



**HAL**  
open science

# Kolmogorov Entropy for Convergence Rate in Incomplete Functional Time Series: Application to Percentile and Cumulative Estimation in High Dimensional Data

Ouahiba Litimein, Fatimah Alshahrani, Salim Bouzebda, Ali Laksaci, Boubaker Mechab

## ► To cite this version:

Ouahiba Litimein, Fatimah Alshahrani, Salim Bouzebda, Ali Laksaci, Boubaker Mechab. Kolmogorov Entropy for Convergence Rate in Incomplete Functional Time Series: Application to Percentile and Cumulative Estimation in High Dimensional Data. *Entropy*, 2023, 25 (7), pp.1108. 10.3390/e25071108. hal-04170382

**HAL Id: hal-04170382**

**<https://hal.science/hal-04170382v1>**

Submitted on 1 Feb 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Article

# Kolmogorov Entropy for Convergence Rate in Incomplete Functional Time Series: Application to Percentile and Cumulative Estimation in High Dimensional Data

Ouahiba Litimein <sup>1</sup>, Fatimah Alshahrani <sup>2</sup>, Salim Bouzebda <sup>3,\*</sup>, Ali Laksaci <sup>4</sup> and Boubaker Mechab <sup>1</sup>

<sup>1</sup> Laboratory of Statistics and Stochastic Processes, University of Djillali Liabes, BP 89, Sidi Bel Abbes 22000, Algeria; ouahiba.litimein@univ-sba.dz (O.L.); mechaboub@yahoo.fr (B.M.)

<sup>2</sup> Department of Mathematical Sciences, College of Science, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia; fmalshahrani@pnu.edu.sa

<sup>3</sup> Laboratoire de Mathématiques Appliquées de Compiègne (L.M.A.C.), Université de Technologie de Compiègne, 60200 Compiègne, France

<sup>4</sup> Department of Mathematics, College of Science, King Khalid University, Abha 62529, Saudi Arabia; alikfa@kku.edu.sa

\* Correspondence: salim.bouzebda@utc.fr

**Abstract:** The convergence rate for free-distribution functional data analyses is challenging. It requires some advanced pure mathematics functional analysis tools. This paper aims to bring several contributions to the existing functional data analysis literature. First, we prove in this work that Kolmogorov entropy is a fundamental tool in characterizing the convergence rate of the local linear estimation. Precisely, we use this tool to derive the uniform convergence rate of the local linear estimation of the conditional cumulative distribution function and the local linear estimation conditional quantile function. Second, a central limit theorem for the proposed estimators is established. These results are proved under general assumptions, allowing for the incomplete functional time series case to be covered. Specifically, we model the correlation using the ergodic assumption and assume that the response variable is collected with missing at random. Finally, we conduct Monte Carlo simulations to assess the finite sample performance of the proposed estimators.

**Keywords:** functional data; missing data; small ball probability; local linear modeling; Kolmogorov entropy; almost complete (a.co.) convergence

**MSC:** 62G20; 62M10; 62N01; 62P30; 60F05



**Citation:** Litimein, O.; Alshahrani, F.; Bouzebda, S.; Laksaci, A.; Mechab, B. Kolmogorov Entropy for Convergence Rate in Incomplete Functional Time Series: Application to Percentile and Cumulative Estimation in High Dimensional Data. *Entropy* **2023**, *25*, 1108. <https://doi.org/10.3390/e25071108>

Academic Editor: Nikolai Leonenko

Received: 22 June 2023

Revised: 18 July 2023

Accepted: 20 July 2023

Published: 24 July 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Statistical problems associated with the study of functional random variables, that is, variables with values in an infinite-dimensional space, have garnered increasing attention in the statistical literature over the past few decades. The abundance of data measured on increasingly fine temporal/spatial grids, as is the case in meteorology, medicine, satellite imagery, and numerous other research fields, has inspired the development of this research theme. Thus, the statistical modeling of these data as random functions resulted in a number of difficult theoretical and numerical research concerns; we may refer to [1–6] for parametric and nonparametric models. For the latest contributions in FDA and its related topics, one can refer to [7–14].

Quantile regression has emerged as a significant statistical technique for data analysis since Koenker and Bassett's [15] seminal work. But, concerns about quantile crossing and model misspecification [16–18] have led to the development of the nonparametric estimation of conditional quantile functions [19,20]. This estimate originates from [21], who proved the convergence using the probability of the empirical conditional law function; refer to [22,23]. This technique is an alternative to mean regression, and it possesses many

desirable properties, such as being more efficient than mean regression when the data follow a heavy-tailed distribution, and it is frequently used to characterize the entire conditional distribution of the response. Conditional percentile estimation has been studied extensively by several researchers [24–26]. However, most of them used the kernel method approach. For instance, Ferraty et al. [4] have studied the uniform convergence of such an estimation in the analogously distributed and independent case. However, in the same case, Samanta [27] acquired the almost complete (a.co.) convergence of the conditional percentile estimation.

Studies of local linear estimation conditional quantile function (LLECQF) are still limited. For example, Messaci et al. [28] have studied the local linear estimation (LLE) of the conditional quantile function (CQF) by reversing the estimator shown in [29]. Al-Awadhi et al. [30] have proved the a.co. convergence and the asymptotic law of the CQF by considering an estimator based on the  $L_1$  approach. Despite its importance in functional data analysis, the LLE has various merits over the kernel technique. Generally, this method can reduce the bias properties of the kernel approach [31,32]. Furthermore, the LLE has been lately introduced in the FDA by [33]. The latter concentrates on the LLE of the curve regression while the insert variable is a Hilbert class. Barrientos-Marin et al. [34] have proposed the LLE of the nonparametric regression operator and studied their asymptotic properties. In particular, this operator can be applied to functional covariable.

In this article, we are interested in studying the local linear estimation of the conditional cumulative distribution function (LLECCDF) and LLECQF under the assumption of ergodicity. We are interested in the uniform a.co. convergence of the constructed sequences using the Kolmogorov entropy function. In addition to Kolmogorov entropy, other information measures such as Kullback–Leibler divergence have been considered for a convergence rate study of estimators in multivariate time series modelling; refer to [35,36].

From a practical point of view, the ergodic framework is an essential condition in statistical physics, number theory, Markov chains, and other fields. The concept of ergodicity is fundamental in the research of stochastic processes. Note also that one of the arguments invoked by [37] motivating the introduction of the concept of ergodicity is that, for certain classes of processes, it can be much easier to prove ergodic properties rather than the mixing condition. Hence, the ergodicity hypothesis seems to be the most naturally adapted and provides a better framework to study data series such as those generated by noisy chaos.

In their discussion, [38] provided an example of processes that are ergodic but not mixing, which may be summarized as follows: let  $\{(T_i, \lambda_i) : i \in \mathbb{Z}\}$  be a strictly stationary process such that  $T_i \mid \mathcal{T}_{i-1}$  is a Poisson process with a parameter  $\lambda_i$ , where  $\mathcal{T}_i$  is the  $\sigma$ -field generated by  $(T_i, \lambda_i, T_{i-1}, \lambda_{i-1}, \dots)$ . Assume that  $\lambda_i = f(\lambda_{i-1}, T_{i-1})$ , where  $f : [0, \infty) \times \mathbb{N} \rightarrow (0, \infty)$  is a given function. This process does not mix in general (see Remark 3 of [39]). It is known that any sequence  $(\varepsilon_i)_{i \in \mathbb{Z}}$  of i.i.d. random variables is ergodic. Hence, it is immediately clear that  $(Y_i)_{i \in \mathbb{Z}}$  with  $Y_i = \vartheta(\dots, \varepsilon_{i-1}, \varepsilon_i, (\varepsilon_{i+1}, \varepsilon_{i+2}, \dots))$ , for some Borel-measurable function  $\vartheta(\cdot)$ , is also ergodic; see Proposition 2.10 on page 54 in [40]. Under the condition of ergodicity, [41] have studied the conditional quantile for ergodic data by considering an iterative model. However, under the same conditions, the authors in [42] have considered the nonparametric estimation of quantile for censored data.

All these studies were concerned with the complete-data situation. This work investigates the question of when data are missing at random (MAR); for instance, see [43]. In contemporary statistics, missing data are pervasive, posing a significant obstacle for various applications. The concept of missing data in statistics occurs when a data value of the variable is not stored in the observation. For example, missingness may occur when our data are compiled from sources that have measured various variables; for instance, in the healthcare industry, the data routinely collected on patients may vary between clinics and hospitals. Among numerous other reasons are sensor failure, data censorship, privacy concerns, pharmaceutical tracing tests, and reliability tests. It can occur in any experimental setting where contamination of the treatment or subject mortality is possible. This topic

has been extensively examined in numerous statistical problems; see [43,44] for a thorough overview. In the present work, we only observe  $Y$  in cases where some indicator  $B$  equals one and the indicator  $B$  is conditionally independent of  $Y$  given  $X$ . This assumption is useful when information in the form of covariate data is available to explain the missingness; refer to [45–47]. The first studies on MAR are presented by [48]. They established an approximation of the regression operator and studied their asymptotic consistency when the curve regressor is observed and the interest response is missing at random. However, in the ergodic data case, Ling et al. [49] have studied the asymptotic distribution of the estimator, which is proposed in [48]; also, refer to [50,51].

The main objective of this paper is to evaluate the convergence velocity of some functional estimators through the Kolmogorov entropy function. More precisely, we focus on the local linear smoothing of the distribution function and its inverse, the quantile function. The constructed estimators' asymptotic properties are evaluated when the data are correlated as functional ergodic time series data and the response variable is observed under the MAR structure. The efficiency of these estimators is uniformly specified using the entropy metric, allowing us to assess the impact of the functional path of the data. In particular, the Kolmogorov entropy gives a trade-off between the data's sparsity and the approximation's efficiency. Moreover, the Kolmogorov entropy explores the functional space's topological structure and its spectral property. Thus, stating the uniform consistency concerning the Kolmogorov entropy is more beneficial than the classical pointwise case. Although the uniform convergence of LLECCDF and LLECQF in the functional ergodic time series (FETS) structure is purely mathematically challenging, the obtained results are also pivotal for many applied issues, such as the smoothing parameter choice, bootstrapping, and the single index modeling. The second challenging issue of this contribution concerns the MAR feature of the response variable. Our results can be used to determine an estimator of the unconditional distribution of the scalar response, even if it is not completely observed. All these challenging issues will be discussed using specific examples in Section 6. In addition to the uniform consistency, we prove the asymptotic normality of the LLECCDF, which is important to provide a confidence interval comparable with the predictive interval deduced from LLECQF. Once again, this prediction using the predictive subset is also primordial in the context of incomplete functional time series data. Finally, we point out that, to our best knowledge, this problem of uniform consistency of local linear approximation under MAR and FETS structures was open up to the present, giving the main motivation to our paper.

The rest of the paper is organized as follows. In Section 2, we state the formal setup and define the estimators. More precisely, Section 2.1 is devoted to the LLECCDF estimator, while Section 2.2 introduces the LLECQF estimator. The convergence rate of the two approximation sequences is established in Section 3. In Section 4, we derive the limiting distribution of the proposed estimators. In Section 5, we conduct Monte Carlo simulations to assess the finite sample performance of the proposed estimators. Section 6 is devoted to highlighting the principal features of our contribution. In Section 7, we give some concluding remarks. All the proofs are gathered in the last section.

## 2. Model and Estimators

### 2.1. LLECCDF: Numerical Approximation of CCDF-Model

Let  $\{(X_i, Y_i) : 1 \leq i \leq n\}$  be a sequence of stationary ergodic functional random variables identically distributed as  $(X, Y)$ , where  $X$  takes values in a some semi-metric abstract space  $\mathcal{F}$  with a semi-metric  $d(\cdot, \cdot)$  and  $Y$  takes values in  $\mathbb{R}$ . For reader convenience, we introduce some details defining the ergodic property of processes and its link with the mixing one. Let  $\{X_n, n \in \mathbb{Z}\}$  be a stationary sequence. Consider the backward field  $\mathcal{A}_n = \sigma(X_k : k \leq n)$  and the forward field  $\mathcal{B}_m = \sigma(X_k : k \geq m)$ . The sequence is strongly mixing if

$$\sup_{A \in \mathcal{A}_0, B \in \mathcal{B}_n} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| \rightarrow 0, \text{ as } n \rightarrow \infty.$$

The sequence is ergodic if, for any two measurable sets  $A, B$ ,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \left| \mathbb{P}(A \cap \tau^{-k} B) - \mathbb{P}(A)\mathbb{P}(B) \right| = 0,$$

where  $\tau$  is the time evolution or shift transformation taking  $X_k$  into  $X_{k+1}$ . We shall also use the same symbol  $\tau$  to denote the induced set transformation, which takes, for example, sets  $B \in \mathcal{B}_m$  into sets  $\tau B \in \mathcal{B}_{m+1}$ ; for instance, see [52]. The naming of strong mixing in the above definition is more stringent than what is ordinarily referred to (when using the vocabulary of measure preserving dynamical systems) as strong mixing, namely to that  $\lim_{n \rightarrow \infty} \mathbb{P}(A \cap \tau^{-n} B) = \mathbb{P}(A)\mathbb{P}(B)$  for any two measurable sets  $A, B$ ; see, for instance [52,53] and more recent references [54–60]. Hence, strong mixing implies ergodicity, whereas the inverse is not always true (see, e.g., Remark 2.6 in page 50 in connection with Proposition 2.8 in page 51 in [40]). For every  $x \in \mathcal{F}$ , the function of conditional law  $CFD(y|x)$  of  $Y$  when  $X = x$  is defined by

$$CFD(y|x) = \mathbb{P}(Y \leq y | X = x).$$

The LLECCDF is obtained by assuming for every  $z$  in the vicinity of  $x$

$$CDF(y|z) = \beta_1 + \beta_2 \alpha(z, x) + o(\alpha(z, x)) \quad \text{with} \quad \alpha(z, z) = 0, \tag{1}$$

where  $\alpha(\cdot, \cdot)$  is a bilinear locating function such that

- (i) For all  $x' \in \mathcal{C}_{\mathcal{F}}$ ,  $C'|\delta(x, x')| \leq |\alpha(x', x)| \leq C|\delta(x, x')|$ ;
- (ii) For all  $z_1, z_2 \in \mathcal{C}_{\mathcal{F}}$ ,  $|\alpha(z_1, x) - \alpha(z_2, x)| \leq C'|\delta(z_1, z_2)|$ ,

where  $\delta(\cdot, \cdot)$  is a bilinear function such that  $|\delta(\cdot, \cdot)| = d(\cdot, \cdot)$ . In the rest of the paper, we suppose that the CDF is of the  $C^1$ -class with respect to  $y$  and its derivative is the conditional density function denoted by  $cdf(\cdot)$ . However, in the case of a missing response, one has an incomplete sample of size  $n$  from  $(X, Y, B)$ , which is usually denoted by  $\{(X_i, Y_i, B_i) : 1 \leq i \leq n\}$ , where  $B_i = 1$  if  $Y_i$  is observed, and  $B_i = 0$  otherwise. The Bernoulli random variable  $B$  is supposed to be such that

$$P(x) = \mathbb{P}(B = 1 | X = x) = \mathbb{P}(B = 1 | X = x, Y = y)$$

is a continuous function. Under this smoothing consideration, we define the LLECCDF of  $CDF(\cdot | \cdot)$  by finding the minimizers  $(\widehat{\beta}_1, \widehat{\beta}_2)$  of

$$\min_{(\beta_1, \beta_2) \in \mathbb{R}^2} \sum_{i=1}^n \left( \mathcal{J} \left( \frac{y - Y_i}{\lambda_J} \right) - \beta_1 - \beta_2 \alpha(X_i, x) \right)^2 \mathcal{Ker} \left( \frac{\delta(x, X_i)}{\lambda_K} \right) B_i, \tag{2}$$

where  $\lambda_K := \lambda_{K,n}$  and  $\lambda_J := \lambda_{J,n}$  are bandwidth parameters and  $\mathcal{J}(\cdot)$  and  $\mathcal{Ker}(\cdot)$  are, respectively, distribution and kernel functions. The explicit solution to this minimization is

$$\widehat{\beta}_1 = \widehat{CDF}(y|x) = \frac{\sum_{j=1}^n \Gamma_j(x) \mathcal{Ker}_j(x) B_j \mathcal{J}_j(y)}{\sum_{j=1}^n \Gamma_j(x) B_j \mathcal{Ker}_j(x)},$$

with

$$\begin{aligned} \Gamma_j(x) &= \sum_{i=1}^n \alpha_i^2(x) \mathcal{Ker}_i(x) B_i - (\sum_{i=1}^n \alpha_i(x) \mathcal{Ker}_i(x) B_i) \alpha_j, \\ \mathcal{Ker}_j(x) &= \mathcal{Ker} \left( \frac{\delta(x, X_j)}{\lambda_K} \right) \end{aligned}$$

and

$$\mathcal{J}_j(y) = \mathcal{J} \left( \frac{y - Y_j}{\lambda_J} \right), \quad \alpha_j(x) = \alpha(X_j, x).$$

