REVSTAT – Statistical Journal
Volume 20, Number 5, October 2022, 605–631
https://doi.org/10.57805/revstat.v20i5.391

Single Index Regression Model for Functional Quasi-Associated Time Series Data

Authors: Salim Bouzebda ✉
– Alliance Sorbonne Université, Université de Technologie de Compiègne, L.M.A.C., Compiègne, France
salim.bouzebda@utc.fr

Ali Laksaci
– Department of Mathematics, College of Science, King Khalid University, Abha, 61413, Saudi Arabia
alilak@yahoo.fr

Mustapha Mohammedi
– Université Djillali Liabès, BP 89, 22000, Sidi Bel Abbès, Algérie
mustapha.mohammedi@gmail.com

Received: March 2020 Revised: February 2021 Accepted: February 2021

Abstract:

- The mixing condition is often considered to modeling the functional time series data. Alternatively, in this work we consider the problem of nonparametric estimation of the regression function in Single Functional Index Model (SFIM) under the quasi-association dependence condition. The main result of this work is the establishment of the asymptotic properties of the estimator, such as the almost complete convergence rates. Furthermore, the asymptotic normality of the constructed are obtained under some mild conditions. We finally discuss how to apply our result to construct the confidence intervals. Finally, the finite-sample performances of the model and the estimation method are illustrated using the analysis of simulated data.

Keywords:

- *single functional index model; functional Hilbert space; kernel regression estimation; mixing; weak dependence; quasi-associated variables; almost complete convergence; asymptotic normality.*

AMS Subject Classification:

- 62G05, 62G08, 62L12, 62G20.

✉ Corresponding author.

1. INTRODUCTION

The statistical study of single index models have been investigated and developed by several authors from a practical and theoretical point of view. The case of a vector explanatory variable was studied by [19] and [20]. The single index models are very popular in the econometric community because it respond two important preoccupations. The first concerns dimension reduction since this type of model makes it possible to provide a solution to the problem of the curse of dimensionality, in the sense that pure nonparametric models are highly affected by dimensionality effects while semiparametric ideas are more appealing candidates. The second is related to the interpretability of the index θ introduced in these models, for more details on refer to [8], [18] and [3] for an overview on methodological issues on FDA. Therefore, the single functional index model accumulate the advantages of single index model, and inherits the potential of the functional linear model in terms of applications. The interested reader, for the semiparametric and the nonparametric functional models, may refer [17], [24, 25], [27] and [7] for survey on the topics.

The modelization of functional data, has been developed intensively. The motivation of such statistical analysis is justified by the recent technological development of the measuring instruments that offers the opportunity to observe phenomena in an increasingly accurate way, but this accuracy obviously generates a large amount of data observed over a finer grid, which can be considered as observations varying over a continuum. The most theoretical results are obtained under independence condition. However, in practice, it is rarely that we have an independent identically distributed observations of functional nature. The functional time series presents the more realistic situation. Thus it is really crucial to study the functional statistical models when the usual independence condition on the statistical sample is relaxed. In this paper, we consider the problem of the nonparametric estimation of the regression function in single functional index model when the data are weakly dependant.

Usually the dependence structure is modelled with the strong mixing hypothesis, in this paper we focus in some more general correlation, that is the quasi-associated condition. The latter has been introduced for real valued random fields by [5], which generalizes the positively associated variables introduced in [13].

From practical point of view, this kind of data has great importance in practice, in particular, in reliability theory, mathematical physics and in percolation theory (see, for instance, [28]) for more discussion on the practical interest of these random variables. Moreover, from the theoretical point of view, the concept of quasi-association correlation can be viewed as a particular case of the weak dependence condition for real-valued stochastic processes introduced by [12] which allows treating the mixing condition and association correlation in a unified approach.

Noting that the single index model is a semi-parametric regression model, thus, it couples the advantages of both parametric and nonparametric regression models. Because of these advantages, it has received an increasing amount of attention in the nonparametric regression literature. Key references on this topic in multivariate statistic are [21] and [20] for previous results and [30] for more recent advances and references.

However, in the literature of functional statistic, the single functional index model is strictly limited in the case where the data is functional (a curve). The first result in this context, was given by [15]. They obtained the almost complete convergence of the regression function $r(\cdot)$ in the independent and identically distributed (i.i.d.) case. The generalization of this result to the dependent case has been studied by [26]. [29] uses a Bayesian method to estimate the bandwidths in the kernel form error density and regression function, under an autoregressive error structure, and according to empirical studies, the author considered that the single functional index model gives improved estimation and prediction accuracies compared to any nonparametric functional regression considered. [27] have proposed a new automatic and location-adaptive procedure for estimating regression in a Functional Single-Index Model (FSIM) based on k-Nearest Neighbours ideas. Motivated by the analysis of imaging data, [23] proposed a novel functional varying-coefficient single-index model to carry out the regression analysis of functional response data on a set of covariates of interest. This method represents a new extension of varying-coefficient single-index models for scalar responses collected from cross-sectional and longitudinal studies. By simulation and real data analysis, the authors demonstrated the advantages of the proposed estimate. [31] have considered the problem of predicting the real-valued response variable using explanatory variables containing both multivariate random variable and random curve. The authors considered the functional partial linear single-index model in order to treat the multivariate random variable as linear part and the random curve as functional single-index part, respectively.

The concept of quasi-association for random variables taking its values in a Hilbert space has been investigated by [10], and obtained some limit theorems for this type of variables. More recently, [11] studied the asymptotic normality of regression function under quasi-associated data when the explanatory variable takes its values in a Hilbert space.

The main purpose of the present paper is to establish the asymptotic properties of the estimator $\widehat{r}_\theta(\cdot)$, when the variables are functional quasi-associated and in single index structure, such as the almost complete convergence rates. Furthermore, the asymptotic distribution is obtained under some mild conditions.

We point out that the mixing and the association concern two distinct classes of processes but not disjoint and offer two complementary approaches to study the dependence. Moreover, the functional quasi-associated data analysis has great importance in various domains such as the reliability theory or the statistical mechanics. Furthermore, it should be noted that the dependence condition considered here allow to avoid the widely used strong mixing condition which is very easy to verified in practice.

The rest of this work is organized as follows. In Section 2, we describe the single index regression model for functional data and in the quasi-associated framework, the next section is devoted to the introduction of the notation and hypotheses needed to state our main results. In Section 4, we will establish our main results of the almost complete convergence of the kernel estimators and the asymptotic normality under non restrictive conditions. In Section 5.2, we discuss the impact of our contribution in practice application of our results for the construction of the confidence interval. In Section 6 we perform a short simulation study to show that our proposed model works well for finite samples. To avoid interrupting the flow of the presentation, all mathematical developments are relegated to the Section 7.

2. MODEL AND ESTIMATOR

We start by giving a definition of quasi-association adapted to the functional framework. In the real valued random fields, [5] define the quasi-association dependence in the Definition 2.1 and it adapted to functional random variables in the Definition 2.2 given in [10] as follows.

Definition 2.1. A sequence $(X_n)_{n\in\mathbb{N}}$ of r.v.'s is said to be quasi-associated, if for any disjoint subsets I and J of $\mathbb{N}$ and all bounded Lipschitz functions $f_1\colon \mathbb{R}^{|I|} \to \mathbb{R}$ and $f_2\colon \mathbb{R}^{|J|} \to \mathbb{R}$ satisfying:

$$\left|\mathrm{Cov}\Big(f_1(X_i,\, i\in I),\, f_2(X_j,\, j\in J)\Big)\right| \leq \mathrm{Lip}(f_1)\ \mathrm{Lip}(f_2) \sum_{i\in I}\sum_{j\in J}\big|\mathrm{Cov}(X_i,X_j)\big|\,, \tag{2.1}$$

where $|I|$ denotes cardinality of a finite set I, and the Lipschitz of a function $f(\cdot)$ is defined by

$$\mathrm{Lip}(f) = \sup_{x\neq y} \frac{\big|f(x)-f(y)\big|}{\|x-y\|_1}\,, \qquad \text{with} \quad \|(x_1,...,x_k)\|_1 = \sum_{k=1}^{n} |x_k|\,.$$

Definition 2.2. A sequence $(X_i)_{i\in\mathbb{N}}$ of r.v.'s taking values in a Hilbert space H is called quasi-associated relative to an orthonormal basis $\{e_p\colon p\geq 1\}$ of H, if for any $p\geq 1$, $(\langle X_i,e_1\rangle,...,\langle X_i,e_p\rangle)_{i\in\mathbb{N}}$ is a sequence of random vectors quasi-associates.

Now, we consider a sequence of quasi-associated random variables $\{(X_i,Y_i)\}_{i\in\mathbb{N}}$ identically distributed as (X,Y), which are valued in $H\times\mathbb{R}$, where H is a separable real Hilbert space with inner product $\langle\cdot,\cdot\rangle$ and a orthonormal basis $\{e_p\colon p\geq 1\}$. We consider the semi-metric $d_\theta(\cdot,\cdot)$ associated to the single-index $\theta\in H$ defined by $\forall\, u,v\in H$:

$$d_\theta(u,v) := |\langle\theta, u-v\rangle|\,.$$

The purpose of this paper is to study the estimation of the nonparametric regression of Y given $\langle\theta,X\rangle$ structure, denoted by

$$r(\langle\theta, X_i = x\rangle) = \mathbb{E}\big(Y \mid \langle\theta, X_i = x\rangle\big)\,. \tag{2.2}$$

Such structure suppose that the explanation of Y from X is done through an fixed functional index θ in Θ. Now, we suppose that exists a $\theta\in\Theta\subset H$ where the observations $(X_i,Y_i)_{i=1,...,n}$ are related by the following relation:

$$Y_i = r(\langle\theta,X_i\rangle) + \varepsilon_i\,, \qquad \forall\, i=1,...,n\,, \tag{2.3}$$

where $r(\cdot)$ is a real function, and for $i=1,...,n$, ε_i is a real random variable such that $\mathbb{E}(\varepsilon_i \mid X_i)=0$. We consider that the single functional index model is identifiable, i.e., if the regression function is differentiable and if $\langle\theta,e_1\rangle = 1$, where e_1 is the first element of an orthonormal basis of H. Then, if $r_1(\langle\theta_1,x\rangle) = r_2(\langle\theta_2,x\rangle)$ implies that $r_1\equiv r_2$ and $\theta_1\equiv\theta_2$.

This hypothesis that we consider is demonstrated by [15] once we have the differentiability of the regression operator $r(\cdot)$. For more details on the problem of identifiability of the single functional index model, one can refer to the last reference. The kernel estimator $\widehat{r}_\theta(\cdot)$ of regression operator $r_\theta(\cdot) = r(\langle\theta,\cdot\rangle)$ is defined by

$$\widehat{r}_{\theta,n}(x) = \frac{\sum_{i=1}^{n} Y_i K_i(x)}{\sum_{i=1}^{n} K_i(x)}, \qquad \text{for all} \quad x \in H\,, \tag{2.4}$$

where $K_i(x) := K\Big(\frac{\langle\theta, x-X_i\rangle}{h_n}\Big)$ is the kernel function and h_n is the bandwidth parameter decreases to zero as n goes to infinity.

