Strong uniform consistency rates of conditional density estimation in the single functional index model for Functional Data Under Random Censorship

Authors: Nadia Kadiri
– Statistics and Stochastic Processes Laboratory, University of Sidi Bel Abbes, Algeria (nad.kad06@yahoo.com)

Mustapha Meghnafi
– Departement of Mathematics, Faculty of Exact Sciences, University of Béchar, Algeria (megnafi3000@yahoo.fr)

Abbes Rabhi
– Laboratory of Mathematics, University of Sidi Bel Abbes, Algeria (rabhi_abbes@yahoo.fr)

Abstract:

- The main objective of this paper is to investigate the estimation of conditional density function based on the single-index model in the censorship model when the sample is considered as an independent and identically distributed (i.i.d.) random variables. First of all, a kernel type estimator for the conditional density function (*cond-df*) is introduced. Afterwards, the asymptotic properties are stated when the observations are linked with a single-index structure. The pointwise almost complete convergence and the uniform almost complete convergence (with rate) of the kernel estimate of this model are established. As an application the conditional mode in functional single-index model is presented. Finally, a simulation study is carried out to evaluate the performance of this estimate.

Key-Words:

- *Conditional density; Functional single-index process; Functional random variable; Nonparametric estimation; Small ball probability.*

AMS Subject Classification:

- 62G05, 62G99, 62M10.

1. Introduction

Over the last two decades, functional data modeling was highly considered in the statistical literature. The new generation of electronic devices is now allowing practitioners to have access to data continuously (over time and/or space). This change in the data structure raised several challenging statistical problems in analyzing curve-type data. In practice, one can observe functional data in several fields such as climatology, stock market analysis, signal processing, satellite images analysis, etc. For an overview of the functional data analysis , the readers can refer to Ramsay and Silverman [30, 31], Masry [26], Ferraty and Vieu [16], Cuevas [8], Hsing and Eubank [21], Goia and Vieu [18] and the references therein.

Modeling of the relationship between two concomitant variables is one of the most relevant tasks in functional data analysis. In this paper, we are interested in using semi-parametric approach to model the conditional density of a real-valued response variable given an infinite dimensional (functional) covariate. A dimension reduction approach, based on single index model, is used in this paper to estimate the conditional mode whenever the response variable is affected by a right censorship phenomenon.

The problem of estimating the conditional density function has taken considerable attention in the past for both independent and dependent data. Conditional density estimation of a scalar response given a scalar/multivariate covariate has been widely used to estimate some characteristic features of a data set, such as the conditional mode, and gained considerable interest in the statistical literature. For completely observed data, several nonparametric approaches have been proposed. Samanta and Thavaneswaran [32] showed that, under some regularity conditions, the kernel estimator of the conditional mode function was consistent and asymptotically normally distributed. Mehra *et al.* [27] established the law of iterated logarithm (LIL). Under random censoring, Ould-Saïd and Cai [28] established the uniform strong consistency of a nonparametric estimator of the censored conditional mode function, in the i.i.d case using a step function for the interest random variable. For their part, Khardani *et al.* [22] obtained the strong consistency with rate and asymptotic normality. Ould-Saïd [29] constructed a kernel estimator of the conditional quantile under an i.i.d. censorship model and established its strong uniform convergence rate. For the censored dependent case, Khardani *et al.* [23] obtained the strong consistency with rate for the α-mixing framework. The asymptotic normality of the conditional mode estimator for the censored dependent case was proved by Khardani *et al.* [24].

Many authors are interested in the estimation of the conditional mode of a scalar response given a functional covariate. The kernel-type estimators of some characteristics of the conditional cumulative distribution function and the successive derivatives of the conditional density were introduced by Ferraty *et al.* [12]. Some asymptotic properties were established with a particular application to the conditional mode and conditional quantiles. An application to a chemometrical data set coming from food industry is also presented. The uniform strong consistency with rates and the asymptotic normality for the kernel conditional mode estimator were obtained by Ezzahrioui and Ould-Saïd [10] in the i.i.d. case. The asymptotic normality, under α-mixing conditions, of the kernel conditional quantile estimator, was established by Ezzahrioui and Ould-Saïd [11].

In multivariate statistics, where the vector of covariates belongs to a high dimensional but finite space, single index model represents one of the well-known semi-parametric models which allows to reduce the dimensionality of the covariate space and, at the same time, gives flexibility in describing the relationship between the response and the covariate through an unknown link function. Indeed, single index model reduces the curse of dimensionality effect known in pure nonparametric estimation methods and it is always seen as a reasonable compromise between nonparametric and parametric models. Consequently, reducing the dimensionality can be of great interest in practice. For instance, it allows to increase the prediction accuracy and to improve the interpretability of the relationship between a response variable with a vector of covariates. For more details about the advantages of single index models in finite dimensional space setting, the reader can be referred to [19], [20], [33] and the references therein. In our infinite dimensional purpose, we use the terminology *functional nonparametric*, where the word *functional* referees to the infinite dimensionality of the data and where the word *nonparametric* referees to the infinite dimensionality of the model. Such *functional nonparametric* statistics is also called *doubly infinite dimensional* (see Ferraty and Vieu [15], for more details).

The extension of the single index model to the functional data framework was introduced first in Ferraty *et al.* [14] to estimate semi-parametrically the regression operator where the response variable is real-valued and the covariate is a functional random variable. The single functional index model (SFIM) assumes that a functional explanatory variable acts on a scalar response only through its projection on one functional direction. The SFIM was intensively extended to estimate several statistical parameters describing the shape of the conditional distribution. For instance, Aït-Saidi *et al.* [1] were interested in using SFIM to estimate the regression operator and suggest to use a cross-validation procedure allowing the estimated the unknown link function as well as the unknown functional index. Attaoui [4] and Attaoui and Ling [6] studied, respectively, the estimation of the conditional density and the conditional cumulative distribution function based on a SFIM and assuming that the data satisfy a strong mixing condition. Bouchentouf *et al.* [7] were interested in the semi-parametric estimation of the hazard function. Goia and Vieu [17] presented a methodology allowing to approximate in a semi-parametric way the unknown regression operator through a single index approach and by taking possible structural changes into account. Furthermore, Ling *et al.* [25] obtained the asymptotic normality of the conditional density estimator and the conditional mode estimator for the α-mixing dependence functional time series data.

The main contribution of this work, is to establish the pointwise almost complete convergence and the uniform almost complete convergence (with rate) of the conditional density estimator in the single functional index model in i.i.d case under random censorship, this result will be applied to obtain the convergence rates of the conditional mode estimator. Moreover, we prove the asymptotic normality of the estimators of conditional density function and conditional mode. The layout of the paper is as follows, Section 1 presents the functional nonparametric framework. In Section 2 we treat the almost complete convergence, while in section 3 the uniform version is studied. The asymptotic normality is given in section 4, and a simulation study is provided in section 5. Finally, all the proofs of the theoretical results are given in section 6.

1.1. The functional nonparametric framework

Consider a random pair (X,T) where T is valued in $\mathbb{R}$ and X is valued in some infinite dimensional Hilbertian space $\mathcal{H}$ with scalar product $< \cdot,\cdot >$. Let $(X_i,T_i)_{i=1,\ldots,n}$ be the statistical sample of pairs which are identically distributed like (X,T), but not necessarily independent. X is called functional random variable *f.r.v.*

As example, in the classical regression case, the important parameter whose one assumed existence is the regression function of Y knowing the covariate X, denoted $r(x) = \mathbb{E}(Y|X = x)$, $X, Y \in \mathbb{R}^d \times \mathbb{R}$. For this model, the non-parametric method considers only regularity assumptions on the function r. Obviously, this method has some drawbacks. One can cite the problem of curse of dimensionality. This problem appears when the number of regressors d increases, the rate of convergence of the nonparametric estimator r which is supposed k times differentiable is $\mathcal{O}(n^{-k/2k+d})$ deteriorate. The second drawback is the lack of means to quantify the effect of each explanatory variable. To alleviate in these drawbacks, an alternative approach is naturally provided by the semi-parametric model which supposes the introduction of a parameter on the regressors. Assume that the conditional expectation of T given X is done through a fixed functional index θ in $\mathcal{H}$, such that by writing than the regression function is of the form

$$\mathbb{E}_\theta(T|X) = \mathbb{E}(T| < X, \theta >= x).$$

This model was introduced by Ferraty *et al.* [14] and we can refer to Attaoui *et al.* [3] for details. From this model, let $f(\theta, \cdot, x)$ be the conditional density of Y given $< X, \theta >=< x, \theta >$ for $x \in \mathcal{H}$, which also shows the relationship between X and Y but it often unknown.

Let $(T_i)_{i\geq 1}$ be a sequence of independent and identically distributed (i.i.d.) random variables, and assume that they form a strictly stationary sequence of lifetimes. Suppose that there exists a sample of i.i.d. censoring random variable (r.v) $(C_i)_{i\geq 1}$ with common unknown continuous distribution function (df).

In the censored framework, the observed random variables are the triplets (Y_i, δ_i, X_i) with

$$Y_i = \min\{T_i, C_i\} \textit{ and } \delta_i = \mathbf{1}_{T_i \leq C_i}.1 \leq i \leq n,$$

where both of T_i and C_i are expected to exhibit some kind of dependence which ensures the identifiability of the model.

In biomedical case studies, it is assumed that C_i and (T_i, X_i) are independent, this condition is plausible whenever the censoring is independent of the patient's modality.

The Kernel estimator $f_n(\theta, \cdot, x)$ of $f(\theta, \cdot, x)$ is defined by :

$$f_n(\theta, t, x) = \frac{h_H^{-1} \sum_{i=1}^{n} K\left(h_K^{-1}(< x - X_i, \theta >)\right) H\left(h_H^{-1}(t - T_i)\right)}{\sum_{i=1}^{n} K\left(h_K^{-1}(< x - X_i, \theta >)\right)}, \tag{1.1}$$

where the functions K and H are kernels and $h_K = h_{K,n}$ (resp. $h_H = h_{H,n}$) a sequence of positive real numbers.

The Kernel type estimator of the conditional density $f(\theta, \cdot, x)$ adapted for censorship model, can be reformulated from the expression (1.1) as follows :

$$\widetilde{f}(\theta, t, x) = \frac{h_H^{-1} \sum_{i=1}^{n} \frac{\delta_i}{\bar{G}(Y_i)} K\left(h_K^{-1}(< x - X_i, \theta >)\right) H\left(h_H^{-1}(t - Y_i)\right)}{\sum_{i=1}^{n} K\left(h_K^{-1}(< x - X_i, \theta >)\right)},. \tag{1.2}$$

In practice $\bar{G}(\cdot) = 1 - G(\cdot)$ is unknown, then using Kaplan and Meier (1958) estimator, $\bar{G}_n(\cdot)$ will be given as

$$\bar{G}_n(t) = 1 - G_n(t) = \begin{cases} \prod_{i=1}^{n} \left(1 - \frac{1 - \delta_{(i)}}{n - i + 1}\right)^{\mathbf{1}_{\{Y_{(i)} \leq t\}}}, & \text{if } t \leq Y_{(n)}; \\ 0, & \text{if } t > Y_{(n)}. \end{cases} \tag{1.3}$$

where $Y_{(1)} < Y_{(2)} < \ldots < Y_{(n)}$ are the order statistics of Y_i and $\delta_{(i)}$ is the non-censoring indicator corresponding to $Y_{(i)}$.

Therefore, estimator of the conditional density function $f(\theta, \cdot, x)$ is given by

$$\widehat{f}(\theta, t, x) = \frac{h_H^{-1} \sum_{i=1}^{n} \frac{\delta_i}{\bar{G}_n(Y_i)} K\left(h_K^{-1}(< x - X_i, \theta >)\right) H\left(h_H^{-1}(t - Y_i)\right)}{\sum_{i=1}^{n} K\left(h_K^{-1}(< x - X_i, \theta >)\right)}. \tag{1.4}$$

2. Asymptotic study

2.1. Pointwise almost complete rate of convergence

In the following, for any $x \in \mathcal{H}$, let N_x be a fixed neighborhood of x and $\mathcal{S}_{\mathbb{R}}$ is a fixed compact of $\mathbb{R}^+$. Denote $B_\theta(x,h) = \{f \in \mathcal{H} : \ 0 < |<x-f,\theta>| < h\}$ the ball of center x and radius h. Assume that $(C_i)_{i\geq 1}$ are independent and $\tau_G < \infty$ where $\tau_G := \sup\{t : G(t) < 1\}$ and let τ be a positive real number such that $\tau < \tau_G$.