Furthermore, we can write

$$\widehat{CDF}(y | x) = \frac{\widehat{CDF}_N^x(y)}{\widehat{CDF}_D(x)},$$

where

$$\widehat{CDF}_N^x(y) := \frac{1}{n\mathbb{E}(\Gamma_1(x)\mathcal{Ker}_1(x))} \sum_{j=1}^n \Gamma_j(x)\mathcal{Ker}_j(x)\mathcal{J}_j(y)B_j,$$

and

$$\widehat{CDF}_D(x) := \frac{1}{n\mathbb{E}(\Gamma_1(x)\mathcal{Ker}_1(x))} \sum_{j=1}^n \Gamma_j(x)\mathcal{Ker}_j(x)B_j.$$

The first main contribution of this work is a precise convergence rate of the approximation  $\widehat{CDF}(\cdot | \cdot)$  uniformly over a non-necessary compact subset  $\mathcal{C}_{\mathcal{F}}$  of  $\mathcal{F}$ . We use the notation

$$Ba(x, h) = \{x' \in \mathcal{F} : d(x', x) \leq h\}.$$

**Definition 1.** Let  $\mathcal{C}_{\mathcal{F}}$  be a subset of a semi-metric space  $\mathcal{F}$ , and let  $\epsilon > 0$  be given. A finite set of points  $x_1, x_2, \dots, x_N$  in  $\mathcal{F}$  is called an  $\epsilon$ -net for  $\mathcal{C}_{\mathcal{F}}$  if

$$\mathcal{C}_{\mathcal{F}} \subset \bigcup_{k=1}^N Ba(x_k, \epsilon).$$

The quantity  $\psi_{\mathcal{C}_{\mathcal{F}}}(\epsilon) = \log(N_{\epsilon}(\mathcal{C}_{\mathcal{F}}))$ , where  $N_{\epsilon}(\mathcal{C}_{\mathcal{F}}) =: d_n$  is the minimal number of open balls in  $\mathcal{F}$  of radius  $\epsilon$  which is necessary to cover  $\mathcal{C}_{\mathcal{F}}$ , is called Kolmogorov’s  $\epsilon$ -entropy of the set  $\mathcal{C}_{\mathcal{F}}$ .

This concept was introduced by Kolmogorov in the mid-1950s (refer to [61]). It serves as a measure of the complexity of a set, indicating that high entropy implies that a significant amount of information is required to accurately describe an element within a given tolerance,  $\epsilon$ . Consequently, the selection of the topological structure (specifically, the choice of semi-metric) plays a crucial role when examining asymptotic results that are uniform over a subset,  $\mathcal{C}_{\mathcal{F}}$  of  $\mathcal{F}$ . In particular, we subsequently observe that a well-chosen semi-metric can enhance the concentration of the probability measure for the functional variable,  $X$ , while minimizing the  $\epsilon$ -entropy of the subset,  $\mathcal{C}_{\mathcal{F}}$ . Ferraty and Vieu [6] emphasized the phenomenon of concentration of the probability measure for the functional variable by calculating small ball probabilities in different standard scenarios; refer to [62]. For readers interested in these concepts (entropy and small ball probabilities) and/or the utilization of Kolmogorov’s  $\epsilon$ -entropy in dimensionality reduction problems, we recommend referring to [63] or/and [64], respectively.

Let  $(u_n)$  for  $n \in \mathbb{N}$ , be a sequence of real r.v.s. We say that  $(u_n)$  converges almost-completely (a.co.) toward zero if, and only if, for all  $\epsilon > 0$ ,

$$\sum_{n=1}^{\infty} \mathbb{P}(|u_n| > \epsilon) < \infty.$$

Moreover, we say that the rate of the almost-complete convergence of  $(u_n)$  toward zero is of order  $v_n$  (with  $v_n \rightarrow 0$ ), and we write  $u_n = O_{a.co.}(v_n)$  if, and only if, there exists  $\epsilon > 0$  such that

$$\sum_{n=1}^{\infty} \mathbb{P}(|u_n| > \epsilon v_n) < \infty.$$

This kind of convergence implies both the almost-sure convergence and the convergence in probability.

### 2.2. LLECQF: Numerical Approximation of CQF-Model

The second approximation concerns the CQF-Model of order  $p$ , for  $p \in (0, 1)$ , denoted as  $CQF_p(x)$ :

$$CDF(CQF_p(x)|x) = p.$$

The natural estimator is obtained by inverting the LLECCDF. However, unlike the local constant estimator, the LLECCDF is not monotone and invertible. To overcome this issue, we use the robust definition of the  $CQF_p(x)$ :

$$CQF_p(x) = \arg \min_t \mathbb{E}[L_p(Y - t) | X = x], \tag{3}$$

where the scoring function  $L_p(y) = |y|(p - \mathbb{1}_{[y < 0]})$ , with  $\mathbb{1}_S$  as the indicator of  $S$ . Once again, the LLECQF is obtained via a smoothing approximation of the model locally in the neighborhood of the location  $x$ . Therefore, we suppose that  $CQF_p(x)$  is such that

$$CQF_p(z) = \beta_1 + \beta_2\alpha(z, x) + o(\alpha(z, x)). \tag{4}$$

Therefore,  $\beta_1$  and  $\beta_2$  are estimated as

$$\arg \min_{(\beta_1, \beta_2) \in \mathbb{R}^2} \sum_{j=1}^n L_p(Y_j - \beta_1 - \beta_2\alpha(X_j, x)) \mathcal{K}er_j(x) B_j. \tag{5}$$

So, the LLECQF of  $CQF_p(x)$  is denoted by

$$\widehat{CQF_p}(x) = \widetilde{\beta}_1,$$

where

$$(\widetilde{\beta}_1, \widetilde{\beta}_2) = \arg \min_{(\beta_1, \beta_2) \in \mathbb{R}^2} \sum_{j=1}^n L_p(Y_j - \beta_1 - \beta_2\alpha(X_j, x)) \mathcal{K}er_j(x) B_j.$$

Once again, our main focus is to establish the convergence rate using the Kolmogorov entropy and the uniform-Lipschitzian condition: For all  $(y_1, y_2) \in \mathcal{N}_y^2$  and  $x_1, x_2 \in \mathcal{C}_\mathcal{F}$ , we have

$$\left| CDF^{(j)}(y_1|x_1) - CDF^{(j)}(y_1|x_2) \right| \leq C_2 \left( |\delta(x_2, x_1)|^{k_2} + |y_2 - y_1|^{k_1} \right), \tag{6}$$

for  $j = 0, 1$  and  $C_2, k_1, k_2 > 0$ .

### 3. Uniform Convergence Rate

Let  $\mathfrak{F}_i$  and  $\mathcal{G}_i$  ( $i = 1, \dots, n$ ) be the  $\sigma$ -algebras generated by  $((Y_1, X_1), \dots, (Y_i, X_i))$  and  $((Y_1, X_1), \dots, (Y_i, X_i), X_{i+1})$ , respectively. Let  $\mathcal{N}_y$  denote a fixed neighborhood of a point  $y$  in  $\mathbb{R}$ . Define

$$\psi_x(\lambda_1, \lambda_2) = \mathbb{P}(\lambda_2 \leq \delta(X, x) \leq \lambda_1).$$

The function  $\psi_x(\cdot, \cdot)$  plays a similar role to small ball probability as in [34]. Finally, let  $C$  and  $C'$  be some generic constants that are strictly positive. To obtain the convergence rate in this functional ergodic MAR scheme, we consider the following assumptions:

(H1) The first-order derivative of  $\psi(\cdot)$  exists and is bounded in  $\mathcal{C}_\mathcal{F}$ . The function  $\psi(\cdot)$  is such that

$$0 < C'\psi(\zeta) \leq \inf_{x \in \mathcal{C}_\mathcal{F}} \mathbb{P}(X \in \mathcal{B}a(x, \zeta)) \leq \sup_{x \in \mathcal{C}_\mathcal{F}} \mathbb{P}(X \in \mathcal{B}a(x, \zeta)) \leq C\psi(\zeta).$$

(H2) There exists a non-random function  $\psi_i(\cdot)$  satisfying

- (i)  $0 < C\psi_i(\zeta) < \mathbb{P}(X_i \in \mathcal{B}a(x, \zeta) | \mathfrak{F}_{i-1}) \leq C'\psi_i(\zeta)$ ,
- (ii)  $\frac{1}{n\psi(\zeta)} \sum_{i=1}^n \psi_i(\zeta) \rightarrow 1, \text{ a.co.}$ ,
- (iii) There exists a non-random function  $\Psi(\cdot)$  satisfying

$$\forall t \in [-1, 1], \quad \lim_{\lambda_K \rightarrow 0} \frac{\psi(t\lambda_K)}{\psi(\lambda_K)} = \Psi(t).$$

(H3) The distribution function  $\mathcal{J}(\cdot)$  and the kernel  $\mathcal{K}er(\cdot)$  fulfill the following:

(i)  $\mathcal{Ker}(\cdot)$  is a Lipschitz on its support  $[-1, 1]$  satisfying

$$D = \begin{pmatrix} \mathcal{Ker}(1) - \int_{-1}^1 \mathcal{Ker}'(t)\Psi(t)dt & \mathcal{Ker}(1) - \int_{-1}^1 (t\mathcal{Ker}(t))'\Psi(t)dt \\ \mathcal{Ker}(1) - \int_{-1}^1 (t\mathcal{Ker}(t))'\Psi(t)dt & \mathcal{Ker}(1) - \int_{-1}^1 (t^2\mathcal{Ker}(t))'\Psi(t)dt \end{pmatrix},$$

is a positive definite matrix.

(ii) The function  $\mathcal{J}(\cdot)$  has a derivative satisfying

$$\int |t|^{k_1} \mathcal{J}^{(1)}(t)dt < \infty.$$

(iii) We almost surely have  $\mathbb{E}[\mathcal{J}_j(y)|\mathcal{G}_{j-1}] = \mathbb{E}[\mathcal{J}_j(y)|X_j]$ .

(H4) The real sequence  $d_n$  associated with  $r_n = O\left(\frac{\ln n}{n}\right)$  satisfies

$$\frac{(\ln n)^2}{n\psi(\lambda_K)} < \ln d_n < \frac{n\psi(\lambda_K)}{\ln n} \text{ and } \sum_{n=1}^{\infty} d_n^{1-\varrho} < \infty \text{ for some } \varrho > 1.$$

(H5) The bandwidth  $\lambda_K$  is linked to  $\alpha(\cdot)$  and  $\psi(\cdot)$  by

(i)

$$\forall n > n_0 \quad - \frac{1}{\psi(\lambda_K)} \int_{-1}^1 \psi(z\lambda_K, \lambda_K) \frac{d}{dz} \left( z^2(\mathcal{Ker}(z)) \right) dz > C;$$

(ii)  $\lim_{n \rightarrow \infty} \lambda_K = 0$ , and  $\lim_{n \rightarrow \infty} \lambda_H = 0$ , and  $\lim_{n \rightarrow \infty} \frac{\ln n}{n\psi(\lambda_K)} = 0$ ;

(iii)

$$\lambda_K \int_{Ba(x, \lambda_K)} \alpha(u, x) d\mathbf{P}(u) = O\left( \int_{Ba(x, \lambda_K)} \alpha^2(u, x) d\mathbf{P}(u) \right),$$

where  $\mathbf{P}$  is the law of  $X$ .

We first recall that our assumptions are not restrictive and may be considered as standard in the functional local linear analysis context. They are similar to those used in the local linear estimation of the quantile regression in [30]. In particular, Assumption (H1) concerns the usual concentration property of the functional variable. It is well documented that this property allows us to explore the functionality nature of the data. Assumption (H3)(iii) is a Markov-type condition and characterizes the conditional moments. It is satisfied when considering, for instance, the model

$$\mathbb{1}_{(Y \leq y)} = m_\psi(X_i) + \epsilon_i,$$

where  $(\epsilon_i)$  is a square-integrable process independent of  $(X_i)$ . Finally, Assumptions (H3)(i)-(ii), (H4), and (H5) are technical conditions, similar to those used by [30].

The following theorem gives the uniform, almost complete convergence of  $\widehat{CDF}(y | x)$  with the rate.

**Theorem 1.** Under Conditions (H1)–(H5), we have

$$\sup_{x \in \mathcal{C}_F} \left| \widehat{CDF}(y | x) - CDF(y | x) \right| = O\left(\lambda_K^{k_1} + \lambda_J^{k_2}\right) + O_{a.co.} \left( \left( \frac{\ln d_n}{n\psi(\lambda_K)} \right)^{\frac{1}{2}} \right). \tag{7}$$

The proof of Theorem 1 is postponed until the last section. The following theorem gives the uniform, almost complete convergence of  $\widehat{CQF}_\alpha(x)$  with the rate.



**Theorem 2.** Under Conditions (H1)–(H5) and if  $\text{cdf}(CQF_\alpha(x)|x) > 0$ , we have

$$\sup_{x \in \mathcal{C}_F} \left| \widehat{CQF}_\alpha(x) - CQF_\alpha(x) \right| = O\left(\lambda_K^{\min(k_2, k_2)}\right) + O_{a.co.}\left(\left(\frac{\ln d_n}{n\psi(\lambda_K)}\right)^{\frac{1}{2}}\right). \tag{8}$$

The proof of Theorem 2 is postponed until the last section.

#### 4. Asymptotic Normality

The second asymptotic result concerns the asymptotic law of the sequence  $\widehat{CDF}(y | x)$ . To do that, we enhance Assumption (H5) by assuming that the kernel  $\mathcal{Ker}(\cdot)$  satisfies (H4) and has a first derivative  $\mathcal{Ker}'(\cdot)$  such that

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

Let us now state the following theorem, which gives the central limit theorem of the estimator  $\widehat{CDF}(y | x)$ . Below, we write  $Z \stackrel{D}{=} \mathcal{N}(\mu, \sigma^2)$  whenever the random variable  $Z$  follows a normal law with expectation  $\mu$  and variance  $\sigma^2$ ;  $\xrightarrow{D}$  denotes the convergence in distribution.

**Theorem 3.** Under the assumptions of Theorem 1 and if the smoothing parameters  $\lambda_K$  satisfy  $\lim_{n \rightarrow \infty} n\lambda_K^{2k_1} \psi(\lambda_K) = 0$ , then

$$\sqrt{n\psi(\lambda_K)} \left( \widehat{CDF}(y | x) - CDF(y | x) \right) \xrightarrow{D} N(0, V_{JK}(y | x)),$$

where

$$M_i = \mathcal{Ker}^i(1) - \int_{-1}^1 (\mathcal{Ker}^i(u))' \Psi(u) du,$$

for  $i = 1, 2$  and

$$V_{JK}(y | x) = \frac{M_2}{M_1^2} P(x) CDF(y | x) (1 - CDF(y | x)).$$

The proof of Theorem 3 is postponed until the last section.

