchebychv's inequality, we obtain for all $\epsilon > 0$

$$\mathbb{P}(|S_3S_4 - \mathbb{E}(S_3S_4)|) > \epsilon) \leq \frac{\mathbb{E}(|S_3S_4 - \mathbb{E}(S_3S_4)|)}{\epsilon}.$$

and the Cauchy–Schwarz inequality implies that

$$\begin{aligned} &\mathbb{E}(|(S_3 - 1)S_4 - \mathbb{E}((S_3 - 1)S_4)|) \\ &\quad \leq 2\mathbb{E}(|(S_3 - 1)S_4)|) \leq 2\sqrt{\mathbb{E}((S_3 - 1)^2)}\sqrt{\mathbb{E}((S_4)^2)} \end{aligned}$$

Taking into account hypothesis (5) and (6), we get

$$\begin{aligned} \mathbb{E}((S_3 - 1)^2 = \mathrm{Var}(S_3) &= \frac{n}{n^2\mathbb{E}^2(\beta_1K_1)}var(\beta_1K_1) \\ &\leq \frac{1}{nO(h^4\phi^2_{\theta,x}(h_K))}\mathbb{E}(\beta_1^4K_1^2) \\ &= O\left(\frac{1}{nh_H\phi_{\theta,x}(h_K)}\right). \end{aligned}$$

On the other hand

$$\begin{aligned} \mathbb{E}((S_4)^2) &= \frac{nh_H\phi_{\theta,x}(h_K)\mathbb{E}^2(\beta_1K_1)}{\mathbb{E}^2(\Omega_1K_1)} \\ &\quad \times \mathbb{E}\left(\sum_{j=1}^{n}\beta_jK_j(H_j - f(\theta,y,x))\right)^2 \\ &= \frac{nh_H\phi_{\theta,x}(h_K)O(h_K^2\phi^2_{\theta,x}(h_K))}{(n-1)^2O(h_K^4\phi^4_{\theta,x}(h_K))} \\ &\quad \times (n\mathbb{E}(\beta_1K_1(H_1 - f(\theta,y,x))))^2 \\ &\quad + n(n-1)\mathbb{E}^2(\beta_1K_1(H_1 - f(\theta,y,x))) \\ &= o(1) + o(nh_H\phi_{\theta,x}(h_K)) \end{aligned}$$

It remains to show

$$\begin{aligned} &\mathbb{E}(|(S_3 - 1)S_4 - \mathbb{E}((S_3 - 1)S_4)|) \\ &\quad \leq 2\sqrt{\mathbb{E}((S_3 - 1)^2)}\sqrt{\mathbb{E}((S_4)^2)} = o(1) \end{aligned}$$

which implies that

$$\mathbb{E}(|(S_3 - 1)S_4 - \mathbb{E}((S_3 - 1)S_4)|) = o_p(1)$$

Therefore,

$$\begin{aligned} &\mathbb{P}(|S_3S_4 - \mathbb{E}(S_3S_4)|) > \epsilon) \\ &\quad \leq \frac{\mathbb{E}(|S_3S_4 - \mathbb{E}(S_3S_4)|)}{\epsilon} \to 0 \quad \text{as } n \to \infty. \end{aligned}$$

So, to prove (A6), it suffices to show $S_4 - \mathbb{E}(S_4) = o(1)$, while

$$\begin{aligned} &\mathbb{E}(S_4 - \mathbb{E}(S_4))^2 \\ &\quad = \mathrm{Var}(S_4) \\ &\quad = \frac{n^2\phi_{\theta,x}(h_K)\mathbb{E}^2(\beta_1K_1)}{\mathbb{E}^2(\Omega_1K_1)}var(\beta_1K_1(H_1 - f(\theta,y,x))) \end{aligned}$$

We arrive finally at

$$\begin{aligned} &\mathrm{Var}(\beta_1K_1(H_1 - f(\theta,y,x))) \\ &\quad = f(\theta,y,x)\left(\int H^2(t)\mathrm{d}t\right)\mathbb{E}(\beta_1^2K_1^2) \end{aligned}$$

This last result together with (5) and (6) lead directly to

$$\begin{aligned} &\mathbb{E}(S_4 - \mathbb{E}(S_4))^2 \\ &\quad = \frac{n^2\phi_{\theta,x}(h_K)\mathbb{E}^2(\beta_1K_1)}{\mathbb{E}^2(\Omega_1K_1)}f(\theta,y,x)\left(\int H^2(t)\mathrm{d}t\right)\mathbb{E}(\beta_1^2K_1^2) \\ &\quad = f(\theta,y,x)\left(\int H^2(t)\mathrm{d}t\right)o(1), \end{aligned}$$

which allows to finish the proof of Theorem. ■