In order to establish the almost complete (a.co.) convergence of our estimator, we need some regular hypotheses as follows.

(H1) $\forall h > 0, \ \mathbb{P}\left(X \in B_\theta(x,h)\right) = \phi_{\theta,x}(h) > 0$

(H2) The conditional density $f(\theta,t,x)$ satisfies the Hölder condition, i.e., $\forall (x_1,x_2) \in N_x \times N_x$, $\forall (t_1,t_2) \in \mathcal{S}^2_{\mathbb{R}}$,

$$|f(\theta,t_1,x_1) - f(\theta,t_2,x_2)| \leq C_{\theta,x}\left(\|x_1-x_2\|^{b_1} + |t_1-t_2|^{b_2}\right), \ \ b_1 > 0, \ b_2 > 0$$

(H3) H is bounded function, such that $\forall (t_1,t_2) \in \mathbb{R}^2, \ |H(t_1) - H(t_2)| \leq C|t_1-t_2|$ and $\displaystyle\int |t|^{b_2} H(t)dt < \infty$,

(H4) K is a positive bounded function with support $[-1,1]$ such that $\forall u \in (0,1) \quad 0 < K(u)$,

(H5) The bandwidths h_K and h_H satisfy

$$\lim_{n \longrightarrow \infty} h_K = 0, \qquad \frac{\log n}{n h_H\, \phi_{\theta,x}(h_K)} \underset{n\to\infty}{\longrightarrow} 0.$$

• Comments on the hypotheses

Our hypotheses are very standard for the conditional density estimation in single functional index model, which have been adopted by Attatoui *et al.* [3]. Hypotheses (H3) and (H5) are technical conditions and are similar to those done in Ferraty and Vieu [16].

Proposition 2.1. *Under conditions (H1)-(H5), we have as n goes to infinity*

$$\sup_{t \in \mathcal{S}_{\mathbb{R}}} \left|\widehat{f}(\theta,t,x) - f(\theta,t,x)\right| = \mathcal{O}\left(h_K^{b_1} + h_H^{b_2}\right) + \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log n}{n h_H \phi_{\theta,x}(h_K)}}\right). \tag{2.1}$$

Proof of Proposition 2.1: Consider now, for $i = 1, \ldots, n$, in what follows, let's denote:

$$K_i(\theta, x) = K(h_K^{-1}(< x - X_i, \theta >)),\ H_i(t) = H\left(h_H^{-1}(t - Y_i)\right),\ \bar{G}_i = \bar{G}(Y_i),$$

$$\begin{aligned}
\widehat{f}_N(\theta, t, x) &= \frac{1}{n\, h_H\, \mathbb{E}(K_1(\theta, x))} \sum_{i=1}^{n} \frac{\delta_i}{\bar{G}_n(Y_i)} K_i(\theta, x) H_i(t) \\
\widetilde{f}_N(\theta, t, x) &= \frac{1}{n\, h_H\, \mathbb{E}(K_1(\theta, x))} \sum_{i=1}^{n} \frac{\delta_i}{\bar{G}(Y_i)} K_i(\theta, x) H_i(t) \\
\widehat{F}_D(\theta, x) &= \frac{1}{n\, \mathbb{E}(K_1(\theta, x))} \sum_{i=1}^{n} K_i(\theta, x)
\end{aligned}$$

The proof is based on the following decomposition, valid for any $t \in \mathcal{S}_{\mathbb{R}}$:

$$\begin{aligned}
\sup_{t \in \mathcal{S}_{\mathbb{R}}} \left| \widehat{f}(\theta, t, x) - f(\theta, t, x) \right| \leq\ & \frac{1}{\widehat{F}_D(\theta, x)} \sup_{t \in \mathcal{S}_{\mathbb{R}}} \left\{ \left| \widehat{f}_N(\theta, t, x) - \widetilde{f}_N(\theta, t, x) \right| \right\} \\
& + \frac{1}{\widehat{F}_D(\theta, x)} \sup_{t \in \mathcal{S}_{\mathbb{R}}} \left\{ \left| \widetilde{f}_N(\theta, t, x) - \mathbb{E}\widetilde{f}_N(\theta, t, x) \right| \right\} \\
& + \frac{1}{\widehat{F}_D(\theta, x)} \sup_{t \in \mathcal{S}_{\mathbb{R}}} \left\{ \left| \mathbb{E}\widetilde{f}_N(\theta, t, x) - f(\theta, t, x) \right| \right\} \\
& + \frac{f(\theta, t, x)}{\widehat{F}_D(\theta, x)} \sup_{t \in \mathcal{S}_{\mathbb{R}}} \left| 1 - \widehat{F}_D(\theta, x) \right|, \qquad (2.2)
\end{aligned}$$

Finally, the proof of this proposition is a direct consequence of the following intermediate results. □

Lemma 2.1. *Under hypotheses (H1)-(H4), and if*

$$n h_H \phi_{\theta,x}(h_K) \longrightarrow \infty, \qquad \frac{\log n}{n h_H\, \phi_{\theta,x}(h_K)} \underset{n \to \infty}{\longrightarrow} 0.$$

we have

$$\sup_{t \in \mathcal{S}_{\mathbb{R}}} \left\{ \left| \widehat{f}_N(\theta, t, x) - \widetilde{f}_N(\theta, t, x) \right| \right\} = \mathcal{O}_{a.s.}\left(\frac{\log \log n}{n} \right).$$

The following lemma shows the asymptotic bias term of $\widetilde{f}_N(\theta, t, x)$ and $\widehat{f}_D(\theta, x)$ as n tends to infinity.

Lemma 2.2. *Under hypotheses (H1)-(H3), we have as* $n \to \infty$

$$\sup_{t \in \mathcal{S}_{\mathbb{R}}} \left| \mathbb{E}\left[\widetilde{f}_N(\theta, t, x) \right] - f(\theta, t, x) \right| = \mathcal{O}\left(h_K^{b_1} + h_H^{b_2} \right). \qquad (2.3)$$

The following result deals with the variance term of the right-hand side of (2.2) which is expressed by: $\sup_{t\in\mathcal{S}_{\mathbb{R}}}\left\{\left|\widetilde{f}_N(\theta,t,x)-\mathbb{E}\widetilde{f}_N(\theta,t,x)\right|\right\}$. For $\widehat{F}_D(\theta,x)-\mathbb{E}\left[\widehat{F}_D(\theta,x)\right]$ the same arguments will be used with a slight difference.

Lemma 2.3. *Under hypotheses (H1), (H4)-(H5), as n goes to infinity, we have*

$$\widehat{F}_D(\theta,x)-\mathbb{E}\widehat{F}_D(\theta,x)=\mathcal{O}_{a.co.}\left(\sqrt{\frac{\log n}{n\,\phi_{\theta,x}(h_K)}}\right),$$

furthermore, we have

$$\sum_{n=1}^{\infty}\mathbb{P}\left(|\widehat{F}_D(\theta,x)\ \leq 1/2\right)<\infty.$$

Lemma 2.4. *Under conditions of the Proposition 2.1, we have as $n\to\infty$*

$$\sup_{t\in\mathcal{S}_{\mathbb{R}}}\left\{\left|\widetilde{f}_N(\theta,t,x)-\mathbb{E}\widetilde{f}_N(\theta,t,x)\right|\right\}=\mathcal{O}_{a.co.}\left(\sqrt{\frac{\log n}{n\,h_H\,\phi_{\theta,x}(h_K)}}\right).$$

We conclude the proof of the Proposition 2.1 by making use the inequality (2.2), in conjunction with lemmas Lemma 2.1, Lemma 2.2 and Lemma 2.3,Lemma2.4

2.2. Pointwise almost complete rate of convergence

In this section, we will consider the problem of the estimation of the conditional mode in the functional single-index model, denoted by $M_\theta(x)$. For this, we assume that $M_\theta(x)$ satisfies the following uniqueness property

(H6) $\forall\varepsilon_0>0,\exists\eta>0,\forall\varphi:$

$$|M_\theta(x)-\varphi(x)|\geq\varepsilon_0\Longrightarrow|f(\theta,\varphi(x),x)-f(\theta,M_\theta(x),x)|\geq\eta.$$

We estimate the conditional mode $M_\theta(x)$ with a random variable $\widehat{M_\theta}(x)$ such that

$$\widehat{M_\theta}(x)=\arg\sup_{t\in\mathcal{S}_{\mathbb{R}}}\hat{f}(\theta,t,x).$$

The difficulty of the problem is naturally linked with the flatness of the function $f(\theta,t,x)$ around the mode M_θ. This flatness can be controlled by the number of vanishing derivatives at point M_θ, and this parameter will also have a great influence on the asymptotic rates of our estimates. More precisely, we introduce the following additional smoothness condition.

(H7) There exists some integer $j > 1$ such that $\forall x \in \mathcal{S}_{\mathcal{H}}$, the function $f(\theta, \cdot, x)$ is j times continuously differentiable w.r.t t on $\mathcal{S}_{\mathbb{R}}$ with

$$f^{(l)}(\theta, M_\theta(x), x) = 0, \quad if \quad 1 \leq l < j$$

and $f^{(j)}(\theta, \cdot, x)$ is uniformly continuous on $\mathcal{S}_{\mathbb{R}}$ such that

$$f^{(j)}(\theta, M_\theta(x), x) \neq 0,$$

where $f^{(j)}(\theta, \cdot, x)$ is the j^{th} order derivative of the conditional density $f(\theta, \cdot, x)$.

Theorem 2.1. *Under hypotheses of proposition2.1 and if the conditional density $f(\theta, \cdot, x)$ satisfies (H6) and (H7), then we get*

$$|\widehat{M}_\theta(x) - M_\theta(x)| = \mathcal{O}\left(h_K^{\frac{b_1}{j}} + h_H^{\frac{b_2}{j}}\right) + \mathcal{O}_{a.co.}\left(\left(\frac{\log n}{n\, h_H\, \phi_{\theta,x}(h_K)}\right)^{\frac{1}{2j}}\right) \tag{2.4}$$

Proof of Theorem 2.1: By the Taylor expansion of $f(\theta, t, x)$ in neighborhood of $M_\theta(x)$, we get

$$\widehat{f}(\theta, \widehat{M}_\theta(x), x) = f(\theta, M_\theta(x), x) + \frac{f^{(j)}(\theta, M_\theta^*(x), x)}{j!}(\widehat{M}_\theta(x) - M_\theta(x))^j, \tag{2.5}$$

where $M_\theta^*(x)$ is between $M_\theta(x)$ and $\widehat{M}_\theta(x)$.