Clearly, this asymptotic result has many applications in practice. In particular, it can be used to build a confidence interval for the true value of  $CDF(y | x)$ . The latter is obtained by estimating the asymptotic variance  $V_{JK}(y | x)$  using a plug-in approach. Indeed,  $M_1$  and  $M_2$  are estimated via

$$\widehat{M}_1 = \frac{1}{n\psi(\lambda_K)} \sum_{i=1}^n \mathcal{Ker}_i(x), \quad \widehat{M}_2 = \frac{1}{n\psi(\lambda_K)} \sum_{i=1}^n \mathcal{Ker}_i^2(x),$$

and

$$\widehat{P}(x) = \frac{\sum_{i=1}^n B_i \mathcal{Ker}_i(x)}{\sum_{i=1}^n \mathcal{Ker}_i(x)}.$$

Therefore, the natural estimator of the asymptotic variance  $V_{JK}(y | x)$  is

$$\widehat{V}_{JK}(y | x) = \frac{\widehat{M}_2}{\widehat{M}_1^2} \widehat{P}(x) \widehat{CDF}(y | x) (1 - \widehat{CDF}(y | x)).$$

Under this consideration, the  $(1 - \zeta)$  confidence interval for  $CDF(y | x)$  is

$$\widehat{CDF}(y | x) \pm t_{1-\zeta/2} \times \left( \frac{\widehat{V}_{JK}(y | x)}{n\psi(\lambda_K)} \right)^{1/2},$$

where  $t_{1-\zeta/2}$  denotes the  $1 - \zeta/2$  quantile of the standard normal distribution. We note that the function  $\psi(\cdot)$  does not appear in the calculation of the confidence interval since it will be simplified.

## 5. Numerical Results

In this computational study, we illustrate the three fundamental axes of our topic that are the functional structure, the asymptotic normality, and the local linear smoothing. In the first illustration, we examine the functional structure's impact on the constructed estimators' convergence rate. To cover the general features of our study, the ergodicity, and the MAR features, we generate an artificial functional time series using the Hilberthian autoregressive processes. Of course, this kind of process's linearity allows us to incorporate our theoretical assumption concerning the ergodicity structure. To do this, we employ the routine-code *fts.rar* from the R package *freedom.fda*. The latter has a nice feature that is based on the dynamic functional principal component analysis (DFPCA); for instance, see [65]. In this empirical study, we use DFPCA to generate the ergodic function using specific basis functions. Specifically, we have used the Fourier basis functions (FBF) to obtain the functional  $(X_i)$ . Formally,

$$X_i = Y(X_{i-1}) + \varepsilon_i,$$

where  $Y$  is an operator with kernel  $\psi(\cdot, \cdot)$  and  $\varepsilon_i$  is the white noise. The kernel operator is expressed by

$$Y(X(\cdot)) = \int_0^1 \psi(t, s)X(s)ds.$$

In the cited routine code  $Y$  is constructed from FBF  $\{u_j : j = 1, \dots, d\}$  by taking  $(\psi_{ij})_{ij} = (\langle \Psi(u_i), u_j \rangle)$  as the corresponding matrix. The function *op.norms* controls the degree of dependency in these functional observations for practical purposes. The functional observed regressors are plotted in Figure 1.

Specifically, such a sample was obtained by taking arbitrary values allowing it to cover various degrees of ergodicity. Secondly, the variable of interest is generated using the regression equation

$$Y = \exp\left\{-\int_0^1 \frac{1}{1+X^2(t)}dt\right\} + \varepsilon,$$

where  $\varepsilon$  is drawn from  $N(0, 0.5)$ . With this consideration, the conditional law of  $Y$  given  $X$  is a normal distribution with a mean equal

$$\exp\left\{-\int_0^1 \frac{1}{1+X^2(t)}dt\right\}.$$

The missing feature is controlled using the conditional probability of observation:

$$P(X = x) = \text{expit}\left(\frac{2\gamma}{\pi} \int_0^{2\pi} x^2(t)dt\right),$$

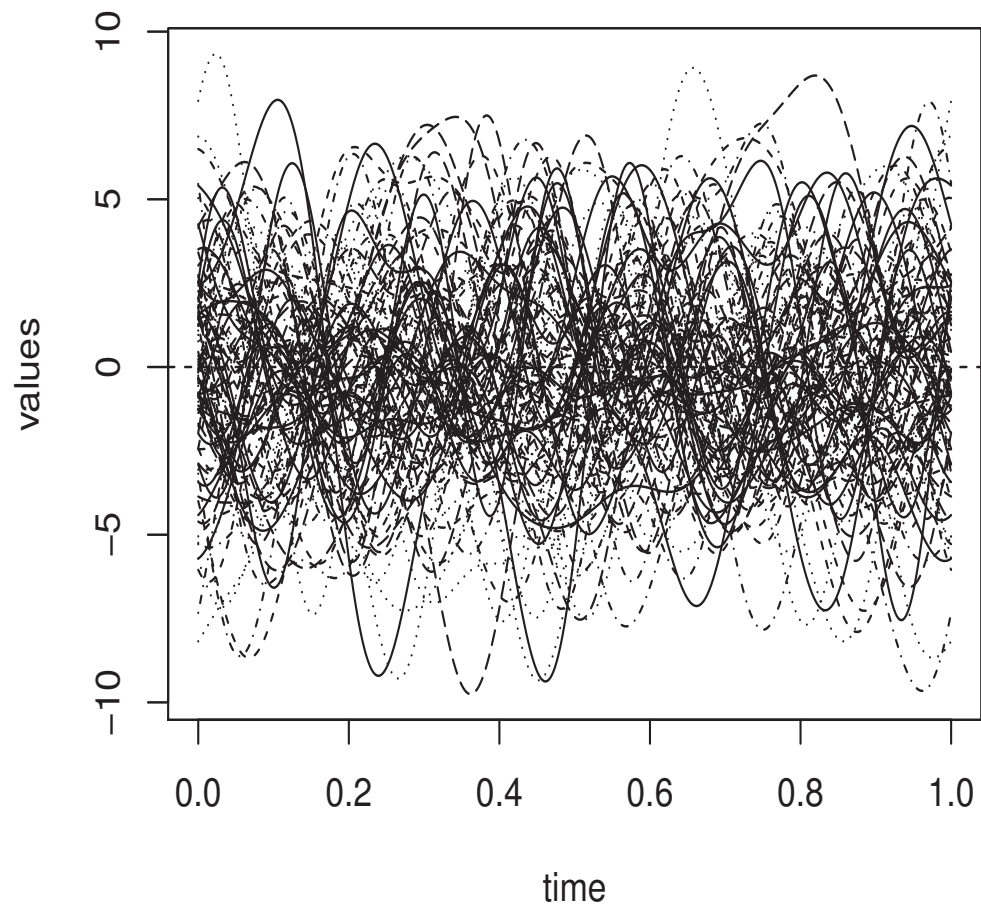
where

$$\text{expit}(u) = e^u / (1 + e^u).$$

Under this consideration, the scalar  $\gamma$  controls the missing rate. Once again, we simulate with several values of  $\gamma$  to evaluate this characteristic's influences in the estimators. The sequences LLECCDF and LLECQF are computed using the  $(-1, 1)$ -quadratic kernel and the locally cross-validation on the number of neighborhoods using the mean square error (MSE)

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n \left( Y_i - \widehat{CQF}_{0.5}^{-i}(X_i) \right)^2,$$

with  $\widehat{CQF}_{0.5}^{-i}$  referring to the estimation leave-one-out-curve of the conditional median. Clearly, the Kolmogorov entropy is an important factor of the topological structure. The latter is controlled through the locating functions  $\alpha$  and  $\delta$ . For the sake of shortness, we simulate  $\alpha = \delta$  equal to the  $L_2$ -distance using three between the  $q$ th derivatives of the curves based on the basis function, as well as the PCA-semi metric associated with the eigenfunctions of the empirical version covariance operator. Exactly, we compute this metric using the  $m$  first one. The MSE was evaluated over various values of  $q$  and  $m$ .



**Figure 1.** A sample of 100 curves.

In Tables 1 and 2, we summarize the MSE of both estimators for various values of mentioned parameters, the level of dependency (*op.norms*), the missing rate  $\gamma$ , and the metric parameter  $q$  or  $m$ .

It is clear that the behavior of the estimator is strongly impacted by different parameters of the estimators, including the dependency-level, the missing rate, and the topological structure. Without surprise, the topological structure has an important role. In a sense, the variability of the MSE as a function of the metric type is more important than the other parameters. All in all, the effectiveness of the estimator is also affected by the quantities (*op.norms*),  $\gamma$ , and  $q$  or  $m$ . In conclusion, we can say that the computational study confirms the theoretical statement that the convergence rate is strongly affected by the topological structure of the functional regressor.

**Table 1.** MSE error for LLECCDF.

The Metric Type	<i>op.norms</i>	$\gamma$	<i>q or m</i>	MSE
Fourier Basis Functions	0.98	5	2	0.196
			0	0.193
		0.5	2	0.184
			0	0.187
		0.05	2	0.156
			0	0.155
	0.48	5	2	0.172
			0	0.174
		0.5	2	0.168
			0	0.161
		0.05	2	0.149
			0	0.151
B-spline Basis Functions	0.48	5	2	0.266
			0	0.264
		0.5	2	0.257
			0	0.259
		0.05	2	0.251
			0	0.248
	0.08	5	2	0.243
			0	0.242
		0.5	2	0.239
			0	0.237
		0.05	2	0.227
			0	0.230
PCA-metric	0.98	5	3	0.408
			1	0.407
		0.5	3	0.398
			1	0.394
		0.05	3	0.387
			1	0.376
	0.08	5	3	0.143
			1	0.292
		0.5	3	0.295
			1	0.283
		0.05	3	0.284
			1	0.279

The second illustration concerns the quality of the asymptotic normality result in Theorem 3. Specifically we aim to examine the behavior of the asymptotic distribution with respect to the degree of correlation, as well as the missing rate. For this purpose, we repeat the previous sampling processes independently  $m$  times, and for each time, we calculate the quantity:

$$\sqrt{\frac{n\psi(\lambda_K)\widehat{M}_1^2}{\widehat{M}_2\widehat{P}(x)CDF(y|x)(1-CDF(y|x))}} \left( \widehat{CDF}(y|x) - CDF(y|x) \right). \quad (9)$$

Recall that the true conditional law of  $Y$  given  $X$  is known justifying the use of the true conditional cumulative distribution function of  $CDF(y|x)$ . Observe also that the estimation of the function  $\psi(\lambda_K)$  is not necessary. It will be simplified using the definition of  $\widehat{M}_1$  and  $\widehat{M}_2$  (see their definitions in Section 4).

**Table 2.** MSE error for LLECQF.

The Metric Type	<i>op.norms</i>	$\gamma$	<i>q or m</i>	MSE
Fourier Basis Functions	0.98	5	2	0.091
			0	0.093
		0.5	2	0.081
			0	0.085
		0.05	2	0.056
			0	0.054
	0.08	5	2	0.051
			0	0.04
		0.5	2	0.033
			0	0.039
		0.05	2	0.031
			0	0.023
B-spline Basis Functions	0.98	5	2	0.128
			0	0.123
		0.5	2	0.121
			0	0.122
		0.05	2	0.117
			0	0.114
	0.48	5	2	0.109
			0	0.107
		0.5	2	0.099
			0	0.095
		0.05	2	0.085
			0	0.074
PCA-metric	0.48	5	3	0.104
			1	0.108
		0.5	3	0.102
			1	0.101
		0.05	3	0.099
			1	0.097
	0.08	5	3	0.086
			1	0.092
		0.5	3	0.081
			1	0.078
		0.05	3	0.064
			1	0.061

Now, the  $m$ -sample of the quantity (9) is calculated using the ideas of the first illustration concerning the construction of  $\widehat{CDF}(y | x)$ . Moreover, the estimators  $\widehat{M}_1$ ,  $\widehat{M}_2$ , and  $\widehat{P}(x)$  are obtained in the same manner. Specifically, we use the same bandwidth, the same kernel, and the metric associated the FBF with  $q = 0$ . Of course, the  $m$ -sample of (9) is drawn for a fixed location curve  $x = X_{i_0}$  and fixed point  $y = Y_{i_0}$ . The index  $i_0$  is randomly chosen independently from the sampling process. Furthermore, the behavior of the asymptotic distribution of the quantity (9) is examined by estimating the density of the obtained  $m$ -sample and we compare it to the density of the standard normal distribution. In order to evaluate the effect of the dependency degree and the missing rate on the accuracy of the asymptotic normality, we perform our sampling process using various values of *op.norms* and  $\gamma$ . Exactly, we keep the same value as that of the first illustration. Finally, we plot in Figures 2–5 the estimated density, obtained via the routine code *density* with  $m = 120$ , against the density of  $N(0, 1)$ . The continuous line represents the estimated density, and the dashed line represents the true density.

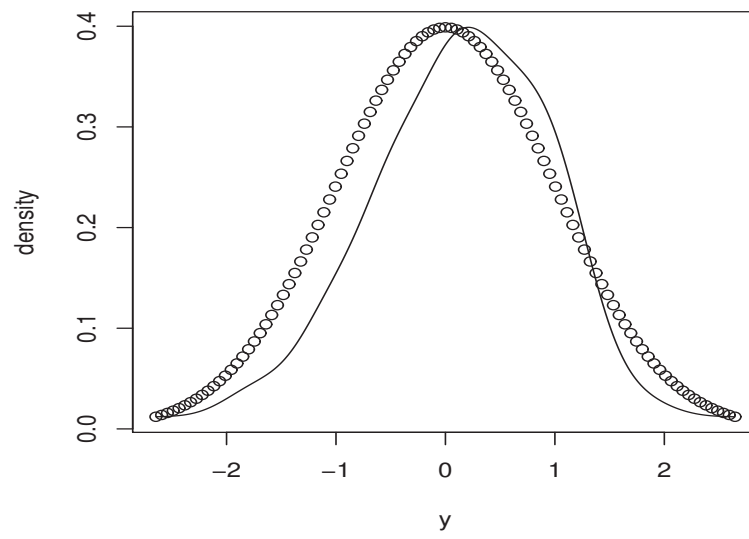


Figure 2. Case  $(op.norms, \gamma) = (0.98, 5)$ .

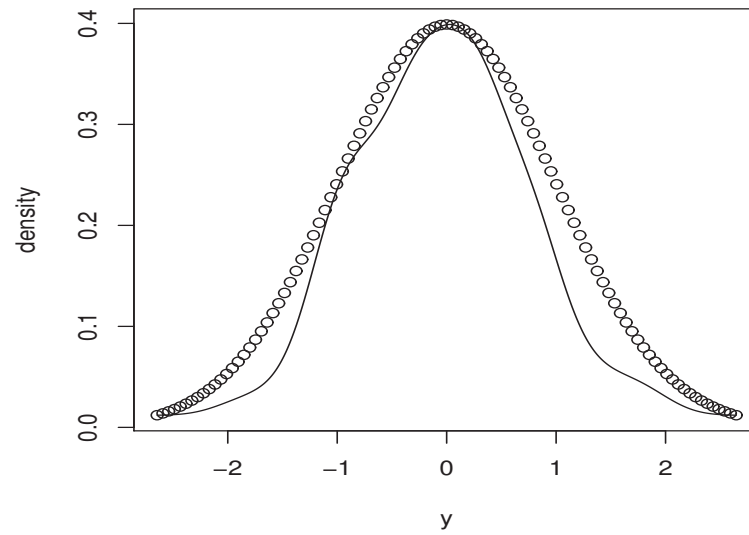


Figure 3. Case  $(op.norms, \gamma) = (0.98, 0.5)$ .

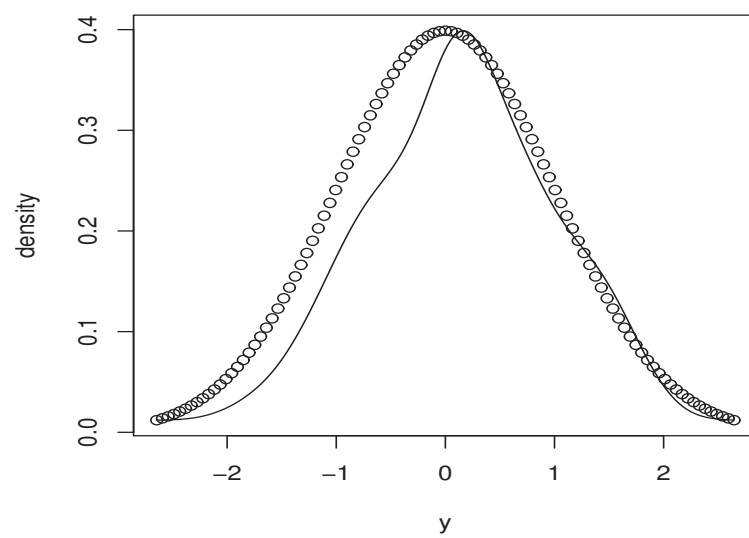


Figure 4. Case  $(op.norms, \gamma) = (0.48, 5)$ .