3. ASSUMPTIONS AND NOTATION

In the sequel, we will denote by C and/or C' some strictly positive constants and by λ_r the covariance coefficient defined as:

$$\lambda_r := \sup_{s\geq r} \sum_{|i-j|\geq s} \lambda_{i,j}\,,$$

where

$$\lambda_{i,j} = \sum_{k\geq 1}\sum_{l\geq 1} \big|\mathrm{Cov}(X_i^k, X_j^l)\big| + \sum_{k\geq 1}\big|\mathrm{Cov}(X_i^k, Y_j)\big| + \sum_{l\geq 1}\big|\mathrm{Cov}(Y_i, X_j^l)\big| + \big|\mathrm{Cov}(Y_i, Y_j)\big|\,,$$

with $X_i^p := \langle X_i, e_p\rangle$. In our analysis, we shall assume the following assumptions:

(H₁) Let $E_i(x) := \langle\theta, x-X_i\rangle$ so that $E_i(x)$ is a real-valued random variable,

$$G_\theta(x,h_n) := \mathbb{P}\big(|E_i(x)| \leq h_n\big) > 0\,,$$

and $G_\theta(x,\cdot)$ is differentiable at 0.

(H₂) The random pair $\{(X_i, Y_i),\ i \in \mathbb{N}\}$ is quasi-associated such that:

(i) The covariance coefficient satisfies

$$\lambda_k \leq C e^{-ak} \quad \text{for some } a > 0\,, \qquad C > 0\,;$$

(ii) The process $(X_i)_i$ satisfies

$$\max_{i\neq j}\Big\{\mathbb{P}\Big(|E_i| \leq h_n,\, |E_j| \leq h_n\Big)\Big\} := \psi_\theta(x,h_n) > 0\,,$$

where $\psi_\theta(x,\cdot)$ is differentiable at 0;

(iii) The response observations $(Y_i)_i$ are such that, almost surely

$$\forall\, i \neq j \quad \mathbb{E}\big(|Y_i Y_j| \mid X_i, X_j\big) \leq C < \infty$$
$$\text{and} \quad \mathbb{E}\big(|Y|^p \mid X = x\big) \leq C < \infty \quad \text{for } p > 4\,.$$

($\mathbf{H_3}$) For all $u, v \in H$ we have

$$\left|r_\theta(u) - r_\theta(v)\right| \leq C\left|\langle\theta, u-v\rangle\right|^\beta, \qquad \text{for certain} \quad \beta > 0\,.$$

($\mathbf{H_4}$) The kernel $K(\cdot)$ is a Lipschitzian function on $[0,1]$ such that

$$C\,\mathbb{1}_{[0,1]}(t) < K(t) < C'\,\mathbb{1}_{[0,1]}(t)\,.$$

($\mathbf{H_5}$) There exists a sequence of positive real numbers δ_n such that

$$\begin{cases} \delta_n^{p-2}\,\chi_\theta^{(p-4)/2p}(x,h_n) \,\to\, 0\,, \\ \displaystyle\sum_n n\,\delta_n^{-p} < \infty\,, \end{cases}$$

where $\chi_\theta(x,h_n) = \max\big(\psi_\theta(x,h_n), G_\theta^2(x,h_n)\big)$ and p is given in (H_2).

Some comments on the assumptions

All the assumptions are standard in this context of semiparametric functional data analysis. The concentration property of the explanatory variable in small balls under single index topological structure is defined in the assumption (H_1). The quasi-association features of the underlying functional time series is explored through the condition (H_2). It covers the three fundamental aspects of the considered process. The correlation's level of the data is quantified by the geometric form of the covariance coefficient λ_k, while the local dependency of the data is expressed by the function $\psi_\theta(x,h_n)$ allowing to emphasize the functional component of the time series $(X_i)_i$. It should be noted that the conditional moments integrability in (H_2)(iii) is usual in the regression data analysis. It was used by [16] for the nonparametric case and by [1] in the single functional index case. It is less restrictive than the exponential version assumed by [10]. Finally, let us mention that the hypothesis (H_3) is used to control the regularity condition of the link function with respect the single index. This kind of assumption is needed to evaluate the bias in the asymptotic results of this paper, while the conditions (H_4) and (H_5) are classical technical assumptions in NFDA.

4. MAIN RESULTS

4.1. The almost consistency

Our aim is to establish the almost complete convergence (a.co.)[1] of $\widehat{r}_\theta(x)$ to $r_\theta(x)$, and the main result is given by the following theorem.

Theorem 4.1. *Under the assumptions* (H$_1$)–(H$_5$)*, we have, as* $n \to \infty$,

$$\widehat{r}_{\theta,n}(x) - r_\theta(x) \;=\; O_{\text{a.co.}}\left(h_n^\beta + \sqrt{\frac{\chi_\theta^{1/2}(x,h_n)\,\log n}{n\,G_\theta^2(x,h_n)}} \right). \tag{4.1}$$

Let

$$\widehat{r}_{\theta,0}(x) := \frac{1}{n\,\mathbb{E}\,K_1(x)} \sum_{i=1}^n K_i(x) \quad \text{and} \quad \widehat{r}_{\theta,1}(x) := \frac{1}{n\,\mathbb{E}\,K_1(x)} \sum_{i=1}^n Y_i\,K_i(x)\,. \tag{4.2}$$

Let us consider the following decomposition:

$$\begin{aligned}
\widehat{r}_{\theta,n}(x) - r_\theta(x) &= \frac{\widehat{r}_{\theta,1}(x)}{\widehat{r}_{\theta,0}(x)} - r_\theta(x) \\
&= \frac{1}{\widehat{r}_{\theta,0}(x)} \Big[\big(\widehat{r}_{\theta,1}(x) - \mathbb{E}(\widehat{r}_{\theta,1})\big) - \big(r_\theta(x) - \mathbb{E}(\widehat{r}_{\theta,1})\big)\Big] - \frac{r_\theta(x)}{\widehat{r}_{\theta,0}(x)}\big(\widehat{r}_{\theta,0} - 1\big) \\
&= \frac{1}{\widehat{r}_{\theta,0}(x)} \Big[\big(\widehat{r}_{\theta,1}(x) - \widehat{r}_{\theta,2}(x)\big) + \big(\widehat{r}_{\theta,2}(x) - \mathbb{E}(\widehat{r}_{\theta,2})\big)\Big] \\
&\quad + \frac{1}{\widehat{r}_{\theta,0}(x)} \Big[\big(\mathbb{E}(\widehat{r}_{\theta,2}(x)) - \mathbb{E}(\widehat{r}_{\theta,1})\big) - \big(r_\theta(x) - \mathbb{E}(\widehat{r}_{\theta,1})\big)\Big] - \frac{r_\theta(x)}{\widehat{r}_{\theta,0}(x)}\big(\widehat{r}_{\theta,0} - 1\big)\,,
\end{aligned}$$

where

$$\widehat{r}_{\theta,2}(x) := \frac{1}{n\,\mathbb{E}\,K_1(x)} \sum_{i=1}^n \widehat{Y}_i\,K_i(x)\,. \tag{4.3}$$

The real variable Y response is not necessarily bounded. For this, we introduce the truncated random variable $\widehat{Y}$, defined by $\widehat{Y}_i = Y_i\,\mathbb{1}_{\{|Y_i| \le \delta_n\}}$. The proof of the Theorem 4.1 is based on the following Lemmas.

[1] We say that the sequence $(\Theta_n)_n$ converges a.co. to zero, if and only if

$$\forall\,\tau > 0\,, \qquad \sum_{n\ge1} \mathbb{P}\big(|\Theta_n| > \tau\big) < \infty\,.$$

Furthermore, we say that $\Theta_n = O_{\text{a.co.}}(\theta_n)$ if there exists $\tau_0 > 0$ such that

$$\sum_{n\ge1} \mathbb{P}\big(|\Theta_n| > \tau_0\,\theta_n\big) < \infty\,.$$

Lemma 4.1 (See [22]). *Let $X_1, \dots, X_n$ be real random variables such that $\mathbb{E} X_i = 0$ and $\mathbb{P}\big(|X_i| \le M\big) = 1$, for all $i = 1, \dots, n$ and some $M < \infty$. Let*

$$\sigma_n^2 = \operatorname{Var}\left(\sum_{i=1}^{n} X_i\right).$$

Assume, furthermore, that there exist $K < \infty$ and $\beta > 0$ such that, for all u-tuplets $(s_1, \dots, s_u)$ and all v-tuplets $(t_1, \dots, t_v)$ with $1 \le s_1 \le \cdots \le s_u \le t_1 \le \cdots \le t_v \le n$, the following inequality is fulfilled:

$$\left|\operatorname{Cov}\big(X_{s_1} \dots X_{s_u}, X_{t_1} \dots X_{t_v}\big)\right| \le K^2 M^{u+v-2}\, v\, e^{-\beta(t_1 - s_u)}.$$

Then,

$$\mathbb{P}\left(\left|\sum_{i=1}^{n} X_i\right| \ge t\right) \le \exp\left\{-\frac{t^2/2}{A_n + B_n^{\frac{1}{3}}\, t^{\frac{5}{3}}}\right\},$$

for $A_n \le \sigma_n^2$ and

$$B_n = \left(\frac{16\, n K^2}{9 A_n (1 - e^{-\beta})} \vee 1\right)\left(\frac{2(K \vee M)}{1 - e^{-\beta}}\right).$$

Lemma 4.2. *Under the assumptions (H$_1$)–(H$_5$), we have, as $n \to \infty$,*

$$\left|\widehat{r}_{\theta,2}(x) - \mathbb{E}(\widehat{r}_{\theta,2})\right| = O_{\text{a.co.}}\left(\sqrt{\frac{\chi_\theta^{1/2}(x, h_n) \log n}{n\, G_\theta^2(x, h_n)}}\right). \tag{4.4}$$

Lemma 4.3. *Under the assumptions (H$_1$), (H$_2$)(i,ii)–(H$_5$), we have, as $n \to \infty$,*

$$\left|\widehat{r}_{\theta,0}(x) - 1\right| = O_{\text{a.co.}}\left(\sqrt{\frac{\chi_\theta^{1/2}(x, h_n) \log n}{n\, G_\theta^2(x, h_n)}}\right). \tag{4.5}$$

Lemma 4.4. *Under the assumptions of Lemma 4.3, we have, as $n \to \infty$,*

$$\exists\, \eta > 0 \quad \textit{such that} \quad \sum_{i=1}^{n} \mathbb{P}\Big(\left|\widehat{r}_{\theta,0}(x)\right| < \eta\Big) < \infty. \tag{4.6}$$

Lemma 4.5. *Under the assumptions (H$_1$), (H$_4$)–(H$_5$), we have, as $n \to \infty$,*

$$\left|r_\theta(x) - \mathbb{E}(\widehat{r}_{\theta,1})\right| = O\big(h_n^{\beta}\big). \tag{4.7}$$

Lemma 4.6. *Under the assumptions (H$_1$), (H$_3$)–(H$_5$), we have, as $n \to \infty$,*

$$\left|\mathbb{E}(\widehat{r}_{\theta,2}) - \mathbb{E}(\widehat{r}_{\theta,1})\right| = O\left(\sqrt{\frac{\chi_\theta^{1/2}(x, h_n) \log n}{n\, G_\theta^2(x, h_n)}}\right). \tag{4.8}$$

Lemma 4.7. *Under the assumptions (H$_1$), (H$_2$)(iii)–(H$_5$), we have, as $n \to \infty$,*

$$\left|\widehat{r}_{\theta,1}(x) - \widehat{r}_{\theta,2}(x)\right| = O_{\text{a.co.}}\left(\sqrt{\frac{\chi_\theta^{1/2}(x, h_n) \log n}{n\, G_\theta^2(x, h_n)}}\right). \tag{4.9}$$

4.2. The asymptotic normality

Now, we study the asymptotic normality of $\widehat{r}_\theta(x)$. To do that, we assume that the function

$$\varphi_\theta(x) := \mathbb{E}\big(Y_1^2 \mid \langle\theta, X_1 = z\rangle\big)\,, \qquad z \in H\,,$$

exists and is uniformly continuous in some neighborhood of z. Moreover, we modify slightly the assumptions (H$_1$), (H$_4$) and (H$'_5$) is required:

(H$'_1$) The concentration property (H$_1$) holds. Moreover, there exists a function $\beta_x(\cdot)$ such that

$$\forall\, s \in [0,1]\,, \quad \lim_{h_n\to 0} G_\theta(x, s\,h_n)/G_\theta(x,h_n) = \beta_x(s)\,.$$

(H$'_4$) The kernel $K(\cdot)$ satisfies (H$_3$) and is a differentiable function on $]0,1[$ with derivative $K'(\cdot)$ such that $-\infty < C < K'(\cdot) < C' < 0$.