Combining the last equality with the fact that

$$|\widehat{f}(\theta, \widehat{M}_\theta(x), x) - f(\theta, M_\theta(x), x)| \leq 2 \sup_{t \in \mathcal{S}_{\mathbb{R}}} |\widehat{f}(\theta, t, x) - f(\theta, t, x)|$$

allow to write :

$$|\widehat{M}_\theta(x) - M_\theta(x)|^j \leq \frac{j!}{f^{(j)}(\theta, M_\theta^*, x)} \sup_{t \in \mathcal{S}_{\mathbb{R}}} |\widehat{f}(\theta, t, x) - f(\theta, t, x)|.$$

Using the second part of (H7) we obtain that,

$$\exists c > 0, \quad \sum_{n=1}^{\infty} \mathbb{P}\left(f^{(j)}(\theta, M_\theta^*, x) < c\right) < \infty.$$

So, we would have

$$|\widehat{M_\theta}(x) - M_\theta(x)|^j = \mathcal{O}_{a.co.}\left(\sup_{t\in\mathcal{S}_\mathbb{R}}\left|\widehat{f}(\theta,t,x) - f(\theta,t,x)\right|\right). \tag{2.6}$$

Finally, Theorem 2.1 can be deduced from proposition 2.1 □

Theorem 2.2. *Under the hpotheses of Proposition 2.1 thus we have,*

$$\widehat{M_\theta}(x) - M_\theta(x) \underset{n\to\infty}{\longrightarrow} 0, \ a.co. \tag{2.7}$$

Proof of Theorem 2.2: Because the continuity of the function $f(\theta,t,x)$, we have, for all $\varepsilon > 0$, $\exists\eta(\varepsilon) > 0$ such that

$$|f(\theta,t,x) - f(\theta,M_\theta(x),x)| \leq \eta(\varepsilon) \Longrightarrow |t - M_\theta(x)| \leq \varepsilon.$$

Therefore, for $t = \widehat{M_\theta}(x)$,

$$\mathbb{P}\left(|\widehat{M_\theta}(x) - M_\theta(x)| > \varepsilon\right) \leq \mathbb{P}\left(|f(\theta,\widehat{M_\theta}(x),x) - f(\theta,M_\theta(x),x)| > \eta(\varepsilon)\right).$$

Then, according to theorem, $\widehat{M_\theta} - M_\theta$ go almost completely to 0, as n goes to infinity. □

3. Uniform almost complete convergence and rate of convergence

In this section, we devote the result of the uniform version of Proposition2.1. The study of the uniform consistency is a crucial tool for studying the asymptotic properties of all estimates of the functional index if is unknown. In the multivariate case, the uniform consistency is a standard extension of the pointwise one, nevertheless, in the studied case, it requires some additional tools and topological conditions (see Ferraty *et al.* [13]). Consequently, coupled with the conditions introduced antecedently, we need the following ones. Firstly, consider

$$\mathcal{S}_\mathcal{H} \subset \bigcup_{k=1}^{d_n^{\mathcal{S}_\mathcal{H}}} B_\theta(x_k, r_n) \ \text{ and } \ \Theta_\mathcal{H} \subset \bigcup_{q=1}^{d_n^{\Theta_\mathcal{H}}} B_\theta(\theta_q, r_n) \tag{3.1}$$

with x_k (resp. θ_q) $\in \mathcal{H}$ and $r_n, d_n^{\mathcal{S}_\mathcal{H}}, d_n^{\Theta_\mathcal{H}}$ are sequences of positive real numbers which tend to infinity as n goes to infinity and suppose that $d_n^{\mathcal{S}_\mathcal{H}}$, $d_n^{\Theta_\mathcal{H}}$ are the minimal numbers of open balls with radius r_n in $\mathcal{H}$, which are required to cover $\mathcal{S}_\mathcal{H}$ and $\Theta_\mathcal{H}$. Moreover, the following assumptions are also satisfied.

(A1) There exists a differentiable function $\phi(\cdot)$ such that $\forall x \in \mathcal{S}_{\mathcal{H}}$ and $\forall \theta \in \Theta_{\mathcal{H}}$,

$$0 < C\phi(h) \leq \phi_{\theta,x}(h) \leq C'\phi(h) < \infty \text{ and } \exists \eta_0 > 0,\ \forall \eta < \eta_0,\ \phi'(\eta) < C,$$

(A2) The kernel K satisfy (H4) and Lipschitz's condition holds

$$|K(x) - K(y)| \leq C\|x - y\|,$$

(A3) The conditional density $f(\theta, t, x)$ satisfies the uniform Hölder condition, i.e, $\forall (t_1, t_2) \in \mathcal{S}_{\mathbb{R}} \times \mathcal{S}_{\mathbb{R}}, \forall (x_1, x_2) \in \mathcal{S}_{\mathcal{H}} \times \mathcal{S}_{\mathcal{H}}$ and $\forall \theta \in \Theta_{\mathcal{H}}$,

$$|f(\theta, t_1, x_1) - f(\theta, t_2, x_2)| \leq C\left(\|x_1 - x_2\|^{b_1} + |t_1 - t_2|^{b_2}\right),$$

(A4) For some $\nu \in (0,1)$, $\lim\limits_{n\to\infty} n^{\nu} h_H = \infty$, and for $r_n = \mathcal{O}\left(\frac{\log n}{n}\right)$, the sequences $d_n^{\mathcal{S}_{\mathcal{H}}}$ and $d_n^{\Theta_{\mathcal{H}}}$ satisfy:

$$\frac{(\log n)^2}{n h_H \phi(h_K)} < \log d_n^{\mathcal{S}_{\mathcal{H}}} + \log d_n^{\Theta_{\mathcal{H}}} < \frac{n h_H \phi(h_K)}{\log n},$$

and

$$\sum_{n=1}^{\infty} n^{(3\gamma+1)/2} (d_n^{\mathcal{S}_{\mathcal{H}}} d_n^{\Theta_{\mathcal{H}}})^{1-\beta} < \infty, \quad for \quad some \quad \beta > 1.$$

In what follows, denote

$$\Lambda_i(x,\theta) = \frac{1}{h_K \phi(h_K)} \mathbf{1}_{B_\theta(x,h) \cup B_\theta(x_{k(x)},h)}(X_i),$$

$$\Omega_i(x,\theta) = \frac{1}{h_K \phi(h_K)} \mathbf{1}_{B_\theta(x_{k(x)},h) \cup B_{\theta_{q(\theta)}}(x_{k(x)},h)}(X_i),$$

$$\Delta_i\left(x_{k(x)}, \theta_{q(\theta)}\right) = \frac{K\left(h_K^{-1} < x_{k(x)} - X_i, \theta_{q(\theta)} >\right)}{\mathbb{E}K\left(h_K^{-1} < x_{k(x)} - X_i, \theta_{q(\theta)} >\right)},$$

and

$$\begin{aligned}\Gamma_i(x_{k(x)}, v_{k_t}, \theta_{q(\theta)}) \;=\; & \frac{1}{h_H} \frac{K\left(h_K^{-1} < x_{k(x)} - X_i, \theta_{q(\theta)} >\right)}{\mathbb{E}K\left(h_K^{-1} < x_{k(x)} - X_i, \theta_{q(\theta)} >\right)} H\left(h_H^{-1}(v_{k_t} - Y_i)\right) \\ & - \frac{1}{h_H} \mathbb{E}\left(\frac{K\left(h_K^{-1} < x_{k(x)} - X_i, \theta_{q(\theta)} >\right)}{\mathbb{E}K\left(h_K^{-1} < x_{k(x)} - X_i, \theta_{q(\theta)} >\right)} H\left(h_H^{-1}(v_{k_t} - Y_i).\right)\right)\end{aligned}$$

Remark 3.1. Note that Assumptions (A1) and (A3) are, respectively, the uniform version of (H1) and (H2). Assumptions (A1) and (A4) are linked with the topological structure of the functional variable, see Ferraty *et al.* [13].

Theorem 3.1. *Under Assumptions (A1)-(A4), we have, as n goes to infinity*

(3.2)
$$\sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{x\in\mathcal{S}_{\mathcal{H}}}\sup_{t\in\mathcal{S}_{\mathbb{R}}}\left|\widehat{f}(\theta,t,x)-f(\theta,t,x)\right| = \mathcal{O}\left(h_K^{b_1}\right)+\mathcal{O}\left(h_H^{b_2}\right)+\mathcal{O}_{a.co.}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{F}}}+\log d_n^{\Theta_{\mathcal{F}}}}{n\,h_H\,\phi(h_K)}}\right).$$

Corollary 3.1. *Under the assumptions of Theorem 3.1 and hypotheses (H6)-(H7), we have*

$$\sup_{x\in\mathcal{S}_{\mathcal{H}}}|\widehat{M_\theta}(x)-M_\theta(x)|^j = \mathcal{O}(h_K^{b_1})+\mathcal{O}(h_H^{b_2})+\mathcal{O}_{a.co}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{F}}}+\log d_n^{\Theta_{\mathcal{F}}}}{n\,h_H\,\phi(h_K)}}\right). \quad (3.3)$$

The proof of Theorem 3.1 and Corollary3.1 can be completed by the following lemmas.

Lemma 3.1. *Under assumptions (A1), (A3) and (A4), we have as* $n \longrightarrow \infty$

$$\sup_{\theta\in\Theta_{\mathcal{F}}}\sup_{x\in\mathcal{S}_{\mathcal{F}}}\left|\widehat{F}_D(\theta,x)-\mathbb{E}\widehat{F}_D(\theta,x)\right| = \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{F}}}+\log d_n^{\Theta_{\mathcal{F}}}}{n\phi(h_K)}}\right)$$

Corollary 3.2. *Under assumptions of Lemma 3.1, we have*

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

Lemma 3.2. *Under assumptions (A1),(A3) and (H3), we have as n goes to infinity*

$$\sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{x\in\mathcal{S}_{\mathcal{H}}}\sup_{t\in\mathcal{S}_{\mathbb{R}}}\left|f(\theta,t,x)-\mathbb{E}\widehat{f}_N(\theta,t,x)\right| = \mathcal{O}\left(h_K^{b_1}\right)+O\left(h_H^{b_2}\right). \quad (3.4)$$

Lemma 3.3. *Under the assumptions of Theorem 3.1, we have as n goes to infinity*

$$\sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{x\in\mathcal{S}_{\mathcal{H}}}\sup_{t\in\mathcal{S}_{\mathbb{R}}}\left|\widetilde{f}_N(\theta,t,x)-\mathbb{E}\widetilde{f}_N(\theta,t,x)\right| = \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{F}}}+\log d_n^{\Theta_{\mathcal{F}}}}{n\,h_H\,\phi(h_K)}}\right).$$

4. Asymptotic Normality

In this section, the asymptotic normality of the conditional density and the conditional mode are established. Therefore, further assumptions are required. Assume that

(N1) There exists a function $\xi^{\theta,x}$, such that

$$\forall u \in [0,1], \ \lim_{h \longrightarrow 0} \frac{\phi_{\theta,x}(uh)}{\phi_{\theta,x}(h)} = \lim_{h \longrightarrow 0} \xi_h^{\theta,x}(u) = \xi_0^{\theta,x}(u).$$

(N2) The bandwidth h_H satisfies,

$$nh_H^3 \phi_{\theta,x}^3(h_K) \longrightarrow 0, \quad as \quad n \longrightarrow \infty.$$

(N3) The *df* of the censored random variable, G has a bounded first derivative $G^{'}$.

(N4) The conditional density function $f(\theta, t, x)$ satisfies: $\exists \beta_0 > 0, \forall (t_1, t_2) \in \mathcal{S}_{\mathbb{R}} \times \mathcal{S}_{\mathbb{R}}$,

$$|f^{(q)}(\theta, t_1, x) - f^{(q)}(\theta, t_2, x)| \leq C(|t_1 - t_2|^{\beta_0}), \quad \forall q = 1, 2.$$

- (N5)] $H^{'}$ and $H^{''}$ are bounded respectively with

$$\int (H^{'}(t))^2 dt < \infty, \quad \int |t|^{\beta_0} H(t) dt < \infty$$

Theorem 4.1. *Under assumptions (H1)-(H5) and (N1)-(N3) for all $x \in \mathcal{H}$, we have*

$$\sqrt{\frac{nh_H \phi_{\theta,x}(h_K)}{\sigma^2(\theta, t, x)}} \left(\widehat{f}(\theta, t, x) - f(\theta, t, x)\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1), \quad as \quad n \longrightarrow \infty,$$

where

$$\sigma^2(\theta, t, x) = \frac{a_2(\theta, x) f(\theta, t, x)}{(a_1(\theta, x))^2 \bar{G}(t)} \int_{\mathbb{R}} H^2(u) du,$$

with

$$a_l(\theta, x) = K^l(1) - \int_0^1 (K^l)^{'}(u) \xi_0^{\theta,x}(u) du, \ l = 1, 2$$