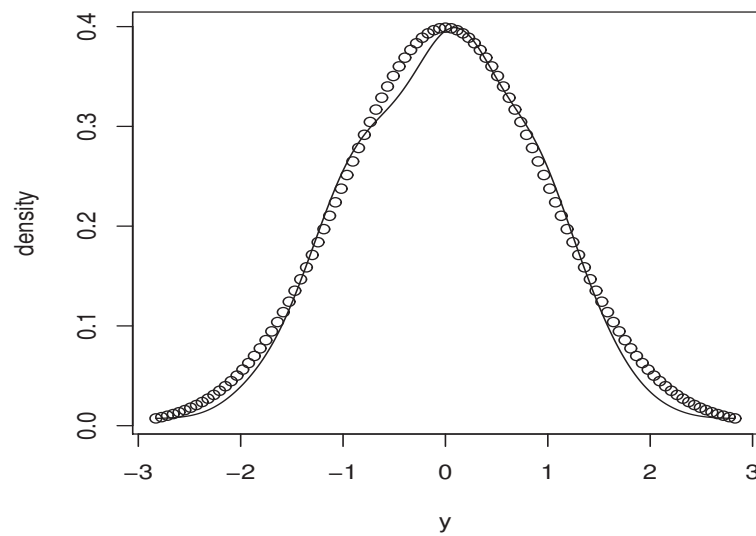


Figure 5. Case  $(op.norms, \gamma) = (0.48, 0.5)$ .

Once again, this empirical analysis confirms the theoretical development. In a sense, the estimation approach is strongly impacted by the degree of dependency, as well as the missing rate. Typically, even if the curve of the estimated density is relatively close to the normal density, the accuracy of the asymptotic normality is significantly varied with respect to the values of  $op.norms$  and  $\gamma$ . It appears that the effect of the missing rate is more important than the degree of dependency. This conclusion is justified by the fact that the missing rate impacts both the bias and the variance terms, whereas the dependency feature impacts only the variance part. To confirm this statement, we report in Table 3 the bias quantities of the different values of  $(op.norms)$  and  $\gamma$ . The bias term is

$$Bias = \frac{1}{m} \sum_{j=1}^m \widehat{CDF}^j(Y_{i_0} | X_{i_0}) - CDF(Y_{i_0} | X_{i_0}),$$

where  $\widehat{CDF}^j(Y_{i_0} | X_{i_0})$  is the LLECCDF obtained with the  $j^{th}$  sample. The result of Table 3 shows that the effect of the data correlation is too small compared with the variability with respect to the missing rate.

Table 3. Bias error for LLECQF.

$op.norms$	$\gamma$	Bias
0.98	5	0.121
	0.5	0.093
0.48	5	0.109
	0.5	0.091

The third illustration concerns the used estimation method that is the local linear approach. More precisely, we compare it with the classical kernel method. We concentrate on this part in the second model, which is the quantile regression. It is well known that this kind of model has a many scopes of application. One of the important application areas is the prediction problem. At this stage, the quantile regression can be used as single-point predictor when  $p = 0.5$  or as a predictive interval  $[CQF_{p/2}, CQF_{1-p/2}]$ . The latter ensures the existence of the response variable  $Y$  in  $[CQF_{p/2}(x), CQF_{1-p/2}(x)]$  with a probability equal to  $1 - p$ . In order to show the easy implantation of the estimator  $\widehat{CQF}_\alpha(\cdot)$  and to highlight its advantages over the classical kernel method studied by [4], we compare the efficiency of both estimation methods on the construction of the  $[CQF_{p/2}(x), CQF_{1-p/2}(x)]$  interval. Undoubtedly, the performance of any predictive interval is measured using two factors that are the coverage probability and the length of the interval. However, for

the sake of brevity, we focus in this third illustration only on the coverage probability factor that measures the belonging percentage to the approximated predictive interval. For this aim, we employ the same sampling process as the second illustration, and for each sampling time  $j$ , we split the observations into learning and testing sample. We determine  $[C\bar{Q}F_{p/2}(x_0), C\bar{Q}F_{1-p/2}(x_0)]$  for all point  $x_0$  in the testing sample, and  $C\bar{Q}F$  means either the local linear or kernel estimator of the quantile  $CQF$ . For the computation aspect, we use the routine code `funopare.quantile.lcv` for the kernel method, whereas that for the local linear method is obtained by minimizing (3) on a regular grid of 100 points, from  $(0.9 * \min(Y_i), 1.1 * \max(Y_i))$ . To the end, let us point out that we have used the same selection strategies as the previous illustrations, namely, the same bandwidth, the same kernel, and the same FBF metric. In order to give a comprehensive comparison, we examine the efficiency of the predictive interval processing for various values of  $p$ . Exactly, we examine 30 values of  $p_i$  in  $(0, 1)$ . Such values are compared with the coverage probabilities of both estimators in the four previous situations covering the strong and weak dependencies, as well as the two levels of missing rates. The results are presented in Figures 6–9 for which we plot the different values of  $q_i = 1 - p_i$  versus the coverage probability  $CP_i$ .

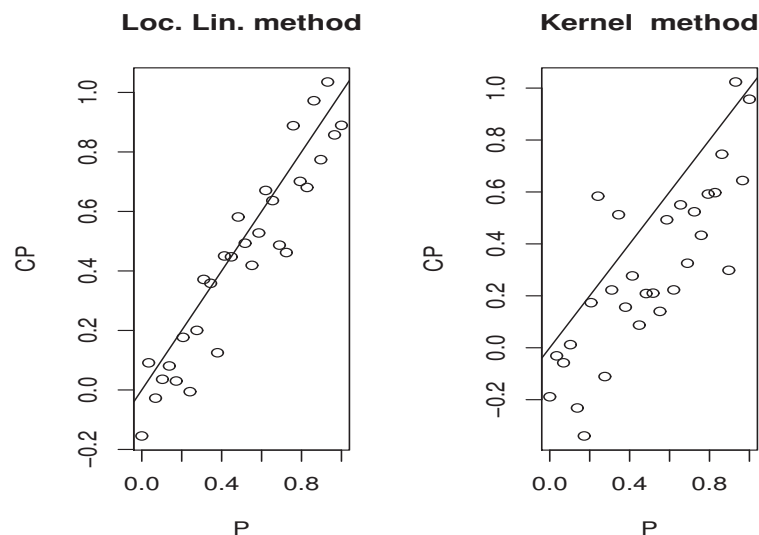


Figure 6. Case  $(op.norms, \gamma) = (0.98, 5)$ .

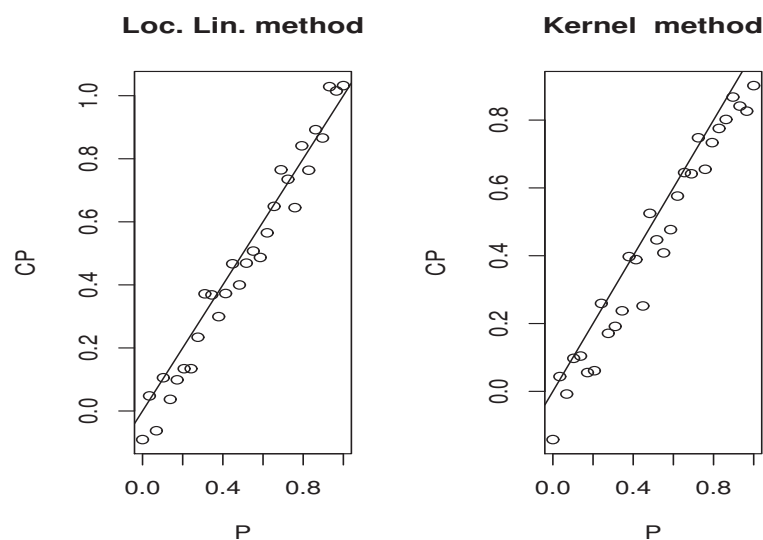


Figure 7. Case  $(op.norms, \gamma) = (0.98, 0.5)$ .



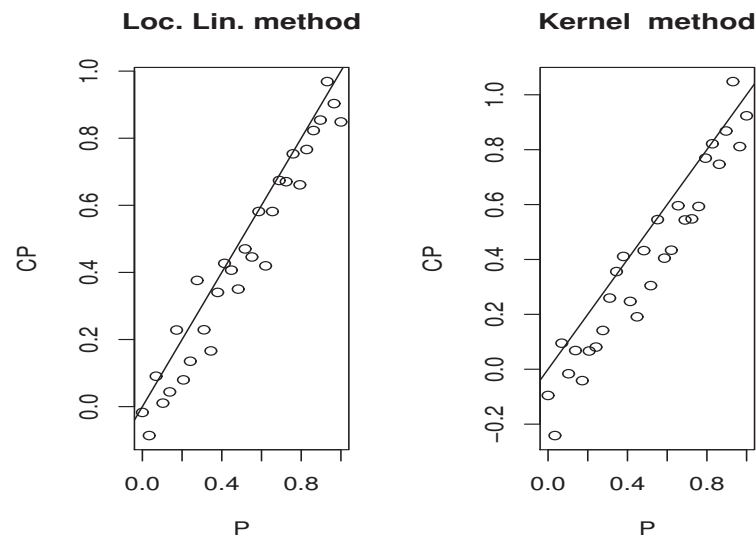


Figure 8. Case  $(op.norms, \gamma) = (0.48, 5)$ .

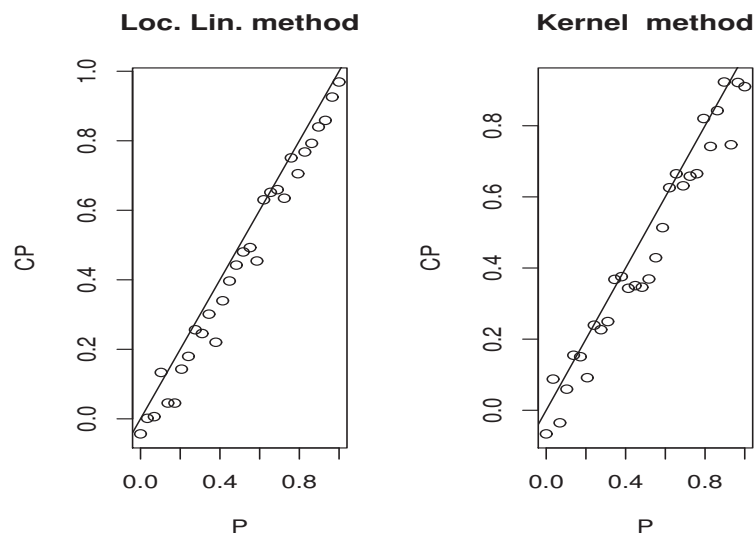


Figure 9. Case  $(op.norms, \gamma) = (0.48, 0.5)$ .

Once again, the estimation quality is strongly affected by the two principal features that are the dependency and the missing phenomena. However, it seems that the local linear algorithm is more robust and more preferment in the sense that its behavior is more stable than the kernel method. To confirm this statement, we summarise in Table 4, the absolute coverage probability error defined by

$$ACPE = \frac{1}{30} \sum_{i=1}^{30} |CP_i - q_i|,$$

where  $CP_i$  represents the coverage probability of the predictive interval associated with the threshold  $\alpha_i$ .

**Table 4.** Coverage probability error.

Method	op.norms	$\gamma$	ACPE
Local linear estimator	0.98	5	0.051
		0.5	0.042
	0.48	5	0.044
		0.5	0.037
Kernel method	0.98	5	0.19
		0.5	0.098
	0.48	5	0.13
		0.5	0.067

**6. Discussion and Comments**

- *Gap between the pointwise and uniform convergence in FDA:* We specify the convergence rate over some known functional structures to highlight the gap between the pointwise and uniform convergence in functional statistics. The first one concerns the resealed version  $(W_c(t) = W(t/c))_{t \geq 0}$  of Gaussian process  $W(t)_{t \geq 0}$ . If we assume that the spectral measure  $\mu$  of  $W$  such that, for some  $a > 0$ ,

$$\int e^{a|\lambda|} \mu(du\lambda) < \infty,$$

then the Kolmogorov’s  $\epsilon$ -entropy of  $\mathcal{C}_{\mathcal{F}}$ , the unit ball in reproducing kernel Hilbert spaces of the process  $W_c(\cdot)$  as a subset of  $(C([0, 1]), \|\cdot\|_{\infty})$ , is of the order

$$\frac{1}{c} \log^2\left(\frac{1}{\epsilon}\right);$$

for instance, see [66], implying an uniform convergence rate over  $\mathcal{C}_{\mathcal{F}}$  asymptotically equal to

$$O\left(\lambda_K^{k_1} + \lambda_J^{k_2}\right) + O_{a.co.} \left( \left( \frac{\ln^2 n}{n\psi(\lambda_K)} \right)^{\frac{1}{2}} \right).$$

Secondly, if we put  $\mathcal{C}_{\mathcal{F}}$  as the closed unit ball of Cameron–Martin spaces associated with the covariance operator of the standard stationary Ornstein–Uhlenbeck process defined by

$$Cov(s, t) = \exp(-a|s - t|), \quad a > 0,$$

the Kolmogorov’s  $\epsilon$ -entropy of this subset with respect to the norm of the Sobolev space  $W^{1,2}([0, 1])$  is of order

$$\left( \sqrt{(2a)} \left( \frac{1}{\pi\epsilon} \right) \right);$$

for instance, refer to [66], and the convergence rate is

$$O\left(\lambda_K^{k_1} + \lambda_J^{k_2}\right) + O_{a.co.} \left( \left( \frac{\ln n}{n\psi(\lambda_K)} \right)^{\frac{1}{2}} \right).$$

Thirdly, it is shown in [67] that any closed ball in a Sobolev space  $W^{1,1}([0, T])$  endowed with the norm  $L^1([0, T])$  has a Kolmogorov’s  $\epsilon$ -entropy of the order

$$\frac{1}{\epsilon} \log \frac{1}{\epsilon},$$

implying a uniform convergence rate asymptotically equal to

$$O(\lambda_K^{k_1} + \lambda_J^{k_2}) + O_{a.co.} \left( \left( \frac{\ln^2 n}{n\psi(\lambda_K)} \right)^{\frac{1}{2}} \right).$$

So, it is clear that the uniform convergence rate differs among functional subsets. However, in the nonfunctional case, where all norms are equivalent, the  $\epsilon$ -entropy for any compact subset in  $\mathbb{R}$  is of the order  $\log\left(\frac{1}{\epsilon}\right)$ , allowing us to keep the usual convergence rate in the finite-dimensional case, that is

$$O(\lambda_K^{k_1} + \lambda_J^{k_2}) + O_{a.co.} \left( \left( \frac{\ln n}{n\lambda_K} \right)^{\frac{1}{2}} \right),$$

which also coincides with the pointwise convergence rate. In conclusion, unlike the finite-dimensional case, there is a real gap between the pointwise and the uniform convergence in FDA. Thus, treating uniform consistency in the FDA is a challenging question; refer to [9,68–70].