(H$'_5$) There exists a sequence of positive real numbers γ_n such that

$$\begin{cases} \gamma_n\,\chi_\theta(x,h_n) \to 0\,, \\ n^{3/2}\,\chi_\theta^{p/p-2}(x,h_n) \to 0\,. \end{cases}$$

Theorem 4.2. *Under the assumptions* (H$'_1$)–(H$_2$), (H$_3$), (H$'_4$), (H$'_5$) *and if*

$$n\,h_n^{2\beta}\,G_\theta(x,h_n) \longrightarrow 0\,,$$

we have, for all $x \in \mathcal{A}$,

$$\sqrt{n\,G_\theta(x,h_n)}\,\big(\widehat{r}_{\theta,n}(x) - r_\theta(x)\big) \xrightarrow{\mathcal{D}}{}^{2}\; N\big(0, \sigma_\theta^2(x)\big)\,, \qquad \text{as} \;\; n\to\infty\,, \tag{4.10}$$

where

$$\sigma_\theta^2(x) = \frac{\beta_2\big(\varphi_\theta(x) - r_\theta^2(x)\big)}{\beta_1^2}\,,$$

with

$$\beta_j = -\int_0^1 (K^j)'(s)\,\beta_x(s)\,ds\,, \qquad \text{for} \;\; j = 1,2\,,$$

and

$$\mathcal{A} = \Big\{x\in H\colon\, \sigma_\theta^2(x) \neq 0\Big\}\,.$$

We can use the same decomposition as in the proof of Theorem 4.1, where δ_n is replaced by γ_n in $\widehat{r}_{\theta,2}(x)$. Observe that the consistency of $\widehat{r}_{\theta,0}$ to 1 is shown in Lemma 4.3 and, under the consideration $n\,h_n^{2\beta}G_\theta(x,h_n) \longrightarrow 0$, we get

$$\sqrt{n\,G_\theta(x,h_n)}\,\big(r_\theta(x) - \mathbb{E}(\widehat{r}_{\theta,1})\big) \longrightarrow 0\,.$$

[2] $\xrightarrow{\mathcal{D}}$ denotes the convergence in distribution.

Moreover, by straightforward modification of the proofs of Lemmas 4.7 and 4.6, we obtain, under (H_5'),

$$\sqrt{n\, G_\theta(x,h_n)}\, \left|\widehat{r}_{\theta,1}(x) - \widehat{r}_{\theta,2}(x)\right| \longrightarrow 0\,, \qquad \text{in probability},$$

and

$$\sqrt{n\, G_\theta(x,h_n)}\, \big(\mathbb{E}(\widehat{r}_{\theta,2}) - \mathbb{E}(\widehat{r}_{\theta,1})\big) \longrightarrow 0\,.$$

So, all it remains to show is the following intermediate lemma.

Lemma 4.8. *Under the hypotheses of Theorem 4.2, we have, as $n \longrightarrow \infty$,*

$$\sqrt{n\, G_\theta(x,h_n)}\, \Big(\widehat{r}_{\theta,2}(x) - r_\theta(x)\, \widehat{r}_{\theta,0}(x) - \mathbb{E}\big(\widehat{r}_{\theta,2}(x) - r_\theta(x)\, \widehat{r}_{\theta,0}(x)\big)\Big) \xrightarrow{\mathcal{D}} N\big(0, \sigma_\theta^2(x)\big)\,. \tag{4.11}$$

5. DISCUSSION AND APPLICATIONS

5.1. On the weak functional time series data analysis

The functional time series data analysis is one of the most important subject in functional data analysis (FDA). It is motivated by the rarity of the independent identically distributed observations functional observations in practice. The functional time series presents the more realistic situation. At this stage, the most of the existing studies on functional dependent data are developed under mixing assumption, namely, strong mixing framework. However, in this contribution, we investigate functional semiparametric regression under weak dependency condition of the quasi-associated correlation. From theoretical point of view this consideration allows to increase the scope of application of the proposed functional model. Indeed, it is well known that the mixing conditions are very hard to check and there exists lot of usual process fail to verify the mixing assumption. [4] have listed a numerous process, we quote, for instance, Bernoulli shifts class, Markov processes driven by discrete innovations and the AR(1) process with $\rho < 1/2$ and Bernoulli innovation among others. Thus we can say that the important feature of our study is to analyse the functional time series data without the mixing assumption. In addition we point out that our study generalize also the classical association (negative or positive). Thus the quasi-associated functional time series data is sufficiently weak to cover a large class of weak functional time series data. Finally, let us precise that our theoretical development explore the dependence structure of the data through the convergence rate. The latter contains the additional $\chi_\theta(x,h_n)$ that is control the local dependency of the data. It is clear that this dependency condition impact significantly the convergence rate of the estimator compared to the independent situation. Of course the independent case is more fast than the dependent one.

5.2. Application to the confidence intervals

The purpose of a confidence interval is to supplement the functional estimate at a point with information about the uncertainty in this estimate. It is a direct application of the Central Limit Theorem (CLT). In order to provide a confidence interval for the regression function in single functional model, we need first to propose a consistent estimator of the variance $\sigma_\theta^2(x)$. A natural consistent estimator of this variance is obtained by estimating the parameters involved in this quantity such as $(\beta_j)_{j=1,2}$ and $\varphi_\theta(\cdot)$. A natural estimator of β_j is

$$\widehat{\beta}_j = \frac{1}{n\,G_\theta(x,h)} \sum_{i=1}^{n} K^j\left(\frac{\langle \theta, x - X_i\rangle}{h_n}\right), \qquad j = 1,2, \tag{5.1}$$

while the Nadaraya–Waston type estimator $\varphi(\cdot)$ is

$$\widehat{\varphi}_n(x) = \frac{\sum_{i=1}^{n} Y_i^2\, K\left(\frac{\langle \theta, x - X_i\rangle}{h_n}\right)}{\sum_{i=1}^{n} K\left(\frac{\langle \theta, x - X_i\rangle}{h_n}\right)}. \tag{5.2}$$

Consequently, by combining the Equations (5.1), (5.2) with the definition of $\widehat{r}_{\theta,n}(x)$ consistent estimator of $\sigma_\theta^2(x)$ denoted by $\widehat{\sigma}_\theta^2(x)$, it follows that the asymptotic confidence band at asymptotic level $1-\alpha$ for $r_\theta(x)$ is

$$\widehat{r}_{\theta,n}(x) \pm \mathcal{U}_{1-\frac{\alpha}{2}} \left(\frac{\widehat{\sigma}_\theta^2(x)}{n\,G_\theta(x,h)}\right)^{\frac{1}{2}}. \tag{5.3}$$

Let us note the $\left(\frac{\widehat{\sigma}_\theta^2(x)}{n\,G_\theta(x,h)}\right)$ is essay to compute and does not require the estimation of $G_\theta(x,h)$. The latter will be removed by a simple manipulation.

5.3. On the applicability of the SFIM

From theoretical point of view, it is well known that the single index model is one of the most important additive models used to improve the convergence rate of the nonparametric approach. This model keeps this feature in functional statistics. However, the applicability of this model in practice requires an additional works that is the determination of the functional index θ and the smoothing parameter h which are often unknown in practice. This issue has been widely addressed in the nonfunctional case, but, remains not fully explored in the functional statistics. The readers interested by this topics can refer to [29] and the references therein (for recent advances in this topic). Thus, the estimation of the functional index and/or the bandwidth h_n in the quasi-associated functional time series case is an important prospect of the present contribution. As preliminary step, we present in this paragraph some selector rules compatible with our context of the functional time series data analysis. The first one is the Least Squares Cross-Validation (LSCV) rule, defined by

$$(\widehat{\theta}, \widehat{h}) = \arg\min_{\substack{h_n \in H_n \\ \theta \in \Theta}} \frac{1}{n} \sum_{i=1}^{n} \left(Y_i - \widehat{r}_{\theta,n}^{-i}(X_i)\right)^2, \tag{5.4}$$

where $\widehat{r}_{\theta,n}^{-i}$ is the leave-one-out estimator of $\widehat{r}_{\theta,n}$. This kind of cross-validation is widely used in the nonparametric prediction problems to select the bandwidth parameter in the kernel smoothing. It was popularized in semi-parametric functional data analysis by [1]. The second one is the Maximum Likelihood Cross-Validation (MLCV) rule, expressed by

$$(5.5)\qquad (\widehat{\theta}, \widehat{h}) = \underset{\substack{h_n \in H_n \\ \theta \in \Theta}}{\arg\min} \frac{1}{n} \sum_{i=1}^{n} \log \widehat{f}\big(Y_i | \widehat{r}_{\theta,n}^{-i}(X_i)\big),$$

where $\widehat{f}(\cdot|\cdot)$ is the estimator of the conditional density of Y given $\langle \theta, X \rangle$. This criterion can be viewed as generalization of the rule (5.4) when the conditional distribution is Gaussian. Of course in practice we must optimize these rule over finite subset Θ of index. Similarly to [1], we propose to select the optimal index from the following subset:

$$\Theta = \Theta_n = \left\{ \theta \in H,\ \theta = \sum_{i=1}^{k} c_i e_i,\ \|\theta\| = 1,\ \text{and } \exists\, j \in [1, k] \text{ such that } \langle \theta, e_j \rangle > 0 \right\},$$

where $(e_i)_{i=1,\ldots,k}$ is finite basis functions of the Hilbert subspace spanned by the covariates $(X_i)_i$ and $(c_i)_i$ some real calibrated constants allowing to insure the identifiability of the model. The common way is to choose the $(c_i)_i$ with calibration from the subset $\{-1, 0, 1\}$. Finally let us point both rules (5.4) and (5.5) we can take H_n as the subsets of the p-quantiles of the vector distance $D = D_{ij} = \|X_i - X_j\|$.