" $\xrightarrow{\mathcal{D}}$ " *means the convergence in distribution*

Proof: In order to establish the asymptotic normality of $\widehat{f}(\theta,t,x)$, we need further notations and definitions. First we consider the following decomposition

$$\begin{aligned}
\widehat{f}(\theta,t,x) - f(\theta,t,x) &= \frac{\widehat{f}_N(\theta,t,x)}{\widehat{F}_D(\theta,x)} - \frac{a_1(\theta,x)f(\theta,t,x)}{a_1(\theta,x)} \\
&= \frac{1}{\widehat{F}_D(\theta,x)}\Big(\widehat{f}_N(\theta,t,x) - \mathbb{E}\widehat{f}_N(\theta,t,x)\Big) \\
&\quad - \frac{1}{\widehat{F}_D(\theta,x)}\Big(a_1(\theta,x)f(\theta,t,x) - \mathbb{E}\widehat{f}_N(\theta,t,x)\Big) \\
&\quad + \frac{f(\theta,t,x)}{\widehat{F}_D(\theta,x)}\Big(a_1(\theta,x) - \mathbb{E}\widehat{F}_D(\theta,x)\Big) \\
&\quad - \frac{f(\theta,t,x)}{\widehat{F}_D(\theta,x)}\Big(\widehat{F}_D(\theta,x) - \mathbb{E}\widehat{F}_D(\theta,x)\Big) \\
&= \frac{1}{\widehat{F}_D(\theta,x)}\left(A_n(\theta,t,x) + B_n(\theta,t,x)\right).
\end{aligned}$$

□

Where

$$\begin{aligned}
A_n(\theta,t,x) &= \frac{1}{nh_H\mathbb{E}K_1(\theta,x)}\sum_{i=1}^{n}\{\left(\frac{\delta_i}{\bar{G}(Y_i)}H_i(t) - h_H f(\theta,t,x)\right)K_i(\theta,x) \\
&\quad - \mathbb{E}\left[\left(\frac{\delta_i}{\bar{G}(Y_i)}H_i(t) - h_H f(\theta,t,x)\right)K_i(\theta,x)\right]\} \\
&= \frac{1}{nh_H\mathbb{E}K_1(\theta,x)}\sum_{i=1}^{n}N_i(\theta,t,x).
\end{aligned}$$

It follows that,

$$\begin{aligned}
nh_H\phi_{\theta,x}(h_K)Var(A_n(\theta,t,x)) &= \frac{\phi_{\theta,x}(h_K)}{h_H(\mathbb{E}K_1(\theta,x))^2}Var(N_1(\theta,t,x)) \\
&= V_n(\theta,t,x),
\end{aligned}$$

and

$$B_n(\theta,t,x) = a_1(\theta,x)f(\theta,t,x) - \mathbb{E}\widehat{f}_N(\theta,t,x) + f(\theta,t,x)(a_1(\theta,x) - \mathbb{E}\widehat{F}_D(\theta,x)).$$

Then, the proof of Theorem 4.1 can be deduced from the following Lemmas.

Lemma 4.1. *Under conditions of Theorem 4.1 ,we have*

$$\sqrt{nh_H\phi_{\theta,x}(h_K)}A_n(\theta,t,x) \xrightarrow{\mathcal{D}} \mathcal{N}(0,\sigma^2(\theta,t,x))$$

where $\sigma^2_{\theta,x}$ is given in Theorem 4.1.

Lemma 4.2. *Under assumptions (H1)-(H5) and (N1)-(N2), we have as* $n \longrightarrow \infty$,

$$\sqrt{nh_H\phi_{\theta,x}(h_K)}B_n(\theta,t,x) \longrightarrow 0 \quad in \quad probability.$$

Corollary 4.1. *If the assumptions (H1)-(H7) as well as (N1)-(N5) hold, then, we have :*

$$\sqrt{\frac{nh_H^3\phi_{\theta,x}(h_K)}{\sigma_1^2(\theta,x)}}(\widehat{M}_\theta(x) - M_\theta(x)) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1), \quad as \quad n \longrightarrow \infty, \tag{4.1}$$

where

$$\sigma_1^2(\theta,t,x) = \frac{a_2(\theta,x)f(\theta,M_\theta(x),x)}{(a_1(\theta,x)f^{(2)}(\theta,M_\theta(x),x))^2\bar{G}(t)}\int_{\mathbb{R}} H'^2(u)du.$$

5. Simulation Study

This section aims at illustrating our study which the forecast via the conditional mode. More precisely, we will compare our model CFSIM (1.4) (censored functional single index model) with CNPFDA (5.1) (censored nonparametric functional data analysis) in censored data.

$$\widehat{f}_n(t,x) = \frac{h_H^{-1}\sum_{i=1}^{n}\frac{\delta_i}{\bar{G}_n(Y_i)}K\left(h_K^{-1}d(x,X_i)\right)H\left(h_H^{-1}(t-Y_i)\right)}{\sum_{i=1}^{n}K\left(h_K^{-1}d(x,X_i)\right)}. \tag{5.1}$$

Note that all the routines for functional data used in this implementation (developed in R/S-Plus software) are available on the website https://www.math.univ-toulouse.fr/staph/npfda/.

We consider a diffusion process on the interval $[0,1]$:

$$X_i(t) = cos(\pi b_i t) + a_i t^2, \quad i = 1,\dots,200; \quad t \in [0,1] \tag{5.2}$$

where a_i are uniformly distributed on $[0,1]$ ($a \sim \mathcal{U}(0,1)$) and b_i are standard normal distribution ($b \sim \mathcal{N}(0,1)$). we carry out the simulation with a 200 sample of the curves X(t) (see Figure 1).

5.1. Estimating the single index in practice

The single index θ is unknown and has to be estimated. In practice this parameter can be selected by cross-validation approach (see Aït Saidi *et al.* [2]).

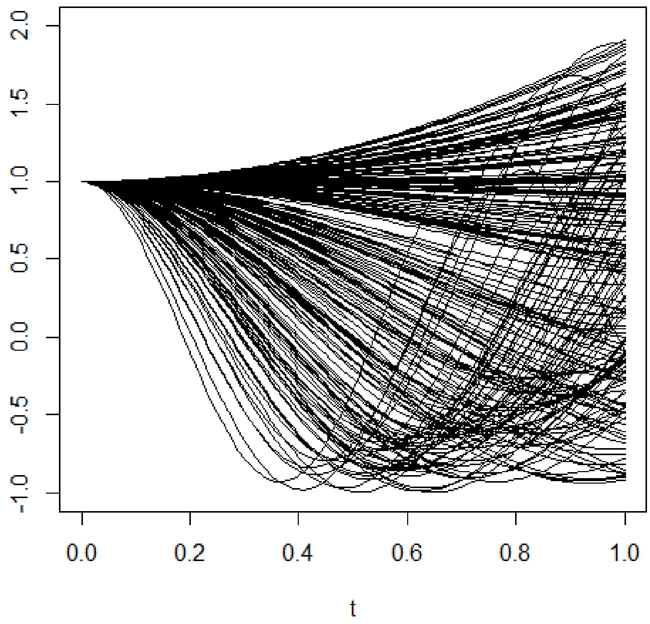

Figure 1: The curves $X_{i=1,\ldots,200}(t), \quad t \in [0,1]$.

We simulates the single functional index model as follows, first, we choose the functional parameter θ.

So for $\mathcal{L} = \{1, \ldots, 200\}$, the best approximation of θ is to estimate the eigenfunctions of the covariance operator $\mathbb{E}[(X' - \mathbb{E}(X')) < X', \cdot >_{\mathcal{H}}]$ by its empirical covariance $\frac{1}{\mathcal{L}} \sum_{i \in \mathcal{L}} (X_i' - \mathbb{E}(X'))^t (X_i' - \mathbb{E}(X'))$ [5]. (Fig2,3 and 4) shows the discretization of the two first eigenfunction, twenty and all the eigenfunctions $\theta_i(t)$ respectively.

Taking θ^* the first eigenfunction corresponding to the first higher eigenvalue, and compute the inner product $< \theta^*, X_1 >, \ldots, < \theta^*, X_{200} >$,Then simulate the response variables $T_i = r(< \theta^*, X_i >) + \epsilon$,

where $r(< \theta^*, X_i >) = \exp(10(< \theta^*, X_i > -0.05))$ and ϵ generate independently from a centered gaussian of variance equal to 0.05 times the empirical variance of $r(< \theta^*, X_i >)$.

We simulate n i.i.d rv $C_i, i = 1, \ldots, n$ with the exponential distribution $\mathcal{E}(1, 5)$. Noting that the computation of those estimators are based on the observed data $(X_i, Y_i, \delta_i)_{i=1,\ldots,n}$, where $Y_i = min(T_i, C_i)$ and $\delta_i = \mathbf{1}_{\{T_i \leq C_i\}}$. On the

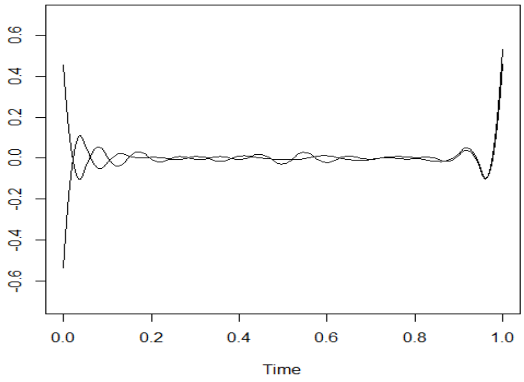

Figure 2: The curves $\theta_{i=1,2}(t), \quad t \in [0,1]$.

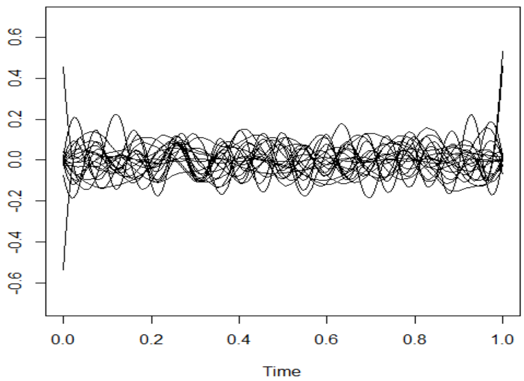

Figure 3: The curves $\theta_{i=1,\ldots,20}(t), \quad t \in [0,1]$.

other hand, we choose the quadratic kernels defined by :

$$K(u) = \frac{3}{2}(1-u^2)\mathbf{1}_{(0,1)}(u)$$

and

$$H(t) = \frac{3}{4}(1-t^2)\mathbf{1}_{(-1,1)}(t)$$

then, taking into account of the smoothness of the curves $X_i(t)$, we choose for the CNPFDA model the semi-metric in $\mathcal{H}$:

$$d(x_i, x_j) = \sqrt{\int_0^1 \left(x_i^{'}(t) - x_j^{'}(t)\right)^2 dt}, \qquad x_i, x_j \in \mathcal{H}$$

For the bandwidths $h_H \sim h_K =: h$ is automatically selected by the procedure of the cross-validation method on the k-nearest neighbors ([16]).

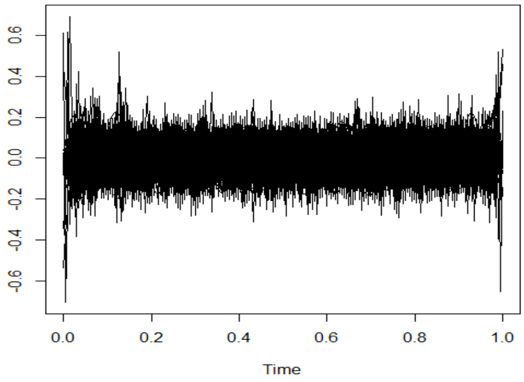

Figure 4: The curves $\theta_{i=1,\dots,200}(t), \quad t \in [0,1]$.

In our simulation, sample sizes are $n = 200$, we take it into two parts, one is a learning sample of 150 observations and the others 50 observations are a test sample. Then using the learning sample to compute the estimator of $\widehat{Y}_i = \widehat{M_{\theta^*}}(X_i)$ and $\widehat{Y}_{ni} = \widehat{M}(X_i)$ for i={151,. . . ,200}.

Finally we show the results by plotting the true values versus the predicted values for the MSE under censored data for both estimators 1.4 and 5.1 which are defined as :

$$CFSIM.MSE = \frac{1}{50}\sum_{i=151}^{200}(Y_i - \widehat{Y}_i)^2,$$

$$CNPFDA.MSE = \frac{1}{50}\sum_{i=151}^{200}(Y_i - \widehat{Y}_{ni})^2$$

respectively.

By, Figures 5 and 6, we can say that both estimators on weak censored rates 3.5% works almost as well as if we had the complete data-set. To show how the different censored rates (CRs) effects the prediction results, we present some CRs and their corresponding MSE for CFSIM and CNPFDA. Two sample sizes are considered n=200 and 300, and for each sample size different censoring rates are taken CR=5%, 16%, 27% and 54%, we carried 100 independent replications of the experiment, then we computed the average of mean squared error. These quantities are presented in Table (1).