- *The effect of the basis function on the Kolmogorov’s entropy:* Similarly to the previous paragraph, Kolmogorov’s entropy is also affected by the spectral decomposition of the functional variable over specific basis functions. Of course, this relationship is justified by the fact that both tools have similar interpretations. In a sense, the Kolmogorov’s entropy of a given subset  $\mathcal{C}_{\mathcal{F}}$  is the number of binary bits of information required to describe any  $x \in \mathcal{C}_{\mathcal{F}}$  with error  $\epsilon$ , while the spectral decomposition of  $x \in \mathcal{C}_{\mathcal{F}}$  given a basis functions  $(f_i)$  can be viewed as the reconstruction of  $x \in \mathcal{C}_{\mathcal{F}}$ . Furthermore, the minimal cardinal of elements of the basis  $(f_i)$ , sufficient to reconstruct any  $x \in \mathcal{C}_{\mathcal{F}}$  with error  $\epsilon$ , is so-called the sampling  $\epsilon$ -entropy of  $\mathcal{C}_{\mathcal{F}}$ ; for instance, see [71]. Theorem 4.1 in this cited work provides, under some general conditions, that the class of bounded piecewise  $C^k$  smooth functions as a subspace of  $L^2$  has a Kolmogorov’s  $\epsilon$ -entropy of the order

$$\left( \frac{1}{\epsilon} \right)^{1/k}.$$

In contrast, Kolmogorov’s  $\epsilon$ -entropy of the class of periodic real analytic functions on  $[-\pi, \pi]$ , as a subspace of  $C^0([-\pi, \pi])$  is

$$\log^2\left(\frac{1}{\epsilon}\right).$$

As a consequence of this statement (see Corollary 4.2 in [71]), the sampling  $\epsilon$ -entropy of the class spanned by the B-splines basis function is greater than the sampling  $\epsilon$ -entropy of the class of periodic real analytic functions that is spanned for the Fourier basis function. In conclusion, in practice, the estimator’s accuracy is also affected by the choice of basis functions and the cardinal of the basis used in the metric.

- *Estimation of the unconditional distribution of the MAR response:* One of the fundamental applications of conditional modeling in the MAR structure is the possibility of reconstructing the feature of the MAR variable. In our context, we use the fact that

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{E}[CDF(y|X)].$$

Thus, the natural estimator of the cumulative distribution function is

$$\widehat{F}_Y(y) = \frac{1}{n} \sum_{i=1}^n \widehat{CDF}^{-i}(y | X_i),$$

where  $\widehat{CDF}^{-i}(y | X_i)$  is the estimator  $\widehat{CDF}(\cdot | \cdot)$  without the observation  $(X_i, Y_i, B_i)$ . An alternative estimator of  $\widehat{F}_Y(y)$  can be obtained by replacing only the missing observation with the conditional expectation. Specifically, the second estimator is expressed by

$$\widetilde{F}_Y(y) = \frac{1}{n} \sum_{i=1}^n \left( B_i \mathbb{1}_{[Y_i < y]} + (1 - B_i) \widehat{CDF}^{-i}(y | X_i) \right).$$

We return to [48] for more ideas to construct other estimators. At this stage, the uniform consistency obtained in this paper is an important preliminary tool to derive the root  $n$ -consistency of this kind of estimator; we refer to [48] for the regression case.

### 7. Concluding Remarks

The convergence rate for a free-distribution functional data analysis is challenging. It requires some advanced tools for functional analysis in pure mathematics. This paper gives several contributions to the existing literature on functional data analysis. First, this paper demonstrated that the Kolmogorov entropy is an essential tool for describing the convergence rate of local linear estimation (LLE). To determine the uniform convergence rate of the LLE of the conditional cumulative distribution function (LLECCDF) and the LLE conditional quantile function (LLECCQF), we have used this device. Second, a central limit theorem is established for the proposed estimators. These results are demonstrated under general assumptions that permit coverage of the case of incomplete functional time series. Specifically, we model the dependence using the ergodic assumption and assume that the response variable is missing at random (MAR). Finally, we have evaluated the finite sample performance of the proposed estimators using Monte Carlo simulations. In addition to the previous issues, the present paper opens some important paths for the future. It will be natural to consider in a future investigation of the functional kNN local linear approach quantile regression estimators to obtain an alternative estimator that benefits from the advantages of both methods, the local linear method, and the kNN approach. Extending nonparametric functional concepts to local stationary processes is a relatively underdeveloped field. It would be intriguing to extend our work to the case of the functional local stationary process, which requires nontrivial mathematics; however, doing so would be far beyond the scope of this paper.

### 8. Proofs

This section contains the proof of our results. The notation introduced previously will be employed in the following. All of the proofs rely on applying the exponential inequality of the martingale difference. The proofs are quite lengthy; we limit ourselves to the main arguments.

#### 8.1. Proof of the Main Theorems

**Proof of Theorem 1.** For (7), we write

$$\begin{aligned} \widehat{CDF}_N^x(y) &:= \frac{1}{n \mathbb{E}(\Gamma_1(x) \mathcal{K}er_1(x))} \sum_{j=1}^n \mathbb{E}(\Gamma_j(x) \mathcal{K}er_j(x) \mathcal{J}_j(x) B_j | \mathfrak{F}_{j-1}), \\ \widehat{CDF}_D(x) &:= \frac{1}{n \mathbb{E}(\Gamma_1(x) \mathcal{K}er_1(x))} \sum_{j=1}^n \mathbb{E}(\Gamma_j(x) \mathcal{K}er_j(x) B_j | \mathfrak{F}_{j-1}). \end{aligned}$$

Then, the statement (7) is based on the following decomposition:

$$\begin{aligned} &\widehat{CDF}(y | x) - CDF(y | x) \\ &= B_n(y | x) + \frac{1}{\widehat{CDF}_D(x)} \left[ (B_n(y | x) + CDF(y | x)) A_n(y | x) + R_n(y | x) \right], \quad (10) \end{aligned}$$

where

$$B_n(y | x) = \frac{C\bar{D}F_N^x(y)}{C\bar{D}F_D(x)} - CDF(y | x),$$

$$A_n(y | x) = C\bar{D}F_D(x) - \widehat{C\bar{D}F}_D(x),$$

and

$$R_n(y | x) = \widehat{C\bar{D}F}_N^x(y) - C\bar{D}F_N^x(y).$$

So, all that remains is to demonstrate Lemmas 1–3:

**Lemma 1.** Under Conditions (H1)–(H5), we have, as  $n \rightarrow \infty$ ,

$$\sup_{x \in \mathcal{C}_{\mathcal{F}}} |B_n(y | x)| = O(\lambda_K^{b_1} + \lambda_J^{b_2}).$$

**Lemma 2.** Under Conditions (H1)–(H5), we have, as  $n \rightarrow \infty$ ,

$$\sup_{x \in \mathcal{C}_{\mathcal{F}}} |R_n(y | x)| = O_{a.co.} \left( \sqrt{\frac{\ln d_n}{n\psi(\lambda_K)}} \right).$$

**Lemma 3.** Under Conditions (H1)–(H5), we have, as  $n \rightarrow \infty$ ,

$$\sup_{x \in \mathcal{C}_{\mathcal{F}}} |A_n(y | x)| = O_{a.co.} \left( \sqrt{\frac{\ln d_n}{n\psi(\lambda_K)}} \right),$$

and

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

□

**Proof of Theorem 2.** For (8), we define for vector  $\vec{\delta} = \begin{pmatrix} c \\ d \end{pmatrix}$ , we put

$$\Phi(\vec{\delta}) = \alpha - \mathbb{1}_{Y_j \leq (c + \beta_1) + (\lambda_K^{-1}d + \beta_2)\alpha_j}'$$

and we employ Lemma 3 in [30] for

$$V_n(\vec{\delta}, x) = \frac{1}{n\psi(\lambda_K)} \sum_{j=1}^n \Phi(\vec{\delta}) \begin{pmatrix} 1 \\ \lambda_K^{-1}\alpha_j \end{pmatrix} \mathcal{K}er_j(x) B_j,$$

$$A_n(x) = V_n(\vec{\delta}_0, x) \quad \text{with} \quad \vec{\delta}_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \delta_n = \begin{pmatrix} \widetilde{\beta}_1 - \beta_1 \\ \lambda_K(\widetilde{\beta}_2 - \beta_2) \end{pmatrix}.$$

Thus, (8) is a consequence of Lemmas 4 and 5.

**Lemma 4.** Under Conditions (H1)–(H5), we have, as  $n \rightarrow \infty$ ,

$$\sup_{x \in \mathcal{C}_{\mathcal{F}}} \|A_n(x)\| = O(\lambda_K^{\min(k_2, k_2)}) + O_{a.co.} \left( \left( \frac{\ln d_n}{n\psi(\lambda_K)} \right)^{1/2} \right).$$

**Lemma 5.** Under Conditions (H1)–(H5), we have, as  $n \rightarrow \infty$ ,

$$\begin{aligned} & \sup_{x \in \mathcal{C}_{\mathcal{F}}} \sup_{\|\vec{\delta}\| \leq M} \|V_n(\vec{\delta}, x) + \lambda_0 D\vec{\delta} - A_n(x)\| \\ &= O(\lambda_K^{\min(k_2, k_2)}) + O_{a.co.} \left( \left( \frac{\ln d_n}{n\psi(\lambda_K)} \right)^{1/2} \right), \end{aligned}$$

where

$$\lambda_0 = cdf(CQF_p(x)|x)P(x).$$

□

**Proof of Theorem 3.** Using the decomposition

$$\widehat{CDF}(y | x) - CDF(y | x) = \frac{C_n(y | x) + Q_n(y | x)}{\widehat{CDF}_D(x)} + B_n(y | x), \tag{11}$$

with

$$Q_n(y | x) = R_n(y | x) + CDF(y | x)A_n(y | x),$$

and

$$C_n(y | x) = B_n(y | x)A_n(y | x).$$

Then, the asymptotic normality can be demonstrated via the following Lemmas.

**Lemma 6.** Under the assumptions of Theorem 3, we have, as  $n \rightarrow \infty$ ,

$$\sqrt{\frac{n\psi(\lambda_K)}{V_{JK}(y | x)}} Q_n(y | x) \xrightarrow{D} N(0, 1). \tag{12}$$

□

### 8.2. Proof of the Technical Lemmas

**Proof of Lemma 1.** Writing

$$\sup_{x \in \mathcal{C}_{\mathcal{F}}} |B_n(y | x)| = \frac{\sup_{x \in \mathcal{C}_{\mathcal{F}}} |\tilde{B}_n(y | x)|}{\inf_{x \in \mathcal{C}_{\mathcal{F}}} |\widehat{CDF}_D(x)|},$$

where

$$\tilde{B}_n(y | x) = \widehat{CDF}_N^x(y) - CDF(y | x)\widehat{CDF}_D(x).$$

The latter is

$$\begin{aligned} \tilde{B}_n(y | x) &= \frac{1}{n\mathbb{E}(\Gamma_1(x)\mathcal{K}er_1(x))} \sum_{j=1}^n \{ \mathbb{E}(\Gamma_j(x)\mathcal{K}er_j(x)\mathcal{J}_j B_j | \mathfrak{F}_{j-1}) \\ &\quad - CDF(y | x)\mathbb{E}(\Gamma_j(x)\mathcal{K}er_j(x)B_j | \mathfrak{F}_{j-1}) \} \\ &\leq \frac{1}{n\mathbb{E}(\Gamma_1(x)\mathcal{K}er_1(x))} \\ &\quad \sum_{j=1}^n \{ P(x)\mathbb{E}(\Gamma_j(x)\mathcal{K}er_j(x)|\mathbb{E}[\mathcal{J}_j(y)|X_j] - CDF(y | x)) | \mathfrak{F}_{j-1} \}. \end{aligned} \tag{13}$$

The integration by part gives

$$\mathbb{I}_{Ba(x, \lambda_K)}(X_j) |\mathbb{E}[\mathcal{J}_j | X_j] - CDF(y | x)| \leq C(\lambda_K^{k_1} + \lambda_j^{k_2}). \tag{14}$$

Hence, we obtain

$$\sup_{x \in \mathcal{C}_{\mathcal{F}}} |\tilde{B}_n(y | x)| = O(\lambda_K^{k_1} + \lambda_j^{k_2}) \sup_{x \in \mathcal{C}_{\mathcal{F}}} \widehat{CDF}_D(x).$$

This completes the proof. □

**Proof of Lemma 2.** Firstly, observe that, for all  $x \in \mathcal{C}_{\mathcal{F}}$ ,

$$|\Gamma_j(x)| \leq nC\lambda_K^2 + nC\lambda_K|\alpha_j(x)|. \tag{15}$$

Next, we let  $k(x) = \arg \min_{k \in \{1, 2, \dots, d_n\}} |\delta|(x, x_k)$ . We have

$$\begin{aligned} \sup_{x \in \mathcal{C}_{\mathcal{F}}} |R_n(y | x)| &\leq \underbrace{\sup_{x \in \mathcal{C}_{\mathcal{F}}} |\widehat{CDF}_N^{x_k(x)}(y) - \bar{CDF}_N^{x_k(x)}(y)|}_{Q_1} \\ &\quad + \underbrace{\sup_{x \in \mathcal{C}_{\mathcal{F}}} |\widehat{CDF}_N^x(y) - \widehat{CDF}_N^{x_k(x)}(y)|}_{Q_2} \\ &\quad + \underbrace{\sup_{x \in \mathcal{C}_{\mathcal{F}}} |CDF_N^{x_k(x)}(y) - \bar{CDF}_N^x(y)|}_{Q_3}. \end{aligned}$$

We start by treating  $Q_2$ . By using the boundeness on  $\mathcal{Ker}(\cdot)$  and  $\mathcal{J}(\cdot)$ , we write

$$\begin{aligned} Q_2 &\leq \sup_{x \in \mathcal{C}_{\mathcal{F}}} \frac{1}{n} \sum_{j=1}^n |\mathcal{J}_j(y)| \left| \frac{1}{\mathbb{E}(\Gamma_1(x) \mathcal{Ker}_1(x))} \Gamma_j(x) \mathcal{Ker}_j(x) B_j \mathbb{I}_{Ba(x, \lambda_K)}(X_j) \right. \\ &\quad \left. - \frac{1}{\mathbb{E}(\Gamma_1(x_k(x)) \mathcal{Ker}_1(x_k(x)))} \Gamma_j(x_k(x)) B_j \mathcal{Ker}_j(x_k(x)) \mathbb{I}_{Ba(x_k(x), \lambda_K)}(X_j) \right| \\ &\leq \left( \frac{C}{n^2 \lambda_K^2 \psi(\lambda_K)} \sup_{x \in \mathcal{C}_{\mathcal{F}}} \sum_{j=1}^n |\Gamma_j(x) B_j \mathbb{I}_{Ba(x, \lambda_K)}(X_j)| \times \right. \\ &\quad \left. |\mathcal{Ker}_j(x) - \mathcal{Ker}_j(x_k(x)) \mathbb{I}_{Ba(x_k(x), \lambda_K)}(X_j)| \right) \\ &\quad + \left( \frac{C}{n^2 \lambda_K^2 \psi(\lambda_K)} \sup_{x \in \mathcal{C}_{\mathcal{F}}} \sum_{j=1}^n \mathcal{Ker}_j(x_k(x)) B_j \mathbb{I}_{Ba(x_k(x), \lambda_K)}(X_j) \times \right. \\ &\quad \left. |\Gamma_j(x) \mathbb{I}_{Ba(x, \lambda_K)}(X_j) - \Gamma_j(x_k(x))| \right) \\ &:= F_1 + F_2. \end{aligned}$$

By using the fact that the  $\mathcal{Ker}(\cdot)$  satisfies the Lipschitz condition, we obtain via (15)

$$\begin{aligned} &\left| \Gamma_j(x) \mathbb{I}_{Ba(x, \lambda_K)}(X_j) |\mathcal{Ker}_j(x) - \mathcal{Ker}_j(x_k(x)) \mathbb{I}_{Ba(x_k(x), \lambda_K)}(X_j)| \right| \\ &\leq n C \lambda_K^2 \left( \frac{r_n}{\lambda_K} \mathbb{I}_{Ba(x, \lambda_K) \cap Ba(x_k(x), \lambda_K)}(X_j) \right. \\ &\quad \left. + \mathbb{I}_{Ba(x, \lambda_K) \cap \overline{Ba(x_k(x), \lambda_K)}}(X_j) \right), \end{aligned}$$

from which we infer that

$$\begin{aligned} F_1 &\leq \frac{C r_n}{n \lambda_K \psi(\lambda_K)} \sup_{x \in \mathcal{C}_{\mathcal{F}}} \sum_{j=1}^n \mathbb{I}_{Ba(x, \lambda_K) \cap Ba(x_k(x), \lambda_K)}(X_j) B_j \\ &\quad + \frac{C}{n \psi(\lambda_K)} \sup_{x \in \mathcal{C}_{\mathcal{F}}} \sum_{j=1}^n \mathbb{I}_{Ba(x, \lambda_K) \cap \overline{Ba(x_k(x), \lambda_K)}}(X_j) B_j. \end{aligned}$$