6. A SIMULATION STUDY

This section is devoted to some simulation experiments allowing to highlight the finite sample performance of the proposed SFIM-regression in different situations. This empirical study has two main purposes: The first one is to show the easy implantation of the SFIM in practice and the second one is to control the effect of the principal settings of the study (such as, the dependence's level, the type of the functional index, the smoothing degree of the link functions and the nature of the conditional distribution) in the efficiency of this functional model. For these objects, we simulate a functional time series data using the following SFIM equation:

$$(6.1)\qquad Y_i = r(\langle \theta, X_i \rangle) + \epsilon_i \qquad \text{for} \quad i = 1, \ldots, n = 150,$$

where the ϵ_i's are generated independently according to a normal distribution $\mathcal{N}(0, 1)$. The functional regressors are generated by the following formula:

$$X_i(t) = \cos(W_i\, t) + \sin(W_i + t) + .2(W_i\, t), \qquad t \in [-\pi, +\pi],$$

and W_i is selected random variable. Three levels of dependency are considered that are independent, quasi-associated (weak-dependency) and α-mixing (strong dependency). For the independent case, we take $(W_i)_i$ as sample of $\mathcal{N}(0, 1)$. The quasi-associated case is carried out by generating the process $(W_i)_i$ as non-strong mixing autoregressive of order 1. It obtained by taking the coefficient of the autoregressive $\rho = 0.1$ and the innovation random variable as $\mathrm{Binom}(10, 0.25)$. It is shown in [6] that this kind of process fails to satisfy the α-mixing assumption. However, this process is quasi-associated because it can be treated as linear process with positive coefficients. Concerning the strong dependency, we drown W from an autoregressive of order 1 with $\rho = 0.75$ and the $\chi_2(4)$ as innovation random variable.

The strong mixing property of this kind of process has been proved by [2]. The following Figure 1 shows the shape of $n = 150$ curves X_i's for three situations (independent, quasi-associated and strong dependency). The curves are discretized in the same grid formed by 100 points $[-\pi, \pi]$.

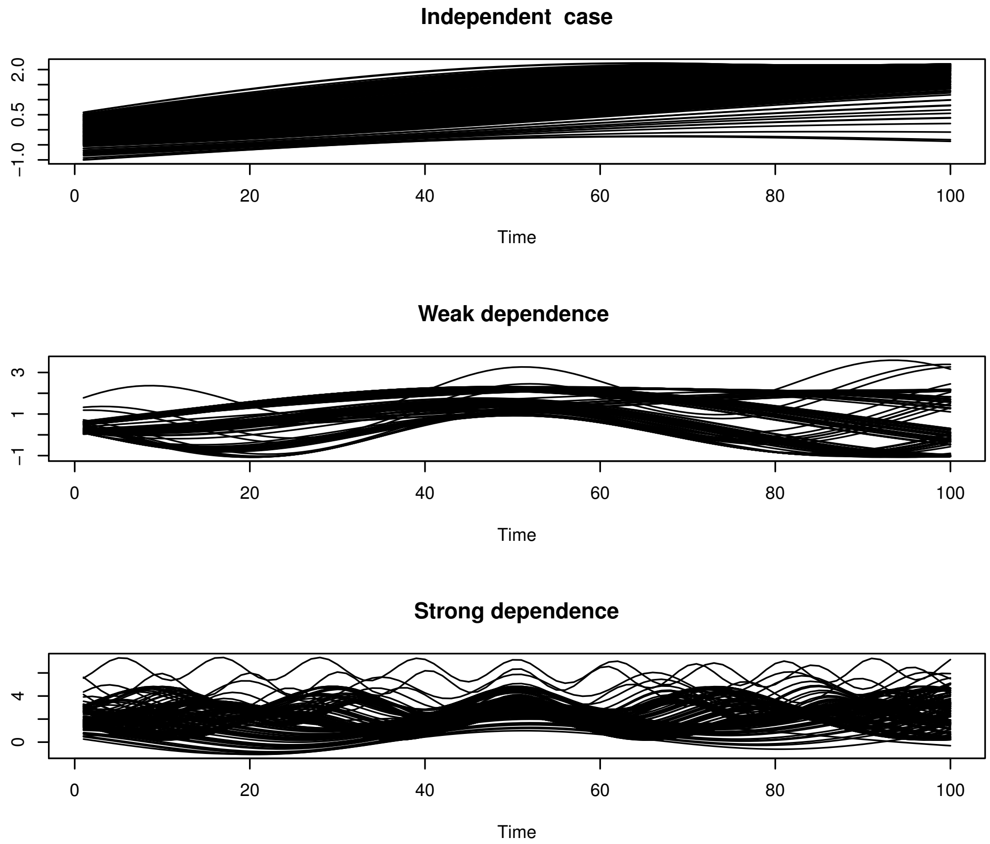

Figure 1: The shape of the regressors in the three cases.

In the first illustration, we control the effect of the degree of dependency on the prediction's quality using the single functional index regression. For this goal, we generate the scalar response Y_i by taking $r_1(x) = 3\log(1+x^2)$ as link function and $\theta_1 = e_1$ is the first element of the Karhunen–Loève basis functions. Explicitly θ_1 is the eigenfunction associated to the first eigenvalue of the covariance operator of the process $(X_i)_i$. It is eligible functionals index because it belongs in the same Hilbert subspace of the functional variable and is an element of Θ_n (see the previous section).

Undoubtedly, the easy implementation of any statistical approach in practice is closely linked to the flexibility of the choice of parameters involved in this approach. At this stage the bandwidth parameter h_n and the functional index θ are the principal parameters of the estimator. In this first illustration, we use the least squares cross-validation rule (5.4) described in the previous section to determine θ. The mentioned rule is optimized over Θ_n associated to the Karhunen–Loève basis functions (for $k = 5$). For sake of brevity, we use the default smoothing parameter h_n of R-package `fda.usc` and quadratic kernel on $(0, 1)$.

The obtained results are given in the Figure 2. The latter gives a global overview on the behaviours of SFIM-predictor with respect the dependence's level. In this figure we plot the true values $(Y_i)_i$ versus the predicted values for the three situations (independent, quasi-associated and strong dependency).

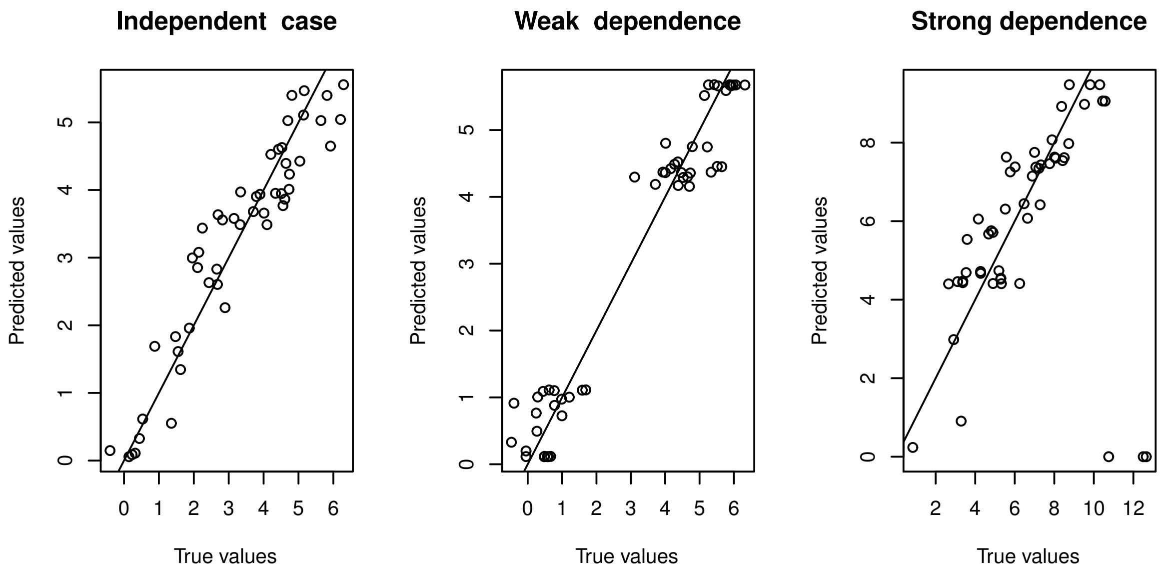

Figure 2: The SFIM-prediction results.

The results are not surprising. The SFIM-predictor has a satisfactory degree of performance. However, its behavior is strongly affected by the correlation of the data. The quality of prediction decreases with the degree of the dependency. The performance of the prediction procedures is tested by comparing the Mean Square Prediction Error defined by:

$$\text{MSPE} = \frac{1}{150}\sum_{i=1}^{150}\Big(Y_i - \widehat{r}_{\widehat{\theta}_1,n}(X_i)\Big)^2, \qquad \widehat{\theta}_1 \text{ being the optimizer of (5.4)}.$$

For this first illustration, we have obtained 0.23 for the independent against 0.92 for the quasi-associated and 1.78 for the strong mixing case.

Now, in order to give comprehensive empirical analysis for this semi-parametric model, we examine, in this second illustration, the impact of the other characteristics (the type of the functional index, the smoothing degree of the link functions and the nature of the conditional distribution) on the SFIM-prediction. More precisely, we compare two link functions (smooth and unsmooth (discontinuous in some points)), two functional indexes (eligible and ineligible) and two conditional distributions (Gaussian and non-Gaussian). This comparison will be carried out for the three previous dependence situations (independent, quasi-associated and strong mixing). We keep the data of the first illustration as perfect situation of the SFIM-prediction (eligible index, smooth link function and Gaussian conditional distribution).

Now, for the other situations, we follow the same algorithm of the first illustration to generate the output observations $(Y_i)_i$. To do that, we simulate with an arbitrary functional index expressed by the normalised function

$$\theta_2(t) = 0.15\, t\, \sin(t)$$

and the link function

$$r_2(x) = r_1(x)\, \mathbb{1}_{[0,.5]} - r_1^2(x)\, \mathbb{1}_{[-1,-.5]}\,.$$

The last factor of SFIM-prediction is the conditional distribution of Y given X. The latter is explicitly given by the distribution of ϵ_i shifted by $r(\langle\theta, x\rangle)$. For this second illustration, we generate the white noise ϵ_i from normal mixture distribution $(0.75)\mathcal{N}(0,1) + 0.25\, N(.5,2)$. To quantify the impact of the conditional distribution on the SFIM-prediction we compare the two selector rules of the functional index (5.4) and (5.5). Of course both rules coincide when the conditional distribution is Gaussian. Finally, we point out that we have used the same kernel and the same bandwidth as in the first illustration and the conditional distribution in the rule (5.5) is computed by the routine `npcdist` in the `R`-package `np`. The results on this comparison study are presented in Table 1. It contains the MSPE for the six scenarios mentioned before.

Table 1: Comparison of the MSPE errors of the SFIM-prediction.