One can observe, that both estimators have a reasonable performance for a lower censored rates, however, they are strongly effected when the percentage of censored rate is high, but the FSIM estimator stay more accurate than the NPFDA one in all cases. And on the other hand, when the sample size increases, the precise of forecast is also increase.

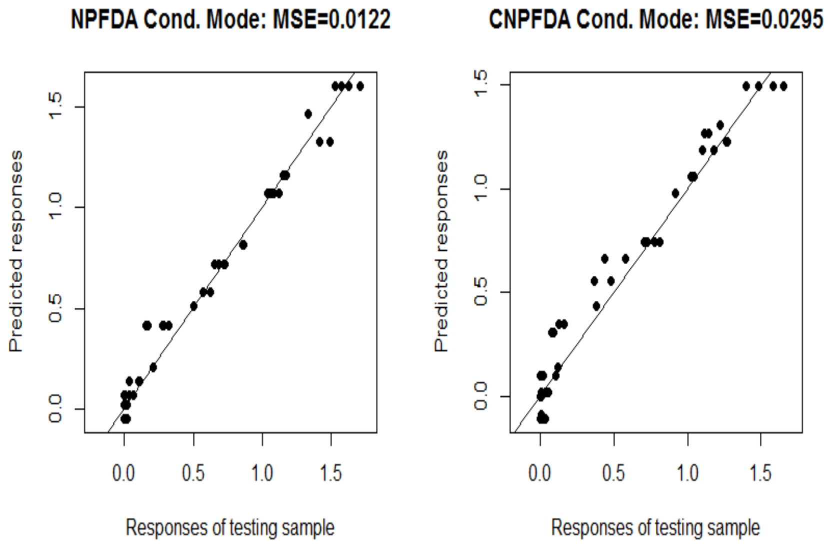

Figure 5: Prediction via the conditional mode by NPFDA for complete data (MSE=0.0122) and censored data with CR$\sim$ 3.5% (MSE=0.0295)

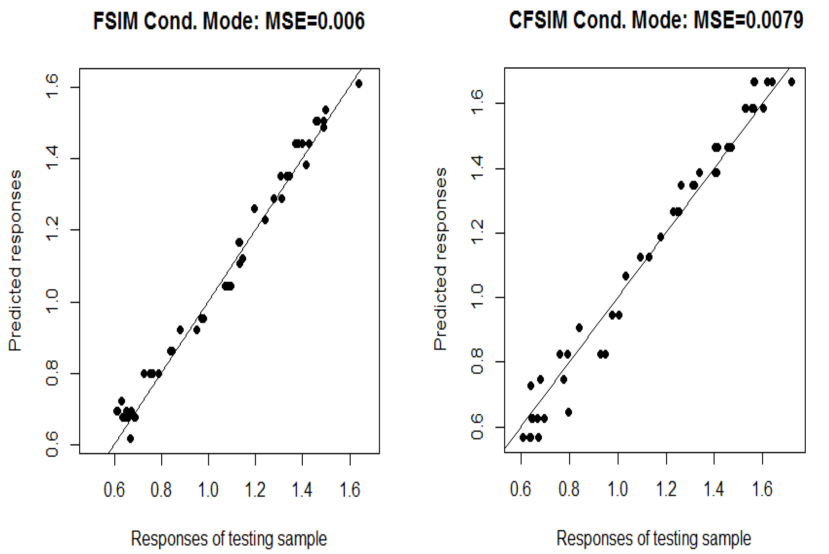

Figure 6: Prediction via the conditional mode by FSIM for complete data (MSE=0.006) and censored data with CR$\sim$ 3.5% (MSE=0.0079)

5.2. Real data example: Peak electricity demand

We evaluate and compare the finite sample performance between a nonparametric functional model with our estimator (the functional single index model).

n	CR%	NPFDA	CNPFDA	FSIM	CFSIM
	5%	0.0122	0.0343	0.0065	0.0168
200	16%		0.0844		0.0554
	27%		0.1274		0.1084
	54%		0.3349		0.3234
	5%	0.0115	0.0303	0.0032	0.0063
300	16%		0.0779		0.0473
	27%		0.1245		0.0952
	54%		0.3165		0.1600

Table 1: MSE comparaison for FSIM and NPFDA

To this end, we apply our method to the data constituting hourly electricity demand for the Rocky Mountain region (WACM) of the United States. The data are daily electricity demands divided into 24 grids, where (each hour of the day corresponds to a grid), from July 2015 to November 2018. The updated version of the data can be found on the site http://www.eia.gov/.

We construct our variables as follow: the observations of our covariate X are the daily electricity demands from 2016 to 2018, $X_i = (x_{i1}, \ldots, x_{i24})$, our sample consists of $n = 1037$ observations. The observations of our response variable Y are $Y_i = \min(\max(X_i), 1408)$, $i = 1, \ldots, n$, where 1408 is the maximum peak of electricity demands in 2015.

In this part, we use Kaplan-Meier's estimator $\overline{G}_n(\cdot)$ as an estimator of $\overline{G}(\cdot)$ to construct our conditional distribution estimator, by taking the variables $(C_i)_i$ as deterministic (all equal to 1408, which is the maximum of the peak observed in 2015).

Since we are performing analysis on a time series spread over 4 years, considering the year 2015 as a base year, and in the simulation we are interested only in the years 2016-2018, we can consider 1408 as a maximum amplitude, that is, any value (or hourly observation) greater than 1408 can be considered as aberrant data. So, on this basis, we built our response variable.

Concerning the estimation of our parameters, we chose $deriv_1$ (the semi-metric based on the first derivatives of the curves) as semi-metric, the kernel $K(\cdot)$ and the cumulative df $H(u)$ are defined in the subsection 5.1. Then, as discussed previously, the optimal bandwidth $h_n = h_{K,n} = h_{H,n}$, are chosen using the cross-validation method on the k-nearest neighbors. Finally, we replace θ by the first eigenfunction corresponding to the first higher eigenvalue of the empirical covariance operator. The curves of the data are represented in Figure 7.

To assess the in-sample estimation accuracy and out-of sample prediction accuracy of the models, we split the original 1037 samples into two samples. The first one (learning set), from 1 to 960, used for the estimation, while the second sample (testing set), from 961 to 1037, is served for the prediction. To measure

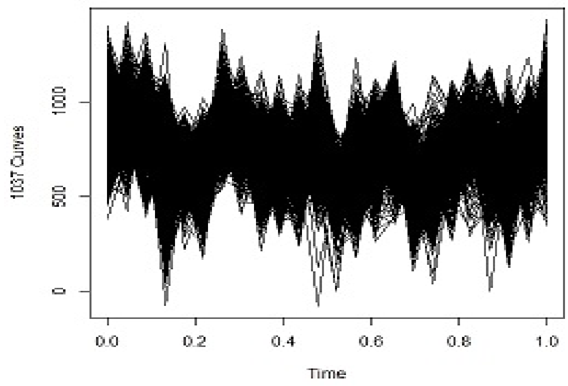

Figure 7: A sample curves $\{X_i(t),\ t \in [0,1]\}_{i=1\ldots,1037}$.

the estimation and prediction accuracies, we evaluate and compare the forecast accuracy using the testing sample, from which we predict responses in the testing sample. To measure the performance of each functional prediction method, we consider the mean square errors (MSE).

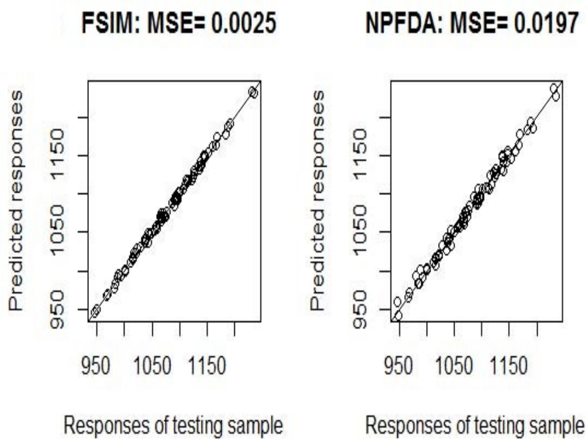

Figure 8: Prediction via the conditional mode by FSIM with error $MSE = 0.0025$ against NPFDA with error $MSE = 0.0197$

After performing the calculations, we find $MSE = 0.0025$ for our estimator, and an $MSE = 0.0197$ for that of NPFDA (see Figure 8). We can therefore conclude that there is an improvement in estimation and prediction accuracies for our model in comparison to the nonparametric functional model.

6. Proofs of technical lemmas

Proof of Lemma 2.1: The proof is similar to that of Lemma 5.2 in [22]. From Equations (1.2) and (1.4), we have

$$\begin{aligned}
|\widehat{f}_N(\theta,t,x)-\widetilde{f}_N(\theta,t,x)| &\leq \frac{h_H^{-1}}{n\mathbb{E}K_1(\theta,x)}\sum_{i=1}^n\Big|\frac{\delta_i}{\bar{G}_n(Y_i)}K_i(\theta,x)H_i(t)-\frac{\delta_i}{\bar{G}(Y_i)}K_i(\theta,x)H_i(t)\Big| \\
&\leq \frac{h_H^{-1}}{n\mathbb{E}K_1(\theta,x)}\sum_{i=1}^n|\delta_iK_i(\theta,x)H_i(t)|\left|\frac{1}{\bar{G}_n(Y_i)}-\frac{1}{\bar{G}(Y_i)}\right| \\
&\leq \frac{h_H^{-1}}{\phi_{\theta,x}(h_K)}\frac{C}{\bar{G}_n(\tau_G)\bar{G}(\tau_G)}\sup_{t\leq\tau_G}{}_{\mathbb{R}}|\bar{G}_n(t)-\bar{G}(t)|\frac{1}{n}\sum_{i=1}^n|K_i(\theta,x)H_i(t)|.
\end{aligned}$$

(6.2)

Since $\bar{G}(t_G)>0$, together with the SLLN and the LIL on the censoring law (see formula (4.28) in Deheuvels and Einmahl [9]), we obtain

$$\sup_{t\leq\tau_G}\left|\bar{G_n}(t)-\bar{G}(t)\right| = O_{a.s.}\left(\frac{\log\log n}{n}\right).$$

We achieve the proof by considering the conditions (H3) and (H4).

□

Proof of Lemma 2.2: We have

$$\begin{aligned}
\mathbb{E}\widetilde{f}_N(\theta,t,x)-f(\theta,t,x) &= \frac{1}{h_H\,\mathbb{E}K_1(x,\theta)}\mathbb{E}\Big(\frac{\delta_i}{\bar{G}(Y_i)}K_i(x,\theta)H_i(t)\Big)-f(\theta,t,x) \\
(6.3)\qquad &= \frac{1}{h_h\,\mathbb{E}K_1(x,\theta)}\mathbb{E}\Big(K_i(x,\theta)\Big[\mathbb{E}\Big(\frac{\delta_i}{\bar{G}(Y_i)}H_i(t)|<X_1,\theta>\Big)-h_H\,f(\theta,t,x)\Big]\Big).
\end{aligned}$$

Using the fact that H is a *cdf* and the use a double conditioning with respect to T_1, we can easily get

$$\begin{aligned}
I &= \mathbb{E}\left(\frac{\delta_i}{\bar{G}(Y_i)}H_i(t)| < X_1,\theta >\right)\\
&= \mathbb{E}\left(\mathbb{E}\left[\frac{\mathbf{1}_{T_1\leq C_1}}{\bar{G}(T_1)}H\left(\frac{t-T_1}{h_H}\right)| < X_1,\theta >, T_1\right]\right)\\
&= \mathbb{E}\left(\frac{1}{\bar{G}(T_1)}H\left(\frac{t-T_1}{h_H}\right)\mathbb{E}\left[\mathbf{1}_{T_1\leq C_1}|T_1\right]| < X_1,\theta >\right)\\
&= \mathbb{E}\left[H\left(\frac{t-T_1}{h_H}\right)| < X_1,\theta >\right]\\
&= \int_{\mathbb{R}} H\left(\frac{t-u}{h_H}\right) f(\theta,u,X_1)du,\\
&= h_H\int_{\mathbb{R}} H(v)f(\theta,t-vh_H,X_1)dv,\\
&= h_H\int_{\mathbb{R}} H(v)\left(f(\theta,t-vh_H,X_1)-f(\theta,t,x)\right)dv + h_H\, f(\theta,t,x)\int_{\mathbb{R}} H(v)dv,
\end{aligned}$$

we can write, because of (H2) and (H3):

$$\begin{aligned}
I &\leq h_H\, C_{x,\theta}\int_{\mathbb{R}} H(v)\left(h_K^{b_1}+|v|^{b_2}h_H^{b_2}\right)dv + h_H\, f(\theta,t,x)\\
&= \mathcal{O}\left(h_K^{b_1}+h_H^{b_2}\right)+h_H\, f(\theta,t,x).
\end{aligned}$$

Combining this last result with (6.3) allows us to achieve the proof. $\square$

Proof of Lemma 2.4: Using the compactness of $\mathcal{S}_{\mathbb{R}}$, we can write that $\mathcal{S}_{\mathbb{R}} \subset \bigcup_{k=1}^{\tau_n}(z_k-l_n, z_k+l_n)$ with l_n and τ_n can be chosen such that $l_n = C\tau_n^{-1} \sim Cn^{-\varsigma-1/2}$. Taking $k_t = \arg\min_{\{z_1,\ldots,z_{\tau_n}\}}|t-z_k|$.