Concerning  $F_2$  we put, for  $l = 0, 1$  and  $k = 1, 2$ ,

$$\begin{aligned} T^{k,l} &= \mathbb{I}_{Ba(x_k(x), \lambda_K) \cap Ba(x, \lambda_K)}(X_j) \\ &\quad \times \left| \left( \sum_{i=1}^n \alpha_i^k(x) \mathcal{Ker}_i(x) B_i \right) \alpha_j^l(x) - \left( \sum_{i=1}^n \alpha_i^k(x_k(x)) \mathcal{Ker}_i(x_k(x)) B_i \right) \alpha_j^l(x_k(x)) \right|. \end{aligned}$$

Therefore, we obtain

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

where

$$T_1^{k,l} = \mathbb{I}_{\mathcal{B}a(x_{k(x)}, \lambda_K) \cap \mathcal{B}a(x, \lambda_K)}(X_j) \left( \sum_{i=1}^n | \alpha_i^k(x) | \mathcal{K}er_i(x) B_i \times | \alpha_j^l(x) - \alpha_j^l(x_{k(x)}) | \right),$$

and

$$T_2^{k,l} = \mathbb{I}_{\mathcal{B}a(x_{k(x)}, \lambda_K) \cap \mathcal{B}a(x, \lambda_K)}(X_j) ( | \alpha_j^l(x_{k(x)}) | B_i \times \left| \sum_{i=1}^n (\alpha_i^k(x) \mathcal{K}er_i(x) - \alpha_i^k(x_{k(x)}) \mathcal{K}er_i(x_{k(x)})) \right| ),$$

Making use of the condition (H4), for  $l = 1$ , we have

$$\mathbb{I}_{\mathcal{B}a(x_{k(x)}, \lambda_K) \cap \mathcal{B}a(x, \lambda_K)}(X_j) | \alpha_j(x) - \alpha_j(x_{k(x)}) | \leq Cr_n \mathbb{I}_{\mathcal{B}a(x_{k(x)}, \lambda_K) \cap \mathcal{B}a(x, \lambda_K)}(X_j).$$

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

$$T_1^{k,l} = 0, \tag{16}$$

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

$$T_1^{k,l} \leq nCr_n \lambda_K \mathbb{I}_{\mathcal{B}a(x_{k(x)}, \lambda_K) \cap \mathcal{B}a(x, \lambda_K)}(X_j).$$

Now, for  $T_2^{k,l}$ , we have

$$\begin{aligned} T_2^{k,l} &\leq \mathbb{I}_{\mathcal{B}a(x_{k(x)}, \lambda_K) \cap \mathcal{B}a(x, \lambda_K)}(X_j) \\ &\times \left( \sum_{i=1}^n | \alpha_j^l(x_{k(x)}) | \mathcal{K}er_i(x) B_i \times | \alpha_i^k(x) - \alpha_i^k(x_{k(x)}) | \right) \\ &+ \mathbb{I}_{\mathcal{B}a(x_{k(x)}, \lambda_K) \cap \mathcal{B}a(x, \lambda_K)}(X_j) \\ &\times \left( \sum_{i=1}^n | \alpha_j^l(x_{k(x)}) | | \alpha_i^k(x_{k(x)}) | B_i \times | \mathcal{K}er_i(x) - \mathcal{K}er_i(x_{k(x)}) | \right). \end{aligned}$$

We have

$$\mathbb{I}_{\mathcal{B}a(x_{k(x)}, \lambda_K) \cap \mathcal{B}a(x, \lambda_K)}(X_j) | \alpha_i^2(x) - \alpha_i^2(x_{k(x)}) | \leq Cr_n \lambda_K \mathbb{I}_{\mathcal{B}a(x_{k(x)}, \lambda_K) \cap \mathcal{B}a(x, \lambda_K)}(X_j).$$

This implies that, for  $k = 1, 2$ ,

$$\mathbb{I}_{\mathcal{B}a(x_{k(x)}, \lambda_K) \cap \mathcal{B}a(x, \lambda_K)}(X_j) | \alpha_i^k(x) - \alpha_i^k(x_{k(x)}) | \leq Cr_n \lambda_K^{k-1} \mathbb{I}_{\mathcal{B}a(x_{k(x)}, \lambda_K) \cap \mathcal{B}a(x, \lambda_K)}(X_j).$$

Thus, for  $(l, k) = (0, 2)$ , we have

$$T_2^{k,l} \leq nCr_n \lambda_K \mathbb{I}_{\mathcal{B}a(x_{k(x)}, \lambda_K) \cap \mathcal{B}a(x, \lambda_K)}(X_j), \tag{17}$$

and for  $l = k = 1$

$$T_2^{k,l} \leq nCr_n \lambda_K \mathbb{I}_{\mathcal{B}a(x_{k(x)}, \lambda_K) \cap \mathcal{B}a(x, \lambda_K)}(X_j). \tag{18}$$

Combining (16) and (17), we find that

$$\begin{aligned} \mathbb{I}_{\mathcal{B}a(x_{k(x)}, \lambda_K) \cap \mathcal{B}a(x, \lambda_K)}(X_j) &\left| \left( \sum_{i=1}^n \alpha_i^2(x) \mathcal{K}er_i(x) B_i - \alpha_i^2(x_{k(x)}) \mathcal{K}er_i(x_{k(x)}) B_i \right) \right| \\ &\leq nCr_n \lambda_K \mathbb{I}_{\mathcal{B}a(x_{k(x)}, \lambda_K) \cap \mathcal{B}a(x, \lambda_K)}(X_j), \end{aligned}$$

and

$$\begin{aligned} &\mathbb{I}_{\mathcal{B}a(x_{k(x)}, \lambda_K) \cap \mathcal{B}a(x, \lambda_K)}(X_j) \\ &\times \left| \left( \sum_{i=1}^n \alpha_i(x) \mathcal{K}er_i(x) B_i \right) \alpha_j(x) - \left( \sum_{i=1}^n \alpha_i(x_{k(x)}) \mathcal{K}er_i(x_{k(x)}) B_i \right) \alpha_j(x_{k(x)}) \right| \\ &\leq nCr_n \lambda_K \mathbb{I}_{\mathcal{B}a(x_{k(x)}, \lambda_K) \cap \mathcal{B}a(x, \lambda_K)}(X_j). \end{aligned}$$



Thus, we obtain

$$F_2 \leq \frac{Cr_n}{n\lambda_K\psi(\lambda_K)} \sup_{x \in \mathcal{C}_{\mathcal{F}}} \sum_{j=1}^n \mathbb{I}_{\mathcal{B}a(x, \lambda_K) \cap \mathcal{B}a(x_{k(x)}, \lambda_K)}(X_j) B_j + \frac{C}{n\psi(\lambda_K)} \sup_{x \in \mathcal{C}_{\mathcal{F}}} \sum_{j=1}^n \mathbb{I}_{\mathcal{B}a(x, \lambda_K) \cap \overline{\mathcal{B}a(x_{k(x)}, \lambda_K)}}(X_j) B_j.$$

Consequently, we infer that

$$Q_2 \leq C \sup_{x \in \mathcal{C}_{\mathcal{F}}} (Q_{2.1} + Q_{2.2} + Q_{2.3}),$$

where

$$\begin{aligned} Q_{2.1} &= \frac{C}{n\psi(\lambda_K)} \sum_{j=1}^n \mathbb{I}_{\mathcal{B}a(x_{k(x)}, \lambda_K) \cap \overline{\mathcal{B}a(x, \lambda_K)}}(X_j) B_j, \\ Q_{2.2} &= \frac{Cr_n}{n\lambda_K\psi(\lambda_K)} \sum_{j=1}^n \mathbb{I}_{\mathcal{B}a(x, \lambda_K) \cap \mathcal{B}a(x_{k(x)}, \lambda_K)}(X_j) B_j, \\ Q_{2.3} &= \frac{C}{n\psi(\lambda_K)} \sum_{j=1}^n \mathbb{I}_{\mathcal{B}a(x, \lambda_K) \cap \overline{\mathcal{B}a(x_{k(x)}, \lambda_K)}}(X_j) B_j. \end{aligned}$$

Now, we apply an exponential inequality for the difference martingale random variables of  $Z_j$  that are defined by

$$Z_j = \begin{cases} \frac{1}{\psi(\lambda_K)} [\mathbb{I}_{\mathcal{B}a(x_{k(x)}, \lambda_K) \cap \overline{\mathcal{B}a(x, \lambda_K)}}(X_j)] B_j & \text{for } Q_{1.1}, \\ \frac{r_n}{\lambda_K\psi(\lambda_K)} [\mathbb{I}_{\mathcal{B}a(x, \lambda_K) \cap \mathcal{B}a(x_{k(x)}, \lambda_K)}(X_j)] B_j & \text{for } Q_{1.2}, \\ \frac{1}{\psi_K} [\mathbb{I}_{\mathcal{B}a(x, \lambda_K) \cap \overline{\mathcal{B}a(x_{k(x)}, \lambda_K)}}(X_j)] B_j & \text{for } Q_{1.3}. \end{cases}$$

Keep in mind that

$$\begin{aligned} Z_j &= O\left(\frac{1}{\psi(\lambda_K)}\right), \\ \mathbb{E}[Z_j | \mathfrak{F}_{j-1}] &= O\left(\frac{r_n}{\psi(\lambda_K)}\right) \\ \mathbb{E}[Z_j^2 | \mathfrak{F}_{j-1}] &= O\left(\frac{r_n}{\psi(\lambda_K)^2}\right). \end{aligned}$$

Therefore, we infer

$$Q_{2.1} = O\left(\frac{r_n}{\psi(\lambda_K)}\right) + O_{a.co.}\left(\sqrt{\frac{r_n \ln n}{n\psi(\lambda_K)^2}}\right),$$

and

$$Q_{2.2} = O_{a.co.}\left(\sqrt{\frac{\ln d_n}{n\psi(\lambda_K)}}\right).$$

We deduce

$$Q_2 = O_{a.co.}\left(\sqrt{\frac{\ln d_n}{n\psi(\lambda_K)}}\right),$$

and  $\forall \varepsilon > 0$ , we obtain

$$\begin{aligned} & \mathbb{P}\left(Q_1 > \varepsilon \sqrt{\frac{\ln d_n}{n\psi(\lambda_K)}}\right) \\ &= \mathbb{P}\left(\max_{k \in 1, \dots, d_n} \left| \widehat{CDF}_N^{x_k(x)}(y) - \bar{CDF}_N^{x_k(x)}(y) \right| > \varepsilon\right) \\ &\leq d_n \max_{k \in 1, \dots, d_n} \mathbb{P}\left(\widehat{CDF}_N^{x_k(x)}(y) - \bar{CDF}_N^{x_k(x)}(y) > \varepsilon \sqrt{\frac{\ln d_n}{n\psi(\lambda_K)}}\right). \end{aligned}$$

Let

$$\widehat{CDF}_N^{x_k(x)}(y) - \bar{CDF}_N^{x_k(x)}(y) = \frac{1}{n\mathbb{E}(\Gamma_1(x)\mathcal{Ker}_1(x))} \sum_{j=1}^n \mathcal{S}_j,$$

where

$$\mathcal{S}_j = \Gamma_j(x_{k(x)})\mathcal{Ker}_j(x_{k(x)})\mathcal{J}_j(y)B_j - \mathbb{E}(\Gamma_j(x_{k(x)})\mathcal{Ker}_j(x_{k(x)})\mathcal{J}_j(y)B_j | \mathfrak{F}_{j-1}).$$

We readily obtain

$$\mathbb{E}(\mathcal{S}_j^2 | \mathfrak{F}_{j-1}) \leq 2C' n^2 \lambda_K^4 \psi_j(\lambda_K).$$

Now, for all  $\varepsilon > 0$ , we infer that

$$\begin{aligned} & \mathbb{P}\left(\left| \widehat{CDF}_N^{x_k(x)}(y) - \bar{CDF}_N^{x_k(x)}(y) \right| > \varepsilon \sqrt{\frac{\ln d_n}{n\psi(\lambda_K)}}\right) \\ &\leq \mathbb{P}\left(\left| \frac{1}{n\mathbb{E}(\Gamma_1\mathcal{Ker}_1)} \sum_{j=1}^n \mathcal{S}_j \right| > \varepsilon \sqrt{\frac{\ln d_n}{n\psi(\lambda_K)}}\right) \\ &\leq 2 \exp^{-C_0 \varepsilon^2 \ln d_n}. \end{aligned}$$

Choosing  $\varepsilon$  for which  $C_0 \varepsilon^2 = \varsigma$ , we then infer

$$d_n \max_{k \in 1, \dots, d_n} \mathbb{P}\left(\left| \widehat{CDF}_N^{x_k(x)}(y) - \bar{CDF}_N^{x_k(x)}(y) \right| > \varepsilon \sqrt{\frac{\ln d_n}{n\psi(\lambda_K)}}\right) \leq C' d_n^{1-\varsigma}.$$

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

$$Q_1 = O_{a.co.} \left( \sqrt{\frac{\ln d_n}{n\psi(\lambda_K)}} \right).$$

For  $Q_3$ , we have

$$Q_3 \leq \mathbb{E} \left( \sup_{x \in \mathcal{C}_{\mathcal{F}}} \left| \widehat{CDF}_N^{x_k(x)}(y) - \bar{CDF}_N^{x_k(x)}(y) \right| \mid \mathfrak{F}_{j-1} \right).$$

We follow the proof for  $Q_1$  to obtain

$$Q_3 = O_{a.co.} \left( \sqrt{\frac{\ln d_n}{n\psi(\lambda_K)}} \right),$$

which achieves the demonstration of Lemma 2.  $\square$

**Proof of Lemma 3.** Obviously, the first statement is deduced by taking  $\mathcal{J}_j = 1$  in Lemma 2, while for the second result, we use

$$\inf_{x \in \mathcal{C}_{\mathcal{F}}} | \widehat{CDF}_D(x) | \leq \frac{\inf_{x \in \mathcal{C}_{\mathcal{F}}} P(x)}{2},$$

meaning that there exists  $x \in \mathcal{C}_{\mathcal{F}}$  such that

$$\begin{aligned} P(x) - \widehat{CDF}_D(x) &\geq \frac{\inf_{x \in \mathcal{C}_{\mathcal{F}}} P(x)}{2} \\ \implies \sup_{x \in \mathcal{C}_{\mathcal{F}}} | P(x) - \widehat{CDF}_D(x) | &\geq \frac{\inf_{x \in \mathcal{C}_{\mathcal{F}}} P(x)}{2}. \end{aligned}$$

So, we have

$$\begin{aligned} &\mathbb{P} \left( \inf_{x \in \mathcal{C}_{\mathcal{F}}} | \widehat{CDF}_D(x) | \leq \frac{1}{2} \right) \\ &\leq \mathbb{P} \left( \sup_{x \in \mathcal{C}_{\mathcal{F}}} | P(x) - \widehat{CDF}_D(x) | \geq \frac{\inf_{x \in \mathcal{C}_{\mathcal{F}}} P(x)}{2} \right). \\ &\leq \mathbb{P} \left( \sup_{x \in \mathcal{C}_{\mathcal{F}}} | \mathbb{E}(\widehat{CDF}_D(x)) - \widehat{CDF}_D(x) | \geq \frac{\inf_{x \in \mathcal{C}_{\mathcal{F}}} P(x)}{2} \right). \end{aligned}$$

Consequently,

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

Hence, the proof is complete.  $\square$

**Proof of Lemma 4.** Notice that

$$\sup_{x \in \mathcal{C}_{\mathcal{F}}} \|A_n(x)\| \leq \sup_{x \in \mathcal{C}_{\mathcal{F}}} \|A_n(x) - \bar{A}_n(x)\| + \sup_{x \in \mathcal{C}_{\mathcal{F}}} \|\bar{A}_n(x)\|,$$

where

$$\bar{A}_n(x) := \frac{1}{n\psi(\lambda_K)} \sum_{j=1}^n \mathbb{E} \left( \Phi(\vec{\delta}_0) \left( \frac{1}{\lambda_K^{-1} \alpha_j} \right) \mathcal{K}er_j(x) B_j | \mathfrak{F}_{j-1} \right).$$