Dependency case	Conditional distribution	SFIM		CV-rule	
		Index	Function	LSCV	MLCV
Independent	Gaussian	Eligible	Smooth	0.23	0.24
		Ineligible	Smooth	0.71	0.76
		Eligible	Discontinuous	0.57	0.64
		Ineligible	Discontinuous	1.23	1.36
	Normal Mixture	Eligible	Smooth	0.41	0.33
		Ineligible	Smooth	0.93	0.67
		Eligible	Discontinuous	0.79	0.62
		Ineligible	Discontinuous	1.56	0.95
Quasi-associated	Gaussian	Eligible	Smooth	0.92	0.97
		Ineligible	Smooth	1.62	1.71
		Eligible	Discontinuous	1.27	1.29
		Ineligible	Discontinuous	2.09	2.14
	Normal Mixture	Eligible	Smooth	1.18	0.97
		Ineligible	Smooth	1.54	1.18
		Eligible	Discontinuous	1.41	1.07
		Ineligible	Discontinuous	2.14	1.92
Strong mixing	Gaussian	Eligible	Smooth	1.78	1.88
		Ineligible	Smooth	2.23	2.34
		Eligible	Discontinuous	2.17	2.25
		Ineligible	Discontinuous	2.57	2.59
	Normal Mixture	Eligible	Smooth	1.93	1.57
		Ineligible	Smooth	2.37	2.05
		Eligible	Discontinuous	2.18	1.93
		Ineligible	Discontinuous	2.68	2.15

The simulation results of Table 1 show that the prediction is strongly affected by the different features of the data (dependence degree) as well as the model (the smoothing property of the link function). This statement incorporates the theoretical result that relates the convergence rate of the estimator to the correlation of the data and the regularity assumption of the model. In addition the choice of the functional index impact also the performance of the prediction by the SFIM. In particular the two rules (5.4) and (5.5) are equivalent when the conditional distribution is Gaussian while the selector criterion (5.5) is more adequate for the mixture case. Overall, both criterion give a satisfactory level of accuracy even in the critical situation when the index is illegible and the link function is discontinuous.

7. PROOFS OF THE INTERMEDIATE RESULTS

This section is devoted to the proofs of our results. The previously presented notation continues to be used in the following.

Proof of Lemma 4.2

The proof of this lemma is based on inequality given in Lemma 4.1 on the variables

$$\widehat{\Delta}_i(x) := \frac{1}{n\,\mathbb{E}\big(K_1(x)\big)}\left[\widehat{Z}_i - \mathbb{E}(\widehat{Z}_i)\right], \qquad i=1,...,n\,,$$

where $\widehat{Z}_i = \widehat{Y}_i K_i(x)$, and we have

$$\begin{aligned}
\mathbb{E}(\widehat{\Delta}_i) \;&=\; 0\,,\\
\|\widehat{\Delta}_i\|_\infty \;&\le\; \frac{2\,\delta_n}{n\,G_\theta(x,h_n)}\;\|K\|_\infty\,,\\
\mathrm{Lip}(\widehat{\Delta}_i) \;&\le\; 2\,\mathrm{Lip}(K)\,\frac{\delta_n}{n\,G_\theta(x,h_n)\,h_n}\,,\\
\widehat{r}_{\theta,2}(x) - \mathbb{E}\big(\widehat{r}_{\theta,2}(x)\big) \;&=\; \sum_{i=1}^n \widehat{\Delta}_i\,.
\end{aligned}$$

We start by evaluating the covariance term $\mathrm{Cov}\big(\widehat{\Delta}_{s_1},...,\widehat{\Delta}_{s_u},\widehat{\Delta}_{t_1},...,\widehat{\Delta}_{t_v}\big)$, for all $(s_1,...,s_u)\in\mathbb{N}^u$ and $(t_1,...,t_u)\in\mathbb{N}^v$ with $1\le s_1\le\cdots\le s_u\le t_1\le\cdots\le t_v\le n$. If $m=t_1-s_u=0$, using the fact that, for all $p>0$,

$$\mathbb{E}\big(K_1^p(x)\big) = O\big(G_\theta(x,h_n)\big)\,,$$

and under the second part of (H$_2$)(iii), we readily obtain

$$\begin{aligned}
\Big|\mathrm{Cov}\big(\widehat{\Delta}_{s_1}...\widehat{\Delta}_{s_u},\widehat{\Delta}_{t_1}...\widehat{\Delta}_{t_v}\big)\Big| \;&\le\; \left(\frac{1}{n\,\mathbb{E}\big(K_1(x)\big)}\right)^{u+v}\mathbb{E}\Big(\big|\widehat{Z}_{s_1}...\widehat{Z}^2_{s_u}...\widehat{Z}_{t_v}\big|\Big)\\
&\le\; \left(\frac{C\,\delta_n\,\|K\|_\infty}{n\,G_\theta(x,h_n)}\right)^{u+v}\mathbb{E}\big(Y^2_{s_u}K^2_{s_u}\big)\\
&\le\; \left(\frac{C\,\delta_n}{n\,G_\theta(x,h_n)}\right)^{u+v}G_\theta(x,h_n)\,.
\end{aligned}$$

If $m = t_1 - s_u > 0$, by quasi-association of the sequence $(\widehat{Z}_n)$, we infer that

$$\begin{aligned}
\Big|\mathrm{Cov}\big(\widehat{\Delta}_{s_1}\dots\widehat{\Delta}_{s_u},\widehat{\Delta}_{t_1}\dots\widehat{\Delta}_{t_v}\big)\Big| &\le 4\left(\frac{\delta_n\,\mathrm{Lip}(K)}{n\,G_\theta(x,h_n)\,h_n}\right)^2\left(\frac{2\,\delta_n\,\|K\|_\infty}{n\,G_\theta(x,h_n)}\right)^{u+v-2}\sum_{i=1}^{u}\sum_{j=1}^{v}\lambda_{s_i,t_j}\\
&\le C^{u+v}\left(\frac{\mathrm{Lip}(K)}{h_n}\right)^2\left(\frac{\delta_n}{n\,G_\theta(x,h_n)}\right)^{u+v}(u\wedge v)\,\lambda_{t_1-s_u}\\
&\le C^{u+v}\left(\frac{\mathrm{Lip}(K)}{h_n}\right)^2\left(\frac{\delta_n}{n\,G_\theta(x,h_n)}\right)^{u+v}v\,e^{-am}. \qquad (7.1)
\end{aligned}$$

On the other hand, making use of the first part of the condition (H_2)(iii) we may write

$$\begin{aligned}
\Big|\mathrm{Cov}\big(\widehat{\Delta}_{s_1}\dots\widehat{\Delta}_{s_u},\widehat{\Delta}_{t_1}\dots\widehat{\Delta}_{t_v}\big)\Big| &\le \left(\frac{C\,\delta_n\|K\|_\infty}{n\,G_\theta(x,h_n)}\right)^{u+v-2}\Big|\mathrm{Cov}\big(\widehat{\Delta}_{s_u},\widehat{\Delta}_{t_1}\big)\Big|\\
&\le \left(\frac{C\,\delta_n\|K\|_\infty}{n\,G_\theta(x,h_n)}\right)^{u+v-2}\left(\Big|\mathbb{E}\big(\widehat{\Delta}_{s_u}\widehat{\Delta}_{t_1}\big)\Big|+\mathbb{E}\big|\widehat{\Delta}_{s_u}\big|\,\mathbb{E}\big|\widehat{\Delta}_{t_1}\big|\right)\\
&\le \left(\frac{C\,\delta_n\|K\|_\infty}{n\,G_\theta(x,h_n)}\right)^{u+v-2}\left(\frac{C}{n\,G_\theta(x,h_n)}\right)^2\delta_n^2\,\chi_\theta(x,h_n)\,.
\end{aligned}$$

It follows that

$$\Big|\mathrm{Cov}\big(\widehat{\Delta}_{s_1},\dots,\widehat{\Delta}_{s_u},\widehat{\Delta}_{t_1},\dots,\widehat{\Delta}_{t_v}\big)\Big| \le C^{u+v}\left(\frac{\delta_n}{n\,G_\theta(x,h_n)}\right)^{u+v}\chi_\theta(x,h_n)\,. \qquad (7.2)$$

Moreover, by multiplying a τ-power of (7.1) and $(1-\tau)$-power of (7.2) for some $\frac{1}{4}<\tau<\frac{1}{2}$, we obtain an upper-bound of the covariance as follows for $1\le s_1\le\dots\le s_u\le t_1\le\dots\le t_v\le n$:

$$\Big|\mathrm{Cov}\big(\widehat{\Delta}_{s_1}\dots\widehat{\Delta}_{s_u},\widehat{\Delta}_{t_1}\dots\widehat{\Delta}_{t_v}\big)\Big| \le C^{u+v}\left(\frac{\delta_n}{n\,G_\theta(x,h_n)}\right)^{u+v}\left(\frac{\mathrm{Lip}(K)}{h_n}\right)^{2\tau}\left(\sqrt{\chi_\theta(x,h_n)}\right)^{2(1-\tau)}v\,e^{-a\tau m}.$$

So, by (H_5), we have

$$\Big|\mathrm{Cov}\big(\widehat{\Delta}_{s_1}\dots\widehat{\Delta}_{s_u},\widehat{\Delta}_{t_1}\dots\widehat{\Delta}_{t_v}\big)\Big| \le \left(\frac{C\,\delta_n}{n\,G_\theta(x,h_n)}\right)^{u+v-2}\left(\frac{C\,\delta_n}{n\,G_\theta(x,h_n)}\right)^2\sqrt{\chi_\theta(x,h_n)}\;v\,e^{-a\tau m},$$

where

$$M_n=\frac{C\,\delta_n}{n\,G_\theta(x,h_n)}\qquad\text{and}\qquad K_n=\frac{C\,\chi_\theta^{1/4}(x,h_n)\,\delta_n}{n\,G_\theta(x,h_n)}\,.$$

It remains to calculate $\mathrm{Var}\left(\sum_{i=1}^{n}\widehat{\Delta}_i\right)$:

$$\begin{aligned}
\mathrm{Var}\left(\sum_{i=1}^{n}\widehat{\Delta}_i\right) &= \left(\frac{1}{n\,\mathbb{E}\big(K_1(x)\big)}\right)^2\sum_i\sum_j\mathrm{Cov}\big(\widehat{Z}_i,\widehat{Z}_j\big)\\
&= \left(\frac{1}{n\,\mathbb{E}\big(K_1(x)\big)}\right)^2\left[n\,\mathrm{Var}\big(\widehat{Z}_1\big)+\sum_i\sum_{j\neq i}\mathrm{Cov}\big(\widehat{Z}_i,\widehat{Z}_j\big)\right]\\
&= \left(\frac{1}{n\,\mathbb{E}\big(K_1(x)\big)}\right)^2\left[n\,T_1+T_{ij}\right].
\end{aligned}$$

Now, under the assumption (H$_5$), we obtain for the first term:

$$\begin{aligned} T_1 = \mathrm{Var}\big(\widehat{Z}_1\big) &= \mathbb{E}\big(\widehat{Y}_1^2 K_1^2(x)\big) - \Big(\mathbb{E}\big(\widehat{Y}_1 K_1(x)\big)\Big)^2 \\ &\le \mathbb{E}\big(Y_1^2 K_1^2(x)\big) \\ &\le \mathbb{E}\Big(K_1^2(x)\,\mathbb{E}\big(Y_1^2|X\big)\Big) \\ &\le C\,\mathbb{E}\big(K_1^2(x)\big)\,. \end{aligned}$$