Thus, we have the following decomposition:

$$\begin{aligned}
\frac{1}{\widehat{F}_D(\theta,x)}\sup_{t\in\mathcal{S}_{\mathbb{R}}}\left|\widetilde{f}_N(\theta,t,x)-\mathbb{E}\widetilde{f}_N(\theta,t,x)\right| &\leq \frac{1}{\widehat{F}_D(\theta,x)}\sup_{t\in\mathcal{S}_{\mathbb{R}}}\left|\widetilde{f}_N(\theta,t,x)-\widehat{f}_N(\theta,t_k,x)\right|\\
&+\frac{1}{\widehat{F}_D(\theta,x)}\sup_{t\in\mathcal{S}_{\mathbb{R}}}\left|\widehat{f}_N(\theta,t_k,x)-\mathbb{E}\widehat{f}_N(\theta,t_k,x)\right|\\
&+\frac{1}{\widehat{F}_D(\theta,x)}\sup_{t\in\mathcal{S}_{\mathbb{R}}}\left|\mathbb{E}\widehat{f}_N(\theta,t_k,x)-\mathbb{E}\widetilde{f}_N(\theta,t,x)\right|\\
&\leq T_1+T_2+T_3
\end{aligned}$$

- On the one hand, as the first and the third terms can be treated in the same manner, we deal only with first term. Making use of (H3) we get

$$\begin{aligned}
\sup_{t\in\mathcal{S}_{\mathbb{R}}}\left|\widetilde{f}_N(\theta,t,x)-\widehat{f}_N(\theta,t_k,x)\right| &\leq \frac{1}{n\,h_H\,\mathbb{E}K_1(\theta,x)}\sup_{t\in\mathcal{S}_{\mathbb{R}}}\sum_{i=1}^{n}\left|\frac{\delta_i}{\bar{G}(Y_i)}H_i(t)-\frac{\delta_i}{\bar{G}_n(Y_i)}H_i(t_k)\right||K_i(\theta,x)| \\
&\leq \frac{C}{n\,h_H\,\mathbb{E}K_1(\theta,x)}\sup_{t\in\mathcal{S}_{\mathbb{R}}}\frac{|t-t_k|}{h_H}\times\left(\sum_{i=1}^{n}K_i(\theta,x)\left(\frac{1}{\bar{G}(Y_i)}-\frac{1}{\bar{G}_n(Y_i)}\right)\right) \\
&\leq \frac{Cl_n}{h_H^2\bar{G}_n(\tau_G)\bar{G}(\tau_G)}\sup_{t\in\mathcal{S}_{\mathbb{R}}}|G_n(t)-G(t)|\widehat{F}_D(\theta,x)
\end{aligned}$$

Using $l_n=n^{-\varsigma-1/2}$ we obtain

$$T_1 \;\leq\; \frac{Cn^{-\varsigma-1/2}}{h_H^2\bar{G}_n(\tau_G)\bar{G}(\tau_G)}\left(\frac{\log n\log n}{n}\right)^{1/2},$$

and note that, because of (H2)-(i), we have

$$\frac{l_n}{h_H^2}=o\left(\sqrt{\frac{\log n}{n\,h_h\,\phi_{\theta,x}(h_K)}}\right).$$

Thus, for n large enough, we have

$$T_1 \;=\; \mathcal{O}_{a.co}\left(\sqrt{\frac{\log n}{n\,h_H\,\phi_{\theta,x}(h_K)}}\right).$$

Following similar arguments, we can write

$$T_3 \;\leq\; T_1.$$

- Concerning T_2, let us consider $\varepsilon\;=\;\epsilon_0\sqrt{\dfrac{\log n}{n\,h_H\,\phi_{\theta,x}(h_K)}}$. Since for all $\epsilon_0>0$, we have that

$$\begin{aligned}
\mathbb{P}\left(\sup_{t\in\mathcal{S}_{\mathbb{R}}}\left|\widehat{f}_N(\theta,t_k,x)-\mathbb{E}\widehat{f}_N(\theta,t_k,x)\right|>\varepsilon\right) &\leq \mathbb{P}\left(\max_{k\in\{1\ldots\tau_n\}}\left|\widehat{f}_N(\theta,t_k,x)-\mathbb{E}\widehat{f}_N(\theta,t_k,x)\right|>\varepsilon\right) \\
&\leq \tau_n\max_{k\in\{1\ldots\tau_n\}}\mathbb{P}\left(\left|\widehat{f}_N(\theta,t_k,x)-\mathbb{E}\widehat{f}_N(\theta,t_k,x)\right|>\varepsilon\right).
\end{aligned}$$

Applying Berstain's exponential inequality to :

$$\Psi_i=\frac{1}{h_H\,\mathbb{E}K_1(x,\theta)}\Big[\frac{\delta_i}{\bar{G}(Y_i)}K_i(x,\theta)H_i(t_k)-\mathbb{E}\Big(\frac{\delta_i}{\bar{G}(Y_i)}K_i(x,\theta)H_i(t_k)\Big)\Big].$$

Firstly, it follows from the fact that the Kernels K and H are bounded, we get

$$\begin{aligned}\mathbb{P}\left(\left|\widehat{f}_N(\theta,t_k,x)-\mathbb{E}\widehat{f}_N(\theta,t_k,x)\right|>\varepsilon\right) &\leq \mathbb{P}\left(\frac{1}{n}\left|\sum_{i=1}^{n}\Psi_i\right|>\varepsilon\right)\\ &\leq 2n^{-c\varepsilon_0^2}.\end{aligned}$$

Finally, by choosing ε_0 large enough, the proof can be concluded by the use of the Borel-Cantelli lemma. the result can be easily deduced. □

Proof of Lemma 3.3: For all $x \in \mathcal{S}_{\mathcal{H}}$ and $\forall\theta \in \Theta_{\mathcal{H}}$, we set

$k(x) = \arg \min_{k\in\{1\ldots d_n^{\mathcal{S}_{\mathcal{H}}}\}} \|x - x_k\|$ and $q(\theta) = \arg \min_{m\in\{1\ldots d_n^{\Theta_{\mathcal{H}}}\}} \|\theta - \theta_q\|$ and by the compact property of $\mathcal{S}_{\mathbb{R}} \subset \mathbb{R}$, we have $\mathcal{S}_{\mathbb{R}} \subset \bigcup_{k=1}^{\tau_n}(v_k - l_n, v_k + l_n)$ with l_n and τ_n can be selected such as $l_n = \mathcal{O}\left(\tau_n^{-1}\right) = \mathcal{O}\left(n^{-(3\varsigma+1)/2}\right)$. Taking $k_t = \arg \min_{\{v_1,\ldots,v_{\tau_n}\}} |t - v_k|$.

Let us consider the following decomposition

$$\begin{aligned}\sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{x\in\mathcal{S}_{\mathcal{H}}}\sup_{t\in\mathcal{S}_{\mathbb{R}}}\left|\widetilde{f}_N(\theta,t,x)-\mathbb{E}\left(\widetilde{f}_N(\theta,t,x)\right)\right| \leq& \sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{x\in\mathcal{S}_{\mathcal{H}}}\sup_{t\in\mathcal{S}_{\mathbb{R}}}\Big\{\left|\widetilde{f}_N(\theta,t,x)-\widetilde{f}_N(\theta,t,x_{k(x)})\right|\\ &+\left|\widetilde{f}_N(\theta,t,x_{k(x)})-\widetilde{f}_N(\theta_{q(\theta)},t,x_{k(x)})\right|\\ &+\left|\widetilde{f}_N(\theta_{q(\theta)},t,x_{k(x)})-\widetilde{f}_N(\theta_{q(\theta)},v_{k_t},x_{k(x)})\right|\\ &+\left|\widetilde{f}_N(\theta_{q(\theta)},v_{k_t},x_{k(x)})-\mathbb{E}\left(\widetilde{f}_N(\theta_{q(\theta)},v_{k_t},x_{k(x)})\right)\right|\\ &+\left|\mathbb{E}\left(\widetilde{f}_N(\theta_{q(\theta)},v_{k_t},x_{k(x)})\right)-\mathbb{E}\left(\widetilde{f}_N(\theta_{q(\theta)},t,x_{k(x)})\right)\right|\\ &+\left|\mathbb{E}\left(\widetilde{f}_N(\theta_{q(\theta)},t,x_{k(x)})\right)-\mathbb{E}\left(\widehat{f}_N(\theta,t,x_{k(x)})\right)\right|\\ &+\left|\mathbb{E}\left(\widetilde{f}_N(\theta,t,x_{k(x)})\right)-\mathbb{E}\left(\widetilde{f}_N(\theta,t,x)\right)\right|\Big\}\\ \leq& \Psi_1+\Psi_2+\Psi_3+\Psi_4+\Psi_5+\Psi_6+\Psi_7.\end{aligned}$$

- Concerning Ψ_3 and Ψ_5; by conditions (H3) and (A4), boundness of K, we obtain

$$\begin{aligned}
\left|\widetilde{f}_N(\theta_{q(\theta)}, t, x_{k(x)}) - \widetilde{f}_N(\theta_{q(\theta)}, v_{k_t}, x_{k(x)})\right| &\leq \frac{1}{nh_H\,\mathbb{E}K_1(\theta,x)} \sup_{t\in\mathcal{S}_{\mathbb{R}}} \sum_{i=1}^{n} \left|\frac{\delta_i}{\bar{G}(Y_i)} K_i(\theta_{q(\theta)}, x_{k(x)})\right| \\
&\qquad \left| H\left(\frac{t-Y_i}{h_H}\right) H\left(\frac{v_{k_t}-Y_i}{h_H}\right)\right| \\
&\leq \sup_{t\in\mathcal{S}_{\mathbb{R}}} C\frac{|t-v_{k_t}|}{h_H^2}\Big(\frac{1}{n\mathbb{E}(K_1(\theta_{q(\theta)}, x_{k(x)}))} \\
&\qquad \sum_{i=1}^{n}\left|K_i(\theta_{q(\theta)}, x_{k(x)})\frac{1}{\bar{G}(Y_i)}\right|\Big) \\
&\leq \frac{Cl_n}{\phi(h_K)h_H^2} = \mathcal{O}\left(\frac{l_n}{h_H^2\phi(h_K)}\right).
\end{aligned}$$

Now, the fact that $\lim_{n\to\infty} n^{\nu}h_H = \infty$, and choosing $l_n = n^{-(3\nu+1)/2}$ and using the second part of (A4), imply that

$$\frac{l_n}{h_H^2\phi(h_K)} = o\left(\sqrt{\frac{logn}{nh_H\phi(h_K)}}\right)$$

as $n \to \infty$, therefore, it follows

$$\Psi_5 \leq \Psi_3 = \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{H}}} d_n^{\Theta_{\mathcal{H}}}}{nh_H\phi(h_K)}}\right).$$