So, all that remains is to prove that

$$\sup_{x \in \mathcal{C}_{\mathcal{F}}} \|A_n(x) - \bar{A}_n(x)\| = O_{a.co.} \left( \left( \frac{\ln d_n}{n\psi(\lambda_K)} \right)^{1/2} \right) \tag{19}$$

and

$$\sup_{x \in \mathcal{C}_{\mathcal{F}}} \|\bar{A}_n(x)\| = O \left( \lambda_K^{\min(k_1, k_2)} \right). \tag{20}$$

The proof of (19) follows the same lines as the Lemma 2, while (20) is based on the same idea as in Lemma 2 using the fact that

$$\mathbb{I}_{\mathcal{B}a(x, \lambda_K)}(X_1) |CDF(CQF_p(x)|x) - CDF(\beta_1 + \beta_2 \alpha_1 | X_1)| \leq C \left( \lambda_K^{\min(k_1, k_2)} \right).$$

This completes the proof of Lemma 4.  $\square$

**Proof of Lemma 5.** We prove that

$$\sup_{x \in \mathcal{C}_{\mathcal{F}}} \sup_{\|\delta\| \leq M} \|V_n(\vec{\delta}, x) - A_n(x) - (\bar{V}_n(\vec{\delta}, x) - \bar{A}_n(x))\| = O_{a.co.} \left( \sqrt{\frac{\log n}{n\psi(\lambda_K)}} \right), \tag{21}$$

and

$$\sup_{x \in \mathcal{C}_{\mathcal{F}}} \sup_{\|\delta\| \leq M} \|(\bar{V}_n(\vec{\delta}, x) - \bar{A}_n(x)) + cdf(CQF_p(x)|x)D\delta\| = O\left(\lambda_K^{\min(k_1, k_2)}\right), \tag{22}$$

where

$$\bar{V}_n(\vec{\delta}, x) := \frac{1}{n\psi(\lambda_K)} \sum_{j=1}^n \mathbb{E} \left( \Phi(\vec{\delta}) \left( \frac{1}{\lambda_K^{-1} \alpha_j} \right) \mathcal{K}er_j(x) B_j | \mathfrak{F}_{j-1} \right).$$

In order to show (21), we keep the definition of  $k(x)$  as in Lemma 2, we use the compactness of the ball  $B(0, M)$  in  $\mathbb{R}^2$ , and we write

$$B(0, M) \subset \bigcup_{j=1}^{d_n} B(\delta_j, l_n), \quad \delta_j = \begin{pmatrix} c_j \\ d_j \end{pmatrix} \text{ and } l_n = d_n^{-1} = 1/\sqrt{n}.$$

Then, we take  $j(\delta) = \arg \min_j |\delta - \delta_j|$ . Similarly to Lemma 2 we write

$$\begin{aligned} & \sup_{x \in \mathcal{C}_{\mathcal{F}}} \sup_{\|\delta\| \leq M} \|V_n(\vec{\delta}, x) - A_n(x) - (\bar{V}_n(\vec{\delta}, x) - \bar{A}_n(x))\| \\ & \leq \underbrace{\sup_{x \in \mathcal{C}_{\mathcal{F}}} \sup_{\|\delta\| \leq M} \|V_n(\vec{\delta}, x) - A_n(x) - (V_n(\vec{\delta}, x_{k(x)}) - A_n(x_{k(x)}))\|}_{T_1} \\ & \quad + \underbrace{\sup_{x \in \mathcal{C}_{\mathcal{F}}} \sup_{\|\delta\| \leq M} \|V_n(\vec{\delta}, x_{k(x)}) - V_n(\vec{\delta}_{j(\delta)}, x_{k(x)})\|}_{T_2} \\ & \quad + \underbrace{\sup_{x \in \mathcal{C}_{\mathcal{F}}} \sup_{\|\delta\| \leq M} \|V_n(\vec{\delta}_{j(\delta)}, x_{k(x)}) - A_n(x_{k(x)}) - (\bar{V}_n(\vec{\delta}_{j(\delta)}, x_{k(x)}) - \bar{A}_n(x_{k(x)}))\|}_{T_3} \\ & \quad + \underbrace{\sup_{x \in \mathcal{C}_{\mathcal{F}}} \sup_{\|\delta\| \leq M} \|\bar{V}_n(\vec{\delta}_{j(\delta)}, x_{k(x)}) - \bar{V}_n(\vec{\delta}, x_{k(x)})\|}_{T_4} \\ & \quad + \underbrace{\sup_{x \in \mathcal{C}_{\mathcal{F}}} \sup_{\|\delta\| \leq M} \|\bar{V}_n(\vec{\delta}, x_{k(x)}) - \bar{A}_n(x_{k(x)}) - (\bar{V}_n(\vec{\delta}, x) - \bar{A}_n(x))\|}_{T_5}. \end{aligned}$$

We treat  $T_1$  and  $T_2$  as  $Q_2$  in Lemma 2. Meanwhile,  $T_4$  and  $T_5$  are evaluated as in  $Q_3$ . Finally, we use the idea of  $Q_1$  to evaluate  $T_3$ . The statement (22) is a consequence of

$$\lambda_K^{-a} \mathbb{E}[\alpha_i^{-a} \mathcal{K}er_i^b] = \psi(\lambda_K) \left( \mathcal{K}er^b(1) - \int_{-1}^1 (u^a \mathcal{K}er^b(u))' \Psi(u) du \right) + o(\psi(\lambda_K)).$$

Hence, the proof is complete.  $\square$

**Proof of Lemma 6.** For a fixed  $x$ , we put all  $j = 1, \dots, n$ ,  $(\Gamma_1(x), \mathcal{K}er_1(x)) = (\Gamma_1, \mathcal{K}er_1)$ . Next, let us denote

$$\eta_{n,j} = \frac{\sqrt{n\psi(\lambda_K)}}{n\mathbb{E}(\Gamma_1 \mathcal{K}er_1)} (\mathcal{J}_j - CDF(y | x)) \Gamma_j \mathcal{K}er_j B_j,$$

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

$$\sqrt{n\psi(\lambda_K)}Q_n(\mathbf{y} | x) = \sum_{j=1}^n \xi_{n,j}. \tag{23}$$

Therefore, to demonstrate Lemma 6, we have to prove the following:

$$\sum_{j=1}^n \mathbb{E}(\xi_{n,j}^2|\mathfrak{F}_{j-1}) \xrightarrow{\mathbb{P}} V_{JK}(\mathbf{y} | x), \tag{24}$$

and

$$\forall \varepsilon > 0 \quad n\mathbb{E}(\xi_{n,j}^2 \mathbb{I}_{\{|\xi_{n,j}^2| > \varepsilon\}}) = o(1). \tag{25}$$

For (24), we have

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

Thus, it suffices to show that

$$\lim_{n \rightarrow \infty} \sum_{j=1}^n (\mathbb{E}(\eta_{n,j}|\mathfrak{F}_{j-1}))^2 = 0 \text{ in probability,} \tag{26}$$

and

$$\lim_{n \rightarrow \infty} \sum_{j=1}^n \mathbb{E}(\eta_{n,j}^2|\mathfrak{F}_{j-1}) = V_{JK}(\mathbf{y} | x) \text{ in probability.} \tag{27}$$

For (26), we use Lemma 5 of [72] to conclude

$$\begin{aligned} |\mathbb{E}(\eta_{n,j}|\mathfrak{F}_{j-1})| &= \frac{\sqrt{n\psi(\lambda_K)}}{n\mathbb{E}(\Gamma_1 \mathcal{K}er_1)} |\mathbb{E}(B_j(\mathcal{J}_j - CDF(\mathbf{y} | x))\Gamma_j \mathcal{K}er_j|\mathfrak{F}_{j-1})| \\ &\leq C\sqrt{n\psi(\lambda_K)}(\lambda_K^{b_1} + \lambda_j^{b_2}) \frac{1}{n\psi(\lambda_K)} (P(x) + o(1)). \end{aligned}$$

Thus, we have

$$\sum_{j=1}^n (\mathbb{E}(\eta_{n,j}|\mathfrak{F}_{j-1}))^2 = O_{a.co.}(n\psi(\lambda_K)(\lambda_K^{b_1} + \lambda_j^{b_2})^2).$$

For Equation (27), we have

$$\begin{aligned} &\sum_{j=1}^n \mathbb{E}(\eta_{n,j}^2|\mathfrak{F}_{j-1}) \\ &= \frac{\psi(\lambda_K)}{n(\mathbb{E}(\Gamma_1 \mathcal{K}er_1))^2} \sum_{j=1}^n \mathbb{E}(\Gamma_j^2 \mathcal{K}er_j^2 B_j(\mathcal{J}_j - CDF(\mathbf{y} | x))^2|\mathfrak{F}_{j-1}) \\ &= \frac{\psi(\lambda_K)}{n(\mathbb{E}(\Gamma_1 \mathcal{K}er_1))^2} \sum_{j=1}^n \mathbb{E}(\Gamma_j^2 \mathcal{K}er_j^2 B_j \mathbb{E}[(\mathcal{J}_j - CDF(\mathbf{y} | x))^2|X_j]|\mathfrak{F}_{j-1}). \end{aligned}$$

Since

$$\mathbb{E}[(\mathcal{J}_j - CDF(\mathbf{y} | x))^2|X_j] = \text{Var}[\mathcal{J}_j|X_j] + [\mathbb{E}(\mathcal{J}_j|X_j) - CDF(\mathbf{y} | x)]^2,$$

we have

$$\text{Var}[\mathcal{J}_j|X_j] = \mathbb{E}(\mathcal{J}_j^2|X_j) - (\mathbb{E}(\mathcal{J}_j|X_j))^2. \tag{28}$$

We highlight that

$$\begin{aligned} \mathbb{E}[\mathcal{J}_j|X_j] &= \int_{\mathbb{R}} \mathcal{J}^{(1)}(t)[CDF(\mathbf{y} - \lambda_j t|x) - CDF(\mathbf{y} | x)]dt + CDF(\mathbf{y} | x) \\ &= CDF(\mathbf{y} | x), \end{aligned} \tag{29}$$

and

$$\begin{aligned} \mathbb{E}[\mathcal{J}_j^2|X_j] &= \int_{\mathbb{R}} \mathcal{J}^2\left(\frac{y-z}{\lambda_j}\right)cdf(z|x)dz \\ &= \int_{\mathbb{R}} 2J(t)\mathcal{J}^{(1)}(t)[CDF(y-\lambda_j t|x) - CDF(y|x)]dt \\ &\quad + \int_{\mathbb{R}} 2\mathcal{J}^{(1)}(t)\mathcal{J}(t)CDF(y|x)dt. \end{aligned}$$

Using the fact

$$\int_{\mathbb{R}} \mathcal{J}^{(1)}(t)\mathcal{J}(t)CDF(y|x)dt = CDF(y|x),$$

we readily obtain, as  $n \rightarrow \infty$ ,

$$\mathbb{E}[\mathcal{J}_j^2|X_j] \rightarrow CDF(y|x). \tag{30}$$

Combining (29) with (30), we arrive at

$$\text{Var}[\mathcal{J}_j|X_j] = CDF(y|x)(1 - CDF(y|x)). \tag{31}$$

Concerning  $\beta_{n_2}$ , we obtain from (29) that, as  $n \rightarrow \infty$ ,

$$\beta_{n_2} \rightarrow 0.$$

Thereafter, we obtain

$$\begin{aligned} \Sigma_{j=1}^n \mathbb{E}(\eta_{n,j}^2|\mathfrak{F}_{j-1}) &= \frac{\psi(\lambda_K)}{n(\mathbb{E}(\Gamma_1 \mathcal{K}er_1))^2} \Sigma_{j=1}^n \mathbb{E}(\Gamma_j^2 K_j^2 B_j \beta_{n_1}|\mathfrak{F}_{j-1}) \\ &= \frac{\psi(\lambda_K)}{n(\mathbb{E}(\Gamma_1 \mathcal{K}er_1))^2} \Sigma_{j=1}^n \mathbb{E}(\Gamma_j^2 K_j^2 \mathbb{E}(B_j|X_j)\beta_{n_1}|\mathfrak{F}_{j-1}) \\ &= \frac{\psi(\lambda_K)}{n(\mathbb{E}(\Gamma_1 \mathcal{K}er_1))^2} \Sigma_{j=1}^n \mathbb{E}((P(x) + o(1))\Gamma_j^2 K_j^2 \beta_{n_1}|\mathfrak{F}_{j-1}). \end{aligned}$$

The use of Equation (31) allow us to obtain, as  $n \rightarrow \infty$ ,

$$\begin{aligned} \Sigma_{j=1}^n \mathbb{E}(\eta_{n,j}^2|\mathfrak{F}_{j-1}) &\rightarrow \frac{M_2}{M_1^2} P(x)(1 - CDF(y|x))CDF(y|x) \\ &= V_{JK}(y|x); \end{aligned}$$

this last part completes the proof of our first claim. Concerning (25), we have

$$n\mathbb{E}(\xi_{n,j}^2) \leq 4n\mathbb{E}(\eta_{n,j}^2 \mathbb{I}_{\{|\eta_{n,j}| > \frac{\varepsilon}{2}\}}).$$

By Markov–Hölder’s inequalities, we obtain  $\varepsilon > 0$ :

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

Putting  $a = 1 + \frac{\delta'}{2}$  for  $\delta' > 0$  and by recalling that the function

$$G_{2+\delta'} = \mathbb{E}(|\mathcal{J}_j - CDF(y|x)|^{2+\delta'}|X_j)$$

is continuous, we have

$$\begin{aligned}
4n\mathbb{E}(\eta_{n,j}^2\mathbb{I}_{\{|\eta_{n,j}|>\frac{\varepsilon}{2}\}}) &\leq C\left(\frac{\psi(\lambda_K)}{n}\right)^{\frac{2+\delta'}{2}}\frac{n}{(\mathbb{E}(\Gamma_1\mathcal{K}er_1))^{2+\delta'}} \\
&\quad \times \mathbb{E}(|(\mathcal{J}_j - CDF(y|x))|\Gamma_j\mathcal{K}er_j B_j|^{2+\delta'}) \\
&\leq C\left(\frac{\psi(\lambda_K)}{n}\right)^{\frac{2+\delta'}{2}}\frac{n}{(\mathbb{E}(\Gamma_1\mathcal{K}er_1))^{2+\delta'}} \\
&\quad \times \mathbb{E}(P(x)|\Gamma_j\mathcal{K}er_j|^{2+\delta'}[\mathbb{E}(|\mathcal{J}_j - CDF(y|x)|^{2+\delta'}|X_j)]) \\
&\leq C(P(x) + o(1))\left(\frac{\psi(\lambda_K)}{n}\right)^{\frac{2+\delta'}{2}}\frac{n}{(\mathbb{E}(\Gamma_1\mathcal{K}er_1))^{2+\delta'}} \\
&\quad \times \mathbb{E}(|\Gamma_j\mathcal{K}er_j|^{2+\delta'})G_{2+\delta'} \\
&= O\left((n\psi(\lambda_K))^{\frac{-\delta'}{2}}\right) \rightarrow 0, \text{ as } n \rightarrow \infty.
\end{aligned}$$

Hence, the proof is complete.  $\square$

**Author Contributions:** Formal analysis, F.A.; Validation, O.L. and B.M.; Writing—review and editing, S.B. and A.L. The authors contributed approximately equally to this work. All authors have read and agreed to the final version of the manuscript.