For all $j \ge 1$, we have

$$\mathbb{E}\big(K_1^j(x)\big) = O\big(G_\theta(x,h_n)\big)\,, \tag{7.3}$$

and

$$T_1 = \mathrm{Var}\big(\widehat{Z}_1\big) = O\big(\chi_\theta^{1/2}(x,h_n)\big)\,.$$

We readily obtain that

$$\frac{1}{n\Big(\mathbb{E}\big(K_1(x)\big)\Big)^2}\,T_1 \le \frac{C\,\chi_\theta^{1/2}(x,h_n)}{n\,G_\theta^2(x,h_n)}\,. \tag{7.4}$$

For the second term, we have the following decomposition

$$T_{ij} = \sum_i \sum_{0<|i-j|\le u_n} \mathrm{Cov}\big(\widehat{Z}_i,\widehat{Z}_j\big) + \sum_i \sum_{|i-j|>u_n} \mathrm{Cov}\big(\widehat{Z}_i,\widehat{Z}_j\big) = J_1 + J_2\,,$$

where (u_n) is a sequence of positive integer and

$$\lim_{n\longrightarrow\infty} u_n = \infty\,.$$

Now, under the assumptions (H$_2$), we have

$$\begin{aligned} |J_1| &= \sum_i \sum_{0<|i-j|\le u_n} \big|\mathrm{Cov}\big(\widehat{Z}_i,\widehat{Z}_j\big)\big| \le n\,u_n\left[\max_{i\ne j}\Big|\mathbb{E}\big(K_i(x)\,K_j(x)\big)\Big| + \Big(\mathbb{E}\big(K_1(x)\big)\Big)^2\right] \\ &\le C\,n\,u_n\,\chi_\theta(x,h_n)\,. \end{aligned}$$

Making use of the condition (H$_2$)(i), we infer that

$$\begin{aligned} |J_2| &= \sum_i \sum_{|i-j|>u_n} \big|\mathrm{Cov}\big(\widehat{Z}_i,\widehat{Z}_j\big)\big| \le C\,\delta_n^2\left(\frac{\mathrm{Lip}(K)}{h_n}\right)^2 \sum_i \sum_{|i-j|>u_n} \lambda_{i,j} \\ &\le C\,n\,\delta_n^2\,h_n^{-2}\,e^{-au_n}\,. \end{aligned}$$

This implies that

$$|T_{ij}| \le \sum_{i=1}^{n} \sum_{i\ne j} \big|\mathrm{Cov}\big(\widehat{Z}_i,\widehat{Z}_j\big)\big| \le C\Big(n\,u_n\,\chi_\theta(x,h_n) + n\,\delta_n^2\,h_n^{-2}\,e^{-au_n}\Big)\,.$$

Next, taking

$$u_n = \frac{1}{a}\log\left(\frac{\delta_n^2\,a}{h_n^2\,\chi_\theta(x,h_n)}\right)\,.$$

Observe that (H_5) insure that

$$\sqrt{\chi_\theta(x,h_n)}\,\log(\delta_n) \;\to\; 0\,,$$

which allows to write that

$$T_{ij} \;=\; o\Big(n\,\chi_\theta^{1/2}(x,h_n)\Big) \;\to\; 0\,. \tag{7.5}$$

It follows that

$$\mathrm{Var}\left(\sum_{i=1}^{n}\widehat{\Delta}_i\right) \;=\; O\left(\frac{\chi_\theta^{1/2}(x,h_n)}{n\,G_\theta^2(x,h_n)}\right).$$

The conditions of Lemma 4.1 are verified for

$$\begin{aligned}
K_n \;&=\; \frac{C\,\chi_\theta^{1/4}(x,h_n)\,\delta_n}{n\,G_\theta(x,h_n)}\,, \qquad M_n \;=\; \frac{C\,\delta_n}{n\,G_\theta(x,h_n)}\,,\\
A_n \;&=\; \frac{\chi_\theta^{1/2}(x,h_n)}{n\,G_\theta^2(x,h_n)}\,,\\
B_n \;&=\; \left(\frac{16\,n\,K^2}{9\,A_n\,(1-e^{-\beta})}\vee 1\right)\left(\frac{2\,(K\vee M)}{1-e^{-\beta}}\right) \;=\; \frac{\delta_n}{n\,G_\theta(x,h_n)}\,.
\end{aligned}$$

So, we apply the inequality in [22] to the random variables $\widehat{\Delta}_i$ to infer that

$$\begin{aligned}
\mathbb{P}\left(\Big|\widehat{r}_{\theta,2}(x)-\mathbb{E}\big(\widehat{r}_{\theta,2}(x)\big)\Big| > \varepsilon\sqrt{\frac{\chi_\theta^{1/2}(x,h_n)\,\log n}{n\,G_\theta^2(x,h_n)}}\right) \;&=\; \mathbb{P}\left(\left|\sum_{i=1}^{n}\widehat{\Delta}_i\right| > \varepsilon\sqrt{\frac{\chi_\theta^{1/2}(x,h_n)\,\log n}{n\,G_\theta^2(x,h_n)}}\right)\\
&\le\; \exp\left(\frac{-\varepsilon^2\,\chi_\theta^{1/2}(x,h_n)\,\log n}{n\,G_\theta^2(x,h_n)\,L_\theta(n)}\right),
\end{aligned}$$

where

$$L_\theta(n) \;=\; \left(\frac{\chi_\theta^{1/2}(x,h_n)}{n\,G_\theta^2(x,h_n)} + \left(\frac{\delta_n}{n\,G_\theta^2(x,h_n)}\right)^{\frac{1}{3}}\left(\frac{\chi_\theta^{1/2}(x,h_n)\,\log n}{n\,G_\theta^2(x,h_n)}\right)^{\frac{5}{6}}\right).$$

Then we finally obtain that

$$\begin{aligned}
\mathbb{P}\left(\Big|\widehat{r}_{\theta,2}(x)-\mathbb{E}(\widehat{r}_{\theta,2})\Big| > \varepsilon\sqrt{\frac{\chi_\theta^{1/2}(x,h_n)\,\log n}{n\,G_\theta^2(x,h_n)}}\right) \;&\le\; \exp\left(\frac{-\varepsilon^2\,\log n}{C+\Big(\delta_n^2\,\chi_\theta^{-1/2}(x,h_n)\,\log^5 n\Big)^{\frac{1}{6}}}\right)\\
&\le\; C_1\,\exp\big(-\varepsilon^2\,\log(n)\big)\,.
\end{aligned}$$

The proof is achieved by a suitable choice of ε. □

Proof of Lemma 4.3

The proof of this lemma is similar to the proof of the previous Lemma 4.2. Since $\widehat{Y}_i = 1$, it suffices to replace $\widehat{\Delta}_i$ by

$$\widetilde{\Delta}_i \;=\; \frac{1}{n\,\mathbb{E}\big(K_1(x)\big)}\Big[K_i(x)-\mathbb{E}\big(K_i(x)\big)\Big]\,, \qquad i=1,\dots,n\,.$$

Thus we obtain, under (H$_1$)–(H$_5$),

$$\mathbb{P}\left(\left|\widehat{r}_{\theta,0}(x)-1\right| > \varepsilon\sqrt{\frac{\log n}{n\,G_\theta(x,h_n)}}\right) = \mathbb{P}\left(\left|\sum_{i=1}^{n}\widetilde{\Delta}_i\right| > \varepsilon\sqrt{\frac{\chi_\theta^{1/2}(x,h_n)\,\log n}{n\,G_\theta^2(x,h_n)}}\right)$$
$$\leq C_1'\,\exp\bigl(-\varepsilon^2\log(n)\bigr)\,.$$

Thus the proof is complete. □

Proof of Lemma 4.4

Notice that we have

$$\left\{\left|\widehat{r}_{\theta,0}(x)\right| \leq \frac{1}{2}\right\} \subset \left\{\left|\widehat{r}_{\theta,0}(x)-1\right| > \frac{1}{2}\right\},$$

that implies that

$$\mathbb{P}\left(\left|\widehat{r}_{\theta,0}(x)\right| \leq \frac{1}{2}\right) \leq \mathbb{P}\left(\left|\widehat{r}_{\theta,0}(x)-1\right| > \frac{1}{2}\right).$$

Under the hypothesis (H$_1$)–(H$_5$) and by applying Lemma 4.3, we deduce that

$$\sum_n \mathbb{P}\left(\left|\widehat{r}_{\theta,0}(x)\right| \leq \frac{1}{2}\right) \leq \sum_n \mathbb{P}\left(\left|\widehat{r}_{\theta,0}(x)-1\right| > \frac{1}{2}\right) < \infty\,.$$

Then, for $\eta=\frac{1}{2}$, we have $\sum_n \mathbb{P}\left(\left|\widehat{r}_{\theta,0}(x)\right| \leq \eta\right) < \infty$. Thus the proof is complete. □

Proof of Lemma 4.5

One can easily see that we have

$$\left|r_\theta(x)-\mathbb{E}\bigl(\widehat{r}_{\theta,1}(x)\bigr)\right| = \left|r_\theta(x)-\mathbb{E}\left(\frac{1}{n\,\mathbb{E}\bigl(K_1(x)\bigr)}\sum_{i=1}^{n}Y_i\,K_i(x)\right)\right|$$
$$= \frac{1}{\mathbb{E}\bigl(K_1(x)\bigr)}\left[\left|r_\theta(x)\,\mathbb{E}\bigl(K_1(x)\bigr)-\mathbb{E}\bigl(Y_1\,K_1(x)\bigr)\right|\right]$$
$$= \frac{1}{\mathbb{E}\bigl(K_1(x)\bigr)}\,\mathbb{E}\left[\left(\left|r_\theta(x)-r_\theta(X_1)\right|\right)K_1(x)\right] \leq C\,h_n^\beta\,.$$

This readily implies that we have

$$r_\theta(x)-\mathbb{E}(\widehat{r}_{\theta,1}) = O\bigl(h_n^\beta\bigr)\,.$$

Thus the proof is complete. □

Proof of Lemma 4.6

We first observe that we have

$$\begin{aligned}\left|\mathbb{E}(\widehat{r}_{\theta,2})-\mathbb{E}(\widehat{r}_{\theta,1})\right| &= \frac{1}{n\,\mathbb{E}\,K_1(x)}\left|\mathbb{E}\left(\sum_{i=1}^{n} Y_i\,\mathbb{1}_{\{|Y_i|>\delta_n\}}\,K_i(x)\right)\right| \\ &\leq \mathbb{E}\Big(|Y_1|\,\mathbb{1}_{|Y_1|>\delta_n}\,K_1(x)\Big)\,\Big(\mathbb{E}\big(K_1(x)\big)\Big)^{-1}.\end{aligned}$$