• Concerning Ψ_4, let us consider $\varepsilon = \varepsilon_0\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{F}}} + \log d_n^{\Theta_{\mathcal{F}}}}{nh_H\phi(h_K)}}$. For all $\varepsilon_0 > 0$, we have

$$\begin{aligned}
\mathbb{P}\left(\Psi_4 > \varepsilon\right) &= \mathbb{P}\left(\max_{q\in\{1\ldots d_n^{\Theta_{\mathcal{H}}}\}} \max_{k\in\{1\ldots d_n^{\mathcal{S}_{\mathcal{H}}}\}} \max_{k_t\in\{1,2,\ldots,\tau_n\}} |\Gamma_i - \mathbb{E}\Gamma_i| > \varepsilon\right) \\
&\leq \tau_n d_n^{\mathcal{S}_{\mathcal{H}}} d_n^{\Theta_{\mathcal{H}}} \mathbb{P}\left(|\Gamma_i - \mathbb{E}\Gamma_i| > \varepsilon\right) \qquad (6.4)
\end{aligned}$$

Applying Bernstein's exponential inequality, under (H4), to get $\forall q \leq d_n^{\Theta_{\mathcal{H}}}$, $\forall k \leq d_n^{\mathcal{S}_{\mathcal{H}}}$ and $\forall k_t \leq \tau_n$,

$$\mathbb{P}\left(|\Gamma_i - \mathbb{E}\Gamma_i| > \varepsilon\right) \leq 2(d_n^{\Theta_{\mathcal{H}}} d_n^{\mathcal{S}_{\mathcal{H}}})^{-C\varepsilon_0^2}.$$

Choosing $\tau_n \leq Cn^{(3\varsigma+1)/2}$, we get

$$\mathbb{P}\left(\Psi_4 > \varepsilon\right) \leq C^{'}\tau_n(d_n^{\Theta_{\mathcal{H}}} d_n^{\mathcal{S}_{\mathcal{H}}})^{1-C\varepsilon_0^2}.$$

Putting $C\varepsilon_0^2 = \beta$ and using (A4), to get

$$\Psi_4 = \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_\mathcal{F}} + \log d_n^{\Theta_\mathcal{F}}}{n\, h_H\, \phi(h_K)}}\right). \tag{6.5}$$

• Concerning Ψ_1 and Ψ_2, by assumption (A1), it follows

$$\begin{aligned}
\sup_{\theta\in\Theta_\mathcal{H}}\sup_{x\in\mathcal{S}_\mathcal{H}}\sup_{t\in\mathcal{S}_\mathbb{R}}|\widetilde{f}_N(\theta,t,x)-\widetilde{f}_N(\theta,t,x_{k(x)})| &\leq \frac{1}{nh_H\mathbb{E}K_1(\theta,x)}\sup_{\theta\in\Theta_\mathcal{H}}\sup_{x\in\mathcal{S}_\mathcal{H}}\sup_{t\in\mathcal{S}_\mathbb{R}}\sum_{i=1}^n\left|\frac{\delta_i}{\bar{G}(Y_i)}\right||H_i(t)| \\
&\qquad |(K_i(\theta,x)-K_i(\theta,x_k))| \\
&\leq \frac{1}{nh_H\phi(h_K)}\sup_{x\in\mathcal{S}_\mathcal{H}}\sup_{\theta\in\Theta_\mathcal{H}}\sum_{i=1}^n\left|\Delta_i(x,\theta)-\Delta_i\left(x_{k(x)},\theta\right)\right| \\
&\leq \frac{1}{h_H\phi(h_K)}\sup_{x\in\mathcal{S}_\mathcal{H}}\sup_{\theta\in\Theta_\mathcal{H}}\frac{1}{n}\sum_{i=1}^n\mathbf{1}_{B_\theta(x,h)\cup B_\theta(x_{k(x)},h)}(X_i) \\
&\leq \frac{C}{h_H}\sup_{x\in\mathcal{S}_\mathcal{H}}\sup_{\theta\in\Theta_\mathcal{H}}\frac{1}{n}\sum_{i=1}^n\Lambda_i(x,\theta).
\end{aligned}$$

Therefore, similar to the arguments for (6.5), we can get that

$$\Psi_1 = \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_\mathcal{F}} + \log d_n^{\Theta_\mathcal{F}}}{nh_H\phi(h_K)}}\right).$$

$$\begin{aligned}
\sup_{\theta\in\Theta_\mathcal{H}}\sup_{x\in\mathcal{S}_\mathcal{H}}\sup_{t\in\mathcal{S}_\mathbb{R}}|\widetilde{f}_N(\theta,t,x)-\widetilde{f}_N(\theta_{q(\theta)},t,x_{k(x)})| &\leq \frac{h_H^{-1}}{n\mathbb{E}K_1(\theta,x)}\sup_{\theta\in\Theta_\mathcal{H}}\sup_{x\in\mathcal{S}_\mathcal{H}}\sup_{t\in\mathcal{S}_\mathbb{R}}\sum_{i=1}^n\left|\frac{\delta_i}{\bar{G}(Y_i)}\right||H_i(t)| \\
&\qquad \left|(K_i(\theta,x_{(k)})-K_i(\theta_{q(\theta)},x_{(k)}))\right| \\
&\leq \frac{Ch_H^{-1}}{n\phi(h_K)}\sup_{x\in\mathcal{S}_\mathcal{H}}\sup_{\theta\in\Theta_\mathcal{H}}\sum_{i=1}^n\left|\Delta_i(\theta,x_k)-\Delta_i\left(\theta_{q(\theta)},x_{k(x)}\right)\right| \\
&\leq \frac{Ch_H^{-1}}{\phi(h_K)}\sup_{x\in\mathcal{S}_\mathcal{H}}\sup_{\theta\in\Theta_\mathcal{H}}\frac{1}{n}\sum_{i=1}^n\mathbf{1}_{B_\theta(x_{(k)},h)\cup B_{q(\theta)}(x_{k(x)},h)}(X_i) \\
&\leq \frac{C}{h_H}\sup_{x\in\mathcal{S}_\mathcal{H}}\sup_{\theta\in\Theta_\mathcal{H}}\frac{1}{n}\sum_{i=1}^n\Omega_i(x,\theta).
\end{aligned}$$

Similar to the deduce of (6.5), it yields

$$\Psi_2 = \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_\mathcal{F}} + \log d_n^{\Theta_\mathcal{F}}}{nh_H\phi(h_K)}}\right).$$

On the other hand, since $\Psi_7 \leq \Psi_1$ and $\Psi_6 \leq \Psi_2$, it also leads to

$$\Psi_6 = \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_\mathcal{F}} + \log d_n^{\Theta_\mathcal{F}}}{nh_H\phi(h_K)}}\right),$$

and

$$\Psi_7 = \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_\mathcal{F}} + \log d_n^{\Theta_\mathcal{F}}}{nh_H\phi(h_K)}}\right).$$

Then the proof of Lemma 3.3 can be completed. $\square$

Proof of Lemma 4.1:

$$\begin{aligned} V_n(\theta,t,x) &= \frac{\phi_{\theta,x}(h_K)}{h_H(\mathbb{E}K_1(\theta,x))^2}\mathbb{E}\left[K_1^2(\theta,x)\left(\frac{\delta_1}{\bar{G}(Y_1)}H_1(t) - h_h f(\theta,t,x)\right)^2\right] \\ &\stackrel{(6.6)}{=} \frac{\phi_{\theta,x}(h_K)}{h_H(\mathbb{E}K_1(\theta,x))^2}\mathbb{E}\left[K_1^2(\theta,x)\mathbb{E}\left(\left(\frac{\delta_1}{\bar{G}(Y_1)}H_1(t) - h_H f(\theta,t,x)\right)^2 / <\theta, X_1>\right)\right]. \end{aligned}$$

Using the definition of conditional variance, we have

$$\mathbb{E}\left[\left(\frac{\delta_1}{\bar{G}(Y_1)}H_1(t) - h_H f(\theta,t,x)\right)^2 | <\theta, X_1>\right] = J_{1n} + J_{2n},$$

where

$$\begin{aligned} J_{1n} &= Var\left(\frac{\delta_1}{\bar{G}(Y_1)}H_1(t) | <\theta, X_1>\right) \\ J_{2n} &= \left[\mathbb{E}\left(\frac{\delta_1}{\bar{G}(Y_1)}H_1(t) | <\theta, X_1>\right) - h_H f(\theta,t,x)\right]^2. \end{aligned}$$

- Concerning J_{1n}

$$\begin{aligned} J_{1n} &= \mathbb{E}\left(\frac{\delta_1}{\bar{G}^2(Y_1)}H_1^2(t) | <\theta, X_1>\right) - \mathbb{E}\left(\frac{\delta_1}{\bar{G}(Y_1)}H_1(t) | <\theta, X_1>\right)^2 \\ &= J_1 + J_2. \end{aligned}$$

As for J_1, by the property of double conditional expectation and by changing variables, we get that,

$$
\begin{aligned}
J_1 &= \mathbb{E}\left[\mathbb{E}\left(\frac{\delta_1}{\bar{G}^2(Y_1)}H_1^2\left(\frac{t-Y_1}{h_H}\right)| < \theta, X_1 >, T_1\right)\right]\\
&= \mathbb{E}\left(\frac{1}{\bar{G}^2(T_1)}H_1^2\left(\frac{t-T_1}{h_H}\right)\mathbb{E}[\mathbf{1}_{T_1\leq C_1}|T_1]| < \theta, X_1 >\right)\\
&= \mathbb{E}\left(\frac{1}{\bar{G}(T_1)}H_1^2\left(\frac{t-T_1}{h_H}\right)| < \theta, X_1 >\right)\\
&= \int_{\mathbb{R}}\frac{1}{\bar{G}(v)}H_1^2\left(\frac{t-v}{h_H}\right)f(\theta, v, X_1)dv\\
&= \int_{\mathbb{R}}\frac{1}{\bar{G}(t-uh_H)}H_1^2(u)dF(\theta, t-uh_H, X_1).
\end{aligned}
$$

By the first order Taylor's expansion of the function $\bar{G}^{-1}(\cdot)$ around zero, one gets $J_1 = \int_{\mathbb{R}}\frac{1}{\bar{G}(t)}H_1^2(u)dF(\theta, t-uh_H, X_1) + \frac{h_H^2}{\bar{G}(t)^2}\int_{\mathbb{R}}uH_1^2\bar{G}^{(1)}(t^*)f(\theta, t - uh_H, X_1)du + o(1)$,

where t^* is between t and $t - uh_H$.