**Funding:** This research project was funded by (1) Princess Nourah bint Abdulrahman University Researchers Supporting Project Number (PNURSP2023R358), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia, and (2) The Deanship of Scientific Research at King Khalid University through the Research Groups Program under grant number R.G.P. 1/177/44.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The authors thank the Editors, the Associate Editor, the Guest Editor Y. Slaoui, and the six anonymous reviewers for their valuable comments and suggestions, which substantially improved the quality of an earlier version of this paper. The authors thank and extend their appreciation to the funders of this project: Princess Nourah bint Abdulrahman University Researchers Supporting Project Number (PNURSP2023R358), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia; The Deanship of Scientific Research at King Khalid University through the Research Groups Program under grant number R.G.P. 1/177/44.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Ramsay, J.O.; Silverman, B.W. *Applied Functional Data Analysis: Methods and Case Studies*; Springer Series in Statistics; Springer: New York, NY, USA, 2002; p. x+190 [[CrossRef](#)]
2. Bosq, D. *Linear Processes in Function Spaces: Theory and Applications*; Lecture Notes in Statistics; Springer: New York, NY, USA, 2000; Volume 149, p. xiv+283. [[CrossRef](#)]
3. Aneiros, G.; Bongiorno, E.G.; Cao, R.; Vieu, P. (Eds.) In *Functional Statistics and Related Fields, Proceedings of the 4th International Workshop on Functional and Operatorial Statistics (IWFOs 2017), A Coruña, Spain, 15–17 June 2017*; Springer: Cham, Switzerland, 2017; p. xxiv+288.
4. Ferraty, F.; Laksaci, A.; Vieu, P. Estimating some characteristics of the conditional distribution in nonparametric functional models. *Stat. Inference Stoch. Process.* **2006**, *9*, 47–76. [[CrossRef](#)]
5. Bongiorno, E.G.; Goia, A.; Salinelli, E.; Vieu, P. (Eds.) In *Contributions in Infinite-Dimensional Statistics and Related Topics, Proceedings of the 3rd International Workshop on Functional and Operatorial Statistics (IWFOs'2014), Stresa, Italy, 19–21 June 2014*; Società Editrice Esculapio: Bologna, Italy, 2014; p. xviii+280.
6. Ferraty, F.; Vieu, P. *Nonparametric Functional Data Analysis: Theory and Practice*; Springer Series in Statistics; Springer: New York, NY, USA, 2006; p. xx+258.

7. Litimein, O.; Laksaci, A.; Mechab, B.; Bouzebda, S. Local linear estimate of the functional expectile regression. *Stat. Probab. Lett.* **2023**, *192*, 109682. [\[CrossRef\]](#)
8. Bouzebda, S.; Laksaci, A.; Mohammedi, M. Single index regression model for functional quasi-associated time series data. *Revstat* **2022**, *20*, 605–631.
9. Bouzebda, S.; Nezzal, A. Uniform consistency and uniform in number of neighbors consistency for nonparametric regression estimates and conditional  $U$ -statistics involving functional data. *Jpn. J. Stat. Data Sci.* **2022**, *5*, 431–533. [\[CrossRef\]](#)
10. Bouzebda, S.; Chaouch, M. Uniform limit theorems for a class of conditional  $Z$ -estimators when covariates are functions. *J. Multivar. Anal.* **2022**, *189*, 104872. [\[CrossRef\]](#)
11. Almanjahie, I.M.; Bouzebda, S.; Kaid, Z.; Laksaci, A. Nonparametric estimation of expectile regression in functional dependent data. *J. Nonparametr. Stat.* **2022**, *34*, 250–281. [\[CrossRef\]](#)
12. Almanjahie, I.M.; Bouzebda, S.; Chikr Elmezouar, Z.; Laksaci, A. The functional  $k$ NN estimator of the conditional expectile: uniform consistency in number of neighbors. *Stat. Risk Model.* **2022**, *38*, 47–63. [\[CrossRef\]](#)
13. Mohammedi, M.; Bouzebda, S.; Laksaci, A. The consistency and asymptotic normality of the kernel type expectile regression estimator for functional data. *J. Multivar. Anal.* **2021**, *181*, 104673. [\[CrossRef\]](#)
14. Wu, C.; Ling, N.; Vieu, P.; Liang, W. Partially functional linear quantile regression model and variable selection with censoring indicators MAR. *J. Multivar. Anal.* **2023**, *197*, 105189. [\[CrossRef\]](#)
15. Koenker, R.; Bassett, G., Jr. Regression quantiles. *Econometrica* **1978**, *46*, 33–50. [\[CrossRef\]](#)
16. Chernozhukov, V.; Fernández-Val, I.; Galichon, A. Quantile and probability curves without crossing. *Econometrica* **2010**, *78*, 1093–1125. [\[CrossRef\]](#)
17. Angrist, J.; Chernozhukov, V.; Fernández-Val, I. Quantile regression under misspecification, with an application to the U.S. wage structure. *Econometrica* **2006**, *74*, 539–563. [\[CrossRef\]](#)
18. Yu, K.; Jones, M.C. Local linear quantile regression. *J. Am. Stat. Assoc.* **1998**, *93*, 228–237. [\[CrossRef\]](#)
19. Chaudhuri, P. Global nonparametric estimation of conditional quantile functions and their derivatives. *J. Multivar. Anal.* **1991**, *39*, 246–269. [\[CrossRef\]](#)
20. Koenker, R.; Ng, P.; Portnoy, S. Quantile smoothing splines. *Biometrika* **1994**, *81*, 673–680. [\[CrossRef\]](#)
21. Stone, C.J. Consistent nonparametric regression. *Ann. Stat.* **1977**, *5*, 595–645. [\[CrossRef\]](#)
22. Bouzebda, S.; Taachouche, N. On the variable bandwidth kernel estimation of conditional  $U$ -statistics at optimal rates in sup-norm. *Phys. A* **2023**, *625*, 129000. [\[CrossRef\]](#)
23. Bouzebda, S. General tests of conditional independence based on empirical processes indexed by functions. *Jpn. J. Stat. Data Sci.* **2023**, *6*, 115–177. [\[CrossRef\]](#)
24. Crambes, C.; Gannoun, A.; Henchiri, Y. Support vector machine quantile regression approach for functional data: Simulation and application studies. *J. Multivar. Anal.* **2013**, *121*, 50–68. [\[CrossRef\]](#)
25. Dabo-Niang, S.; Kaid, Z.; Laksaci, A. Asymptotic properties of the kernel estimate of spatial conditional mode when the regressor is functional. *ASTA Adv. Stat. Anal.* **2015**, *99*, 131–160. [\[CrossRef\]](#)
26. Wang, K.; Lin, L. Variable selection in semiparametric quantile modeling for longitudinal data. *Commun. Stat. Theory Methods* **2015**, *44*, 2243–2266. [\[CrossRef\]](#)
27. Samanta, M. Nonparametric estimation of conditional quantiles. *Stat. Probab. Lett.* **1989**, *7*, 407–412. [\[CrossRef\]](#)
28. Messaci, F.; Nemouchi, N.; Ouassou, I.; Rachdi, M. Local polynomial modelling of the conditional quantile for functional data. *Stat. Methods Appl.* **2015**, *24*, 597–622. [\[CrossRef\]](#)
29. Demongeot, J.; Laksaci, A.; Rachdi, M.; Rahmani, S. On the local linear modelization of the conditional distribution for functional data. *Sankhya A* **2014**, *76*, 328–355. [\[CrossRef\]](#)
30. Al-Awadhi, F.A.; Kaid, Z.; Laksaci, A.; Ouassou, I.; Rachdi, M. Functional data analysis: Local linear estimation of the  $L_1$ -conditional quantiles. *Stat. Methods Appl.* **2019**, *28*, 217–240. [\[CrossRef\]](#)
31. Fan, J.; Gijbels, I. *Local Polynomial Modelling and Its Applications*; Monographs on Statistics and Applied Probability; Chapman & Hall: London, UK, 1996; Volume 66, p. xvi+341.
32. Rachdi, M.; Laksaci, A.; Demongeot, J.; Abdali, A.; Madani, F. Theoretical and practical aspects of the quadratic error in the local linear estimation of the conditional density for functional data. *Comput. Stat. Data Anal.* **2014**, *73*, 53–68. [\[CrossRef\]](#)
33. Baíllo, A.; Grané, A. Local linear regression for functional predictor and scalar response. *J. Multivar. Anal.* **2009**, *100*, 102–111. [\[CrossRef\]](#)
34. Barrientos-Marin, J.; Ferraty, F.; Vieu, P. Locally modelled regression and functional data. *J. Nonparametr. Stat.* **2010**, *22*, 617–632. [\[CrossRef\]](#)
35. Contreras-Reyes, J.E. Information quantity evaluation of multivariate SETAR processes of order one and applications. *Stat. Pap.* **2023**. [\[CrossRef\]](#)
36. Bouzebda, S.; Keziou, A. New estimates and tests of independence in semiparametric copula models. *Kybernetika* **2010**, *46*, 178–201.
37. Leucht, A.; Neumann, M.H. Degenerate  $U$ - and  $V$ -statistics under ergodicity: Asymptotics, bootstrap and applications in statistics. *Ann. Inst. Stat. Math.* **2013**, *65*, 349–386. [\[CrossRef\]](#)
38. Leucht, A.; Neumann, M.H. Dependent wild bootstrap for degenerate  $U$ - and  $V$ -statistics. *J. Multivar. Anal.* **2013**, *117*, 257–280. [\[CrossRef\]](#)



39. Neumann, M.H. Absolute regularity and ergodicity of Poisson count processes. *Bernoulli* **2011**, *17*, 1268–1284. [[CrossRef](#)]
40. Bradley, R.C. *Introduction to Strong Mixing Conditions*; Kendrick Press: Heber City, UT, USA, 2007; Volume 3, p. xii+597.
41. Benziadi, F.; Laksaci, A.; Tebboune, F. Recursive kernel estimate of the conditional quantile for functional ergodic data. *Commun. Stat. Theory Methods* **2016**, *45*, 3097–3113. [[CrossRef](#)]
42. Chaouch, M.; Khardani, S. Randomly censored quantile regression estimation using functional stationary ergodic data. *J. Nonparametr. Stat.* **2015**, *27*, 65–87. [[CrossRef](#)]
43. Rubin, D.B. Inference and missing data. *Biometrika* **1976**, *63*, 581–592. [[CrossRef](#)]
44. Josse, J.; Reiter, J.P. Introduction to the special section on missing data. *Stat. Sci.* **2018**, *33*, 139–141. [[CrossRef](#)]
45. Müller, U.U. Estimating linear functionals in nonlinear regression with responses missing at random. *Ann. Stat.* **2009**, *37*, 2245–2277. [[CrossRef](#)]
46. Wang, D.; Chen, S.X. Empirical likelihood for estimating equations with missing values. *Ann. Stat.* **2009**, *37*, 490–517. [[CrossRef](#)]
47. Chen, X.; Hong, H.; Tarozzi, A. Semiparametric efficiency in GMM models with auxiliary data. *Ann. Stat.* **2008**, *36*, 808–843. [[CrossRef](#)]
48. Ferraty, F.; Sued, M.; Vieu, P. Mean estimation with data missing at random for functional covariables. *Statistics* **2013**, *47*, 688–706. [[CrossRef](#)]
49. Ling, N.; Liang, L.; Vieu, P. Nonparametric regression estimation for functional stationary ergodic data with missing at random. *J. Stat. Plan. Inference* **2015**, *162*, 75–87. [[CrossRef](#)]
50. Ling, N.; Cheng, L.; Vieu, P.; Ding, H. Missing responses at random in functional single index model for time series data. *Stat. Pap.* **2022**, *63*, 665–692. [[CrossRef](#)]
51. Ling, N.; Kan, R.; Vieu, P.; Meng, S. Semi-functional partially linear regression model with responses missing at random. *Metrika* **2019**, *82*, 39–70. [[CrossRef](#)]
52. Rosenblatt, M. A strong mixing condition and a central limit theorem on compact groups. *J. Math. Mech.* **1967**, *17*, 189–198. [[CrossRef](#)]
53. Rosenblatt, M. Uniform ergodicity and strong mixing. *Z. Wahrscheinlichkeitstheorie Verwandte Geb.* **1972**, *24*, 79–84. [[CrossRef](#)]
54. Bouzebda, S.; Didi, S.; El Hajj, L. Multivariate wavelet density and regression estimators for stationary and ergodic continuous time processes: Asymptotic results. *Math. Methods Stat.* **2015**, *24*, 163–199. [[CrossRef](#)]
55. Bouzebda, S.; Didi, S. Multivariate wavelet density and regression estimators for stationary and ergodic discrete time processes: Asymptotic results. *Commun. Stat. Theory Methods* **2017**, *46*, 1367–1406. [[CrossRef](#)]
56. Bouzebda, S.; Didi, S. Additive regression model for stationary and ergodic continuous time processes. *Commun. Stat. Theory Methods* **2017**, *46*, 2454–2493. [[CrossRef](#)]
57. Bouzebda, S.; Didi, S. Some asymptotic properties of kernel regression estimators of the mode for stationary and ergodic continuous time processes. *Rev. Mat. Complut.* **2021**, *34*, 811–852. [[CrossRef](#)]
58. Bouzebda, S.; Chaouch, M.; Didi Biha, S. Asymptotics for function derivatives estimators based on stationary and ergodic discrete time processes. *Ann. Inst. Stat. Math.* **2022**, *74*, 737–771. [[CrossRef](#)]
59. Didi, S.; Bouzebda, S. Wavelet Density and Regression Estimators for Continuous Time Functional Stationary and Ergodic Processes. *Mathematics* **2022**, *10*, 4356. [[CrossRef](#)]
60. Didi, S.; Alharby, A.; Bouzebda, S. Wavelet Density and Regression Estimators for Functional Stationary and Ergodic Data: Discrete Time. *Mathematics* **2022**, *10*, 3433. [[CrossRef](#)]
61. Kolmogorov, A.N.; Tihomirov, V.M.  $\varepsilon$ -entropy and  $\varepsilon$ -capacity of sets in function spaces. *Usp. Mat. Nauk* **1959**, *14*, 3–86.
62. Ferraty, F.; Laksaci, A.; Tadj, A.; Vieu, P. Rate of uniform consistency for nonparametric estimates with functional variables. *J. Stat. Plan. Inference* **2010**, *140*, 335–352. [[CrossRef](#)]
63. Kuelbs, J.; Li, W.V. Metric entropy and the small ball problem for Gaussian measures. *J. Funct. Anal.* **1993**, *116*, 133–157. [[CrossRef](#)]
64. Nicolieris, T.; Yatracos, Y.G. Rates of convergence of estimates, Kolmogorov’s entropy and the dimensionality reduction principle in regression. *Ann. Stat.* **1997**, *25*, 2493–2511. [[CrossRef](#)]
65. Hörmann, S.; Horváth, L.; Reeder, R. A functional version of the ARCH model. *Econom. Theory* **2013**, *29*, 267–288. [[CrossRef](#)]
66. van der Vaart, A.; van Zanten, H. Bayesian inference with rescaled Gaussian process priors. *Electron. J. Stat.* **2007**, *1*, 433–448. [[CrossRef](#)]
67. Shirikyan, A. Euler equations are not exactly controllable by a finite-dimensional external force. *Phys. D* **2008**, *237*, 1317–1323. [[CrossRef](#)]
68. Bouzebda, S.; Nemouchi, B. Uniform consistency and uniform in bandwidth consistency for nonparametric regression estimates and conditional  $U$ -statistics involving functional data. *J. Nonparametr. Stat.* **2020**, *32*, 452–509. [[CrossRef](#)]
69. Bouzebda, S.; Nezzal, A.; Zari, T. Uniform Consistency for Functional Conditional  $U$ -Statistics Using Delta-Sequences. *Mathematics* **2023**, *11*, 161. [[CrossRef](#)]
70. Bouzebda, S.; Soukarieh, I. Non-Parametric Conditional  $U$ -Processes for Locally Stationary Functional Random Fields under Stochastic Sampling Design. *Mathematics* **2023**, *11*, 16. [[CrossRef](#)]

71. Batenkov, D.; Friedland, O.; Yomdin, Y. Sampling, metric entropy, and dimensionality reduction. *SIAM J. Math. Anal.* **2015**, *47*, 786–796. [[CrossRef](#)]
72. Ayad, S.; Laksaci, A.; Rahmani, S.; Rouane, R. Local linear modelling of the conditional distribution function for functional ergodic data. *Math. Model. Anal.* **2022**, *27*, 360–382. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.