The Hölder's inequality allows to write that, for $\alpha=\frac{p}{2}$ and β such that $\frac{1}{\alpha}+\frac{1}{\beta}=1$,

$$\begin{aligned}\Big|\mathbb{E}\big(\widehat{r}_{\theta,2}(x)\big)-\mathbb{E}\big(\widehat{r}_{\theta,1}(x)\big)\Big| &\leq \frac{1}{G_\theta(x,h_n)}\;\mathbb{E}^{1/\alpha}\big[|Y^\alpha|\,\mathbb{1}_{\{Y\geq\delta_n\}}\big]\;\mathbb{E}^{1/\beta}\big[K_1^\beta\big] \\ &\leq \frac{1}{G_\theta(x,h_n)}\;\delta_n^{-1}\,\mathbb{E}^{1/\alpha}\big[|Y^p|\big]\,G_\theta^{1/\beta}(x,h_n) \\ &\leq C\,\delta_n^{-1}\,G_\theta^{(1-\beta)/\beta}(x,h_n)\,.\end{aligned}$$

Hence, we obtain from (H_5) that

$$\Big|\mathbb{E}\big(\widehat{r}_{\theta,2}(x)\big)-\mathbb{E}\big(\widehat{r}_{\theta,1}(x)\big)\Big| = o\left(\sqrt{\frac{\chi_\theta^{1/2}(x,h_n)\,\log n}{n\,G_\theta^2(x,h_n)}}\right).$$

Thus the proof is complete. □

Proof of Lemma 4.7

By (H_5) and we apply the Markov's inequality to show that, $\forall\,\epsilon>0$,

$$\begin{aligned}\mathbb{P}\Big(\big|\widehat{r}_{\theta,1}(x)-\widehat{r}_{\theta,2}(x)\big|>\epsilon\Big) &= \mathbb{P}\left(\frac{1}{n\,\mathbb{E}\,K_1(x)}\left|\sum_{i=1}^{n} Y_i\,\mathbb{1}_{\{|Y_i|>\delta_n\}}\,K_i(x)\right|>\epsilon\right) \\ &\leq n\,\mathbb{P}\big(|Y_1|>\delta_n\big) \;\leq\; n\,\delta_n^{-p}\,\mathbb{E}\big(|Y|^p\big) \;\leq\; C\,n\,\delta_n^{-p}\,.\end{aligned}$$

Since

$$\sum_{n\geq1} n\,\delta_n^{-p}<\infty\,,$$

then there exists $\epsilon_0>0$, such that

$$\sum_{n\geq1}\mathbb{P}\left(\big|\widehat{r}_{\theta,1}(x)-\widehat{r}_{\theta,2}(x)\big|>\epsilon_0\sqrt{\frac{\chi_\theta^{1/2}(x,h_n)\,\log n}{n\,G_\theta^2(x,h_n)}}\right)<\infty\,, \tag{7.6}$$

which completes the proof of the lemma. □

Proof of Lemma 4.8

Let us introduce the following sum $S_n = \sum_{i=1}^{n} L_{ni}$, where

$$L_{ni} = \frac{\sqrt{G_\theta(x,h_n)}}{\sqrt{n}\ \mathbb{E}\big(K_1(x)\big)} \left(\big(\widehat{Y}_i - r_\theta(x)\big) K_i(x) - \mathbb{E}\Big(\big(\widehat{Y}_i - r_\theta(x)\big) K_i(x) \Big) \right).$$

Therefore

$$S_n = \sqrt{n\, G_\theta(x,h_n)} \left(\big(\widehat{r}_{\theta,2}(x) - r_\theta(x)\big) \widehat{r}_{\theta,0}(x) - \mathbb{E}\big(\widehat{r}_{\theta,2}(x) - r_\theta(x)\big) \widehat{r}_{\theta,0}(x) \right).$$

Thus, our claimed result is now

$$S_n \to \mathcal{N}\big(0, \sigma_\theta^2(x)\big)\,. \tag{7.7}$$

To do that, we use the basic technique of [9] for which we split S_n into

$$S_n = T_n + T_n' + \zeta_k\,,$$

with

$$T_n = \sum_{j=1}^{k} \eta_j \qquad \text{and} \qquad T_n' = \sum_{j=1}^{k} \xi_j\,,$$

where

$$\eta_j := \sum_{i \in I_j} L_{ni}\,, \qquad \xi_j := \sum_{i \in J_j} L_{ni}\,, \qquad \zeta_k := \sum_{i=k(p+q)+1}^{n} L_{ni}\,,$$

with

$$I_j = \Big\{ (j-1)(p+q)+1,\ \dots,\ (j-1)(p+q)+p \Big\}\,,$$
$$J_j = \Big\{ (j-1)(p+q)+p+1,\ \dots,\ j(p+q) \Big\}\,,$$

and $p = p_n$, $q = q_n$ two sequences of natural numbers tending to ∞, such that

$$p = O\big(G_\theta^{-1}(x,h_n)\big)\,, \qquad q = o(p) \qquad \text{and} \qquad k = \left\lfloor \frac{n}{p+q} \right\rfloor,$$

where $\lfloor \cdot \rfloor$ stands for the integer part. Firstly, observe that we have $\frac{kq}{n} \to 0$, and $\frac{kp}{n} \to 1$, $\frac{q}{n} \to 0$, which imply that $\frac{p}{n} \to 0$ as $n \to \infty$. Now, our asymptotic normality results are a consequence of the following statements:

$$\mathbb{E}(T_n')^2 + \mathbb{E}(\zeta_k)^2 \to 0 \tag{7.8}$$

and

$$T_n \to \mathcal{N}\big(0, \sigma_\theta^2(x)\big)\,. \tag{7.9}$$

For (7.8), we write

$$\mathbb{E}(T_n')^2 = k\ \mathrm{Var}(\xi_1) + 2 \sum_{1 \le i < j \le k} \big|\mathrm{Cov}(\xi_i, \xi_j)\big|$$

and

$$\operatorname{Var}(\xi_1) \;\le\; q\operatorname{Var}(L_{n1}) + 2\sum_{1\le i<j\le q}\big|\operatorname{Cov}(L_{ni},L_{nj})\big|\,.$$

Similarly to (7.4), we infer that

$$\operatorname{Var}(L_{n1}) = O\big(n^{-1}\big)\,,$$

which implies that

$$k\,q\operatorname{Var}(L_{n1}) = O\bigg(\frac{kq}{n}\bigg) \to 0\,, \qquad \text{as} \quad n\to\infty\,.$$

On the other hand, we use the same arguments as those used in (7.5) to conclude that

$$k\sum_{1\le i<j\le q}\big|\operatorname{Cov}(L_{ni},L_{nj})\big| = o\bigg(\frac{kq}{n}\bigg) \to 0\,, \qquad \text{as} \quad n\to\infty\,. \tag{7.10}$$

Thus, the limit of the first term of $\mathbb{E}(T_n')^2$ is equal to 0. Next, by using stationarity, we can write

$$\begin{aligned}
\sum_{1\le i<j\le k}\big|\operatorname{Cov}(\xi_i,\xi_j)\big| \;&=\; \sum_{l=1}^{k-1}(k-l)\,\big|\operatorname{Cov}(\xi_1,\xi_{l+1})\big| \\
&\le\; k\sum_{l=1}^{k-1}\big|\operatorname{Cov}(\xi_1,\xi_{l+1})\big| \\
&\le\; k\sum_{l=1}^{k-1}\;\sum_{(i,j)\in J_1\times J_{l+1}}\operatorname{Cov}(L_{ni},L_{nj})\,.
\end{aligned}$$

It is clear that, for all $(i,j)\in J_1\times J_j$, we have $|i-j|\ge p+1>p$, and then

$$\begin{aligned}
\sum_{1\le i<j\le k}\big|\operatorname{Cov}(\xi_i,\xi_j)\big| \;&\le\; k\,\frac{C\,\gamma_n^2}{n\,h_n^2\,G_\theta(x,h_n)}\sum_{i=1}^{p}\sum_{\substack{j=2p+q+1,\\ |i-j|>p}}^{k(p+q)}\lambda_{i,j} \\
&\le\; \frac{C\,k\,p\,\gamma_n^2}{n\,h_n^2\,G_\theta(x,h_n)}\,\lambda_p \\
&\le\; \frac{C\,\gamma_n^2}{G_\theta^3(x,h_n)}\,e^{-ap} \;\to\; 0\,.
\end{aligned}$$

Finally, we get

$$\mathbb{E}(T_1')^2 \to 0 \qquad \text{as} \quad n\to\infty\,.$$

Since $\big(n-k(p+q)\big)\le p$, we have by the same manner

$$\begin{aligned}
\mathbb{E}(\zeta_k)^2 \;&\le\; \big(n-k(p+q)\big)\operatorname{Var}(L_{n1}) + 2\sum_{1\le i<j\le n}\big|\operatorname{Cov}(L_{ni},L_{nj})\big| \\
&\le\; p\operatorname{Var}(L_{n1}) + 2\sum_{1\le i<j\le n}\big|\operatorname{Cov}(L_{ni},L_{nj})\big| \\
&\le\; \frac{C\,p}{n} + o(1)\,.
\end{aligned}$$

Hence,

$$\mathbb{E}(\zeta_k)^2 \to 0\,, \qquad \text{as} \quad n\to\infty\,.$$

So, it remains to proof the asymptotic normality (7.9). The proof is standard. Indeed, it is based in the following assertions

$$\left|\mathbb{E}\Big(e^{it\sum_{j=1}^{k}\eta_j}\Big)-\prod_{j=1}^{k}\mathbb{E}\big(e^{it\eta_j}\big)\right|\to 0\,, \tag{7.11}$$

and

$$k\,\mathrm{Var}(\eta_1)\to\sigma_\theta^2(x)\,,\qquad k\,\mathbb{E}\big(\eta_1^2\,\mathbb{1}_{\{\eta_1>\epsilon\sigma_\theta(x)\}}\big)\to 0\,. \tag{7.12}$$

To prove (7.11), notice that

$$\begin{aligned}
&\left|\mathbb{E}\Big(e^{it\sum_{j=1}^{k}\eta_j}\Big)-\prod_{j=1}^{k}\mathbb{E}\big(e^{it\eta_j}\big)\right|\\
&\quad\le\left|\mathbb{E}\Big(e^{it\sum_{j=1}^{k}\eta_j}\Big)-\mathbb{E}\Big(e^{it\sum_{j=1}^{k-1}\eta_j}\Big)\mathbb{E}\big(e^{it\eta_k}\big)\right|+\left|\mathbb{E}\Big(e^{it\sum_{j=1}^{k-1}\eta_j}\Big)-\prod_{j=1}^{k-1}\mathbb{E}\big(e^{it\eta_j}\big)\right|\\
&\quad=\left|\mathrm{Cov}\Big(e^{it\sum_{j=1}^{k-1}\eta_j},\,e^{it\eta_k}\Big)\right|+\left|\mathbb{E}\Big(e^{it\sum_{j=1}^{k-1}\eta_j}\Big)-\prod_{j=1}^{k-1}\mathbb{E}\big(e^{it\eta_j}\big)\right| \qquad (7.13)
\end{aligned}$$

and, successively, we have

$$\begin{aligned}
&\left|\mathbb{E}\Big(e^{it\sum_{j=1}^{k}\eta_j}\Big)-\prod_{j=1}^{k}\mathbb{E}\big(e^{it\eta_j}\big)\right|\\
&\quad\le\left|\mathrm{Cov}\Big(e^{it\sum_{j=1}^{k-1}\eta_j},\,e^{it\eta_k}\Big)\right|+\left|\mathrm{Cov}\Big(e^{it\sum_{j=1}^{k-2}\eta_j},\,e^{it\eta_{k-1}}\Big)\right|+\cdots+\left|\mathrm{Cov}\big(e^{it\eta_2},e^{it\eta_1}\big)\right|. \qquad (7.14)
\end{aligned}$$