Under assumptions (N3) and using hypothesis (H2), we get

$$
\frac{h_H^2}{\bar{G}^2(t)}\int_{\mathbb{R}}uH_1^2\bar{G}^{(1)}(t^*)f(\theta, t-uh_H, X_1)du = o(h_H^2).
$$

Indeed

$\frac{h_H^2}{\bar{G}^2(t)}\int_{\mathbb{R}}uH_1^2\bar{G}^{(1)}(t^*)f(\theta, t-uh_H, X_1)du \leq h_H^2\left(\sup_{u\in\mathbb{R}}|G^{'}(u)||\bar{G}^2(t)\right)\int_{\mathbb{R}}uf(\theta, t-yh_H, x)du.$

On the other hand, by applying (H2) and (H3), we have

$$
\begin{aligned}
\int_{\mathbb{R}}\frac{1}{\bar{G}(t)}H_1^2(u)dF(\theta, t-uh_H, X_1) &= h_H\int_{\mathbb{R}}\frac{1}{\bar{G}(t)}H_1^2(u)f(\theta, t-uh_H, X_1)du\\
&\leq \frac{h_H}{\bar{G}(t)}\int_{\mathbb{R}}H_1^2(u)(f(\theta, t-uh_H, X_1) - f(\theta, t, x))du\\
&\quad + \frac{h_H}{\bar{G}(t)}\int_{\mathbb{R}}H_1^2(u)f(\theta, t, x)du\\
&\leq \frac{h_H}{\bar{G}(t)}\left(C_{x,\theta}\int_{\mathbb{R}}H^2(u)\left(h_K^{b_1} + |v|^{b_2}h_H^{b_2}\right)du + f(\theta, t, x)\int_{\mathbb{R}}H^2(u)du\right)\\
&= \mathcal{O}\left(h_k^{b_1} + h_H^{b_2}\right) + \frac{h_H}{\bar{G}(t)}f(\theta, t, x)\int_{\mathbb{R}}H^2(u)du. \qquad (6.7)
\end{aligned}
$$

As for J_2,

$$\begin{aligned}
J_2^{'} &= \mathbb{E}\left(\frac{\delta_1}{\bar{G}(Y_1)}H_1(t)| < \theta, X_1 >\right)\\
&= \mathbb{E}\left[\mathbb{E}\left(\frac{\delta_1}{\bar{G}(Y_1)}H_1\left(\frac{t-Y_1}{h_H}\right)| < \theta, X_1 >, T_1\right)\right]\\
&= \mathbb{E}\left(\frac{1}{\bar{G}(T_1)}H_1\left(\frac{t-T_1}{h_H}\right)\mathbb{E}[\mathbf{1}_{T_1\leq C_1}|T_1]| < \theta, X_1 >\right)\\
&= \mathbb{E}\left(H_1\left(\frac{t-T_1}{h_H}\right)| < \theta, X_1 >\right)\\
&= \int_{\mathbb{R}} H^{(1)}\left(\frac{t-v}{h_H}\right)f(\theta, t, X_1)dv.
\end{aligned}$$

Moreover, we have by changing variables

$$J_2^{'} = h_H\int_{\mathbb{R}} H(u)(f(\theta, t-uh_H, X_1 - f(\theta,t,x))du + h_H f(\theta,t,x)\int_{\mathbb{R}} H(u)du,$$

the last equality is due to the fact that H is a probability density.
Thus, we have:

$$J_2^{'} = \mathcal{O}\left(h_k^{b_1} + h_H^{b_2}\right) + h_H f(\theta,t,x).$$

Finally we get $J_2 \underset{n\,\longrightarrow\,\infty}{\longrightarrow} 0.$

As for J_{2n}, by (H1)-(H3), we obtain that

$$J_{2n} \longrightarrow 0, \quad as \quad n\longrightarrow\infty. \tag{6.8}$$

Meanwhile, by (H1)-(H3) and (N3), it follows that

$$\frac{\phi_{\theta,x}(h_K)\mathbb{E}K_1^2(\theta,x)}{\mathbb{E}^2K_1(\theta,x)} \underset{n\longrightarrow\infty}{\longrightarrow} \frac{a_2(\theta,x)}{(a_1(\theta,x))^2},$$

which leads to combining equations 6.6 and 6.7

$$V_n(\theta,t,x) \longrightarrow \frac{a_2(\theta,x)}{(a_1(\theta,x))^2}\frac{f(\theta,t,x)}{\bar{G}(t)}\int_{\mathbb{R}} H^2(u)du \tag{6.9}$$

□

Proof of Lemma 4.2: We have

$$\begin{aligned}
\sqrt{nh_H\phi_{\theta,x}(h_K)}B_n(\theta,t,x) &= \frac{\sqrt{nh_H\phi_{\theta,x}(h_K)}}{\widehat{F}_D(\theta,x)}\{\mathbb{E}\widehat{f}_N(\theta,t,x) - a_1(\theta,x)f(\theta,t,x)\\
&\quad + f(\theta,t,x)(a_1(\theta,x) - \mathbb{E}\widehat{F}_D(\theta,x))\}.
\end{aligned}$$

Firstly, observed that the results below

(6.10)
$$\frac{1}{\phi_{\theta,x}(h_K)}\mathbb{E}\left[K^l\left(\frac{< x - X_i, \theta >}{h_K}\right)\right] \longrightarrow a_l(\theta, x), \quad as \quad n \longrightarrow \infty, \quad for \quad l = 1, 2,$$

$$\mathbb{E}\left[\widehat{F}_D(\theta, x)\right] \longrightarrow a_1(\theta, x), \quad as \quad n \longrightarrow \infty, \tag{6.11}$$

and

$$\mathbb{E}\left[\widehat{f}_n(\theta, t, x)\right] \longrightarrow a_1(\theta, x)f(\theta, t, x), \quad as \quad n \longrightarrow \infty, \tag{6.12}$$

can be proved in the same way as in Ezzahrioui and Ould Saïd [10] corresponding to their Lemmas 5.1 and 5.2, and then their proofs omitted.

Secondly, on the one hand, making use of (6.10),(6.11) and (6.12) we have

$$\{\mathbb{E}\widehat{f}_N(\theta, t, x) - a_1(\theta, x)f(\theta, t, x) + f(\theta, t, x)(a_1(\theta, x) - \mathbb{E}\widehat{F}_D(\theta, x))\} \underset{n\longrightarrow\infty}{\longrightarrow} 0.$$

On other hand,

$$\frac{\sqrt{nh_H\phi_{\theta,x}(h_K)}}{\widehat{F}_D(\theta, x)} = \frac{\sqrt{nh_H\phi_{\theta,x}(h_K)}\widetilde{f}(\theta, t, x)}{\widehat{F}_D(\theta, x)\widetilde{f}(\theta, t, x)} = \frac{\sqrt{nh_H\phi_{\theta,x}(h_K)}\widetilde{f}(\theta, t, x)}{\widetilde{f}_N(\theta, t, x)}. \tag{6.13}$$

Then using proposition(2.1), it suffices to show that $\frac{\sqrt{nh_H\phi_{\theta,x}(h_K)}}{\widetilde{f}_N(\theta,t,x)}$ tend to zero as n goes to infinity.

Indeed

$$\widetilde{f}_N(\theta, t, x) = \frac{1}{nh_H\mathbb{E}K_1(\theta, x)}\sum_{i=1}^{n}\frac{\delta_i}{\bar{G}(Y_i)}K\left(\frac{< x - X_i, \theta >}{h_K}\right)H\left(\frac{t - Y_i}{h_H}\right).$$

Because $K(\cdot)H(\cdot)$ is continuous with support on [0,1], then by (H3) and (H4) $\exists m = \inf_{[0,1]} K(t)H(t)$ if follows that

$$\widetilde{f}_N(\theta, t, x) \geq \frac{m}{h_H\phi_{\theta,x}(h_K)},$$

which gives

$$\frac{\sqrt{nh_H\phi_{\theta,x}(h_K)}}{\widetilde{f}_N(\theta, t, x)} \leq \frac{\sqrt{nh_H^3\phi_{\theta,x}^3(h_K)}}{m}.$$

Finally, using (N2), completes the proof of Lemma. 4.2 □

Acknowledgement:

The authors would like to express their gratitude to the referees and chief editor for their valuable suggestions which have considerably improved the earlier version of the paper.

REFERENCES

[1] Aït-Saidi, A., Ferraty, F., Kassa, R. (2005). Single functional index model for a time series. *Revue Roumaine de Mathématique Pures et Appliquées*, **50**, 321–330.

[2] Aït-Saidi, A., Ferraty, F., Kassa, R., Vieu, P. (2008). Cross-validated estimation in the single functional index model. *Statistics*, **24**, 475–494.

[3] Attaoui, S., Laksaci, A., E. Ould-Saïd, E. (2011). A note on the conditional density estimate in the single functional index model. *Statist. Probab. Lett. Vol.*, **81**, 45–53.

[4] Attaoui, S. (2014). On the Nonparametric Conditional Density and Mode Estimates in the Single Functional Index Model with Strongly Mixing Data. *Sankhyã: The Indian Journal of Statistics*, **76**, A, 356–378.

[5] Attaoui, S. (2014). Strong uniform consistency rates and asymptotic normality of conditional density estimator in the single functional index modeling for time series data. *J AStA.*, **98**, 257–286

[6] 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**, 485–511.

[7] Bouchentouf, A.A., Djebbouri, T., Rabhi, A., Sabri, K. (2014). Strong uniform consistency rates of some characteristics of the conditional distribution estimator in the functional single-index model. *Appl. Math. (Warsaw)*, **41**, 4, 301–322.

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

[9] Deheuvels, P., Einmahl, J.H.J. (2000). Functional limit laws for the increments of Kaplan-Meier product-limit processes and applications. *Ann. Probab.*, **28**, 3, 1301–1335.

[10] Ezzahrioui, M., Ould Saïd, E. (2008). Asymtotic results of a nonparametric conditional quantile estimator for functional time series. *Comm. Statist. Theory and Methods.*, **37**, 16-17, 2735–2759.

[11] Ezzahrioui, M., Ould Saïd, E. (2010). Some asymtotic results of a nonparametric conditional mode estimator for functional time series data. *Statist. Neerlandica.*, **64**, 171–201.

[12] Ferraty, F., Laksaci, A., Vieu, P. (2006). Estimating some characteristics of the conditional distribution in nonparametric functional models. *Statist. Inference Stoch. Process.*, **9**, 47–76

[13] Ferraty, Laksaci, A., Tadj, A., Vieu, P. (2010). Rate of uniform consistency for nonparametric estimates with functional variables. *J. Statist. Plann. and Inf.*, **140**, 335–352.

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

[15] Ferraty, F., Vieu, P. (2003). *Functional nonparametric statistics: a double infinite dimensional framework*. Recent advanvces and trends in Nonparametric Statistics, Ed. M. Akritas and D. Politis, Elsevier.

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

[17] Goia, A., Vieu, P. (2015). A partitioned Single Functional Index Model. *Computational Statistics*, **30**, 3, 673–692.

[18] Goia, A., and Vieu, P. (2016). An Introduction to Recent Advances in High/Infinite Dimensional Statistics. *J. Mutiv. Anal.*, **146**, 1–6.

[19] Härdle, W., and Marron, J.S. (1985). Optimal Bandwidth Selection in Nonparametric Regression Function Estimation. *Ann. Statist.*, **13**, 1465–1481.

[20] Hristache, M., Juditsky, A., Spokoiny, V. (2001). Direct Estimation of the Index Coeffcient in the Single-index Model. *Ann. Statist.*, **29**, 595–623.

[21] Hsing, T., Eubank, R. (2015). *Theoretical Foundations of Functional Data Analysis, With an Introduction to Linear Operators*. Wiley, Chichester.

[22] Khardani, S., Lemdani, M., Ould Saïd, E. (2010). Some asymptotic properties for a smooth kernel estimator of the conditional mode under random censorship. *Journal of the Korean Statistical Society*, **39**, 455–469.

[23] Khardani, S., Lemdani, M., Ould Saïd, E. (2011). Uniform rate of strong consistency for a smooth Kernel estimator of the conditional mode under random censorship. *J. Statist. Plann. and Inf.*, **141**, 3426–3436

[24] Khardani, S., Lemdani, M., Ould Saïd, E. (2014). On the Central Limit Theorem for a Conditional Mode Estimator of a Randomly Censored Time Series. *Journal of Statistical Theory and Practice*, **8**, 722–742.

[25] Ling, N., Xu, Q. (2012). Asymptotic normality of conditional density estimation in the single index model for functional time series data. *Statistics and Probability Letters*, **82**, 2235–2243.

[26] Masry, E. (2005). Nonparametric Regression Estimation for Dependent Functional Data: Asymptotic Normality. *Stochastic. Process. Appl.*, **115**, 155–177.

[27] Mehra, K.L., Ramakrishnaiah, Y.S., Sashikala, P. (2000). Laws of iterated logarithm and related asymptotics for estimators of conditional density and mode. *Ann. Inst. Math. Stat.*, **52**, 630645.

[28] Ould-Saïd, E., Cai, Z. (2005). Strong uniform consistency of nonparametric estimation of the censored conditional mode function. *Nonparametric Statistics*, **17**, 7, 797–806.

[29] Ould-Saïd, E. (2006). A strong uniform convergence rate of Kernel conditional quantile estimator under random censorship. *Statistics and Probability Letters*, **76**, 579–586.

[30] Ramsay, J., Silverman, B. (2002). *Applied Functional Data Analysis: Methods and Case Studies*. Springer, New York.

[31] Ramsay, J., Silverman, B. (2005), *Functional Aata Analysis (2nded.)* Springer, New York.

[32] Samanta, M., Thavaneswaran, A. (1990). Nonparametric estimation of the conditional mode. *Commun. Stat. Theory Methods.*, **16**, 45154524.

[33] Xia, Y., and Härdle, W. (2006). Semi-parametric Estimation of Partially Linear Single-index Models. *J. Multivar. Anal.*, **97**, 1162–1184.