The use of the quasi-associated propriety permits to write that

$$\left|\mathrm{Cov}\big(e^{it\eta_2},e^{it\eta_1}\big)\right|\le\frac{C\,t^2\,\gamma_n^2}{n\,G_\theta^3(x,h_n)}\sum_{i\in I_1}\sum_{j\in I_2}\lambda_{i,j}\,.$$

Applying this inequality to each term on the right-hand side of (7.14) in order to obtain

$$\begin{aligned}
&\left|\mathbb{E}\Big(e^{it\sum_{j=1}^{k}\eta_j}\Big)-\prod_{j=1}^{k}\mathbb{E}\big(e^{it\eta_j}\big)\right|\\
&\quad\le\frac{C\,t^2\,\gamma_n^2}{n\,G_\theta^3(x,h_n)}\left[\sum_{i\in I_1}\sum_{j\in I_2}\lambda_{i,j}+\sum_{i\in I_1\cup I_2}\sum_{j\in I_3}\lambda_{i,j}+\cdots+\sum_{i\in I_1\cup\cdots\cup I_{k-1}}\sum_{j\in I_k}\lambda_{i,j}\right].
\end{aligned}$$

Observe that for every $2\le l\le k-1$, $(i,j)\in I_l\times I_{l+1}$, we have $|i-j|\ge q+1>q$, then

$$\sum_{i\in I_1\cup\cdots\cup I_{l-1}}\sum_{j\in I_l}\lambda_{i,j}\le p\,\lambda_q\,.$$

Therefore, inequality (7.13) becomes

$$\left|\mathbb{E}\Big(e^{it\sum_{j=1}^{k}\eta_j}\Big)-\prod_{j=1}^{k}\mathbb{E}\big(e^{it\eta_j}\big)\right|\le\frac{C\,t^2\,\gamma_n^2}{n\,G_\theta^3(x,h_n)}\,k\,p\,\lambda_q\le\frac{C\,t^2\,\gamma_n^2}{n\,G_\theta^3(x,h_n)}\,k\,p\,e^{-aq}\;\to\;0\,.$$

Concerning (7.12), we use the same arguments as in to conclude that

$$\lim_{n\to\infty} k\,\mathrm{Var}(\eta_1) \;=\; \lim_{n\to\infty} k\,p\,\mathrm{Var}(L_{n1})\,.$$

On the other hand

$$\mathrm{Var}(L_{n1}) \;=\; \frac{G_\theta(x,h_n)}{n\,\mathbb{E}^2\big(K_1(x)\big)}\;\mathrm{Var}\Big(\big(\widehat{Y}_1-r_\theta(x)\big)\,K_1(x)\Big)\,.$$

It can be written as

$$\begin{aligned}\mathrm{Var}(L_{n1}) \;=\;& \frac{G_\theta(x,h_n)}{n\,\mathbb{E}^2\big(K_1(x)\big)}\left\{\mathbb{E}\Big(K_1^2(x)\,\big(Y_1-r_\theta(x)\big)^2\Big)-\mathbb{E}\Big[K_1^2(x)\,\big(Y_1-r_\theta(x)\big)^2\,\mathbb{1}_{|Y_1|>\gamma_n}\Big]\right\}\\ &-\frac{G_\theta(x,h_n)}{n\,\mathbb{E}^2\big(K_1(x)\big)}\left(\mathbb{E}\Big(K_1(x)\,\big(Y_1-r_\theta(x)\big)\,\mathbb{1}_{|Y_1|<\gamma_n}\Big)\right)^2.\end{aligned}$$

By combining the same ideas used in the proof of Lemma 4.6 to those used by [14], we show that

$$\mathrm{Var}(L_{n1}) \;=\; \frac{\sigma_\theta^2(x)}{n}+o\left(\frac{1}{n}\right). \tag{7.15}$$

Therefore,

$$k\,\mathrm{Var}(\eta_1) \;=\; \frac{k\,p\,\sigma_\theta^2(x)}{n}+o\left(\frac{kp}{n}\right)\;\to\;\sigma_\theta^2(x)\,.$$

For the second part of (7.12), we use the fact that

$$|\eta_1| \;\le\; C\,p\,|L_{n1}| \;\le\; \frac{C\,\gamma_n\,p}{\sqrt{n\,G_\theta(x,h_n)}}\,,$$

and Tchebychev inequality to get

$$\begin{aligned}k\,\mathbb{E}\big(\eta_1^2\,\mathbb{1}_{\{\eta_1>\epsilon\sigma_\theta(x)\}}\big) \;&\le\; \frac{C\,\gamma_n^2\,p^2\,k}{n\,G_\theta(x,h_n)}\;\mathbb{P}\big(\eta_1>\epsilon\,\sigma_\theta(x)\big)\\ &\le\; \frac{C\,\gamma_n^2\,p^2\,k}{n\,G_\theta(x,h_n)}\;\frac{\mathrm{Var}(\eta_1)}{\epsilon^2\,\sigma_\theta^2(x)} \;=\; O\left(\frac{\gamma_n^2\,p^2}{n\,G_\theta(x,h_n)}\right),\end{aligned}$$

which completes the proof. □

ACKNOWLEDGMENTS

The authors are indebted to the Editor-in-Chief, the Associate Editor and the referee for their very valuable comments and suggestions which led to a considerable improvement of the manuscript.

The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work through the Research Groups Program under grant number R.G.P. 2/82/42.

REFERENCES

[1] Ait-Saïdi, A.; Ferraty, F.; Kassa, R. and Vieu, P. (2008). Cross-validated estimations in the single-functional index model, *Statistics*, **42**(6), 475–494.

[2] Andrews, D. (1983). First order autoregressive processes and strong mixing, Cowles Foundation Discussion Papers 664, Cowles Foundation for Research in Economics, Yale University.

[3] Aneiros, G.; Cao, R.; Fraiman, R.; Genest, C. and Vieu, P. (2019). Recent advances in functional data analysis and high-dimensional statistics, *J. Multivariate Anal.*, **170**, 3–9.

[4] Ango Nze, P.; Bühlmann, P. and Doukhan, P. (2002). Weak dependence beyond mixing and asymptotics for nonparametric regression, *Ann. Statist.*, **30**(2), 397–430.

[5] Bulinski, A. and Suquet, C. (2001). Normal approximation for quasi-associated random fields, *Statistics and Probability Letters*, **54**, 215–226.

[6] Chernick, M.R. (1981). A limit theorem for the maximum of autoregressive processes with uniform marginal distributions, *Ann. Probab.*, **9**(1), 145–149.

[7] Chowdhury, J. and Chaudhuri, P. (2019). Nonparametric depth and quantile regression for functional data, *Bernoulli*, **25**(1), 395–423.

[8] Cuevas, A. (2014). A partial overview of the theory of statistics with functional data, *J. Statist. Plann. Inference*, **147**, 1–23.

[9] Doob, J.L. (1953). *Stochastic Processes*, John Wiley & Sons, Inc., New York; Chapman & Hall, Limited, London.

[10] Douge, L. (2010). Théorèmes limites pour des variables quasi-associées Hilbertiennes, *Ann. I.S.U.P.*, **54**(1-2), 51–60.

[11] Douge, L. (2018). Nonparametric regression estimation for quasi-associated Hilbertian processes.

[12] Doukhan, P. and Louhichi, S. (1999). A new weak dependence condition and applications to moment inequalities, *Stochastic Processes and their Applications*, **84**(2), 313–342.

[13] Esary, J.D.; Proschan, F. and Walkup, D.W. (1967). Association of random variables, with applications, *Ann. Math. Statist.*, **38**, 1466–1474.

[14] Ferraty, F.; Mas, A. and Vieu, P. (2007). Nonparametric regression on functional data: inference and practical aspects, *Aust. N. Z. J. Stat.*, **49**(3), 267–286.

[15] Ferraty, F.; Peuch, A. and Vieu, P. (2003). Modèle à indice fonctionnel simple, *C. R. Math. Acad. Sci. Paris*, **336**(12), 1025–1028.

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

[17] Geenens, G. (2011). *A nonparametric functional method for signature recognition.* In "Recent Advances in Functional Data Analysis and Related Topics", Contrib. Statist., pages 141–147. Physica-Verlag/Springer, Heidelberg.

[18] Goia, A. and Vieu, P. (2016). An introduction to recent advances in high/infinite dimensional statistics [Editorial], *J. Multivariate Anal.*, **146**, 1–6.

[19] Härdle, W.; Hall, P. and Ichimura, H. (1993). Optimal smoothing in single-index models, *Ann. Statist.*, **21**(1), 157–178.

[20] Hristache, M.; Juditsky, A. and Spokoiny, V. (2001). Direct estimation of the index coefficient in a single-index model, *Ann. Statist.*, **29**(3), 595–623.

[21] Ichimura, H. (1993). Semiparametric least squares (SLS) and weighted SLS estimation of single-index models, *J. Econometrics*, **58**(1-2), 71–120.

[22] Kallabis, R.S. and Neumann, M.H. (2006). An exponential inequality under weak dependence, *Bernoulli*, **12**(2), 333–350.

[23] Li, J.; Huang, C. and Zhu, H. (2017). A functional varying-coefficient single-index model for functional response data, *J. Amer. Statist. Assoc.*, **112**(519), 1169–1181.

[24] Ling, N. and Vieu, P. (2018). Nonparametric modelling for functional data: selected survey and tracks for future, *Statistics*, **52**(4), 934–949.

[25] Ling, N. and Vieu, P. (2021). On semiparametric regression in functional data analysis, *Wiley Interdiscip. Rev. Comput. Stat.*, **13**(2), e1538, 15.

[26] Masry, E. (2005). Nonparametric regression estimation for dependent functional data: asymptotic normality, *Stochastic Process. Appl.*, **115**(1), 155–177.

[27] Novo, S.; Aneiros, G. and Vieu, P. (2019). Automatic and location-adaptive estimation in functional single-index regression, *J. Nonparametr. Stat.*, **31**(2), 364–392.

[28] Richard E. Barlow, F.P. (1975). *Statistical theory of reliability and life testing: probability models*. In "International Series in Decision Processes Series in Quantitative Methods for Decision Making", Holt, Rinehart and Winston, 1975.

[29] Shang, H.L. (2020). Estimation of a functional single index model with dependent errors and unknown error density, *Comm. Statist. Simulation Comput.*, **49**(12), 3111–3133.

[30] Strzalkowska-Kominiak, E. and Cao, R. (2013). Maximum likelihood estimation for conditional distribution single-index models under censoring, *J. Multivariate Anal.*, **114**, 74–98.

[31] Wang, G.; Feng, X.N. and Chen, M. (2016). Functional partial linear single-index model, *Scand. J. Stat.*, **43**(1), 261–274.